DTIC FILE COPY

AD-A202 790

UNITED STATES AIR FORCE

SUMMER FACULTY RESEARCH PROGRAM

1986

PROGRAM MANAGEMENT REPORT

UNIVERSAL ENERGY SYSTEMS, INC.

VOLUME 3 OF 3

PROGRAM DIRECTOR, UES — PROGRAM MANAGER, AFOSR

SIDNEY G. ZEBERAL — MAJOR RICHARD EDELA

PROGRAM ADMINISTRATOR, UES

SUSAN K. [illegible]

DTIC ELECTE JAN 23 1989 S D

SUBMITTED TO

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

BOLLING AIR FORCE BASE

WASHINGTON, DC

DECEMBER 1986

BEST AVAILABLE COPY

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

89 1 23 004

SECURITY CLASSIFICATION OF THIS PAGE

AD-A202 790

(1)

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION: UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT: APPROVED FOR PUBLIC RELEASE; Distribution Unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S): AFOSR-TR- 87-0309
6a. NAME OF PERFORMING ORGANIZATION: Universal Energy Systems, Inc / 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION: AFOSR/XOT
6c. ADDRESS (City, State and ZIP Code): 4401 Dayton-Xenia Road, Dayton, OH 45432	7b. ADDRESS (City, State and ZIP Code): Building 410, Bolling AFB, DC 20332-6448
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: AFOSR / 8b. OFFICE SYMBOL (If applicable): XOT	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: F49620-85-C0013
8c. ADDRESS (City, State and ZIP Code): Building 410, Bolling AFB, DC 20332	10. SOURCE OF FUNDING NOS.

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.
61102F	3396	D5	

11. TITLE (Include Security Classification): USAF Summer Faculty Research Program - Volume 3 - 1986 Program Management Report

12. PERSONAL AUTHOR(S): Rodney C. Darrah, Susan K. Espy

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
Annual	FROM _____ TO _____	December 1986	

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB. GR.	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

See Attached

DTIC ELECTE JAN 23 1989 S D

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐	UNCLASSIFIED

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE NUMBER (Include Area Code)	22c. OFFICE SYMBOL
Major Richard W. Kopka, Program Manager	202-767-4970	XOT

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE.

89 1 23 004

SECURITY CLASSIFICATION OF THIS PAGE

AFOSR-TR-87-0309

USAF SUMMER FACULTY RESEARCH PROGRAM - 1986

The United States Air Force Summer Faculty Research Program - (USAF-SFRP) is a program designed to introduce university, college, and technical institute faculty members to Air Force research. This is accomplished by the faculty member 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 and research activities. Each assignment is in a subject area and at an Air Force facility mutually agreed upon by the faculty member and the Air Force. In addition to compensation and travel, 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 1986 USAF-SFRP are:

(1) To provide a productive means for scientists and engineers holding advanced degrees to participate in research at various Air Force research activities;

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

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

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

During the summer of 1986, 158 faculty members participated in the USAF-SFRP. These researchers were assigned to approximately 25 USAF laboratories/centers and research activities across the country. A Management Report along with a three volume Technical Report consisting of a compilation of the final reports written by the assigned faculty members describing their summer research efforts is provided by the contractor, Universal Energy Systems.

UNITED STATES AIR FORCE

SUMMER FACULTY RESEARCH PROGRAM

1986

PROGRAM TECHNICAL REPORT

UNIVERSAL ENERGY SYSTEMS, INC.

VOLUME III of III

Program Director, UES
Podney C. Darrah

Program Administrator, UES
Susan K. Espy

Program Manager, AFOSR
Major Richard Kopka

Submitted to

Air Force Office of Scientific Research

Bolling Air Force Base

Washington, DC

December 1986

Accesion For
NTIS CRA&I
DTIC TAB
Unannounced
Justification
By
Distribution /
Availability Codes
Dist | Avail and/or Special
A-1

TABLE OF CONTENTS

SECTION	PAGE
PREFACE	i
LIST OF PARTICIPANTS	ii
PARTICIPANT LABORATORY ASSIGNMENT	xxx
RESEARCH REPORTS	xxxv

PREFACE

The United States Air Force Summer Faculty Research Program (USAF-SFRP) is a program 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 1986 USAF-SFRP are:

(1) To provide a productive means for Scientists and Engineers holding Ph.D. degrees to participate in research at the Air Force Weapons Laboratory;

(2) To stimulate continuing professional association among the Scholars 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 Scientists and Engineers especially as these relate to Air Force technical interests.

During the summer of 1986, 158-faculty members participated. These researchers were assigned to 25 USAF laboratories/centers across the country. This three volume document is a compilation of the final reports written by the assigned faculty members about their summer research efforts.

Pg. XIII

LIST OF 1986 PARTICIPANTS

LIST OF PARTICIPANTS

NAME/ADDRESS	DEGREE, SPECIALTY, LABORATORY ASSIGNED
Dr. John E. Ahlquist Assistant Professor Dept. of Meteorology Florida State University Tallahassee, FL 32306-3034 (904) 644-1558.	Degree: Ph.D., Meteorology, 1981 Specialty: Meteorology Assigned: AFGL
Dr. Rasphal S. Ahluwalia Associate Professor Dept. of Industrial Eng. West Virginia University Morgantown, WV (304) 293-4607	Degree: Ph.D., Systems Eng., 1977 Specialty: Systems Engineering Assigned: AAMRL
Dr. David R. Anderson Assistant Professor Dept. of Chemistry University of Colorado Austin Bluffs Parkway Colorado Springs, CO 80907 (303) 570-9578	Degree: Ph.D., Organic Chemistry, 1978 Specialty: Organic Chemistry Assigned: FJSRL
Dr. David M. Barnhart Professor Dept. of Physical Sciences Eastern Montana College 1500 North 30th Street Billings, MT 59101-0298 (406) 657-2028	Degree: Ph.D., Chemistry, 1964 Specialty: Chemistry Assigned: FJSRL
Dr. Frank P. Battles Professor Mass. Maritime Academy Basic Science Dept. Buzzards Bay, MA 02532 (617) 224-8388	Degree: Ph.D., Physics, 1969 Specialty: Physics Assigned: AFGL
Dr. Georges A. Becus Associate Professor University of Cincinnati Aero. Engineering and Engineering Mechanics Cincinnati, OH 45221 (513) 475-6115	Degree: Ph.D., Engineering Science, 1973 Specialty: Engineering Services Assigned: FDL

Dr. Rex L. Berney
Associate Professor
University of Dayton
Physics Department
300 College Park
Dayton, OH 45469
(513) 299-3012 or 299-2311

Degree: Ph.D., Solid State Physics, 1978
Specialty: Solid State Physics
Assigned: AL

Dr. Albert W. Biggs
Professor
University of Alabama
in Huntsville
Electrical Engineering
Engineering Bldg 263-B
(ECE Dept.)
Huntsville, AL 35899
(205) 895-6459

Degree: Ph.D., Electrical Engineering, 1965
Specialty: Electrical Engineering
Assigned: WL

Dr. Phillip A. Bishop
Assistant Professor
Director of Human Performance
Laboratory
The University of Alabama
P O Box 1967
Area of HPER
University, AL 35486
(205) 348-8370

Degree: Ed.D., Exercise Physiology, 1983
Specialty: Exercise Physiology
Assigned: SAM

Dr. Patricia T. Boggs
Assistant Professor
of Management Science
Wright State University
Dayton, OH 45435
(513) 873-2080 or 2290

Degree: D.B.A., Decision Science, 1984
Specialty: Decision Science
Assigned: HRL/LR

Dr. James A. Brown
Assistant Professor
Tougaloo College
c/o History Dept.
Tougaloo, MS 39174
(601) 957-3623

Degree: M.A., History, 1966
Specialty: History
Assigned: WL

Dr. Clifford G. Burgess
Assistant Professor
Univ. of Southern Mississippi
Computer Science Department
Sou. Station Box 5106
Hattiesburg, MS 39406
(601) 266-4949

Degree: Ph.D., Computer Science, 1985
Specialty: Computer Science
Assigned: SAM

Dr. Jeffrey D. Camm
Assistant Professor
University of Cincinnati
Quantitative Analysis/IS
ML #130
Cincinnati, OH 45221
(513) 475-3621

Degree: Ph.D., Management Science, 1984
Specialty: Management Science
Assigned: BRMC

Dr. Thomas A. Carney
Assistant Professor
Dept. of Meteorology
Florida State University
404 Love Building
Tallahassee, FL 32306
(904) 644-6806

Degree: Ph.D., Meteorology, 1984
Specialty: Meteorology
Assigned: ESC

Dr. George D. Catalano
Associate Professor
Mechanical Engineeering Dept.
Louisiana State University
Baton Rouge, LA 70803-6413
(504) 388-5792

Degree: Ph.D., Aerospace Engineering, 1977
Specialty: Aerospace Engineering
Assigned: AD

Dr. Bor-Chin Chang
Assistant Professor
Electrical Engineering Dept.
Bradley University
Peoria, IL 61625
(309) 676-7611

Degree: Ph.D., Electrical Engineering, 1983
Specialty: Electrical Engineering
Assigned: FDL

Dr. Garvin Chastain
Associate Professor
Boise State University
Dept. of Psychology
1910 University Drive
Boise, ID 83704
(208) 385-2855

Degree: Ph.D., Human Experimental Psychology, 1976
Specialty: Human Experimental Psychology
Assigned: HRL/OT

Dr. Shive K. Chaturvedi
Assistant Professor
Civil Engineering
Ohio State University
470 Hitchcock Hall
Columbus, OH 43210
(614) 422-2617

Degree: Ph.D., Mechanical Engineering, 1979
Specialty: Mechanical Engineering
Assigned: ML

Dr. Hoffman H. Chen
Associate Professor
Grambling State University
Department of Chemistry
Grambling, LA 71245
(318) 274-2260

Degree: Ph.D., Organic Chemistry, 1976
Specialty: Mechanical Engineering
Assigned: SAM

Dr. Lea D. Chen
Assistant Professor
The University of Iowa
Dept. of Mechanical Eng.
Iowa City, IA 52242
(319) 353-5695

Degree: Ph.D., Mechanical Engineering, 1981
Specialty: Organic Chemistry
Assigned: APL

Dr. Wu C. Cheng
Associate Professor
Physics Department
Paine College
1235 Fifteenth Street
Augusta, GA 30910-2799
(404) 722-4471

Degree: Ph.D., Physical Chemistry, 1954
Specialty: Physical Chemistry
Assigned: FJSRL

Dr. John Y. Cheung
Associate Professor
University of Oklahoma
School of Electrical Eng.
and Computer Science
202 W. Boyd, Room 219
Oklahoma City, OK 73170
(405) 325-4324

Degree: Ph.D., Electrical Engineering, 1975
Specialty: Electrical Engineering
Assigned: AL

Dr. Derald Chriss
Assistant Professor
Southern University
Chemistry Dept.
Southern Branch P.O.
Baton Rouge, LA 70813
(504) 771-2000

Degree: M.A., Chemistry, 1981
Specialty: Physical and Analytical Chemistry
Assigned: ML

Dr. Wolfgang Christian
Assistant Professor
Dept. of Physics
Davidson College
Davidson, NC 28036
(704) 892-2000

Degree: Ph.D., Physics, 1975
Specialty: Physics
Assigned: AFGL

Dr. Jacob N. Chung
Associate Professor
Washington State University
Dept. of Mechanical Engineering
Pullman, WA 99164-2920
(509) 335-3222

Degree: Ph.D., Mechanical Engineering, 1979
Specialty: Mechanical Engineering
Assigned: APL

Dr. Brenda J. Claiborne
Assistant Professor
University of Texas, San Antonio
Div. of Life Sciences
San Antonio, TX 78285
(512) 691-4458

Degree: Ph.D., Biology, 1981
Specialty: Biology
Assigned: SAM

Dr. Donald F. Collins
Physics Chairman
Warren Wilson College
Physics Department
Box 5117
701 Warren Wilson Road
Swannanoa, NC 28778
(704) 298-3325

Degree: Ph.D., Physics, 1970
Specialty: Physics
Assigned: AFGL

Dr. William T. Cooper
Assistant Professor
Chemistry Dept.
Florida State University
Tallahassee, FL 32306-3006
(904) 644-6875 or 644-3810

Degree: Ph.D., Chemistry, 1981
Specialty: Chemistry
Assigned: ESC

Dr. Richard H. Cox
Professor
Kansas State University
Center for Human Motor Performance
203 Ahearn
Manhattan, KS 66502
(913) 532-6765

Degree: Ph.D., Motor Learning and Control, 1973
Specialty: Motor Learning & Control
Assigned: HRL/MO

Dr. William Day
Associate Professor
Dept. of Computer Science and Engineering
Auburn University
244 Payne Street
Auburn, AL 36830
(205) 826-4330

Degree: Ph.D., Mathematics, 1969
Specialty: Mathematics
Assigned: RADC

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

Degree: Ph.D., Biochemical Genetics, 1967
Specialty: Biochemistry, Genetics
Assigned: SAM

Dr. Shirshak K. Dhali
Assistant Professor
Dept. of Electrical Engineering
Southern Illinois University
Carbondale, IL 62901
(618) 536-2364

Degree: Ph.D., Electrical Engineering, 1984
Specialty: Electrical Engineering
Assigned: APL

Dr. Lokesh R. Dharani
Assistant Professor
University of Missouri-Rolla
Dept. of Engineering Mechanics
Rolla, MO 65401-0249
(314) 341-4586

Degree: Ph.D., Engineering Mechanics, 1982
Specialty: Engineering Mechanics
Assigned: ML

Dr. Peter J. Disimile
Assistant Professor
Aerospace Engineering and Engineering Mechanics
University of Cincinnati
Mail Location 70
Cincinnati, OH 45221
(513) 475-2936

Degree: Ph.D., Fluid Mechanics, 1984
Specialty: Fluid Mechanics
Assigned: FDL

Dr. Michael L. Doria
Associate Professor of Mechanical Engineering
Valparaiso University
Valparaiso, IN 46383
(219) 464-5104

Degree: Ph.D., Mechanics, 1968
Specialty: Mechanics
Assigned: FDL

Dr. George R. Doyle, Jr.
Associate Professor
University of Dayton
Mechanical Engineering Dept.
300 College Park Drive
Dayton, OH 45469
(513) 229-2995

Degree: Ph.D., Mechanical Engineering, 1973
Specialty: Mechanical Engineering
Assigned: FDL

Dr. Franklin E. Eastep
Professor
Aerospace Engineering
University of Dayton
300 College Park Drive
Dayton, OH 45469
(513) 229-2241

Degree: Ph.D., Aero and Astro, 1968
Specialty: Aero and Astro
Assigned: FDL

Dr. Thaddeus J. Englert
Assistant Professor
University of Wyoming
Dept. of Electrical Eng.
University Station
Laramie, WY 82071
(307) 766-6321

Degree: Ph.D., Physics, 1982
Specialty: Physics
Assigned: FJSRL

Dr. Dennis R. Flentge
Assistant Professor
Cedarville College
Math/Science Department
Box 601
Cedarville, OH 45314
(513) 766-2211

Degree: Ph.D., Physical Chemistry, 1974
Specialty: Physical Chemistry
Assigned: APL

Dr. Mark A. Fulk
Assistant Professor
University of Florida
Computer & Information Sciences
512 Weil Hall
Gainesville, FL 32611
(904) 392-9219

Degree: Ph.D., Computer Science, 1986
Specialty: Computer Science
Assigned: AD

Dr. Patrick T. Gannon, Sr.
Assistant Professor
Dept. of Meteorology
Lyndon State College
Vail Hill
Lyndonville, VT 05851
(802) 626-9371

Degree: Ph.D., Atmospheric Science, 1977
Specialty: Atmospheric Science
Assigned: AFGL

Dr. John K. George
Professor of Chemistry
Mary Washington College
Fredericksburg, VA 22401
(703) 899-4320

Degree: Ph.D., Physical Chemistry, 1976
Specialty: Physical Chemistry
Assigned: FJSRL

Dr. Albert C. Giere
Assistant Professor
Physics Department
Oakwood College
Huntsville, AL 35806
(205) 837-1630

Degree: Ph.D., Mathematical Physics, 1965
Specialty: Mathematical Physics
Assigned: AEDC

Dr. Doris O. Ginn
Associate Professor
Dept. of English and Modern Foreign
1400 John R. Lynch Street
Jackson, MS 39217
(601) 968-2116

Degree: Ph.D., Linguistics, 1979
Specialty: Linguistics
Assigned: HRL/ID

Dr. Thomas A. Gosink
Research Associate Professor
University of Alaska
Geophysical Institute
Fairbanks, AK 99776-0800
(907) 786-1800

Degree: Ph.D., Organic Chemistry, 1965
Specialty: Organic Chemistry
Assigned: SAM

Dr. Raghava G. Gowda
Assistant Professor
Dept. of Computer Science
University of Dayton
300 College Park Drive
Dayton, OH 45469
(513) 229-3808

Degree: A.B.D., Electrical Engineering, 1982
Specialty: Electrical Engineering
Assigned: HRL/LR

Dr. Gerald R. Graves
Assistant Professor
Industrial Engineering
Louisiana State University
3128 CEBA
Baton Rouge, LA 70803
(504) 388-5112

Degree: Ph.D, Industrial Engineering, 1985
Specialty: Industrial Engineering
Assigned: ML

Dr. Ronald L. Greene
Associate Professor
University of New Orleans
Dept. of Physics
3726 Piedmont Drive
New Orleans, LA 70122
(504) 286-6714

Degree: Ph.D., Physics, 1974
Specialty: Physics
Assigned: AL

Dr. William M. Grissom
Instructor
Physics Dept.
Morehouse College
830 Westview Dr., S.W.
Dansby Hall, Suite 131
Atlanta, GA 30314
(404) 681-2800

Degree: MSE Mechanical Engineering, 1978
Specialty: Mechanical Engineering
Assigned: RPL

Dr. William A. Grosky
Associate Professor
Computer Science Dept.
Wayne State University
Detroit, MI 48202
(313) 577-2477

Degree: Ph.D., Engineering and Applied Science, 1971
Specialty: Eng. & Applied Science
Assigned: AL

Dr. Thomas R. Gulledge
Assistant Professor
Quantitative Business Analysis
Louisiana State University
3190 CEBA, QBA
Baton Rouge, LA 70803
(504) 388-2506

Degree: Ph.D., Engineering Management, 1981
Specialty: Engineering Management
Assigned: BRMC

Dr. Ramesh C. Gupta
Professor of Mathematics
University of Maine at Orono
Dept. of Mathematics
Orono, MN 04469
(207) 581-3913

Degree: Ph.D., Mathematical Statistics, 1970
Specialty: Mathematical Statistics
Assigned: SAM

Dr. Fabian C. Hadipriono
Assistant Professor
The Ohio State University
Civil Engineering Dept.
2070 Neil Avenue
Columbus, OH 43210
(614) 422-8518

Degree: Ph.D., Engineering, Civil Engineering, 1982
Specialty: Engineering, Civil Engineering
Assigned: WL

Dr. Frank O. Hadlock
Professor and Chairman
Florida Atlantic University
Dept. of Mathematics
Boca Raton, FL 33431
(305) 393-3342

Degree: Ph.D., Mathematics, 1966
Specialty: Mathematics
Assigned: SAM

Dr. Prabhat Hajela
Assistant Professor
Engineering Sciences
University of Flordia
231 Aero Engineering Bldg.
Gainesville, FL 32611
(904) 392-0961

Degree: Ph.D., Aeronautics and Astronautics, 1982
Specialty: Aeronautics & Astronautics
Assigned: AD

Dr. Patrick R. Hannon
Assistant Professor
Helth & Physical Education
Northern Arizona University
Box 6012, HPR Dept.
Flagstaff, AZ 86011
(602) 523-4122

Degree: Ed.D., Exercise Science, 1980
Specialty: Education, Exercise Science
Assigned: AAMRL

Dr. Donald F. Hanson
Associate Professor
University of Mississippi
Dept. of Electrical Engineering
University, MS 38677
(601) 232-5389

Degree: Ph.D., Electrical Eng. (Electromagnetics), 1976
Specialty: Electrical Engineering, Electromagnetic Electronics
Assigned: RADC

Dr. Gerald F. Harris
Director
Biomedical Engineering
Shriners Hospital
2211 N. Oak Park Avenue
Chicago, IL 60635
(312) 622-5400

Degree: Ph.D., Biomedical Engr., 1981
Specialty: Biomedical Engineering, Biomechanics
Assigned: AAMRL

Dr. Edward J. Hass
Assistant Professor
Franklin and Marshall College
Whitely Psychology Lab.
Box 3003
Lancaster, PA 17604
(717) 291-4202

Degree: Ph.D., Psychology, 1983
Specialty: Psychology
Assigned: HRL/OT

Dr. Doyle E. Hasty
Associate Professor
Engineering/Physics
Motlow State College
P O Box 860
Tullahoma, TN 37330
(615) 455-8511

Degree: M.S., Engineering Administration, 1974
Specialty: Engineering Administration
Assigned: AEDC

Dr. Michael A. Hayes
Assistant Professor
Dept. of Physics
Dartmouth College
Hanover, NH 03755
(603) 646-2973

Degree: Ph.D., Plasma Physics, 1981
Specialty: Plasma Physics
Assigned: AFGL

Dr. James C. Ho
Professor of Physics & Chemistry
Wichita State University
Wichita, KS 67208
(316) 689-3190

Degree: Ph.D., Chemistry, 1966
Specialty: Chemistry
Assigned: APL

Dr. Peter F. Hoffman
Assistant Professor
Dept. of Mathematics
University of Colorado
Box 170, UCD
Denver, CO 80202
(303) 556-4808

Degree: Ph.D., Applied Mathematics, 1985
Specialty: Applied Mathematics
Assigned: AEDC

Dr. Robert R. Hoffman
Assistant Professor
Dept. of Psychology
Adelphi University
Garden City, NY 11530
(516) 663-1055

Degree: Ph.D., Experimental Psychology, 1976
Specialty: Experimental Psychology
Assigned: ESD

Dr. Clifford C. Houk
Professor of Chemistry
Director Industrial Hygiene Program
Ohio University
Dept. of Chemistry
Athens, OH 45701-2979
(614) 594-6205

Degree: Ph.D., Inorganic Chemistry, 1966
Specialty: Inorganic Chemistry
Assigned: OEHL

Dr. Ming-Shing Hung
Associate Professsor
Dept. of Administrative Sciences
Kent State University
Kent, OH 44242
(216) 672-2750

Degree: Doctor of Business Admin., Management Science, 1973
Specialty: Business Administration, Management Science
Assigned: LC

Dr. John M. Jobe
Assistant Professor
Decision Sciences Dept.
Miami University of Ohio
229 Culler Hall
(513) 529-7291

Degree: Ph.D., Statistics, 1984
Specialty: Statistics
Assigned: RADC

Dr. Glen Johnson
Associate Professor
Mechanical and Materials Engineering
Vanderbilt University
Box 8-B
Nashville, TN 37235
(615) 322-0414

Degree: Ph.D., Mechanical Engineering, 1978
Specialty: Mechanical Engineering
Assigned: AEDC

Dr. Betty Jones
Associate Professor
Dir. of the EM Institute
830 Westview Dr., S.W.
Atlanta, GA 30314
(404) 681-2800

Degree: Ph.D., Biology, 1978
Specialty: Biology
Assigned: SAM

Dr. Jeremy C. Jones
Assistant Professor
University of West Florida
Systems Science
USF LIB 630 CS & Eng
Tampa, FL 33620
(813) 974-2114

Degree: M.S., Physics, 1968
M.S., Math, 1965
Specialty: Physics, Math
Assigned: AD

Dr. Marvin S. Keener
Professor of Mathematics
Oklahoma State University
Math Department
Stillwater, OK 74078
(405) 624-5789

Degree: Ph.D., Mathematics, 1970
Specialty: Mathematics
Assigned: AD

Dr. Yong S. Kim
Assistant Professor
Dept. of Civil Engineering
The Catholic Univ. of America
Washington, DC 20064
(202) 635-5163

Degree: Ph.D., Civil Engineering, 1984
Specialty: Civil Engineering
Assigned: ESC

Dr. Joel R. Klink
Professor
Univ. of Wisconsin-Eau Claire
Chemistry Dept.
Eau Claire, WI 54701
(715) 836-5518

Degree: Ph.D., Organic Chemistry, 1964
Specialty: Organic Chemistry
Assigned: RPL

Dr. Stephan E. Kolitz
Assistant Professor
University of Massachusetts
Management Sciences Dept.
Boston, MA 02125
(617) 929-8051

Degree: Ph.D., Operations Research, 1983
Specialty: Operations Research
Assigned: ESD

Dr. Philipp G. Kornreich
Professor
Dept. of Electrical Engineering
and Computer Engineering
Syracuse University
Syracuse, NY 13244
(315) 423-4447

Degree: Ph.D., Electrical Eng., 1967
Specialty: Electrical Engineering
Assigned: RADC

Dr. Mou-Liang Kung
Associate Professor
Math and Computer Science
Norfolk State University
2401 Corprew Avenue
Norfolk, VA 23504
(804) 623-8820

Degree: Ph.D., Math, 1974
M.S., Computer Science, 1985
Specialty: Mathematics
Assigned: RADC

Dr. Charles E. Lance
Assistnat Professor
University of Georgia
Dept. of Psychology
Athens, GA 30602
(404) 542-4439

Degree: Ph.D., Psychology, 1985
Specialty: Psychology
Assigned: HRL/ID

Dr. David I. Lawson
Assistant Professor
Mathematics and Computer Science
Stetson University
Box 8348
DeLand, FL 32720
(904) 734-4121

Degree: M.A., Mathematics, 1968
Specialty: Mathematics
Assigned: AD

Dr. Paul S.T. Lee
Associate Professor
Dept. of Business Admin.
N.C. A&T State University
1601 E. Market Street
Greensboro, NC 27410
(919) 379-7656

Degree: Ph.D., Quantitative Methods and Research Economics, 1973
Specialty: Quantitative Methods and Research Economics
Assigned: FDL

Dr. C. Randal Lishawa
Physics Instructor
Dept. of Chemistry and Physics
Jefferson State University
Birmingham, AL 35243
(205) 853-1200

Degree: Ph.D., Physical Chemistry, 1981
Specialty: Physical Chemistry
Assigned: AFGL

Dr. Cheng Liu
Associate Professor
Dept. of Engr. Tech.
University of North Carolina
Charlotte, NC 28223
(704) 597-4191

Degree: M.S., Civil Engineering, 1963
Specialty: Civil Engineering
Assigned: ESC

Dr. James C. LoPresto
Professor of Astronomy
Edinboro University
Dept. of Physics
Edbinboro, PA 16444
(814) 732-2469

Degree: Ph.D., Astronomy, 1974
Specialty: Astronomy
Assigned: AFGL

Dr. Stephen L. Loy
Assistant Professor
Iowa State University
Management
374 Carver
Ames, IA 50010
(515) 294-8108

Degree: DBA, MIS, 1986
Specialty: Management Information System
Assigned: HRL/LR

Dr. Nancy I. Lyons
Associate Professor
Dept. of Statistics
University of Georgia
Statistics Building
Athens, GA 30602
(404) 542-5232

Degree: Ph.D., Statistics, 1975
Specialty: Statistics
Assigned: LMC

Dr. Robert L. Manicke
Assistant Professor
U.S. Naval Academy
Math Department
Attn: 9E
Annapolis, MD 21402
(301) 267-3603

Degree: Ph.D., Statistics, 1980
Specialty: Statistics
Assigned: AAMRL

Dr. Arthur A. Mason
Professor of Physics
The University of Tennessee
Space Institute
Tullahoma, TN 37388
(615) 455-0631

Degree: Ph.D., Physics, 1963
Specialty: Physics
Assigned: AEDC

Dr. Curtis W. McDonald
Professor of Chemistry
Dept. of Chemistry
Texas Southern University
3100 Cleborne
Houston, TX 77004
(713) 527-7003

Degree: Ph.D., Chemistry, 1962
Specialty: Chemistry
Assigned: OEHL

Dr. Gopal M. Mehrotra
Assistant Professor
Wright State University
Materials Science and
Engineering Program
Systems Engineering
Dayton, OH 45435
(513) 873-2481

Degree: Ph.D., Metallurgy, 1975
Specialty: Metallurgy
Assigned: ML

Dr. Jorge L. Mendoza
Associate Professor
Texas A&M University
Psychology Department
College Station, TX 77843
(409) 845-0880

Degree: Ph.D., Psychology, 1974
Specialty: Psychology
Assigned: HRL/MO

Dr. Shreenivas Moorthy
Associate Professor
Texas A&I University
Electrical Engineering &
Computer Science
Campus Box 192
Kingsville, TN 78363
(512) 595-2004

Degree: Ph.D., Electrical Engineering, 1972
Specialty: Electrical Engineering
Assigned: HRL/LR

Dr. Mary L. Morton-Gibson
Assistant Professor
Physics Department
The Citadel
Charleston, SC 29409
(803) 792-6943

Degree: Ph.D., Physiology and Biophysics, 1970
Specialty: Physiology and Biophysics
Assigned: SAM

Dr. Rex C. Moyer
Associate Professor
Trinity University
Biology Dept.
715 Stadium Drive
San Antonio, TX 78284
(512) 736-7242

Degree: Ph.D., Microbiology, 1965
Specialty: Microbiology
Assgined: SAM

Dr. V. Dakshina Murty
Associate Professor
University of Portland
School of Engineering
Portland, OR 97203
(503) 283-7379

Degree: Ph.D., Engineering Mechanics, 1982
Specialty: Engineering Mechanics
Assigned: FDL

Dr. Richard W. Nau
Professor of Math and Computer Science
Carleton College
Northfield, MN 55057
(607) 663-4361

Degree: Ph.D., Applied Math and Computer Science, 1970
Specialty: Applied Math and Computer Science
Assigned: WL

Dr. Henry Nebel
Associate Professor
Alfred University
Physics Department
Alfred, NY 14802
(607) 871-2208

Degree: Ph.D., Physics, 1967
Specialty: Physics
Assigned: AFGL

Dr. Robert M. Nehs
Associate Professor
Dept. of Mathematical Sciences
Texas Southern University
3100 Cleburne Avenue
Houston, TX 77004
(713) 527-7915

Degree: Ph.D., Mathematics, 1980
Specialty: Mathematics
Assigned: AFGL

Dr. Douglas L. Oliver
Assistant Professor
University of Toledo
Toledo, OH 43606
(419) 537-2885

Degree: Ph.D., Mechanical Engineering, 1985
Specialty: Mechanical Engineering
Assigned: APL

Dr. Philip D. Olivier
Assistant Professor
University of Texas at San Antonio
Division of Engineering
San Antonio, TX 78285
(512) 691-5565

Degree: Ph.D., Electrical Engineering, 1980
Specialty: Electrical Engineering
Assigned: HRL/ID

Dr. Harvey L. Paige
Associate Professor
Alfred University
Dept. of Chemistry
Alfred, NY 14802
(607) 871-2201

Degree: Ph.D., Inorganic Chemistry, 1969
Specialty: Inorganic Chemistry
Assigned: ML

Dr. Parsottam J. Patel
Associate Professor
Meharry Medical College
Dept. of Microbiology
Nashville, TN 37208
(615) 327-6760

Degree: Ph.D., Microbiology, 1976
Specialty: Microbiology
Assigned: SAM

Dr. Robert A. Patsiga
Professor
Dept. of Chemistry
Indiana Univ. of Pennsylvania
Indiania, PA 15705
(412) 357-2210

Degree: Ph.D., Organic and Polymer Chemistry, 1962
Specialty: Organic Polymer Chemistry
Assigned: ML

Dr. Martin A. Patt
Associate Professor
University of Lowell
Dept. of Electrical Engineering
1 University Avenue
Lowell, MA 01854
(617) 452-5000

Degree: M.S., Science, Electrical Engineering, 1964
Specialty: Electrical Engineering
Assigned: AFGL

Dr. Jacqueline G. Paver
Research Associate
Duke University
Biomedical Engineering Dept.
136 Engineering Building
Durham, NC 27706
(919) 684-6185

Degree: Ph.D., Biomedical Engineering, 1984
Specialty: Biomechanical Engineering
Assigned: AAMRL

Dr. Alexandru A. Pelin
Associate Professor
of Computer Science
Florida International University
Dept. of Mathematical Sciences
Tamiami Campus
Miami, FL 33199
(305) 554-3386

Degree: Ph.D., Computer Science, 1977
Specialty: Computer Science
Assigned: WL

Dr. Bernard J. Piersma
Professor of Chemistry
Houghton College
Houghton, NY 14744
(716) 567-2211

Degree: Ph.D., Physical Chemistry, 1965
Specialty: Physical Chemistry
Assigned: FJSRL

Dr. Leonard Price
Professor of Chemistry
Xavier University of Louisiana
Chemistry Department
7325 Palmetto Street
New Orleans, LA 70125

Degree: Ph.D., Organic Chemistry, 1962
Specialty: Organic Chemistry
Assigned: SAM

Dr. Craig G. Prohazka
Assistant Professor
University of Lowell
Dept. of Electrical Engineering
Lowell, MA 01854
(617) 452-5000

Degree: Ph.D., Electrical Engineering, 1981
Specialty: Electrical Engineering
Assigned: RADC

Dr. L. Rai Pujara
Assistant Professor
Electrical Systems Engineering
Wright State University
Dayton, OH 45435
(513) 873-2456

Degree: Ph.D., Mathematics, 1971
Specialty: Electrical Sys. Engineering
Assigned: FDL

Dr. Richard S. Quimby
Assistant Professor
Worcester Polytechnic Institute
Dept. of Physics
100 Institute Road
Worcester, MA 01609
(617) 793-5490

Degree: Ph.D., Physics, 1979
Specialty: Physics
Assigned: RADC

Dr. Singiresu S. Rao
Associate Professor
School of Mechanical Engineering
Purdue University
West Lafayette, IN 47907
(317) 494-9766

Degree: Ph.D., Engineering Design, 1971
Specialty: Engineering Design
Assigned: FDL

Dr. Ralph J. Rascati
Associate Professor
Kennesaw College
Biology Department
Marietta, GA 30061
(404) 529-2878

Degree: Ph.D., Biochemistry, 1975
Specialty: Biochemistry
Assigned: OEHL

Dr. Kuldip S. Rattan
Associate Professor
Dept. of Electrical Systems Engineering
Wright State University
Dayton, OH 45435
(513) 873-2497

Degree: Ph.D., Electrical Engineering, 1975
Specialty: Electrical Engineering
Assigned: AAMRL

Dr. Barbara Rice
Associate Professor
Mathematics Dept.
Alabama A&M University
P O Box 326
Normal, AL 35762
(205) 859-7448

Degree: Ph.D., Mathematics, 1965
Specialty: Mathematics
Assigned: AD

Dr. Dan B. Rinks
Assistant Professor
Louisiana State University
Dept. of Quantitative Business Analysis
3180 CEBA
Baton Rouge, LA 70803
(504) 388-5318

Degree: Ph.D., Quantitative Management Science, 1978
Specialty: Quantitative Mgmt. Science
Assigned: LMC

Dr. William P. Robey
Assistant Professor
Electronics & Computer Technology
Oklahoma State University
202CR
Stillwater, OK 74078
(405) 624-5716

Degree: B.S., Engineering Physics, 1968
Specialty: Engineering Physics
Assigned: AFGL

Dr. Kenneth C. Russell
Professor
Massachusetts Institute of Technology
Materials Science
Nuclear Engineering
Room 13-5066
77 Massachusetts Avenue
Cambridge, MA 02139
(617) 253-3328

Degree: Ph.D., Metall. Engineering, 1964
Specialty: Metallurgy Engineering
Assigned: ML

Dr. Sally A. Sage
Assistant Professor
West Georgia College
Dept. of Math/Computer Science
Carrollton, GA 30118
(404) 834-1380

Degree: M.S., Computer Science, 1979
Specialty: Computer Science
Assigned: AD

Dr. Mo Samimy
Assistant Professor
Ohio State University
Mechanical Engineering Dept.
206 W. 18th Avenue
Columbus, OH 43210
(614) 422-6988

Degree: Ph.D., Mechanical Engineering, 1984
Specialty: Mechanical Engineering
Assigned: APL

Dr. John F. Schaefer
Associate Professor
Dept. of Electrical Engineering
The Citadel
Charleston, SC 29409
(803) 792-4899

Degree: Ph.D., Electrical Engineering, 1965
Specialty: Electrical Engineering
Assigned: ESC

Dr. John R. Schneider
Professor of Physics
University of Dayton
Physics Department
Dayton, OH 45469
(513) 229-1000

Degree: Ph.D., Physics, 1965
Specialty: Physics
Assigned: ML

Dr. Richard M. Schori
Professor of Mathematics
Oregon State University
Math Department
Corvallis, OR 97331
(503) 754-4686

Degree: Ph.D., Mathematics, 1964
Specialty: Mathematics
Assigned: SAM

Dr. William D. Schulz
Professor
Eastern Kentucky University
Dept. of Chemistry
Moore 337, EKU
Richmond, KY 40475
(606) 622-1463

Degree: Ph.D., Analytical Chemistry, 1975
Specialty: Analytical Chemistry
Assigned: ESC

Dr. Meckinley Scott
Professor
University of Alabama
Mathematics Department
Box 1416
University, AL 35486
(206) 348-1985

Degree: Ph.D., Statistics, 1964
Specialty: Statistics
Assigned: AD

Dr. Martin A. Shadday, Jr.
Assistant Professor
University of South Carolina
Mechanical Engineering
College of Engineering
Columbia, SC 29208
(803) 777-7118

Degree: Ph.D., Mechanical Engineering, 1982
Specialty: Mechanical Engineering
Assigned: WL

Dr. Nisar Shaikh
Assistant Professor
University of Nebraska-Lincoln
Dept. of Engineering Mechanics
212 Bancroft Hall
Lincoln, NE 68588-0437
(402) 472-2384

Degree: Ph.D., Applied Mechanics, 1982
Specialty: Applied Mathematics
Assigned: ML

Dr. Dolores C. Shockley
Associate Professor
Meharry Medical College
Dept. of Pharmacology
1005 D.B. Todd Blvd.
Nashville, TN 37208
(615) 327-6510

Degree: Ph.D., Pharmacology, 1955
Specialty: Pharmacology
Assigned: SAM

Dr. William D. Shontz
Associate Professor
Montana State University
Department of Psychology
Bozeman, MT 59717
(406) 994-5180

Degree: Ph.D., Experimental Psychology, 1967
Specialty: Psychology
Assigned: AFHRL/OTE

Dr. William D. Siuru, Jr.
Senior Research Associate
Space and Flight Systems Lab.
University of Colorado at Colorado Springs
1867 Austin Bluffs Parkway
Colorado Springs, CO 80907
(303) 593-3573

Degree: Ph.D., Mechanical Engineering, 1975
Specialty: Mechanical Engineering
Assigned: FJSRL

Dr. Boghos D. Sivazlian
Professor
The University of Florida
Dept. of Industrial and Systems Engineering
303 Weil Hall
Gainesville, FL 32611
(904) 392-1464

Degree: Ph.D., Operations Research, 1966
Specialty: Operations Research
Assigned: AD

Dr. Siavash H. Sohrab
Assistant Professor
Dept. of Mechanical and Nuclear Engineering
Northwestern University
Technical Institute
Evanston, IL 60201
(312) 491-3572

Degree: Ph.D., Engineering Physics, 1981
Specialty: Engineering Physics
Assigned: RPL

Dr. Stuart R. Stock
Assistant Professor
Georgia Institute of Technology
School of Materials Engineering
Atlanta, GA 30332-0245
(404) 894-6882

Degree: Ph.D., Metallurgy, 1983
Specialty: Metallurgy
Assigned: ML

Dr. James E. Sturm
Professor
Dept. of Chemistry #6
Lehigh University
Bethlehem, PA 18015
(215) 861-3477

Degree: Ph.D., Physical Chemistry, 1957
Specialty: Physical Chemistry
Assigned: AFGL

Dr. Edgar C. Tacker
Professor of Electrical Eng.
University of Tulsa
600 S. College
Tulsa, OK 74104
(918) 592-6000

Degree: Ph.D., Electrical Engineering, 1964
Specialty: Electrical Engineering
Assigned: AAMRL

Dr. Nicholas E. Takach
Assistant Professor
Dept. of Chemistry
University of Tulsa
600 S. College
Tulsa, OK 74104
(918) 592-6000

Degree: Ph.D., Chemistry, 1979
Specialty: Chemistry
Assigned: RPL

Dr. Arjun Tan
Assistant Professor
Alabama A&M University
Physics Dept.
Box 447
Normal, AL 35762
(205) 859-7470

Degree: Ph.D., Physics, 1979
Specialty: Physics
Assigned: AEDC

Dr. Robert P. Taylor
Assistant Professor
Mech. and Nuclear Engr. Dept.
Mississippi State University
Drawer ME
Mississippi State, MS 39762
(601) 325-7316

Degree: Ph.D., Mechanical Engineering, 1983
Specialty: Mechanical Engineering
Assigned: APL

Dr. Ken Tomiyama
Assistant Professor
Pennsylvania State University
Dept. of Electrical Engineering
121 E.E. East
University Park, PA 16802
(814) 865-7667

Degree: Ph.D., System Science, 1977
Specialty: System Science
Assigned: AL

Dr. Phillip D. Tomporowski
Assistant Professor
Dept. of Psychology
University of Alabama
Box 2968
University, AL 35486
(205) 348-1936

Degree: Ph.D., Experimental Psychology, 1977
Specialty: Experimental Psychology
Assigned: HRL/MO

Dr. Timothy R. Troutt
Assistant Professor
Washington State University
Dept. of Mechanical Engineering
Sloan Hall 201
Pullman, WA 99164-2920
(509) 335-4375

Degree: Ph.D., Mechanical Engineering, 1978
Specialty: Mechanical Engineering
Assigned: FJSRL

Dr. C. Randall Truman
Assistant Professor
Mechanical Engineering
University of New Mexico
Albuquerque, NM 87131
(505) 277-6296

Degree: Ph.D., Mechanical Engineering, 1983
Specialty: Mechanical Engineering
Assigned: WL

Dr. Roy M. Ventullo
Associate Professor
Dept. of Biology
University of Dayton
300 College Park
Dayton, OH 45469-0001
(513) 229-2503

Degree: Ph.D., Microbiology, 1978
Specialty: Microbiology
Assigned: ESC

Dr. Doris J. Walker-Dalhouse
Director of Independent/
Home-Study Programs
Associate Professor of Reading
Jackson State University
P O Box 17120
Jackson, MS 39217
(601) 968-9684

Degree: Ph.D., Reading Education, 1977
Specialty: Reading Education
Assigned: HRL/ID

Dr. Donald W. Welch
Research Scientist
Texas A&M University
Hyperbaric Laboratory
College Station, TX 77843
(409) 845-4027

Degree: Doctorate, Microbiology, 1985
Specialty: Microbiology
Assigned: WHMC

Dr. Albert R. Wellens
Associate Professor and
Associate Chairman
Dept. of Psychology
University of Miami
P O Box 248185
Coral Gables, FL 33158
(305) 284-2814

Degree: Ph.D., Experimental Social Psychology, 1972
Specialty: Experimental Social Psychology
Assigned: AAMRL

Dr. Stephen T. Welstead
Assistant Professor
University of Alabama
in Huntsville
Dept. of Mathematics
Huntsville, AL 35899
(205) 895-6470

Degree: Ph.D., Applied Mathematics, 1982
Specialty: Applied Mathematics
Assigned: RADC

Dr. Shih-sung Wen
Professor of Psychology
Psychology Department
Jackson State University
1325 J.R. Lynch Street
Jackson, MS 39217
(60') 968-2371

Degree: Ph.D., Educational Psychology, 1971
Specialty: Educational Psychology,
Assigned: SAM

Dr. Stanley J. Whidden
Researcher
JESM Baromedical Research
Institute
4400 Gen. Meyer Avenue, 114
New Orleans, LA 70117
(504) 363-7656

Degree: Ph.D., Physiology, Pharmacology, 1978
Specialty: Hyperbaric Medicine,
Assigned: SAM

Dr. Dennis W. Whitson
Professor of Physics
Indiana Univ. of Pennsylvania
Physics Department
Indiana, PA 15705
(412) 357-2589

Degree: Ph.D., Physics, 1969
Specialty: Physics
Assigned: AL

Dr. Shirley A. Williams
Assistant Professor
Jackson State University
Jackson, MS 39217
(601) 968-2586

Degree: Ph.D., Physiology and Biophysics, 1985
Specialty: Physiology and Biophysics
Assigned: OEHL

Dr. Billy R. Wooten
Professor
Dept. of Psychology
Brown University
Providence, RI 02906
(401) 863-2330

Degree: Ph.D. of Philosophy, Psychology, 1970
Specialty: Philosophy, Psychology
Assigned: HRL/OT

Dr. Daniel W. Yannitell
Associate Professor
Mechanical Engineering Dept.
Louisiana State University
Baton Rouge, LA 70803
(504) 388-5972

Degree: Ph.D., Theoretical and Applied Mechanics, 1967
Specialty: Theoretical and Applied Mechanics
Assigned: RPL

Dr. Tsun-wai G. Yip
Assistant Professor
Aero-Astro Engineering Dept.
Ohio State University
2300 West Case Road
Columbus, OH 43220
(614) 422-1241

Degree: Ph.D., Aero-Astro Engineering, 1984
Specialty: Aeronautics-Astronautics Engineering
Assigned: FDL

Dr. Robert L. Yolton
Associate Professor of Psychophysiology
Pacific University
College of Optometry
Forest Grove, OR 97116
(503) 357-6151

Degree: Ph.D., Psychology, 1975
Doctor of Optometry, 1975
Specialty: Psychology, Optometry
Assigned: AAMRL

Dr. Richard W. Young
Associate Professor
Aerospace Engineering and Engineering Mechanics
University or Cincinnati
ML 70
Cincinnati, OH 45242
(513) 475-3014

Degree: Ph.D., Applied Mechanics, 1975
Specialty: Applied Mechanics
Assigned: FDL

Dr. Ajmal Yousuff
Assistant Professor
Drexel University
Dept. of Mechanical Enginnering and Mechanics
Philadelphia, PA 19104
(215) 895-1868

Degree: Ph.D., Aeronautics, 1983
Specialty: Aeronautics
Assigned: FDL

Dr. David D. Zeigler
Adjunct Faculty
North Texas State University
Biology Department
Denton, TX 76203
(817) 565-3622

Degree: Ph.D., Zoology, 1984
Specialty: Zoology
Assigned: AD

Dr. Henry Zmuda
Assistant Professor
Stevens Institute of Technology
Electrical Engineering Dept.
Castle Point Station
Hoboken, NJ 07030
(201) 420-5507

Degree: Ph.D., Electrical Engineering, 1984
Specialty: Electrical Engineering
Assigned: RADC

Dr. George W. Zobrist
Professor of Computer Science
Dept. of Computer Science
University of Missouri-Rolla
Rolla, MO 65401
(314) 341-4836

Degree: Ph.D., Electrical Engineering, 1965
Specialty: Electrical Engineering
Assigned: AL

PARTICIPANT LABORATORY ASSIGNMENT

C. PARTICIPANT LABORATORY ASSIGNMENT (Page 1)

1986 USAF/UES SUMMER FACULTY RESEARCH PROGRAM

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

1. Lea Chen
2. Jacob Chung
3. Shirshak Dhali
4. Dennis Flentge
5. James Ho
6. Douglas Oliver
7. Mo Samimy
8. Robert Taylor

ARMAMENT LABORATORY (AD)
(Eglin Air Force Base)

1. George D. Catalano
2. Mark A. Fulk
3. Jeremy C. Jones
4. Marvin S. Keener
5. David I. Lawson
6. Barbara C. Rice
7. Sally A. Sage
8. Meckinley Scott
9. Boghos D. Sivazlian
10. David D. Zeigler
11. Probhat Hajela

ARMSTRONG AEROSPACE MEDICAL RESEARCH LABORATORY (AAMRL)
(Wright-Patterson Air Force Base)

1. Rashpal S. Ahluwalia
2. Patrick R. Hannon
3. Gerald F. Harris
4. Robert L. Manicke
5. Jacqueline G. Paver
6. Kuldip S. Rattan
7. Edgar C. Tacker
8. Albert R. Wellens
9. Robert L. Yolton

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

1. Albert C. Giere
2. Doyle E. Hasty
3. Peter E. Hoffman
4. Glen E. Johnson
5. Arthur A. Mason
6. Arjun Tan

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

1. Rex L. Berney
2. John Y. Cheung
3. Ronald L. Greene
4. William Grosky
5. Ken Tomiyama
6. Dennis W. Whitson
7. George W. Zobrist

BUSINESS RESEARCH MANAGEMENT CENTER (BRMC)
(Wright-Patterson Air Force Base)

1. Jeffrey Camm
2. Thomas Gulledge

C. PARTICIPANT LABORATORY ASSIGNMENT (Page 2)

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

1. Robert Hoffman
2. Stephan Kolitz

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

1. Thomas Carney
2. William Cooper
3. Yong Kim
4. Cheng Liu
5. John Schaefer
6. William Schulz
7. Roy Ventullo

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

1. George Becus
2. Bor-Chin Chang
3. Peter Disimile
4. Michael Doria
5. George Doyle
6. Franklin Eastep
7. Paul Lee
8. V. Dakshina Murty
9. L. Pujara
10. Singiresu Rao
11. Tsun-Wai Yip
12. Warren Young
13. Ajmal Yousuff

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

1. David Anderson
2. David Barnhart
3. Wu Cheng
4. Thaddeus Englert
5. John George
6. Bernard Piersma
7. William Siuru
8. Timothy Troutt

GEOPHYSICS LABORATORY (AFGL)
(Hanscom Air Force Base)

1. Jon Ahlquist
2. Frank Battles
3. Wolfgang Christian
4. Donald Collins
5. Patrick Gannon
6. Michael Hayes
7. C. Lishawa
8. James LoPresto
9. Henry Nebel
10. Robert Nehs
11. Martin Patt
12. William Robey
13. James Sturm

HUMAN RESOURCES LABORATORY/ID (HRL/ID)
(Lowry Air Force Base)

1. Doris Ginn
2. Charles Lance
3. Philip Olivier
4. Doris Walker-Dalhouse

HUMAN RESOURCES LABORATORY/LR (HRL/LR)
(Wright-Patterson Air Force Base)

1. Patricia Boggs
2. Raghava Gowda
3. Stephen Loy
4. Shreenivas Moorthy

C. PARTICIPANT LABORATORY ASSIGNMENT (Page 3)

HUMAN RESOURCES LABORATORY/MO (HRL/MO)
(Brooks Air Force Base)

1. Richard Cox
2. Jorge Mendoza
3. Phillip Tomporowski

HUMAN RESOURCES LABORATORY/OT (HRL/OT)
(Williams Air Force Base)

1. Garvin Chastain
2. Edward Hass
3. William Shontz
4. Billy Wooten

LOGISTICS COMMAND (LC)
(Wright-Patterson Air Force Base)

1. Ming-Shing Hung

LOGISTICS MANAGEMENT CENTER (LMC)
(Gunter Air Force Base)

1. Nancy Lyons
2. Dan Rinks

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

1. Shive Chaturvedi
2. Derald Chriss
3. Lokesh Dharani
4. Gerald Graves
5. Gopal Mehrota
6. Harvey Paige
7. Robert Patsiga
8. Kenneth Russell
9. James Schneider
10. Nisar Shaikh
11. Stuart Stock

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

1. Clifford Houk
2. Curtis McDonald
3. Ralph Rascati
4. Shirley Williams

ROCKET PROPULSION LABORATORY (RPL)
(Edwards Air Force Base)

1. William Grissom
2. Joel Klink
3. Siavash Sohrab
4. Nicholas Takach
5. Danial Yannitell

C. PARTICIPANT LABORATORY ASSIGNMENT (Page 4)

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

1. William Day
2. Donald Hanson
3. John Jobe
4. Philipp Kornreich
5. Mou-Liang Kung
6. Craig Prohazka
7. Richard Quimby
8. Stephen Welstead
9. Henry Zmuda

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

1. Vito Del Vecchio
2. Phillip Bishop
3. Clifford Burgess
4. Hoffman Chen
5. Brenda Claiborne
6. Thomas Gosink
7. Ramesh Gupta
8. Frank Hadlock
9. Betty Jones
10. Mary Morton-Gibson
11. Rex Moyer
12. Parsottam Patel
13. Leonard Price
14. Richard Schori
15. Dolores Shockley
16. Shih-sung Wen
17. Stanley Whidden

WEAPONS LABORATORY (WL)
(Kirtland Air Force Base)

1. Albert Biggs
2. James Brown
3. Fabian Hadipriono
4. Richard Nau
5. Alexandru Pelin
6. Martin Shadday
7. C. Truman

WILLFORD HALL MEDICAL CENTER
(Lackland Air Force Base)

1. Donald Welch

RESEARCH REPORTS

RESEARCH REPORTS

1986 SUMMER FACULTY RESEARCH PROGRAM

Technical Report Number	Title	Professor
Volume I		
1	Weather Forecast Evaluation by Decomposition of the Wind Field into Rotational and Divergent Components	Dr. Jon E. Ahlquist
2	Specification of a Vision Based Navigation System for a Mobile Robot	Dr. Rashpal S. Ahluwalia
3	An EPR Study of the Intermediate Radicals Produced in the Catalyzed and Uncatalyzed Thermal Decomposition of Dinitrotoluenes	Dr. David R. Anderson
4	An EPR and HPLC Study of the Mechanisms and Kinetics of TNT with Various Reagents	Dr. David M. Barnhart
5	Atmospheric Turbulence Effects on Optical Beams	Dr. Frank P. Battles
6	Some Issues in the Modeling and Control of Large Flexible Space Structures	Dr. Georges A. Becus
7	Fresnel Drag Unit and Registration Optics for the Ring Laser Gyro	Dr. Rex L. Berney
8	Corrugated Waveguides for Slow Waves	Dr. Albert W. Biggs
9	Work Capacity Increased in High Ambient Temperature Chemical Warfare Environments Through Use of Intermittent Work and Individual Liquid Cooling	Dr. Phillip A. Bishop
10	Factors Impacting Adaptation of Decision Support Systems: An Interdisciplinary Approach	Dr. Patricia T. Boggs
11	American Ballistic Missile Defense, 1955-1979	Dr. James A. Brown

12	Data Management Within the School of Aerospace Medicine	Dr. Clifford G. Burgess
13	Measuring Production Rate in Aircraft Repricing Models	Dr. Jeffrey D. Camm
14	A Comparative Study and Evaluation of Four Atmospheric Dispersion Models with Present or Potential Utility in Air Force Operations	Dr. Thomas A. Carney
15	Turbulent Flow Over an Embedded, Rectangular Cavity	Dr. George D. Catalano
16	Fast Iterative Algorithm for 2-Block H∞ Optimization Problems	Dr. Bor-Chin Chang
17	Effects of Unchanging Clutter on Peripherally-Precued Covert Attention Shifts	Dr. Garvin Chastain
18	Thermo-Mechanical Behavior of High Temperature Composites: A Review	Dr. Shive K. Chaturvedi
19	Structure of Jet Diffusion Flames	Dr. Lea D. Chen
20	Serum Phospholipid and Cholesterol Ester Fatty Acids as Risk Predictors for Coronary Artery Disease	Dr. Hoffman Hor-Fu Chen
21	A Laser Study of Sulfur Hexafluoride	Dr. Wu C. Cheng
22	A Preliminary Study of the Characteristics of Various Digital Signal Processing Techniques in Receivers	Dr. John Yan-Poon Cheung
23	An FTIR Study of the Isomerization of Isoimides	Dr. Derald Chriss
24	Multiphoton Ionization and Infrared Generation in a Cesium Heat Pipe Oven	Dr. Wolfgang Christian
25	A Numerical Simulation of the Liquid-Metal Dual-Latent Heat Packed Bed Thermal Energy Storage System	Dr. Jacob Nan-Chu Chung
26	An Ultrastructural Study of Mossy Fiber Terminals Isolated from the Mammalian Brain	Dr. Brenda J. Claiborne
27	Evaluation and Calibration of the AFGL Ultraviolet Imaging System	Dr. Donald F. Collins

28	A Dispersion-Corrected HPLC/FACP Method for Measuring Sorption Isotherms of Substituted Aromatics on Soil Organic Matter	Dr. William T. Cooper
29	A Study Designed to Enhance the Predictive Validity of the Two-Hand Coordination and Complex Coordination Psychomotor Tests	Dr. Richard H. Cox
30	Truth Maintenance and Learning in Knowledge-Based Systems	Dr. William B. Day
31	Cloning of Mycoplasma Genomic Libraries in E. Coli	Dr. Veto G. DelVecchio
32	A Monte Carlo Simulation of the Electron Motion in Silane and the Ambipolar Diffusion of a Multi-component Plasma	Dr. Shirshak K. Dhali
33	Modeling of Failure Mechanisms in Brittle Matrix High Temperature Composites	Dr. Lokesh R. Dharani
34	The Effects of Surface Roughness on Turbulent Boundary Layer Separation at Hypersonic Speeds	Dr. Peter J. Disimile
35	A Numerical Investigation Into the Acoustic Disturbance of a Laminar Flow Field Over an Airfoil at High Angle of Attack	Dr. Michael L. Doria
36	Computer Aided Engineering Techniques in Aircraft Landing Gear Analysis	Dr. George R. Doyle, Jr.
37	Structural Modification to Enhance the Active Control of Aeroelastic Instabilities	Dr. Franklin E. Eastep
38	Laser-Induced Breakdown of Sulfur-hexafluoride	Dr. Thaddeus J. Englert
39	Carbon Residue Studies with a Microcarbon Residue Tester	Dr. Dennis R. Flentge
40	Symbolic Processing In Automatic Target Recognition	Dr. Mark A. Fulk
41	Response of Downslope and Florida Mesoscale Wind Systems to Physiographic Features	Dr. Patrick Gannon, Sr.

42	Theoretical Studies in: I. Development of a Model for Tribological Studies II. Modes of Decomposition of Stabilized Explosives	Dr. John K. George
43	The Electron Number Density of a Plasma Derived from Measurements of Radar Cross-Sections	Dr. Albert C. Ciere
44	Text Linguistics and the Assessment of Military Rhetoric	Dr. Doris O. Ginn
45	Chemical Defense Detection Devices	Dr. Thomas A. Gosink
46	Structured Techniques for IMIS (Integrated Maintenance Information Systems) Software Development	Dr. Raghava G. Gowda
47	Multiple Processor System Design for AI-Based Process Control	Dr. Gerald R. Graves
48	Wannier Excitons in GaAs-$Ga_{1-x}Al_xAs$ Heterostructures: Magnetic Field Parallel to the Interfaces	Dr. Ronald L. Greene
49	A Feasibility Study of Liquid Rocket Engine Combustion Diagnostics	Dr. William M. Grissom
50	A Unified Approach to the Linear Camera Calibration Problem	Dr. William I. Grosky
51	Measuring Production Rate in Aircraft Repricing Models	Dr. Thomas R. Gulledge
52	Estimation of Relative Risk in Epidemiological Studies	Dr. Ramesh C. Gupta
53	Development of a Rule-Based Expert System for Damage Assessment of Air Force Base Structures	Dr. Fabian C. Hadipriono
Volume II		
54	Simulation of the Cardiac Conduction System	Dr. Frank O. Hadlock
55	A Framework of an Optimum Synthesis Environment for the Hydrocode EPIC-2	Dr. Prabhat Hajela
56	Modeling of Human Body Movement	Dr. Patrick R. Hannon
57	Fields of a Slot Antenna on a Half-Space Fed by Coplanar Waveguide Using the Method of Moments	Dr. Donald F. Hanson

58	Effect of Low Frequency Vibration on Bone Remodelling in the Rhesus Os Calcis	Dr. Gerald F. Harris
59	Mental Rotation and Perspective-Taking Skills In Pilots and Non-Pilots	Dr. Edward J. Hass
60	Revitalization of Operations and Controls for the Turbine Engine Test Cells at the Arnold Engineering Development Center	Dr. Doyle E. Hasty
61	Operation of the Electron Ion Momentum Transfer Instability Mechanism in Moderately Dense Plasmas	Dr. Michael A. Hayes
62	Evaluation of Several High Strength Composite Conductors	Dr. James C. Ho
63	The Locally Implicit Method for Computational Aerodynamics	Dr. Peter F. Hoffman
64	Procedures for Efficiently Extracting the Knowledge of Experts	Dr. Robert R. Hoffman
65	Fluorescent Dye Binding Analysis for the Identification of Asbestos	Dr. Clifford C. Houk
66	Thrust Computing System	Dr. Ming-Shing Hung
67	Estimation and Discrimination Procedures for a New Measure of Maintainability/Reliability	Dr. John M. Jobe
68	Design Synthesis of Nonlinear Systems	Dr. Glen E. Johnson
69	Analysis of FPS Tracking Radar for Error Reduction and Modeling	Dr. Jeremy C. Jones
70	Organophosphate Inhibitors: Repeated Low Dose Effects of Diisopropyfluro-phosphate on Serotinin Receptors in Rat Cortex	Dr. Betty R. Jones
71	Optimal Filtering	Dr. Marvin S. Keener
72	A Preliminary Study for Centrifuge Model Testing of Semihardened Concrete Arches	Dr. Yong S. Kim

73	The Synthesis of Fluorodinitroethy-lnitraminoalkyl Nitrates and Compatibility Studies of GAP-Nitrate and TAEl	Dr. Joel R. Klink
74	Reliability in Satellite Communication Networks	Dr. Stephan E. Kolitz
75	Investigation of Vapor Deposited Aluminum Alloy Films	Dr. Philipp G. Kornreich
76	Modification of Priority Handling Algorithm in the Integrated Node Network	Dr. Mou-Liang Kung
77	Ability, Experience and Task Characteristic Predictors of Performance	Dr. Charles E. Lance
78	Multiaperture Optical Systems and Neural Networks Capable of the Detection of Motion, Speed, Direction and Distance	Dr. David I. Lawson
79	Experimental Design and Transparency Durability Prediction	Dr. Paul S.T. Lee
80	Ion-Molecule Reactions of H_2O^+/H_2O, N_2^+/CO_2, and N^+/CO_2	Dr. C. Randal Lishawa
81	Study of Rutting of Asphalt Pavement Under High Tire Pressure and Temperature	Dr. Cheng Liu
82	Selected Spectral Studies of the Sun	Dr. James C. LoPresto
83	Visual Problem-Structuring and Hemispheric Processes of the Human Brain	Dr. Stephen L. Loy
84	Review and Evaluation of a Refueling Capability Assessment Model	Dr. Nancy I. Lyons
85	Statistical Pattern Recognition Modelling of Visual Perceptions	Dr. Robert L. Manicke
86	An Experimental Design to Verify the AFGL FASCOD2 for Water Vapor and Carbon Dioxide at Low Temperatures and Pressures	Dr. Arthur A. Mason
87	The Determination of Lead in Blood	Dr. Curtis W. McDonald

88	Compatibility of Reinforcement and Matrix Phases in Composite Materials for High-Temperature, Aerospace Applications	Dr. Gopal M. Mehrotra
89	Empirical Confidence Intervals for a Validity Coefficient Under Range Restriction: An Application of the Bootstrap	Dr. Jorge L. Mendoza
90	Human Factors Analysis of a Micro-Computer-Based Maintenance System for Advanced Combat Aircrafts	Dr. Shreenivas Moorthy
91	Evaluation of a Computer Model to Predict Thermal Retinal Damage from LASER Radiation	Dr. Mary L. Morton-Gibson
92	Chlamydomonas Phototaxis as a Simple System for Testing the Effect of Drugs on Vision	Dr. Rex C. Moyer
93	An Investigation of the Utility of Computational Fluid Dynamics in the Prediction of Structural Active Cooling	Dr. V. Dakshina Murty
94	A Model for a Coordinated System of Parallel Expert Systems for Autonomous Satellites	Dr. Richard W. Nau
95	CO_2 (001) Vibrational Temperatures in the 50 to 150 KM Altitude Range	Dr. Henry Nebel
96	A Study of the Finite Element Method in Limited Area Weather Prediction Modeling	Dr. Robert M. Nehs
97	Issues Related to Lithium and Lithium-Hydride Thermal Storage Spheres	Dr. Douglas L. Oliver
98	A Network Tutor Based on the Heuristic of Polya	Dr. Philip D. Olivier
99	Oxidative Stability and Related Studies of Silahydrocarbons	Dr. Harvey L. Paige
100	Cleansing of Bone-Marrow by Lymphokine Activated Killer Cells (LAK-Cells)	Dr. Parsottam J. Patel
101	All-Aromatic Rod-Like Polymers Based on Intramolecular Cycloadditions: Model Compound Study	Dr. Robert A. Patsiga

102	Computer Software Executable Image Efficiency in Real-Time LIDAR Applications	Dr. Martin A. Patt
103	A Biomechanical Study of Anthropomorphic Head-Neck Systems	Dr. Jacqueline G. Paver
104	Automatic Program Generation from Specifications Using PROLOG	Dr. Alexandru A. Pelin
105	Electrochemistry in Room Temperature Molten Salt Systems	Dr. Bernard J. Piersma
106	Effects of Acceleration Stress Upon Blood Lipid Levels	Dr. Leonard Price

Volume III

107	The Inter-Site Communication Services Required by a Distributed Operating System	Dr. Craig G. Prohazka
108	A Computer-Aided Method of Designing Control Systems Incorporating Aircraft Flying Qualities	Dr. L. Rai Pujara
109	Infrared to Visible Light Conversion in Rare Earth Doped Heavy Metal Fluoride Glasses	Dr. Richard S. Quimby
110	Optimization of Actively Controlled Structures Using Multiobjective Programming Techniques	Dr. Singiresu S. Rao
111	Development of a Rapid and Sensitive Assay Procedure for the Detection of the Protozoan Parasite Giardia lamblia in Drinking Water Supplies	Dr. Ralph J. Rascati
112	State Variable Model of the Cardiovascular System and a Controller Design for an Anti-G Suit	Dr. Kuldip S. Rattan
113	A Comparison of Two Mathematical Systems for a Standard Image Algebra	Dr. Barbara S. Rice
114	Multi-Echelon Inventory Models for EOQ Items	Dr. Dan B. Rinks
115	Preliminary Development of a Global Positioning System Package for use in Determining Exact Position of AFGL Research Balloons at Precise Time	Dr. William P. Robey

116	Kinetic Processes in Advanced Alloys	Dr. Kenneth C. Russell
117	Computer Modeling of Infrared Signatures	Dr. Sally A. Sage
118	Swirling Flows in Dump Combustors	Dr. Mo Samimy
119	Results of a Brief Investigation of Two Alternative Energy Systems	Dr. John F. Schaefer
120	Raman Spectroscopic Investigations on Group IIIA Doped Silicon Crystals	Dr. James R. Schneider
121	An Intentional Tutor	Dr. Richard M. Schori
122	Polynuclear Aromatic Hydrocarbons in Particulate Turbine Engine Exhaust and from Combustion of Single Compound Fuels	Dr. William D. Schulz
123	Systems Effectiveness Concerning Vulnerability of Hardened Targets to a Variety of Weapons	Dr. Meckinley Scott
124	Resolution of Laser Beam Intensity Spots by the Target Plate Measurement Technique	Dr. Martin A. Shadday
125	Propagation of Lamb Waves in Fibrous Composite Materials	Dr. Nisar Shaikh
126	Pharmacokinetics of Certain Substances of Abuse: A Review of the Literature	Dr. Dolores C. Shockley
127	Contrast Sensitivity at Low Luminance Levels	Dr. William D. Shontz
128	Aerodynamic Parameters for a Rapidly Pitching Airfoil	Dr. William D. Siuru
129	Aircraft Sortie Effectiveness Model	Dr. Boghos D. Sivazlian
130	Combustion Under Supercritical State and Influence of Radiation on Droplet Combustion	Dr. Siavash H. Sohrab
131	X-Ray Topographic Characterization of Si and GaAs	Dr. Stuart R. Stock
132	Assessment of Maximum Entropy Method Software by Operation on Interferograms of Model Spectra	Dr. James E. Sturm

133	AI and Large-Scale Systems Approaches to Enhanced Situation Awareness in Missile Warning Systems. (f.w)	Dr. Edgar C. Tacker
134	Feasibility Investigation of Single-Step Nitrations of Organometallics by Nitronium Triflate	Dr. Nicholas E. Takach
135	Dissipation of Plasma Cloud Generated by Third Stage Rocket Separation	Dr. Arjun Tan
136	Surface Roughness Effects on Heat Transfer and Skin Friction	Dr. Robert P. Taylor
137	Atmospheric Modeling for Operational Tactical Decision Aid	Dr. Ken Tomiyama
138	Vigilance Behavior of Military Personnel: A Study of Individual Differences	Dr. Phillip Tomporowski
139	An Investigation of Unsteady Vorticity Production by a Pitching Airfoil	Dr. Timothy R. Troutt
140	A Study of Turbulence Models for Predicting Aero-Optic Interactions	Dr. C. Randall Truman
141	Analysis and Fate of Organic Components of Aqueous Film Forming Foams	Dr. Roy M. Ventullo
142	Comprehension and Cohesion of Text	Dr. Doris J. Walker-Dalhouse
143	Effect on Hyperoxia on the Permeability of the Blood-Brain Barrier in the Rat	Dr. Donald W. Welch
144	Multimodal Information Exchange for Individual and Group Problem Solving	Dr. Albert R. Wellens
145	Preliminary Study of an Optical Implementation of the Conjugate Gradient Algorithm	Dr. Stephen T. Welstead
146	Assessing Cognitive Skills Through A Supervisory Control Simulation	Dr. Shih-sung Wen
147	Hyperbaric (3ATA) Oxygen 100% Therapy as an Adjuvant in the Treatment of Resuscitated (Brooke Formula) Guinea Pigs' Burn (3^{o}, 50 BSA) Shock	Dr. Stanley J. Whidden
148	Analytical Computer Modeling of the NPN BICFET Device	Dr. Dennis W. Whitson

149	Women in the Workforce: Occupational Pulmonary Disorders	Dr. Shirley A. Williams
150	Mechanisms of Chromatic Contrast	Dr. Billy R. Wooten
151	Continuum Analysis of Low Pressure Tube and Nozzle Flows	Dr. Daniel W. Yannitell
152	Chemical Kinetics of High Temperature Air for Mach 5 - 14 Flight	Dr. Tsun-wai G. Yip
153	Physiological Correlates of Behavioral Performance on the Mathematical Processing Subtest of the CTS Battery	Dr. Robert L. Yolton
154	Application of Finite Element Analysis to Two Disparate Structural Problems: Thermomechanical Coupling and Optimal Sizing of Truss Members	Dr. Richard W. Young
155	(M,N)-Approximation: A System Simplification Method	Dr. Amjal Yousuff
156	The Compound Eye: An Introduction to the Variety of Visual Capabilities, Goals, and Approaches Found in the Class Insecta	Dr. David D. Zeigler
157	Microwave Impedance Matching for Optical Devices	Dr. Henry Zmuda
158	Evaluation and Analysis of VHSIC Software Tools	Dr. George W. Zobrist

1986 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT SUMMER SUPPORT PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

The Inter-Site Communication Services Required by a Distributed Operating System

Prepared By: Craig G. Prohazka

Academic Rank: Assistant Professor

Department and University: Department of Electrical Engineering, University of Lowell

Research Location: Rome Air Development Center
Directorate of Command and Control
Distributed Systems Branch

USAF Researcher: Thomas F. Lawrence

Date: 10 September 1986

Contract No.: F49620-85-C-0013

The Inter-Site Communication Services Required By a Distributed Operating System

by

Craig G. Prohazka

ABSTRACT

This report identifies the inter-site communication services required by a typical distributed operating system. These services are found to be the abilities to:

a. Transmit a datagram

b. Broadcast a datagram

c. Transmit a datagram to all of a site's neighbors

d. Establish a virtual circuit

e. Establish a full-duplex virtual circuit

f. Establish virtual circuits to all other sites

ACKNOWLEDGMENTS

I would like to thank the Air Force Systems Command and the Air Force Office of Scientific Research for sponsorship of my research. Above all, I would like to thank Tom Lawrence and Tony Newton at RADC for guiding my research and, in particular, for many insightful conversations. The conscientious work of Alan Baginski, who accompanied me to RADC under the Graduate Student Summer Support Program, was greatly appreciated.

I. INTRODUCTION

I received my PhD from the University of Massachusetts at Amherst in the field of computer networking. I am currently a professor at the University of Lowell in Lowell, Massachusetts, where I specialize in computer engineering, specifically, distributed computer systems. At Rome Air Development Center (RADC) I was assigned to investigate distributed operating systems, a topic of great interest to me. My research at RADC was aided by Alan Baginski, a graduate student who accompanied me to RADC under the Graduate Student Summer Support Program.

II. OBJECTIVES OF THE RESEARCH EFFORT

a. Our work this summer was an investigation of the communication services required in a distributed operating system (DOS). Our approach was to first determine the computational services a DOS should provide, then to propose an appropriate DOS structure, and finally to identify the communications services our DOS requires.

b. Before beginning this task, let us motivate the investigation.

III. BACKGROUND

a. Recently, interest in distributed computer systems has been high. Distributed systems differ from centralized ones in that:

(1) Communication delays may be much greater than typical computation times.

(2) The probability of failure of a communication link is not negligible.

b. Interest in distributed systems has been heightened by the steadily declining price of computer hardware. The price has dropped to the level at which significant computational power may be dedicated to an individual. This development benefits the user by providing him convenient access and so results in a tendency for computer systems to become distributed.

IV. DOS COMPUTATIONAL SERVICES

a. This section identifies the computational service which a DOS should provide; for only by analyzing these computational services can we determine the required communication services. Following the DOS computational service specification, we propose a DOS structure which we believe has typical communication needs. Finally, we identify these communication needs.

b. We assume that the DOS runs on top of the constituent operating systems (COSs) controlling the different sites. Before specifying the services a DOS should provide, we must divide up operating system functions between the DOS and the COSs. It seems reasonable to relegate all operating system functions requiring inter-site communication to the DOS and all such functions requiring only intra-site communication to the COSs.

c. In addition, we use the following principle to guide the DOS structure design: a DOS should allow applications to easily take advantage of the distributed nature of the computer system it controls. The potential advantages of a distributed system include the sharing of data over a wide area, parallel processing, and fault tolerance. Thus, a DOS should offer database, entity, and fault management services in order to allow an application to easily take advantage of the distributed system. We will see that all these services require inter-site communication and so are properly assigned to the DOS. Additionally, because a process at one site may wish to use a resource at another, a DOS should offer resource management services. Finally, for military applications, a DOS should provide security management services. In this report, we assume that the DOS does offer security management.

d. We propose that the DOS consist of five routines, called managers. Each manager implements one of the services types listed above.

e. The next section describes a variety of communication facilities supported by standard communication protocols. The following section presents the inter-site communication services necessary for our DOS to implement its services.

V. SOME STANDARD COMMUNICATION FACILITIES

a. The first communication facility supported by many communication protocols is the datagram. It is a packet of bits, all with a common source and a common destination. A datagram is moved through the system as an isolated unit. Included in the datagram are fields which specify the source and destination.

b. A virtual circuit, on the other hand, is a path from source to destination over which a sequence of packets is sent. These packets all follow the same path and so arrive at the destination in the same order as that in which they left the source. The establishment of a virtual circuit requires overhead: for example, the creation of routing tables to direct the packets plowing along each virtual circuit as they leave each intermediate site. However, the virtual circuit is a more efficient communication facility than the datagram for transmitting a large amount of data.

c. The third communication facility is the broadcasting of a datagram. The broadcasting of a datagram is its transmission to all sites of the system.

d. We now present the communication services needed by our DOS.

VI. THE SECURITY MANAGER (SM) Before determining the SM communication needs, we must specify SM operation in more detail. We do that by describing the security policy and security mechanism of our DOS.

a. POLICY

(1) The purpose of a system's security policy is to define the meaning of security within the system. The policy considered in this report is the following. The entities comprising a computer system are grouped into three categories: action, data, and hardware. An action entity (AE) access a data entity (DE) provided by a hardware entity (HE). Action entities include processes and human users. Data entities are files. Hardware entities include main memory, disk drives, tape drives, and communication links.

(2) We use the following shorthand notation. A C data entity (C DE) is a data entity with classification C. Similarly, a C hardware entity (C HE) is a hardware entity with classification C.

(3) Our security policy is the following:

(a) An AE should know of the existence of a DE or HE only if it has a need-to-know.

(b) The classification of a DE should be equal to the classification of the most sensitive data it contains.

(c) A C1 DE should be stored in, or transmitted through, a C2 HE only if C2 $\geq$ C1.

(d) No C1 DE should contain any information about a C2 DE for any C2 $>$ C1. In other words, an AE may receive an access right to a C DE only if: (1) the AE can access the DE using only HEs with classifications greater than or equal to C, and (2) the AE has a need-to-know for the DE. An AE may receive an access right to an HE only if the AE has a need-to-know for the HE.

b. MECHANISM: CAPABILITIES

(1) The function of a system's security mechanism is to control the access of AEs to DEs and HEs. The mechanism considered in this report is capabilities. Specifically, an AE is given a capability to a DE or HE only if the AE has an access right. A capability is a secure pointer. We assume that capabilities are stored in the SM memory partition, which is located in the DOS memory. Thus, an AE may access a capability only via a system call to the DOS.

(2) Figure 1 illustrates the SM memory partition, which contains all the SM routines and the capabilities of all local AEs. SM routines are stored in a main memory partition of the highest possible classification. Note that the capabilities are divided into classifications according to the classification of the DE or HE which the capability points to. This division serves to shorten the search an SM must perform to locate the capability to a

desired DE or HE. Capabilities pointing to C1 entities may be stored only in main memory of classification C2, for any C2 $\geq$ C1. This requirement follows from policy d. Within each classification, capabilities are grouped according to the AEs which possess them. Thus, even if two or more AEs have capabilities to exactly the same entities, these capabilities appear only once. This is made possible by grouping the AEs into a domain which is a set with identical capabilities. Figure 2 illustrates the domain concept.

c. CAPABILITY INDIRECTION In order to allow easy capability revocation and data entity migration, capabilities point indirectly to entities. That is, a capability points to an entity-locator (EL), which is a pointer aimed at the entity itself. Figure 3 illustrates the capability format. The EL is stored in DOS memory at the site containing the entity, as shown in Figure 4.

d. NEED-TO-KNOW

(1) Recall that our security policy demands that an AE have access only to DEs for which it has a need-to-know. This policy is enforced by informing an AE of the existence of a DE or HE only if it has a need-to-know. A human user's need-to-know is determined by the security officer. It is from this source that the human learns of the existence of a DE or HE.

(2) A process' need-to-know is established by

(a) a human user, if the process is created by a human, or

(b) the parent of the process, if the process was forked by another. The SM must establish and enforce rules for assigning needs-to-know to forked processes.

e. CAPABILITY INITIALIZATION

(1) When a process is created, it is assigned to a domain which gives the process a capability to each DE and HE it will need to access. As was true for need-to-know, the domain is chosen either by a human user (if the process was created by a human) or by the process' parent (if it was forked by another

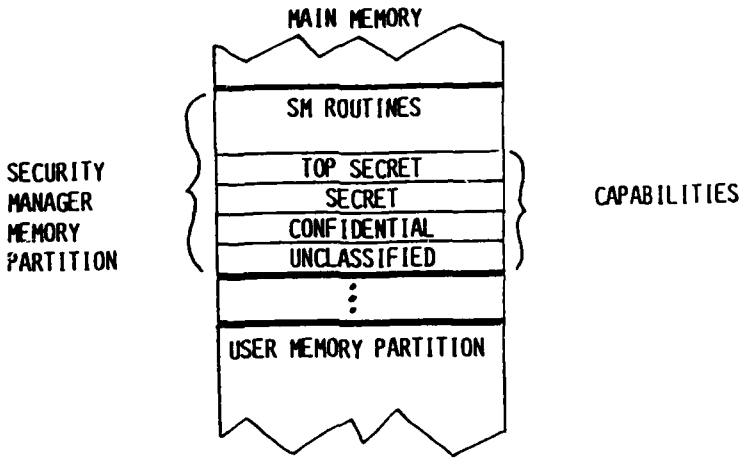


Figure 1. The security manager memory partition.

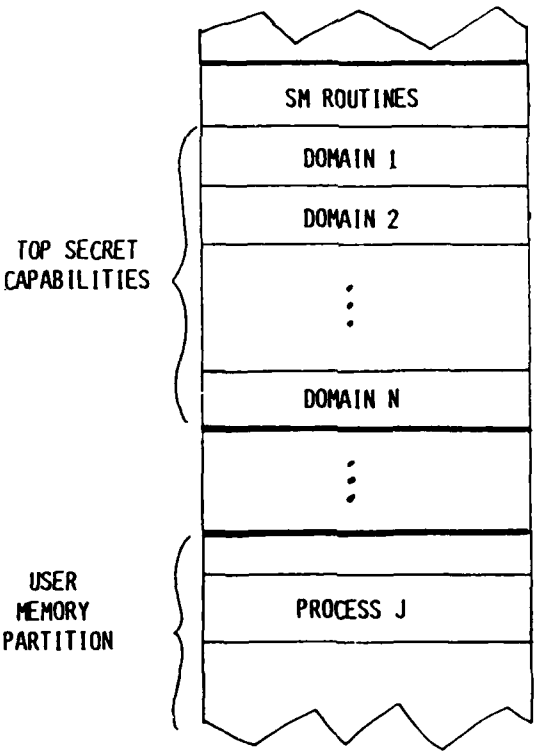


Figure 2. Domains.

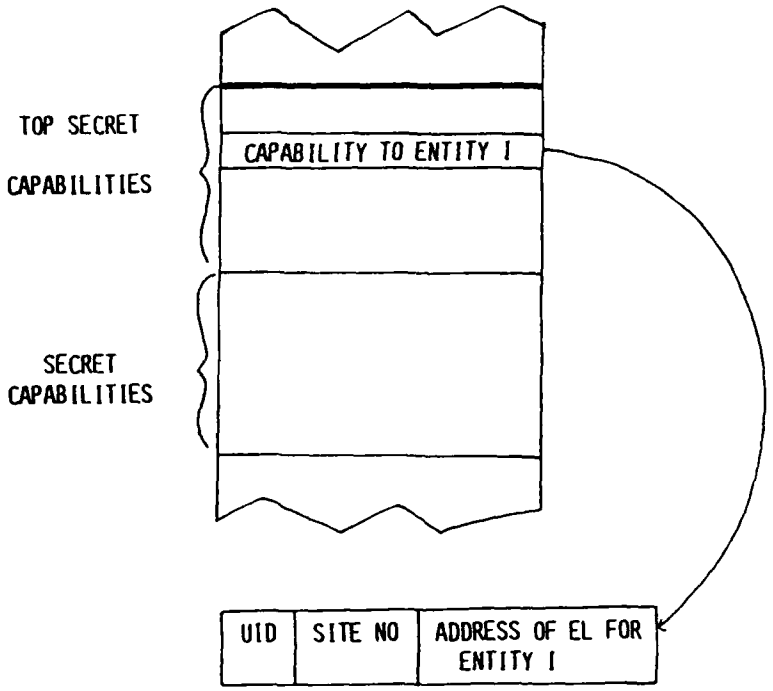


Figure 3. The format of a (top secret) capability.

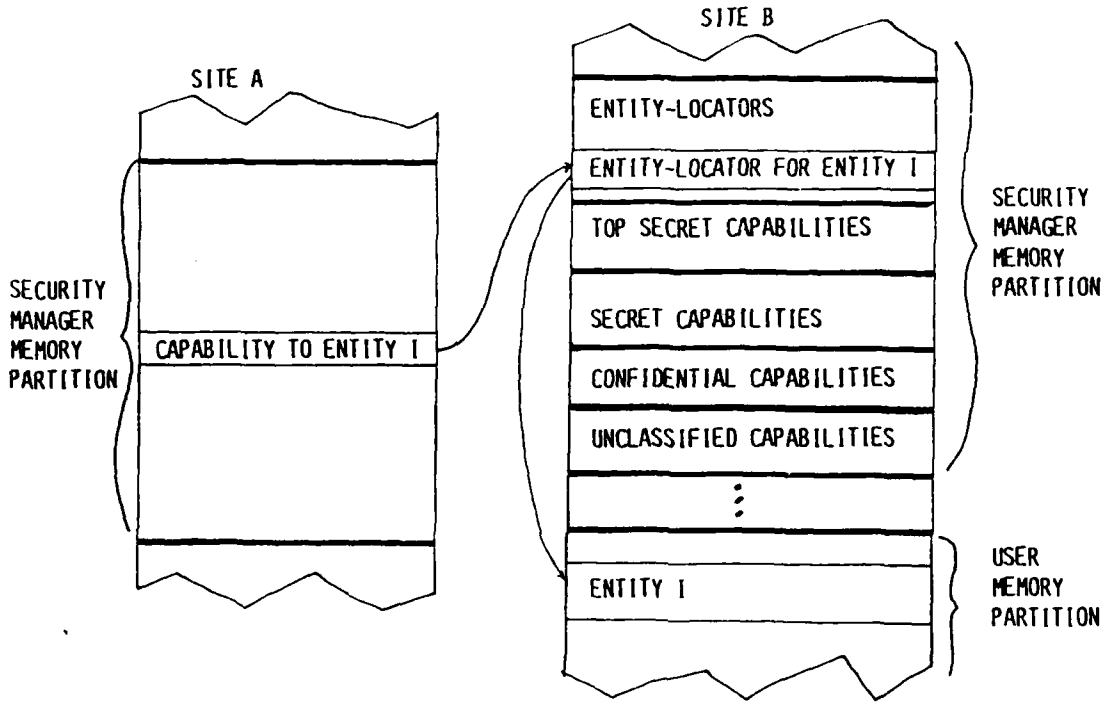


Figure 4. Capability indirection.

process). The SM specifies the rules for this initial capability assignment. Similarly, when a human user logs into the system, he is assigned to a domain.

(2) Each SM maintains a list of capabilities to all local DEs and HEs, stored in DOS memory. When the SM gives an AE a capability to a C DE or a C HE, it simply copies this capability from the list to the appropriate domain in the SM memory partition dedicated to storing classification C capabilities.

f. UPDATING CLASSIFICATIONS AND CAPABILITIES A change in the classification of a DE or HE is reflected in its capabilities. Specifically, if the change is a classification increase, the capabilities of some AEs may need to be revoked. Also, all AEs must be informed of the change, so that they can specify the correct classification in any DOS call to access the DE or HE. Finally, all capabilities to the DE or HE must be moved to the SM memory partition corresponding to the new classification.

g. SM FUNCTIONS AND REQUIRED COMMUNICATION SERVICES

(1) There are three SM functions which require inter-site communication:

(a) Changing the classification of a DE or HE

(b) Deleting a DE

(c) Migrating a DE

(2) We now describe the details of the implementation of these functions. At the same time, we identify the required communication services.

(3) First, suppose the classification of DE D is increased. Assume that AE A had a capability to D. Immediately, A's capability must be revoked to prevent possible compromising of D's data. But how can this be done in spite of the possibly long propagation delay experienced by a message from D's site to A's site? At this point, the purpose of the EL can be understood. All capabilities to D can effectively be revoked at D's site by overwriting the EL

pointing to D! This can be accomplished without sending a long delay message. Next, D's new classification is broadcast to all foreign sites. Typically, the new classification message is short enough to be sent as a datagram. After A's SM receives the datagram, A's SM can determine whether A will be returned a capability to D. If not, A's capability, which points to D's EL is deleted. If so, a capability to D is written into the appropriate SM main memory partition. It points not to the old EL, which as been overwritten, but to a new EL which, in turn, points to D.

(4) Now suppose that the classification of DE D is decreased. Then D's new classification is broadcast to all sites. The SMs at these sites move their capabilities to the appropriate SM memory partition.

(5) Second, suppose that DE D is deleted. Then DE's EL is immediately overwritten with a special address which informs any AE which references D that D has been deleted.

(6) Third, suppose DE D is migrated. The destination SM creates an EL pointing to the new replication of D and broadcasts to all SM a datagram containing the address of the newly created EL. After enough time has elapsed to allow all SMs to modify their capabilities to D, the original version of D and its EL are overwritten.

(7) In summary, the only inter-site communication service required by an SM is the broadcasting of a datagram.

VII. THE RESOURCE MANAGER (RM)

a. This section identifies the inter-site communication services required for our resource managers (RMs) to perform the functions demanded of them. First we specify these functions. Then we propose a particular resource management system to perform the functions. Finally, we identify the communication services this system requires.

b. Our resource management system operates as follows. All the resources at a particular site are managed by that site's resource manager instance. This RM decides which of its resources to allocate to a particular AE. An AE which wishes to access a foreign resource, generates a system call to the local RM which, in turn, sends a request to the foreign site's RM. This request should be short enough to be sent as a datagram. Thus, the first communication service required by an RM is the ability to send a datagram to a particular site. The destination RM uses the AE's priority to help decide whether or not to make the allocation. The RM assigns a priority to each AE at its site based upon a set of criteria established at system generation.

c. The RM at each site maintain statistics on the load presented to its resources. Also, each RM periodically receives from all foreign RMs statistics on all resources at their sites. Typically, the set of statistics for each site is much larger than a datagram. Hence, a virtual circuit should be used. So, the second communication service required by an RM is the ability to establish virtual circuits to all other sites.

d. Since all RMs receive load statistics for local and foreign resources, all the RMs become aware of any possible excessive degree of contention among AEs for resources. If this contention increases to an unacceptable level, the RMs begin to reduce the number of AEs by inhibiting new process creation and new user log-in. Each RM is in charge of limiting the number of AEs at the local site. Thus, no additional intersite communication is needed to implement this function.

e. Finally, our RM uses a hierarchical deadlock detection algorithm (1). This algorithm is called hierarchical because a tree of controllers is used to detect deadlocks. The leaf controllers are all the RMs. Each leaf controller is, in turn, controlled by a higher level controller, which is, in turn, controlled by a higher level controller, etc. The higher level controllers are actually RMs elected to that function. Thus, our resource management system requires the election of non-leaf controllers.

f. The election is implemented through the use of the Bully Algorthim (2). This algorithm assumes that each RM has been assigned a priority. Then it

elects the highest priority as-yet-unelected RM which is currently operating to fill any vacant higher-level controller position. The Bully Algorithm is presented in Figure 5. It requires the ability to broadcast a datagram.

g. The hierarchical deadlock detection algorithm also requires the maintenance of wait-for-graphs. A wait-for graph is constructed as follows. Each vertex vi represents a process pi. A directed edge (vi, vj) connects vi to vj if, and only if, pi is waiting to use a resource currently allocated to pj. A local wait-for graph is maintained at each RM and a regional wait-for graph is maintained at each higher level controller. The latter graphs are needed to express deadlocks involving more than one site. In any case, the RMs must exchange messages to maintain the wait-for graphs. Typically, these messages are short enough to be sent as datagrams. So, the fourth communication service required by an RM is the same as the first; namely, the ability to send a datagram.

h. Suppose the deadlock detection algorithm actually does detect a deadlock. Then the RMs must first agree upon the AE to preempt from its allocated resources, and second, actually preempt it. A modified version of the Bully Algorithm is used to solve the first part of this problem. As was true for non-leaf controller election, this Bully Algorithm implementation assumes that a priority has been assigned to each AE. However, to solve the present problem, the Bully Algorithm must be modified to elect the lowest priority AE as the one to be preempted. In spite of the modification, the use of this algorithm requires the same communication service as the original algorithm: namely, the ability to broadcast a datagram. Finally, the preemption of the chosen AE may be accomplished by a command from the local RM and so requires no inter-site communication.

i. In short, the communication services required by our RMs are the following:

(a) Send a datagram

(b) Establish virtual circuits to all other sites

(c) Broadcast a datagram

1. Process i sends a message to the current controller.

2. It waits for the time-out period, T, for an acknowledgment.

3. If process i receives no ACK, it concludes the controller is down.

4. Process i broadcasts a message to elect itself the new controller. The message specifies its priority.

5. Any process with higher priority receiving this election message responds.

6. If process i receives no such response within time period T, it elects itself controller and broadcasts this fact.

7. If process i receives such a response, it waits for time period T for an election message broadcast from the higher priority process.

8. If no such election message is received, process i concludes that the higher priority process just failed, elects itself controller, and broadcasts this fact.

9. If such an election message is received, process i waits for a message informing it of the newly elected controller.

FIGURE 5. The Bully Algorithm

VIII. THE ENTITY MANAGER (EM)

a. An EM must perform the following functions. First, it must be able to read a DE at a foreign site. In general, this function is performed in two phases. First, the closest replication of the DE is found. Here, the closest replication is the one which can be accessed with the smallest delay. Then this replication is actually read. The EM locates the closest replication by broadcasting to all foreign EMs a message requesting those EMs whose sites have a replication of the DE to respond. We assume that an EM does not know the location of all the replications of a particular DE. For this reason, the EM must broadcast its message to all other EMs. Since this message contains so little information, it may be sent as a datagram. The EM originating the message can deduce the location of the closest replication by noting the source of the first response it receives. Once the EM has located the closest replication, the EM reads it. Typically, the amount of data the DE contains is much greater than the length of a packet. Thus, the EM should establish a virtual circuit to the DE's site, and finally, read the DE. So, the performance of the first function, namely reading a DE at a foreign site, requires that an EM be able to broadcast a datagram and establish a virtual circuit.

b. The second function an EM must perform is to create a DE at a foreign site. This may be accomplished simply by sending a message to the foreign site instructing it to create the DE. The message may be sent as a datagram.

c. The third EM function is to copy a DE to a particular foreign site. For a DE of typical length, this is best accomplished by establishing a virtual circuit to the destination site.

d. The fourth EM function is to update all replication of a DE. This is done by broadcasting a message containing the update to all sites. Typically, the update will be short enough to be sent as a datagram.

e. The fifth function is to delete a DE at a particular foreign site. To perform this function, an EM need only send a datagram containing a delete command to the appropriate site.

f. The sixth function is to delete all replications of a DE. To perform this function, the EM broadcasts a delete command to all sites.

g. The seventh function is to fork an AE at a foreign site. A datagram addressed to the site should suffice.

h. Finally, the eighth function is to terminate an AE at a foreign site. Again, a datagram should suffice.

i. In summary, an EM requires the following communication services:

(1) Broadcast a datagram

(2) Establish a virtual circuit

(3) Send a datagram

IX. THE DATABASE MANAGER (DM)

a. A DM must perform the following functions. First, it must control concurrent access by many AEs to the same DE. Second, it must maintain the consistency of multiple replications of the same DE. Third, it must manage the recovery from partition of a system containing mulitple replications of a DE. That is, if the partition caused inconsistency between different replications, then the DM must resolve the inconsistency upon system reconnection.

b. Before identifying the communication services required by the DMs, we must propose a database management system.

c. Our database management operates as follows. Concurrent access to DEs is controlled by the use of the two-phase protocol.

d. Recall that our DM uses the two-phase protocol for concurrency control. This protocol demands that, in any transaction, all locks precede all unlocks. Any schedule for a set of transactions which obey this protocol is guaranteed to be serializable and hence, correct, in spite of possible concurrency. Note that

this technique for concurrency control requires no inter-site communication. The DM need only ensure that data items are locked and unlocked properly, as demanded by the two-phase protocol.

e. The next DM function is the maintenance of consistency between the replications of a DE. Our DM performs this function in the following way. If a data item in a replicated DE is to be updated, this data item is first locked in all replications, then the items are updated, and finally, the items are unlocked. The performance of this function requires the following communication services. First, the lock command must be broadcast to all sites. This command should fit into a datagram. The updating of the data item requires the broadcasting of a datagram. Finally, the unlock operation also requires the broadcasting of a datagram.

f. The next DM function is recovery from a partition. Our DM performs this function in the following way. Associated with each DE and located at the same site, is a journal containing the history of modifications to the DE since the last back-up. Suppose that our system becomes partitioned in such a way that two replications of a particular DE cannot communicate with each other. If an update is made to one replication, it will not be made to the other. Inconsistency results. The inconsistency cannot be resolved until the system is reconnected. At that time, however, the journals are exchanged and consistency reestablished. Our DM obeys the two-phase commit policy: namely, that a DE update is first written to the journal and then written to the database.

g. Consistency may be reestablished after system reconnection following partition as described next. Each DM transmits a message for each of his DEs. This message contains the DM's site number, the DE's UID, and the latest journal entry for that DE. Suppose a DM discovers that for one of its DEs, the latest journal entry received from a foreign site is different from its latest journal entry. Then the DMs at the two sites exchange the two journals and update their replications to reestablish consistency.

h. The following communication services are required to perform this function. First, the maintenance of the journal requires no additional communication services. Second, after system reconnection following partition,

each DM must broadcast one datagram for each DM it has in common with other sites. If inconsistency is found, the journals must be exchanged. Typically, a journal is much longer that a datagram. Hence, full-duplex (two-way) virtual circuits must be established between all pairs of sites with inconsistent replications of the same DE.

i. The last DM function is the checking of DE replication consistency. If the above functions work properly, consistency should be maintained except during partitions. Nevertheless, it is reasonable to require that our DM have the ability to check consistency. This function may be performed by locking all replications of the particular DE being checked, copying them to a common site, and comparing them. The following communication services are required. The lock command must be broadcast to all DE replications. Then, the replications must be sent to the common site. Typically, a DE is very long compared with a datagram. Hence, all DMs with a replication of the DE must establish a virtual circuit to the common site.

j. In summary, our DM requires the following communication services:

(1) Broadcast a datagram

(2) Establish a full-duplex virtual circuit

(3) Establish a virtual circuit

X. THE FAULT MANAGER (FM)

a. An FM must perform the following functions. First, it must detect failed processors and system partition. This may be accomplished in the following way. Each FM periodically transmits to all its neighbors an I-am-up message. This message is, in turn, transmitted by any receiving FM to all its neighbors, etc. An FM transmits a particular I-am-up message only once. Flooding the system ensures that the I-am-up message is received by all FMs which are connected to the transmitting FM. The I-am-up mesage should fit into a datagram. Thus, the first communication services required by an FM are the ability to broadcast a datagram and the ability to transmit a datagram to all of

a site's neighbors. Suppose an FM does not receive an I-am-up message when it is due. It concludes that either the FM which should have sent the message is down, or system partition has occurred.

b. The second FM function is to restart the processes which had been running at the failed site. After an FM detects the failure, it attempts, by following the Bully Algorithm, to elect its site to be the site at which the processes are restarted. The communication services required to accomplish this are the same as those required for the other implementations of the Bully Algorithm: namely, the ability to broadcast a datagram. After the site has been chosen, process restart takes place. Process execution is reinitialed at the latest checkpoint, which is a snapshot of the state of an executing process. Of course, a process cannot be restarted at a different site unless a checkpoint is stored there before the failure of the original site. Thus, our FM periodically sends a checkpoint for each of its processes to all other FMs. The fact that an FM sends its checkpoints to all other FMs allows any FM to restart the processes at a failed or disconnected site. A checkpoint should fit into a datagram. Thus, the communication services required to accomplish this is the ability to broadcast a datagram.

c. In short, the communication services required by our FM are the ability to broadcast a datagram and the ability to transmit a datagram to all of a site's neighbors.

XI. CONCLUSIONS

Finally, condensing the results of the previous sections, we conclude that our DOS requires the following inter-site communication services:

(1) Transmit a datagram

(2) Broadcast a datagram

(3) Transmit a datagram to all of a site's neighbors

(4) Establish a virtual circuit

(5) Establish a full-duplex virtual circuit

(6) Establish virtual circuits to all other sites.

XII. RECOMMENDATIONS

Future work should address the design of communication protocols which efficiently support the required communication services.

XIII. REFERENCES

a. Menasce, D. and Muntz, R.R., "Locking and Deadlock Detection in Distributed Data Bases." IEEE Transactions on Software Engineering, 5 (1979) 195-202.

b. Garcia-Molina, H., "Election in a Distributed Computing System, "IEEE Transactions on Computers, 31 (1982) 48-59.

c. Peterson, J.L. and Silberschatz, A., Operating System Concepts, Addison-Wesley Publishing Company, 1985.

A COMPUTER-AIDED METHOD OF DESIGNING CONTROL SYSTEMS INCORPORATING AIRCRAFT FLYING QUALITIES

L.R. PUJARA

Dept. of Electrical Systems Engineering

Wright State University

Dayton, Ohio 45435

Sponsored by AFOSR and UES under Project #760 and contract No.F49620-85-C-0013

L. R. Pujara

A Computer-Aided Method of Designing Control Systems Incorporating Aircraft Flying Qualities

ABSTRACT

In this report, aircraft flying qualities have been made the main focal point for the design of control systems of an aircraft. First, "desired transfer functions" satisfying flying qualities criteria as in MIL-F-8785 are synthesized for the longitudinal direction controls for the F-16 and F-104 and then using the frequency matching technique, control systems are designed for these aircrafts such that augumented systems are "close" to the "desired transfer functions". For the lateral direction mode of an aircraft, a generic model of a "desired transfer function matrix" satisfying the flying qualities criteria is proposed. An illustrative example of a "desired transfer function matrix" for the A-7A in the lateral mode has been constructed.

ACKNOWLEDGEMENT

I would like to extend my sincere thanks to Mr. Frank George, Group Leader, Flying Qualities Group, Flight Control Division, Wright-Patterson Air Force Base, Ohio for having provided me with excellent working conditions. I would also like to acknowledge my thanks to Mr. Frank George, Mr. Bob Woodcock, Mr. Tom Gentry, Mr. Chuck Suchomel and Mr. Dave Leggett of the Flying Qualities Group for several useful and stimulating conversations regarding the subject matter of this report. Finally, thanks to AFOSR and UES for having given me this opportunity to participate in the Summer Faculty Research Program.

Introduction

Flight control systems developed for Air Force airplanes must satisfy several specifications, MIL-F-8785[1] being one of the principal ones. It is, therefore, important that flying qualities criteria are kept in focus while designing control systems for aircraft. In this report, the frequency matching technique developed by the author [1,2] is applied to design control systems for the F104 and F16 longitudinal dynamic modes so that the compensated system has desirable flying qualities. The existing software package has been updated to implement the technique. In addition to this, a significant first step has been taken in extending the frequency matching technique to the multivariable case. This has been accomplished by proposing a desired transfer function matrix for the augumented control system of an aircraft in its lateral mode. This desired transfer function matrix incorporates the flying qualities specifications and also takes advantage of the cross coupling between aileron and rudder.

In future research, it is proposed to develop a synthesis procedure for the multivariable case by extending the frequency matching method such that the unknown parameters of the controller will be solutions of linear algebraic equations. Much of the recent research work in the design of multivariable control systems is in the frequency domain [5,6,7,8,9,10] based primarily on optimal control techniques and singular values.

Longitudinal Flying Qualities

Ultimately, it is the pilot who assesses the flying qualities of an aircraft. This is quantified by correlating the pilot's Cooper-Harper scale

1 Military Specification, MIL-F-8785, Flying Qualities of Piloted Airplanes, Anon.

rating with aircraft dynamic parameters such as modal frequencies and damping or response-time constants. The longitudinal flying qualities depend on short period and phugoid responses. In the case of conventional fixed wing aircraft, the period of the phugoid response is very long. But, if the damping of the phugoid mode is small, the pilot will be forced to pay more attention to it. In general, a divergent phugoid is undesirable and should be avoided. Based on the pilot's evaluation, for the short period, a range of values of ξ_{sp} and ω_{sp} is obtained for different tasks to be performed by the aircraft.

This range is then split into "acceptable" and "unacceptable" groups. A typical "desired" transfer function for the longitudinal mode is

$$\frac{q}{\delta_e} = \frac{A\left[s+\frac{1}{T_R}\right]}{s^2+2\xi_{sp}\omega_{sp}s+\omega^2_{sp}} \qquad (1)$$

where q is the pitch rate and δ_e is the elevator* deflection. Appropriate values of A, T_R, ξ_{sp}, ω_{sp} are selected depending upon the aircraft and the mission to be performed. These specifications have been extensively formulated in MIL-F-8785C.

Frequency-Matching Method of Design

The frequency-matching method of designing a single input-single output control system [1,2] seems to be a natural framework for designing control systems to satisfy flying qualities. This is so because the technique is based

*Note: In a conventional airplane, the elevator provides pitch control. In some advanced aircraft, combinations of control effectors may generate an "equivalent elevator" pitch control.

on finding the unknown parameters of a controller by matching the frequency-responses of the compensated closed loop system and a desired transfer function. The flying qualities criteria could be incorporated in determining the "desired" transfer function. Another important feature of the frequency-matching technique is that the values of the controller unknown parameters turn out to be solutions of linear algebraic equations. This should be mentioned as quite an important point in developing a software package to implement the procedure.

For the sake of completeness, the SISO frequency-matching design method [1,2] is briefly reviewed.

Frequency-Matching Method

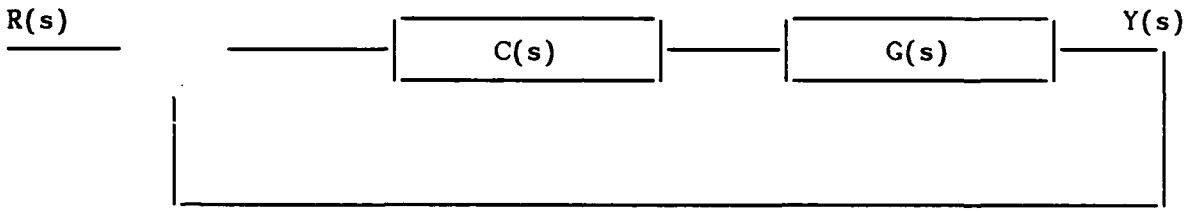


Figure 1

Consider the block diagram of a typical control system as shown in Figure 1 where G(s) is the plant and C(S) is a controller. It is desired to synthesize a controller C(s) so that the overall closed loop transfer function is "close" to a certain "desired transfer function" over a frequency-interval of interest. Suppose D(s) is a "desired transfer function" which has been synthesized to satisfy a given set of specifications - this could include flying qualities specifications.

Suppose

$$C(s) = \frac{\sum_{i=0}^{m} a_i s^i}{1+\sum_{i=1}^{n} b_i s^i} \quad , \qquad m \leq n \qquad (2)$$

where a's and b's are unknown constants.

Suppose

$$G(j\omega) = \frac{L_1(\omega)+jL_2(\omega)}{M_1(\omega)+jM_2(\omega)} \qquad (3)$$

and

$$D(j\omega) = \frac{N_1(\omega)+jN_2(\omega)}{Q_1(\omega)+jQ_2(\omega)} \qquad (4)$$

Then the frequency response of the compensated closed loop transfer function can be written as

$$F(j\omega) = \frac{\left\{\left[\frac{\sum_{i=0}^{m} a_i(j\omega)^i}{1+\sum_{i=1}^{n} b_i(j\omega)^i}\right]\left[\frac{L_1(\omega)+jL_2(\omega)}{M_1(\omega)+jM_2(\omega)}\right]\right\}}{\left\{1+\left[\frac{\sum_{i=0}^{m} a_i(j\omega)^i}{1+\sum_{i=1}^{n} b_i(j\omega)^i}\right]\left[\frac{L_1(\omega)+jL_2(\omega)}{M_1(\omega)+jM_2(\omega)}\right]\right\}} \qquad (5)$$

As in [1,2], the weighted mean square error between the frequency responses of $F(j\omega)$ and $D(j\omega)$ can be written as

$$E = \int_{\omega_1}^{\omega_2} \left| \left[Q_1(\omega)+jQ_2(\omega)\right]\left[L_1(\omega)+jL_2(\omega)\right]\left[\sum_{i=0}^{m} a_i(j\omega)^i\right] \right.$$

$$- \left[N_1(\omega)+jN_2(\omega)\right]\left\{\left[M_1+jM_2\right]\left[1+\sum_{i=1}^{n} b_i(j\omega)^i\right]\right. \qquad (6)$$

$$\left. \left. + \left[L_1(\omega)+jL_2(\omega)\right]\left[\sum_{i=o}^{m} a_i(j\omega)^i\right]\right\} \right|^2 d\omega$$

or

$$E = \int_{\omega_1}^{\omega_2} \left[\left[AA_1-BA_2-CB_1+DB_2\right]^2 + \left[A_1B+AA_2-DB_1-CB_2\right]^2\right] dw$$

where

$$A = L_1\left[Q_1-N_1\right] - L_2\left[Q_2-N_2\right]$$

$$B = L_1\left[Q_2-N_2\right] - L_2\left[Q_1-N_1\right]$$

$$C = N_1M_1 - M_2N_2$$

$$D = N_1M_2 + M_1N_2$$

$$A_1 = a_0 - a_2\omega^2 + a_4\omega^4 - \ldots$$

$$A_2 = a_1\omega - a_3\omega^3 + a_5\omega^5 - \ldots$$

$$B_1 = 1 - b_2\omega^2 + b_4\omega^4 - \ldots$$

$$B_2 = b_1\omega - b_3\omega^3 + b_5\omega^5 - \ldots$$

As in [1,2] minimization of E leads to $x = y^{-1}z$

where

$$
Y = \begin{bmatrix}
T_0 & 0 & -T_2 & 0 & \cdots & P_1 & S_2 & -P_3 & -S_4 & \cdots \\
0 & T_2 & 0 & -T_4 & \cdots & -S_2 & P_3 & S_4 & -P_5 & \cdots \\
-T_2 & 0 & T_4 & 0 & \cdots & -P_3 & -S_4 & P_5 & S_6 & \cdots \\
0 & -T_4 & 0 & T_6 & \cdots & S_4 & -P_5 & -S_6 & P_7 & \cdots \\
 & \cdots & & & & & & & & \\
 & \cdots & & & & & & & & \\
 & \cdots & & & & & & & & \\
P_1 & -S_2 & -P_3 & S_4 & \cdots & Q_2 & 0 & -Q_4 & 0 & \cdots \\
S_2 & P_3 & -S_4 & -P_5 & \cdots & 0 & Q_4 & 0 & -Q_6 & \cdots \\
-P_3 & S_4 & P_5 & -S_6 & \cdots & -Q_4 & 0 & Q_6 & 0 & \cdots \\
-S_4 & -P_5 & S_6 & P_7 & \cdots & 0 & -Q_6 & 0 & Q_8 & \cdots \\
 & \cdots & & & & & & & & \\
 & \cdots & & & & & & & &
\end{bmatrix} \tag{7}
$$

$$
X = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ \vdots \\ b_1 \\ b_2 \\ b_3 \\ b_4 \\ \vdots \end{bmatrix} \qquad Z = \begin{bmatrix} S_0 \\ P_1 \\ -S_2 \\ -P_3 \\ \vdots \\ 0 \\ Q_2 \\ 0 \\ Q_4 \\ \vdots \end{bmatrix} \tag{8}
$$

and

$$T_m = \int_{\omega_1}^{\omega_2} \omega^m (A^2+B^2) d\omega, \quad S_m = \int_{\omega_1}^{\omega_2} \omega^m (AC+BD) d\omega$$

$$P_m = \int_{\omega_1}^{\omega_2} \omega^m (AD-BC) d\omega, \quad Q_m = \int_{\omega_1}^{\omega_2} \omega^m (C^2+D^2) d\omega \tag{9}$$

Applications to Flying Qualities

Example 1

Consider the pitch rate/(elevator deflection) transfer function of an F-104 at M = 0.84 and h = 30,000 ft.

$$\frac{q}{\delta e} = \frac{(-17.8)(s+0.0144)(s+0.432)}{[s^2+1.12056s+12.1104][s^2+0.010404s+0.002601]} \tag{10}$$

From flying qualitites point of view, the desired transfer function D(s) has the following form:

$$D(s) = \frac{K(s+.432)}{s^2+2(\xi_{sp})\omega_{sp}s+(\omega_{sp})^2} \tag{11}$$

Using MIL-F-8785C, we pick $\xi_{sp} = 0.7$, $\omega_{sp} = 4$ rad/sec. as desired values for the F-104.

We pick K such that

$$\frac{K \cdot \left[\frac{z_\alpha}{g}\right]}{\left[\omega_{sp}\right]^2} = \frac{1}{3} \tag{12}$$

where $z_\alpha = v_o z_\omega$, v_o being the speed of the aircraft. From the aircraft data, $z_\omega = -0.504$, $v_o = 846$ ft/sec and, therefore, $z_\alpha = -426$

$$\therefore K = \frac{g \cdot \left[\omega_{sp}\right]^2}{3z_\alpha} = \frac{(32.3)(4)^2}{3(-426)} = -.40438185 \qquad (13)$$

$\therefore$ the desired transfer function D(s) is given by

$$D(s) = -\frac{0.40438185(s+0.432)}{s^2+5.6s+16} \qquad (14)$$

Using equation (6) with $\omega_1 = 1.0$ and $\omega_2 = 2.0$, the first order controller is given by

$$C(s) = \frac{0.001962411452s+0.01241724808}{0.4658282249s+1}$$

$\therefore$ the closed loop compensated transfer function is given by

$$\frac{0.074986s^4+0.507953s^3+0.212274s^2+0.002951s}{s^5+13.4377s^4+67.89249s^3+93.658292s^2+1.141985s+0.241048} \qquad (15)$$

Using the McFIT program for finding the lower order equivalent system with 30 equally spaced points over the frequency interval [0.1, 10], the equivalent system was found to be

$$\frac{6.9960276s+3.59116195}{s^2+9.04897875s+17.7889281} \qquad (16)$$

with a payoff function 4.04 which is quite satisfactory from flying qualities viewpoint. Thus, the control system augumentation for the F-104 done by the frequency matching method has desirable flying qualities.

Example 2

This example is related to the (pitch rate)/stick force of an F-16 at M = 0.8 and h = 20,000 ft. For the bare-airframe, the transfer function is given by

$$\frac{q}{F_s} = -\frac{21.13s(s+0.02598)(s+1.082)}{[s^2+2(0.1426)(0.05702)s+(0.05702)^2][s+4.496][s-2.372]} \times \frac{13}{[s+13]} \tag{17}$$

$$= -\frac{[274.69s^3+304.35103s^2+7.7216348s]}{[s^5+15.140262s^4+17.196687s^3-138.31389s^2-2.1994547s-0.4567531]}$$

where $\frac{13}{s+13}$ is an approximation for the actuator. Note also that it has an unstable pole at 2.372. Using the procedure similar to the one employed in the previous example and consulting MIL-F-8785C, we determine

$$D'(s) = -\frac{0.1(s+1.082)}{s^2+7s+25} \text{ as a}$$

desired transfer function for $\frac{q}{F_s}$.

To maintain the dimensional equivalence requirement [2] of the "desired transfer function" and the closed loop compensated system, we introduce a nondominant pole at -40 in the "desired transfer function" and also adjust the gain to keep the same steady state value. Thus, the "desired closed loop transfer function"

becomes

$$D(s) = -\frac{(4s+4.328)}{s^3+46.9s^2+300.8918s+995.672} \qquad (18)$$

Using equation (6) with $\omega_1 = 0.3$, $\omega_2 = 2.0$, a second order controller is given by

$$\frac{0.000277810136s^2+0.0001807087407s-0.002009433028}{0.03433431074s^2+0.323512251s+1} \qquad (19)$$

$\therefore$ the closed loop compensated system is given by

$$-\frac{2.2225168s^4+9.204348s^3+7.532304s^2+0.189516s}{\left\{s^5+26.934685s^5+252.869072s^4+1064.491645s^3+1720.382638s^2 \right. \\ \left. +31.13077s+5.535239\right\}} \qquad (20)$$

Note that the compensated system is stable. Again, comparing with the existing standards of flying qualities, this control system was found to be quite satisfactory. McFIT error (payoff function) is 1.251.

Multivariable Design

In the frequency matching method for the SISO control systems, we need a "desired transfer function" meeting flying qualities specifications. For extending the technique to multivariable case, we would need a "desired transfer function matrix" satisfying the flying qualitites. In this section, we propose a model of a transfer function matrix with ϕ and β as the outputs and δ_a and δ_r as the inputs. This model transfer function matrix will include desirable

lateral flying qualities.

Lateral Flying Qualities

Lateral directional control of an aircraft is more difficult than longitudinal control. Primarily, this is due to the cross-coupling effects of the control surfaces of rudder and aileron.

A typical bare airframe transfer function matrix of an aircraft with ϕ, the roll angle and β, the sideslip angle as outputs and δ_a, the aileron deflection and δ_r, the rudder deflection as inputs is of the following form:

$$\begin{bmatrix} \phi \\ \beta \end{bmatrix} = \begin{bmatrix} \frac{A_\phi[\xi_\phi, \omega_\phi]}{\Delta} & \frac{A'_\phi\left[\frac{1}{T_{\phi_1}}\right]\left[\frac{1}{T_{\phi_2}}\right]}{\Delta} \\ \frac{A_\beta\left[\frac{1}{T_{\beta_1}}\right]\left[\frac{1}{T_{\beta_2}}\right]\left[\frac{1}{T_{\beta_3}}\right]}{\Delta} & \frac{A'_\beta\left[\frac{1}{T_{\beta_1}}\right]\left[\frac{1}{T_{\beta_2}}\right]\left[\frac{1}{T_{\beta_3}}\right]}{\Delta} \end{bmatrix} \begin{bmatrix} \delta_a \\ \delta_r \end{bmatrix} \quad (21)$$

where

$$\Delta = \left[s+\frac{1}{T_s}\right]\left[s+\frac{1}{T_R}\right]\left[s^2+2\xi_d\omega_d s+\omega_d^2\right] \quad (22)$$

$$[\xi_d, \omega_d] = s^2+2\xi_d\omega_d s+\omega_d^2$$

$$\left[\frac{1}{T}\right] = s+\frac{1}{T}$$

In Δ, $s + \frac{1}{T_s}$ is the spiral mode, $s + \frac{1}{T_R}$ is the roll mode and $s^2 + 2\xi_d\omega_d s + \omega_d^2$ is the dutch roll mode.

Desired Roll Control

(a) Desired $\frac{\phi}{\delta_a}$

From the flying qualities point of view, an ideal $\frac{\phi}{\delta_a} = \frac{K}{s\left[s+\frac{1}{T_R'}\right]}$ (23)

i.e., for good flying qualities, we would like $T_s \to 0$, dutch roll mode cancelling with the numerator quadratic polynomial. The constants K and T_R' are then determined using MIL-F-8785C depending on the aircraft and the task.

(b) Desired $\frac{\phi}{\delta_r}$

Desired $\frac{\phi}{\delta_r}$ depends on the pilot's task. It may be desired to get some roll from rudder deflection i.e., take advantage of the favorable cross coupling. For this, we must retain the dutch roll mode term in the denominator. We propose:

$$\text{Desired } \frac{\phi}{\delta_r} = \frac{K'}{s\left[s+\frac{1}{T_R'}\right]\left[s^2+2\xi_d'\omega_d's+\omega_d'^2\right]} \qquad (24)$$

where the value of T_R' is the same as in "desired" $\frac{\phi}{\delta_a}$ and ξ_d', ω_d' are the same as in "desired" $\frac{\beta}{\delta_r}$ to be discussed later. As a designer's choice, the constant K' is determined so that the steady state response of ϕ due to ramp rudder input is ten percent of the steady state response of $\phi(t)$ due to ramp aileron input, both in desired form.

(c) Desired $\frac{\beta}{\delta_r}$

A proposed form is

$$\frac{\beta}{\delta_r} = \frac{A'_\beta \cdot \frac{1}{T_{\beta_3}}}{s^2+2(\xi'_d)\omega'_d s+\omega'^2_d} \tag{25}$$

where the acceptable values of ξ_d' and ω_d' are chosen to meet the MIL-F-8785C specifications for the aircraft and the task.

(d) Desired $\frac{\beta}{\delta_a}$

It would not be desirable to get any sideslip angle response from aileron deflection. It is, therefore, reasonable to seek $\frac{\beta}{\delta_a} = 0$.

Thus, a typical desired transfer function matrix for an aircraft in the *lateral model will appear as*

$$\begin{bmatrix} \phi \\ \\ \beta \end{bmatrix} = \begin{bmatrix} \frac{K}{s\left[s+\frac{1}{T'_R}\right]} & \frac{K'}{s\left[s+\frac{1}{T'_R}\right]\left[s^2+2\xi'_d\omega'_d s+\omega'^2_d\right]} \\ \\ 0 & \frac{\frac{A'_\beta}{T_{\beta_3}}}{\left[s^2+2\xi'_d\omega'_d s+\omega'^2_d\right]} \end{bmatrix} \begin{bmatrix} \delta_a \\ \\ \delta_r \end{bmatrix} \tag{26}$$

or

$$\begin{bmatrix} \phi \\ \beta \end{bmatrix} = \begin{bmatrix} \dfrac{K(s^2+2\xi_d'\omega_d's+\omega_d'^2)}{s\left[s+\frac{1}{T_R'}\right]\left[s^2+2\xi_d'\omega_d's+\omega_d'^2\right]} & \dfrac{K'}{s\left[s+\frac{1}{T_R'}\right]\left[s^2+2\xi_d'\omega_d's+\omega_d'^2\right]} \\ 0 & \dfrac{A_\beta' s\left[s+\frac{1}{T_R'}\right]}{T_{\beta_3} s\left[s+\frac{1}{T_R'}\right]\left[s^2+2\xi_d'\omega_d's+\omega_d'^2\right]} \end{bmatrix} \begin{bmatrix} \delta_a \\ \delta_r \end{bmatrix} \quad (27)$$

where the constants are determined using the flying qualities specifications and, in some cases, designer's judgement. The above procedure is now illustrated by constructing a "desired transfer function matrix" for the A-7A.

Example of a "Desired Transfer Function Matrix"

We consider the example of the aircraft A-7A operating at M = 0.6 and h = 15,000 ft. Its bare airframe transfer function matrix in the lateral mode is given by

$$\begin{bmatrix} \phi \\ \beta \end{bmatrix} = \begin{bmatrix} \dfrac{17.7(s^2+0.80964s+5.4756}{\Delta} & \dfrac{6.89(s+4.35)(s-4.68)}{\Delta} \\ -\dfrac{0.0065(s+2.21)(s-1.63)(s+23.2)}{\Delta} & \dfrac{0.0537(s-0.00616)(s+2.7)(s+113)}{\Delta} \end{bmatrix} \begin{bmatrix} \delta_a \\ \delta_r \end{bmatrix} \quad (28)$$

where

$$\Delta = (s+0.0435)(s+2.71)(s^2+0.71448s+5.2441).$$

Clearly this system does not meet flying qualities specifications and hence, will have to be augmented. We construct a desired transfer function matrix for this system. We use the specifications as outlined in MIL-F-8785-C.

Desired Roll Control

(a) Desired $\frac{\phi}{\delta_a}$

A-7A is a class IV aircraft and from the specifications, we take

$$T_R' = 1 \tag{29}$$

$$\therefore \frac{\phi}{\delta_a} = \frac{K}{s(s+1)}$$

Pick K so that $\phi(t)\Big|_{t=1s} = 90^\circ$ if the input $\delta_a(t)$ is a unit step.

$\therefore$ for this case, we have

$$\phi(s) = \frac{K}{s^2(s+1)} = K\left[-\frac{1}{s} + \frac{1}{s^2} + \frac{1}{s+1}\right]$$

$$\therefore \phi(t) = K\left[-u(t)+tu(t)+e^{-t}u(t)\right]$$

$$\therefore \phi(1) = K\left[-1+1+e^{-1}\right]$$

Also $\phi(1) = 90^\circ$

$\therefore$ we have $e^{-1}K = \pi/2$ or $K = 4.2698671$

$$\therefore \frac{\phi}{\delta_a} = \frac{4.2698671}{s(s+1)} \tag{30}$$

(b) Desired $\frac{\phi}{\delta_r}$

From desired $\frac{\beta}{\delta_r}$ considerations, we pick $\xi_d' = 0.2$, $\omega_d' = 4.0$.

$$\therefore \text{ desired } \frac{\phi}{\delta_r} = \frac{K'}{s(s+1)(s^2+1.6s+16)} \qquad (31)$$

For $\delta_r = \frac{1}{s^2}$, $\quad \phi_{ss}^{(r)}(t) = \frac{K'}{16}$.

For $\delta_a = \frac{1}{s^2}$, $\quad \phi_{ss}^{(a)}(t) = 4.2698671$.

where $\phi_{ss}^{(r)}$, $\phi_{ss}^{(a)}$ are the steady state values of ϕ with δ_r alone and δ_a alone as inputs respectively.

Since by designer's choice,

$$\phi_{ss}^{(r)} = \left[\frac{1}{10}\right]\phi_{ss}^{(a)}, \text{ we get}$$

$$\frac{K'}{16} = \frac{4.2698671}{10} \quad \text{or } K = 6.8317874$$

$$\therefore \text{ desired } \frac{\phi}{\delta_r} = \frac{6.8317874}{s(s+1)(s^2+1.6s+16)} \qquad (32)$$

(c) Desired $\frac{\beta}{\delta_r}$

As explained in general and with $\xi_d' = 0.2$ and $\omega_d' = 4$ as selected earlier for desired $\frac{\phi}{\delta_r}$ we have

$$\frac{\beta}{\delta_r} = \frac{(0.0537)(113)}{s^2+2(0.2)4s+(4)^2} = \frac{6.0681}{s^2+1.6s+16} \qquad (33)$$

Thus, a desired transfer function matrix for A-7A at M = 0.6 and h = 15,000 ft. is given by

$$\begin{bmatrix} \phi \\ \beta \end{bmatrix} = \begin{bmatrix} \frac{4.26986716}{s(s+1)} & \frac{6.8317874}{s(s+1)(s^2+1.6s+16)} \\ 0 & \frac{6.0681}{s^2+1.6s+16} \end{bmatrix} \begin{bmatrix} \delta_a \\ \delta_r \end{bmatrix} \tag{34}$$

$$= \begin{bmatrix} \frac{4.26986716(s^2+1.6s+16)}{s(s+1)(s^2+1.6s+16)} & \frac{6.8317874}{s(s+1)(s^2+1.6s+16)} \\ 0 & \frac{6.0681s(s+1)}{s(s+1)(s^2+1.6s+16)} \end{bmatrix} \begin{bmatrix} \delta_a \\ \delta_r \end{bmatrix}$$

Summary

In this work, we have used the SISO frequency matching technique to design control systems for aircraft in the longitudinal mode. The design procedure includes, in a natural way, the flying qualities considerations. The existing software was updated for the design puposes. The advantage of the frequency-matching technique is that the unknown parameters of the controller turn out to be solutions of linear algebraic equations. This gives a quick capability to get a controller incorporating flying qualities which could be fine-tuned to meet the overall design-objectives. Two numerical examples, one for F104 and the other for F16, are given to illustrate the procedure.

In addition to this, a "desired transfer function matrix" for the control systems augumentation of an aircraft in its lateral modes has been proposed.

The parameters in this matrix are fixed in accordance with MIL-F-8785C. A numerical example is presented for the A-7A aircraft.

Recommendations for Future Work

The work carried out in this report for SISO systems should be extended to the multivariable case. The first significant step in this direction has already been taken by proposing the model of a "desired transfer function matrix" for the lateral mode augumentation of the control systems of an aircraft. We make the following recommendations:

Short Term Goal

(a) The SISO technique of designing control systems should be extended to the MIMO case such that the flying qualities are naturally incorporated in the process. Under the proposed technique, a general cascade controller as shown in Figure 1 will be synthesized such that the compensated closed loop multivariable system is "close" to the desired transfer function matrix over a frequency interval of interest.

(b) The MIMO technique developed in (a) should be applied to a numerical example of an aircraft in its lateral mode.

(c) Appropriate software should be developed to implement the synthesis procedure taking advantage of the fact that the parameters of the controller are solutions of linear algebraic equations.

Long Term Goal

(a) The question of robustness of the control system designed by the frequency matching method should be explored. Perhaps, good

robustness conditions could be incorporated in the "desired transfer function matrix". This condition will certainly strengthen the technique.

(b) A comparison of the frequency matching technique for MIMO control systems design with other design-techniques should be made. A comparison in the SISO case made by the author is quite favorable for the frequency matching technique.

REFERENCES

1. L.R. Pujara, "A Frequency-Domain Method of Designing a Controller," Proceedings of IEEE Naecon, 1985, pp. 647-650.

2. L.R. Pujara, "A Frequency-Domain Method of Designing a Controller II," Proceedings of IEEE Naecon, 1986, pp.

3. C.F. Chen and Shieh, "An Algebraic Method for Control System Design," International Journal of Control, 1970, Vol. 11, No. 5, pp. 717-739.

4. D.J. Moorhouse and R.J. Woodcock, "Background Information and User Guide for MIL-F-8785C, Military Specifications Flying Qualities of Piloted Airplanes," AFWAL, TR-3109, July, 1982.

5. Hiroyuki, Imai and Makoto Kobayakawa, "Disturbance Attenuation by a Frequency-Shaped Linear-Quadratic-Regulator Method," Journal of Guidance and Control, Vol. 9, No. 4, July-Aug 1986, pp. 397-402.

6. Norman Lehtomaki, G. Stein and Joseph Walls, Jr., "Multivariable Prefilter Design for Command Shaping," AIAA Conference, 1984, pp. 38-43.

7. N.K. Gupta, "Frequency-Shaped Cost Functionals: Extensions of Linear-Quadratic-Gaussian Design Methods, " Journal of Guidane and Control, Vol 3, No. 6, Nov-Dec 1980, pp. 529-535.

8. J.C. Doyle and G. Stein, "Multivariable Feedback Design: Concepts for Classical/Modern Synthesis," IEEE Trans. Automat. Contr., vol. AC-26, no.1, Feb. 1981, pp. 4-16.

9. A.G.J. McFarlane, and I. Postlethwaite,, "The Generalized Nyquist Stability Criterion and Multivariable Root-Loci," Int. J. Control, vol. 23, no. 1, Jan. 77, pp. 81-128.

10. Hsi-Han Yeh, Siva S. Banda and D. Brett Ridgeley, "Stability Robustness Measures Utilizing Structural Information," Int. J. Control, vol. 41, no. 2, Feb. 85, pp. 365-387.

1986 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT SUMMER SUPPORT PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

Infrared to Visible Light Conversion In Rare Earth Doped Heavy Metal Fluoride Glasses

Prepared by:	Richard S. Quimby
Academic Rank:	Assistant Professor
Department and University:	Physics Department, Worcester Polytechnic Institute
Research Location:	Hanscom Air Force Base, Rome Air Development Center, ESM
USAF Researcher:	Dr. Martin G. Drexhage
Date:	August 26, 1986
Contract No:	F49620-85-C-0013

Infrared To Visible Light Conversion In Rare Earth Doped Heavy Metal Fluoride Glasses

by

Richard S. Quimby

ABSTRACT

A phenomena known as "frequency upconversion" or "anti-stokes fluorescence via energy transfer" has been examined in a series of BaF_2-ZnF_2-LuF_3-ThF_4 glasses co-doped with YbF_3 and ErF_3. In this process, (low frequency) infrared photons in the 1 micrometer wavelength region are absorbed by Yb^{3+} ions, with a subsequent transfer of energy to Er^{3+} ions. Decay of the Er^{3+} excited state produces a (high frequency) green fluorescence in the vicinity of 0.55 micrometers. The effect has been extensively studied in (poly)crystalline phosphors, and is observed to be nonlinear in that two infrared photons are needed to generate one green photon. Using both a filtered arc lamp and a Nd-YAG laser, we measured the absolute efficiency of the upconversion process in glasses where the Yb^{3+}/Er^{3+} ratio was between 1:1 and 35:1. The efficiency is defined as the ratio of the green power emitted to the infrared power absorbed by the specimen. The highest efficiencies were obtained in glasses containing (mol%) 11.25 % YbF_3 and 0.75 % ErF_3. In these vitreous heavy metal fluorides, the upconversion efficiency is 3-4 orders of magnitude greater than that observed in equivalently doped oxide glasses, and compares very favorably with results obtained for Yb^{3+}/ $^{3+}$ containing crystals.

Acknowledgements

I would like thank the Air Force Systems Command and the Air Force Office of Scientific Research for sponsorship of my research. I am particularly indebted to Martin Drexhage for originating this project and for his constructive and friendly collaboration over the summer. Others at the Rome Air Development Center laboratory to whom I owe a word of thanks include Mike Suscavage for preparing the fluoride samples, Herb Lipson for supplying the oxide glass samples, and Bob Andrews for valuable technical assistance. Thanks also to Bill Sibley at Oklahoma State University, and Bill Miniscalco and Lenn Andrews at GTE Labs for helpful conversations. Mort Robinson and Slava Pollock supplied the CaF_2 sample which was very helpful in these experiments.

Finally, to the others on the staff at RADC with whom I had contact over the summer, thanks for putting up with the intruder whose opening line was apt to be: "can I borrow this?". All the little borrowed things helped to make this project a success.

I. Introduction

I obtained my Ph. D. in physics at the University of Wisconsin at Madison in 1979, with a thesis on the new technique of photo-acoustic spectroscopy. At Madison I became familiar with the optical spectroscopy of rare-earth and transition-metal ion impurities in solids. Recently I have become interested in the development and application of fiber optics, and it was through this interest that I initially made contact with Dr. Martin Drexhage at RADC - Hanscom AFB. One of the thrusts of Dr. Drexhage's research program at RADC is the utilization of newly developed fluoride glasses as hosts for optically active devices such as IR upconverters and lasers. Because such devices rely on impurity ion dopants for their operation, my background in the physics of rare-earth and transition metal ion spectroscopy nicely complemented Dr. Drexhage's expertise in fluoride glass development.

II. Objectives of the Research Effort

The goal of this project was to investigate from a fundamental point of view the efficiency of upconversion in rare-earth doped heavy metal fluoride glasses. From previous work by Dr. Drexhage and others, it was known that the characteristic vibrational energies in fluoride glasses are lower than in oxide glasses.[1] This results in weaker nonradiative relaxation between rare earth energy levels, and hence higher fluorescence efficiencies.[2] For the same reason, it was expected that upconversion processes in fluoride glasses would be more efficient than in oxide glasses, and perhaps as efficient as in cyrstalline hosts.

III. Theoretical Background

The type of upconversion process considered here is illustrated in Fig. 1. Two impurity ions such as Yb^{3+} and Er^{3+} are co-doped into the glass matrix. The sensitizer ion (Yb^{3+} in this case) absorbs incident IR radiation and transfers its energy

to a neighboring activator ion (Er^{3+}). This first step in the upconversion process leaves the Er^{3+} ion in a long lived intermediate state (labelled state 2), and the Yb^{3+} ion in its ground state. In the second step a second Yb^{3+} ion (or the first one, re-excited by incident IR light) transfers its energy to the previously excited Er^{3+} ion, raising it from level 2 to level 4. The Er^{3+} ion then decays rapidly through nonradiative relaxation to level 3, from which green fluorescence is emitted. The net result is that two IR photons have been converted into one visible photon. Variations of the above process are possible which involve other rare-earth energy levels, and these different schemes have been reviewed by Auzel[3] and Wright[4]. In fluoride hosts it has been found that the green emission process described above is predominant in Yb, Er doped systems.

The upconversion process described above involves a step-wise excitation of discreet energy levels involving ion-ion energy transfer. This should be contrasted with other types of upconversion wuch as frequency doubling, which involve the nonlinear electronic susceptibility of a material. Advantages of the latter include speed and wavelength tunability, whereas the former has the advantage of sensitivity and wide acceptance angle for the incident light.

With a straightforward rate equation calculation one can show that for the upconversion scheme shown in Fig. 1, the emitted green power per unit volume is given by [4]

$$\frac{P_{emitted}}{Vol} = K N_a N_d^2 (\sigma I)^2 \qquad (1)$$

where N_a is the number density of acceptor (Er^{3+}) ions, N_d is the number density of donor (Yb^{3+}) ions, σ is the absorption cross section of the donor ion, and I is the IR light intensity. The parameter K depends on microscopic quantities such as transfer

rates, radiative and nonradiative lifetimes, and energy level splittings. It can also depend on the ionic concentrations N_a and N_d through the processes of concentration quenching and back transfer.

In comparing the upconversion efficiency of different materials, it is useful to have a definition of efficiency which is independent of the level of excitation of the material. This is partially accomplished by dividing the emitted power by the absorbed power:

$$\frac{P_{abs}}{V_{ol}} = N_a \sigma I \qquad (2)$$

$$\eta \equiv \frac{P_{emitted}}{P_{abs}} = K N_a N_d \sigma I \qquad (3)$$

The efficiency defined in this way still depends on the excitation level σI, however, so it is useful to divide η by σI and normalize to a "standardized" value $(\sigma I)_{st}$. The "standardized efficiency" η_{st} is related to the measured efficiency η by

$$\eta_{st} = \frac{(\sigma I)_{st}}{\sigma I} \eta \qquad (4)$$

When the values of η_{st} are compared for different materials, what is being compared is the product $K N_a N_d$, which is independent of the excitation level. For convenience in these measurements we chose a value of $(\sigma I)_{st} = 1.00 \times 10^{-20}$ W. It is important to recognize that the product σI defines the excitation level, rather than the optical intensity I alone. This is because the upconversion efficiency depends on the number density of excited sensitizer ions, which is proportional to σI.

IV. Experimental Procedure

The preparation of heavy-metal fluoride glasses has been discussed elsewhere.[1] For these experiments, a base glass with

the following composition was used: BaF_2 (14 mol %), ZnF_2 (27 %), LuF_3 (27%), ThF_4 (27%), and NaF (5%). The dopants YbF_3 and ErF_3 were substituted for LuF_3 in a series of Yb concentrations, such that the sum of Lu, Yb, and Er concentrations always added up to 27%. The samples were cut into rectangular slabs approximately 1 X 1 X 0.2 cm, and polished on all sides. In addition to these fluoride glasses, samples of Er, Yb doped into a silicate glass, a phosphate glass, and CaF_2 were obtained. The oxide glass samples were rectangular pieces approximately 5 X 5 X 9 mm, and the CaF_2 sample was a rod 3.4 cm long and 0.95 cm in diameter.

Two different light sources were used to excite the samples. The first was a cw Nd:YAG laser which supplied a TEM_{oo} beam of unpolarized light at a wavelength of 1.064 microns. After passing through a 20 cm focal length lense the cw beam power was 240 mW, and was focused on the sample to a spot with a Gaussian beam radius of 90 μm. This resulted in a peak intensity on the sample of 940 W/cm^2. The second source was an arc lamp and filter, which supplied broadband radiation in the range 950 nm to 1010 nm. After spatial filtering and refocusing, the beam power was 2 mW and was focused onto the sample in an area of 0.029 cm^2, resulting in a peak intensity of 0.070 W/cm^2.

The samples were placed at the center of a 10 inch diameter integrating sphere and were illuminated with light incident through one of the ports. Light which passed through the sample without being absorbed or scattered was allowed to exit the sphere through the opposite port. The photodetector was placed in a port at 90^o from the incident beam, and shielded from direct radiation by a baffle. The photodetector was either a Silicon photodiode or an S-20 photomultiplier, depending on the signal level. Scattered IR light was prevented from reaching the photodetector by the use of IR blocking filters. The detected signal

then corresponded only to the emitted green light.

Calibration of the signal was accomplished in one of two ways. The first method was to compare the measured green signal with the photodiode signal abtained by blocking the exit port for the IR light and removing the IR blocking filters in front of the photodiode. Since all of the incident IR light is now trapped in the sphere and detected, and since the absolute power of the incident beam can be easily measured with a calorimetric power meter (a Scientech meter was used), the previously detected green signal can then be calibrated. In this method the relative sensitivity of the photodiode to green vs. infrared light must be known. This was measured independently, using the Scientech meter as a standard.

The second method for calibrating the signal was to direct green light of a known power into the sphere. The light was from the arc lamp and a green 10 nm filter centered at 546 nm, and its absolute power was again measured with the Scientech meter. The two calibration techniques yielded the same efficiency values, within about 20%.

The signal from the photodetector was sent to a lock-in detector for improvement of signal-to-noise ratio. The beam was chopped at a low frequency (10 Hz) in order to achieve a quasi-steady state condition. The signal was independent of chopping frequency for frequencies below 20 Hz.

The power absorbed by each sample was determined by measuring the Yb^{3+} absorption spectrum on a Cary 14 spectrophotometer. A representative scan for a low Yb concentration sample is shown in Fig. 2. Although the absorption at $\lambda = 1.064\ \mu m$ is rather small, it was sufficient in the higher concentration samples to allow a determination of the absorption cross section to an accuracy of about 30%. In the case of arc-lamp and filter excitation, the absorption was high enough that it could be

determined directly by measuring the ratio of transmitted to incident light.

V. Results

According to eq.(1), the green emission should vary as the square of the incident intensity. This was verified experimentally by inserting a variable beam attenuator in the path of the beam. The graph of output power vs. input power on log-log paper is a straight line with a slope of 2.0 ± 0.1, as expected.

The efficiency of upconversion for YAG excitation was measured for a series of fluoride glass samples with Yb^{3+} concentration ranging from 0.75 mol % to 26.25 mol %. For each of these samples the Er^{3+} concentration was 0.75%. The standardized efficiency was calculated for each sample according to eq.(4), with the results shown in Fig. 3 by the dots and the crosses. The A and B samples were cut out from different sections of a larger original sample. It is clear that an optimum concentration range exists for the Yb^{3+}, between 5 and 11 mol %. The flattening of the efficiency curve above 3% Yb is likely due to increasing back-transfer from Er^{3+} to Yb^{3+},[4] while the decrease at still higher Yb^{3+} concentrations is due to increasing Yb-Yb energy transfer with subsequent energy loss to traps.

The effect of varying Er^{3+} concentration was investigated by doubling the Er^{3+} concentration to 1.5 mol %, with the Yb^{3+} concentration at 11.25%. Although eq.(3) predicts the efficiency to be proportional to Er concentration, the square data point in Fig. 3 shows that the efficiency with 1.5% Er is only some 15% higher than the efficiency with 0.75% Er. This is probably due to fluorescence quenching by Er-Er interactions.[2]

As a check to see if nonrepresentative sites with unusually small or large efficiencies were being excited with the YAG pump, the standardized efficiency was measured on the same series of samples using the arc-lamp and filter as the light source. The

results are shown with circles in Fig. 3. It is seen that the standardized efficiencies using these two methods of excitation agree within a factor of two, and furthermore they exhibit the same dependence on Yb concentration. It appears, then, that the ions excited by the YAG laser are representative of the ions excited with broadband illumination, and do not have unusually high or low conversion efficiencies.

The standardized upconversion efficiency of one of the fluoride glass samples (11.25 mol % Yb, 0.75 % Er) is compared to that of the oxide glass and CaF_2 samples in Table I, along with some relevant parameters. The fluoride glass host is within a factor of two as efficient as the CaF_2 crystalline host, whereas the two oxide hosts are some four orders of magnitude lower in efficiency. This confirms our original expectation that the fluoride glasses would make efficient upconverters.

VI. Recommendations

From the results presented here, it is clear that fluoride glasses do indeed make efficient hosts for IR upconversion, at least for the Er, Yb system. The limits on upconversion efficiency appear to be due to back transfer and energy transfer to traps, rather than nonradiative decay by multiphonon emission. One important limitation of the Er, Yb system is wavelength tunability. As can be seen from Fig. 2, the Yb absorbs strongly only between 900 and 1000 nm. It is desirable to have an upconverter which is sensitive to a broader range of wavelengths, particularly at 1.064 μm and 1.5 μm. A promising approach would be to use transition metal ions as the sensitizer, since they have much broader absorptions in the near infrared than do rare-earth ions. Furthermore, the transfer of energy from transition metal ion to rare-earth ion is expected to be very efficient, based on previous work using such transfer to increase pump efficiency for solid state lasers.[5]

It is recommended that this project be extended to study upconversion in transition metal-rare earth doped systems. Specific systems that merit consideration are transition metal ions such as V^{2+}, Cr^{2+}, Ni^{2+}, and Cu^{2+} as the sensitizer, co-doped with Er^{3+} as the activator. The encouraging results obtained for the Er, Yb system lead us to be optimistic about the development of an upconvertor that is both highly efficient and widely tunable.

References

1. Martin G. Drexhage, "Heavy-Metal Fluoride Glasses", in Glass IV, Treatise on Material Science and Technology, Vol. 26, ed. by M. Tomozawa and R. H. Doremus, Academic Press, 1985

2. M. D. Shinn, W. A. Sibley, M. G. Drexhage, and R. N. Brown, "Optical Transitions of Er^{3+} ions in Fluorozirconate Glass", Phys. Rev. B 27, 6635 (1983)

3. F. E. Auzel, "Materials and Devices Using Double-Pumped Phosphers with Energy Transfer", Proc. IEEE, 61, 758 (1973)

4. J. C. Wright, "Up-conversion and Excited State Energy Transfer in Rare Earth Doped Materials", in Topics in Applied Physics, Vol. 15, Springer Science, New York, 1976.

5. D. Pruss, G. Huber, A. Beimowski, "Efficient Cr^{3+} sensitized Nd^{3+} GdScGa-Garnet Laser at 1.06 microns", Appl. Phys. B 28, 355 (1982)

Table I. Standardized efficiency of upconversion of IR light at $\lambda = 1.064\ \mu m$ into green light at 550 nm. $\sigma_{1.064}$ is the Yb^{3+} absorption cross section at $\lambda = 1.064\ \mu m$.

Host	Yb conc. (10^{20} cm^{-3})	Er conc. (10^{20} cm^{-3})	l (mm)	$\sigma_{1.064}$ ($10^{-23} cm^2$)	η_{ss}
fluoride glass	21.4	1.43	2.05	1.0	1.0×10^{-3}
silicate glass	9.3	0.31	5.1	2.1	2×10^{-7}
phosphate glass	12.1	0.40	5.1	2.6	6×10^{-8}
CaF_2 crystal	25	2.5	3.44	1.1	2.1×10^{-3}

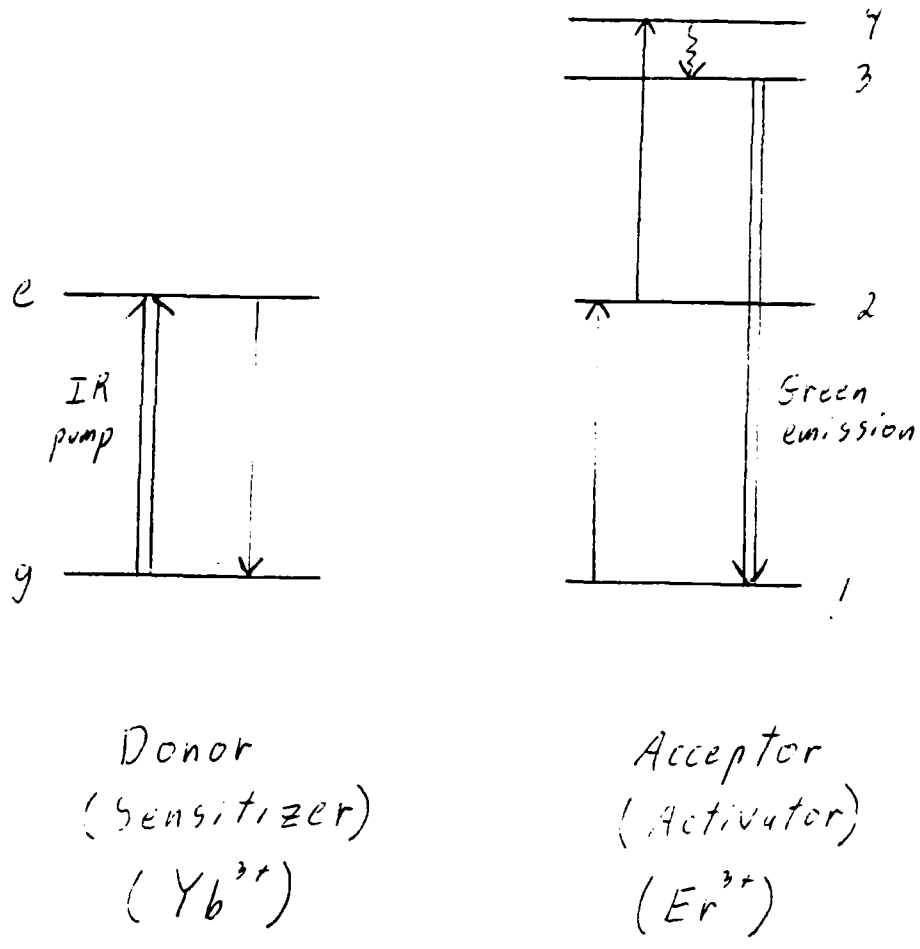


Fig. 1. Energy levels invloved in the upconversion process.

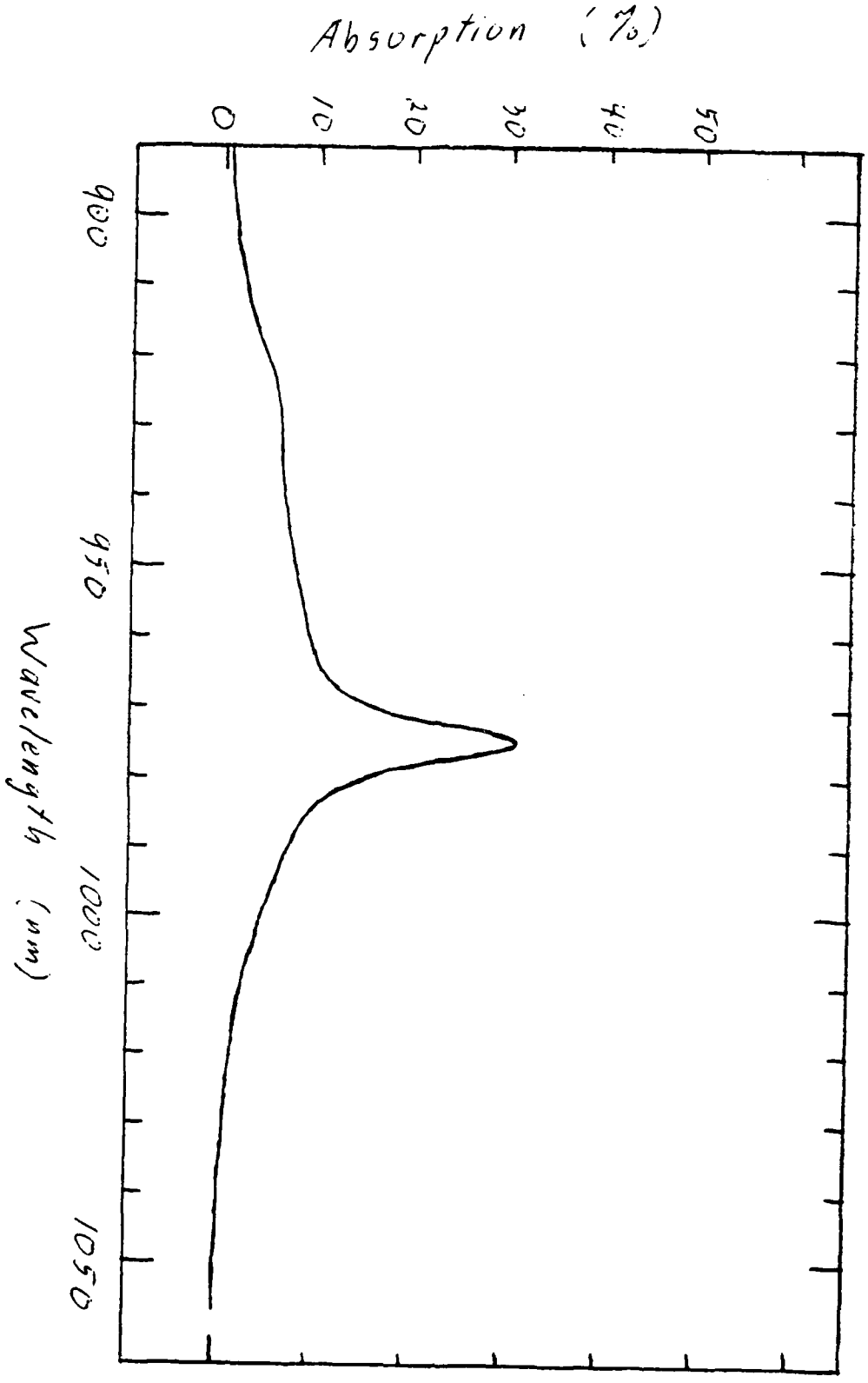


Fig. 2 Yb^{3+} absorption vs. wavelength in the fluoride glass. Sample length is 2.13 mm and the Yb concentration is 0.75 mole% (1.43×10^{20} cm^{-3}).

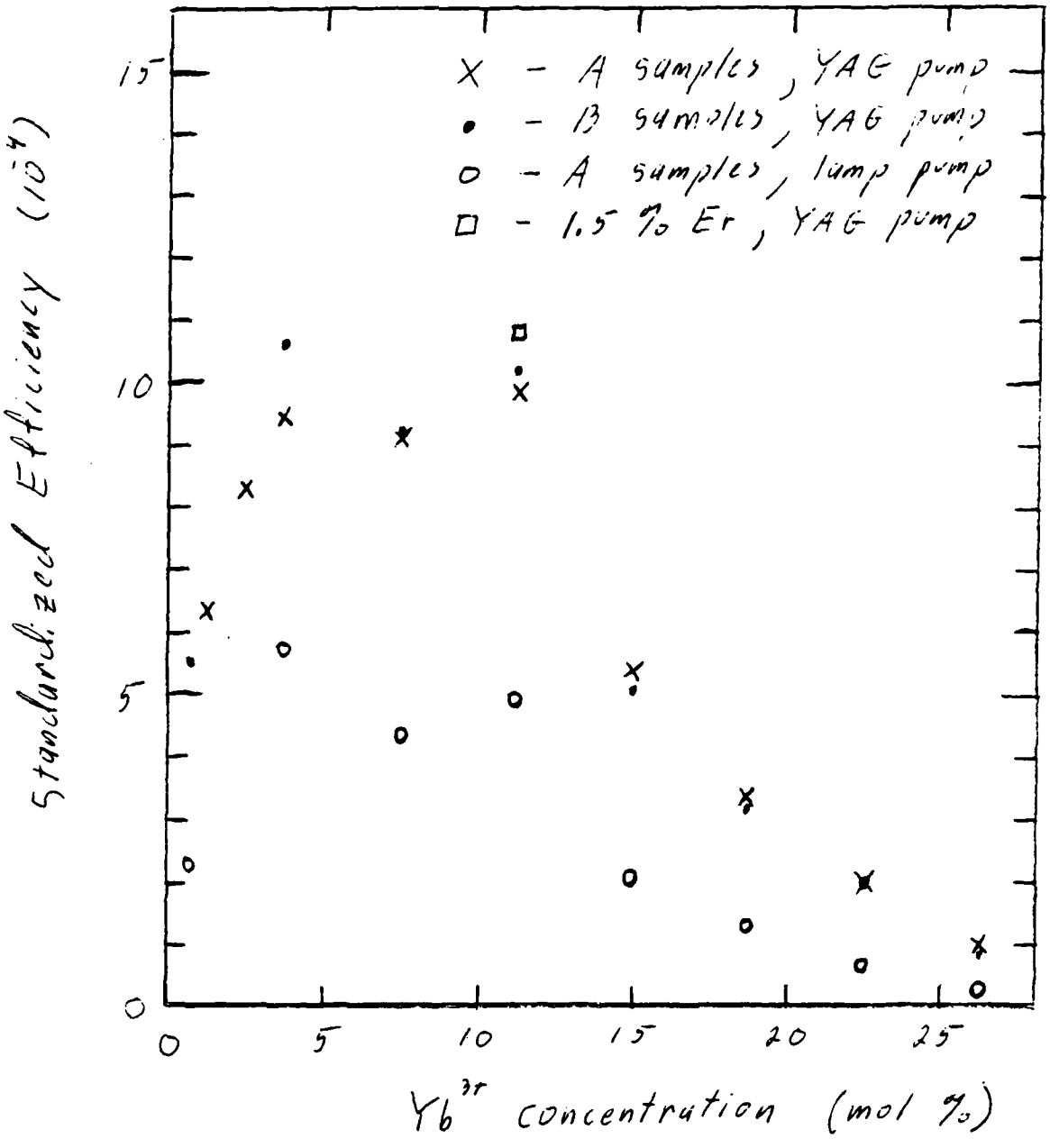


Fig. 3. Efficiency of upconversion, defined as visible power emitted divided by IR power absorbed. The efficiency values given correspond to pumping at line center with an incident intensity of 1 W/cm^2.

1986 USAF-UES Summer Faculty Research Program/

Graduate Student Summer Support Program

Sponsored by the
Air Force Office of Scientific Research
Conducted by the
Universal Energy Systems, Inc.

Final Report

Optimization of Actively Controlled Structures Using Multiobjective Programming Techniques

Prepared by:	Singiresu S. Rao
Academic Rank:	Associate Professor
Department and	School of Mechanical Engineering
University:	Purdue University
Research Location:	Flight Dynamics Laboratory
	Analysis and Optimization Branch
	Wright Patterson Air Force Base, Ohio
USAF Researcher:	Dr. Vipperla B. Venkayya
Date:	August 31, 1986
Contract No.:	F49620-85-C-0013

Optimization of Actively Controlled Structures Using Multiobjective Programming Techniques

by

Singiresu S. Rao

ABSTRACT

The design of minimum weight structures with constraints on the damping parameters of the closed loop system in the design of an active control system is considered using multiobjective optimization techniques. The cross sectional areas of the members are treated as design variables. The structural weight and the controlled system energy are considered as objective functions for minimization. The goal programming and game theory approaches are used for the solution of the multiobjective optimization problems. The feasibility of the approaches is demonstrated through the design of a two-bar and a twelve-bar truss structures.

Acknowledgements

I would like to thank the Air Force Systems Command and the Air Force Office of Scientific Research for sponsorship of my research. In order to be meaningful, scientific research must be conducted in an intellectually stimulating environment. The Analysis and Optimization Branch of the Flight Dynamics Laboratory (FDL) provided this environment. I would like to thank several members of the FDL staff, Dr. V. B. Venkayya, Dr. N. S. Khot and V. A. Tischler, for giving me the opportunity and the guidance necessary for my research.

Finally, I would like to thank my wife, Kamala, and my daughters, Sri and Shobha, for their willingness to relocate and for their support and encouragement throughout the summer.

I. Introduction

I received my Ph.D. from Case Western Reserve University, Cleveland, studying optimal design of aircraft structures under dynamic and aeroelastic constraints. I later was a faculty member at the Indian Institute of Technology, Kanpur, and San Diego State University, San Diego. At present, I am a faculty member in the School of Mechanical Engineering, Purdue University, West Lafayette. I have been doing and guiding research in the areas of structural and mechanical optimum design.

The research problem at the Flight Dynamics Laboratory involved the development of suitable optimization methodologies for the design of actively controlled structures.

The problem under investigation at FDL was similar to the problems I have been studying over the past several years. Because of this similarity I was assigned to work on the optimization of actively controlled structures.

II. Objectives of the Research Effort

The objective of the vibration control project is to design the structure and its control system of a large space satellite to either eliminate structural vibration or reduce it to a desired level within a reasonable time span. An important aspect is to integrate a large order structural optimization with the algorithms for a closed-loop control system so that both can be used to effectively control the large order system.

My individual objectives were:

1. Formulate the problem of optimization of actively controlled structures by identifying all the important objectives.
2. Develop multiobjective programming techniques for the solution of the design problem.

III. Past Work

Large space structures face difficult problems of vibration control. Because of the requirement of low weight, such structures will lack the stiffness and damping necessary for the passive control of vibrations. Hence, current research is directed toward the design of active vibration control systems for such structures. The objective of vibration control is to optimally design the structure and its controls either to eliminate vibration completely or to reduce the mean square response of the system to a desired level within a reasonable span of time.

The structural designer wants to adjust the structural parameters to minimize the mass while achieving the desired frequencies, mode shapes and dynamic response. The control designer, on the other hand, wishes to size the controllers (and possibly select the number and locations of the actuators and sensors) so as to minimize the energy of the vibrating structure.

A great deal of research is currently in progress on designing active vibration control systems for large flexible space structures [1]. In Refs. [2,3], an optimal structure is initially designed to satisfy constraints on weight, strength, displacements and frequency distribution, and then an optimum control system is designed to improve the dynamic response of the structure. In Refs. [4,5], a simultaneous integrated design of the structure and vibration control system is achieved by improving the configuration as well as the control system. A unified approach to achieve satisfaction of eigenspace constraints is presented in

Ref. [6].

The weight of the structure is minimized with constraints on the distribution of the eigenvalues and/or damping parameters of the closed-loop system by Khot, Eastep and Venkayya [7]. Salama [5] and Miller and Shim [8] considered the simultaneous minimization, in structural and control variables, of the sum of structural weight and the infinite horizon linear regulator quadratic control cost. The frequency control, the effect on the dynamic response of flexible structures and the associated computational issues were discussed by Venkayya and his associates [3,9]. The structure/control system optimization problem was formulated by Khot et al. [10] with constraints on the closed-loop eigenvalue distribution and the minimum Frobenions norm of the control gains.

It can be seen that the combined structural/control optimization problem has not been formulated and solved as a multiobjective optimization problem. The present work aims at using goal programming and game theory approaches for the formulation and numerical solution of the problem.

IV. Equations of Motion

The equations of motion of a large space structure with active controls under external forces are given by

$$[M]\ddot{\vec{U}} + [C]\dot{\vec{U}} + [K]\vec{U} = [D]\vec{F} \qquad (1)$$

where [M] is the mass matrix, [C] is the damping matrix and [K] is the stiffness

matrix. These matrices are of order r where r denotes the number of degrees of freedom of the structure. $\vec{U}$ represents the vector of displacements and a dot over a symbol denotes differentiation with respect to time. [D] is the r×p matrix denoting the applied load distribution that relates the control input vector $\vec{F}$ to the coordinate system. The number of components in $\vec{F}$ is assumed to be p. By introducing the coordinate transformation

$$\vec{U} = [\phi]\vec{\eta} \tag{2}$$

where $[\phi]$ is the r×r modal matrix whose columns are the eigenvectors and $\vec{\eta}$ is the vector of modal coordinates, Eq. (1) can be transformed into a system of uncoupled differential equations as

$$[\bar{M}]\ddot{\vec{\eta}} + [\bar{C}]\dot{\vec{\eta}} + [\bar{K}]\vec{\eta} = [\phi]^T[D]\vec{F} \tag{3}$$

where

$$[\bar{M}] = [\ I\ ] \tag{4}$$

$$[\bar{C}] = [\ 2\zeta\omega\ ] \tag{5}$$

$$[\bar{K}] = [\ \omega^2\ ] \tag{6}$$

$\vec{\zeta}$ is the vector of modal damping factors and $\vec{\omega}$ is the vector of natural frequencies of the structure. Equations (3) can be converted into a state space representation by using the transformation

$$\vec{x} = \begin{Bmatrix} \vec{\eta} \\ \dot{\vec{\eta}} \end{Bmatrix} \tag{7}$$

where $\vec{x}$ is the n×1 state variable vector. This gives the state input equation

$$\dot{\vec{x}} = [A]\,\vec{x} + [B]\,\vec{F} \tag{8}$$

where [A] is the n×n plant matrix and [B] is the n×p input matrix, n = 2r

$$[A] = \begin{bmatrix} [\,O\,] & [\,I\,] \\ [\,-\omega^2\,] & [\,-2\eta\omega\,] \end{bmatrix} \tag{9}$$

and

$$[B] = \begin{bmatrix} [O] \\ [\phi]^T[D] \end{bmatrix} \tag{10}$$

The state output equation is given by

$$\vec{y} = [\underset{\sim}{C}]\,\vec{x} \tag{11}$$

where y is the q×1 output vector and $[\underset{\sim}{C}]$ is the q×n output matrix. If the sensors and actuators are co-located, then q = p and

$$[\underset{\sim}{C}] = [B]^T \tag{12}$$

In order to design a controller using a linear quadratic regulator, a performance index J is defined as

$$J = \int_0^t \left(\vec{x}^T\,[Q]\,\vec{x} + \vec{F}^T\,[R]\,\vec{F} \right) dt \tag{13}$$

where [Q] is the state weighting matrix which has to be positive semidefinite and [R] is the control weighting matrix which has to be positive definite. It is possible to control the damping response time, amplitudes of vibration, etc., of

the system by proper selection of the elements of the matrices [Q] and [R]. If [Q] and [R] are chosen as

$$[Q] = \begin{bmatrix} [K] & [O] \\ [O] & [M] \end{bmatrix} \tag{14}$$

and

$$[R] = [D]^T [K]^{-1} [D] \tag{15}$$

then Eq. (13) provides a measure of total system strain, kinetic and potential energies.

The result of minimizing the quadratic performance index and satisfying the state equation gives the state feedback control law [13]

$$\vec{F} = - [G] \vec{x} \tag{16}$$

where [G] is the optimum gain matrix given by

$$[G] = [R]^{-1} [B]^T [P] \tag{17}$$

with [P] representing a symmetric positive definite matrix called the Riccati matrix and is found by solving the following algebraic Riccati equation

$$[A]^T [P] + [P] [A] + [Q] - [P] [B] [R]^{-1} [B]^T [P] = [O] \tag{18}$$

Substituting Eq. (16) into Eq. (8) gives the governing equations for the optimal closed-loop system in the form

$$\dot{\vec{x}} = [A_{cl}] \vec{x} \tag{19}$$

where

$$[A_{cl}] = [A] - [B]\,[G] \tag{20}$$

The eigenvalues of the closed-loop matrix, $[A_{cl}]$, are a set of complex conjugate pairs written as

$$\lambda_i = \bar{\sigma}_i \pm j\,\bar{\omega}_i\,; \qquad i = 1,2,\ldots,n \tag{21}$$

where $j = \sqrt{-1}$. The damping factors ξ_i and the damped frequencies $\bar{\omega}_i$ are related to the complex eigenvalues through

$$\xi_i = -\frac{\bar{\sigma}_i}{\sqrt{\bar{\sigma}_i^2 + \bar{\omega}_i^2}} \tag{22}$$

V. Optimization Problem

The cross-sectional areas of the members of the structure are chosen as the design variables. The weight of the structure, f_1, is given by

$$f_1(\vec{z}) = \sum \rho_i\, A_i\, \ell_i \tag{23}$$

where ρ_i is the weight density, A_i is the cross sectional area, ℓ_i is the length of element i, and $\vec{z}$ is the design vector. For any specified initial condition $\vec{x}_o$, it is well known that

$$\min_{\vec{F}} \frac{1}{2} \int_0^\infty \left(\vec{x}^T\,[Q]\,\vec{x} + \vec{F}^T\,[R]\,\vec{F} \right) dt = \frac{1}{2}\,\vec{x}_o^T\,[P]\,\vec{x}_o \tag{24}$$

and the corresponding optimal control is given by Eq. (16). The minimum value of the quadratic performance index is taken as the second objective function,

$f_2(\vec{z})$, as

$$f_2(\vec{z}) = \vec{x}_o^T \, [P] \, \vec{x}_o \tag{25}$$

Other possible objective functions are the Frobenions norm and the effective damping response time. The Frobenions norm of the control gains, $f_3(\vec{z})$, is given by

$$f_3(\vec{z}) = \text{trace} \left\{ [G]^T \, [R] \, [G] \right\} \tag{26}$$

Essentially, the Frobenions norm of the control gains represents the expected value of the integrand of the quadratic control effort. By minimizing the expected value of the integrand, one can hope to minimize the required control effort. The effective damping response time, $f_4(\vec{z})$, under the action of an initial disturbance $\vec{x}_o = \vec{x}\,(t = 0)$ can be expressed as

$$f_4(\vec{z}) = \frac{\vec{x}_o^T \, [P] \, \vec{x}_o}{\vec{x}_o^T \, [Q] \, \vec{x}_o} \tag{27}$$

The magnitude of f_4 indicates the effect of the control system in reducing vibrations.

Constraints are placed on the closed loop damping ratios as

$$\xi_i - \xi_i^{(o)} = 0 \tag{28}$$

or

$$\xi_i - \xi_i^{(o)} \geq 0 \tag{29}$$

where $\xi_i^{(o)}$ is a specified value. Bounds are placed on the cross-sectional areas of members as

$$A_i^{(\ell)} \leq A_i \leq A_i^{(u)} \tag{30}$$

where the superscripts ℓ and u indicate the lower and upper bound values.

The design problem was formulated as a standard multiobjective optimization problem. The goal programming and game theory techniques were used for solving the problem. These techniques transform the problem into an equivalent single objective optimization problem. The subroutine UNCON, which is based on Powell's algorithm for nonlinear constraints that uses Lagrangian functions [12], has been used for solving the single objective optimization problem.

VI. Illustrative Example

The multiobjective optimization techniques outlined above were applied to two truss structures. The dimensions of the structure were defined in unspecified consistent units. The elastic modulus of the members was assumed to be 1.0 and the density of the structural material was assumed to be 0.001.

Example 1. Two-Bar Truss

The two bar truss shown in Fig. 1 was selected for its simplicity. A nonstructural mass of 2 units was attached at node 2. The actuator and the sensor were located in element 1 connecting nodes 1 and 2. It was observed

that the minimization of f_1 is the same as that of f_3. Similarly, the optimum solution of f_2 was found to be the same as that of f_4. The solutions of the problem obtained with consideration of single and multiple objective functions are shown in Table 1. The minimization of F_1 gave a value of $F_1^* = 49.94$ with the corresponding $F_2 = 127.41$ while the minimization of F_2 yielded $F_2^* = 1.21$ with the corresponding $F_1 = 1110.8$. These results indicate the penalty associated with the other objective functions while minimizing a particular objective function. It can be observed that the results are essentially the same with $p = 1$ and $p = 2$ in goal programming. The damping ratios were observed to be the same in all the cases. Considering all the four objective functions, the game theory approach seemed to yield a good compromise solution.

Example 2. 12-bar Truss (ACOSS-FOUR)

The finite element model of ACOSS-FOUR is shown in Fig. 2. The edges of this tetrahedral truss are 10 units long. The structure has 12 degrees of freedom and 4 masses of 2 units each, which are attached at nodes 1 through 4. The actuators and the sensors are located in six bipods and are assumed to coincide with each other. Thus, the matrix [D] in Eq. (1) would consist of the direction cosines relating the forces in the six bipods with their components in the coordinate directions. Since the sensors and the actuators are co-located, matrices [B] and [C] in Eqs. (8) and (11) would satisfy Eq. (12). The weighting matrix [R] for this case would be of order 6×6. The passive damping parameters $\vec{\zeta}$ in Eq. (9) were assumed to be zero.

At the starting design, the cross-sectional areas of the members were taken to be the same as those assigned by the Charles Stark Draper Lab for their investigation [2]. The characteristics of this design, along with those of the designs found in this investigation, are shown in Table 2. For this structure also, it was observed that the minimization of $F_1(F_2)$ reduces $F_3(F_4)$. The game theory solution appeared to yield the best compromise solution.

VII. Recommendations

1. The multiobjective optimization procedures were applied to simple truss structures. Further work is needed to investigate the behavior of more complex structures.

2. It was observed that the minimization of structural weight (control energy) yielded the same results as the minimization of the Frobenions norm of the control gains (effective damping response time). Additional work is needed to validate this observation.

3. The present study seemed to point out that the game theory approach yields good compromise solutions. Further work is needed to find the nature of solutions given by the game theory.

4. Further multiobjective optimization work is needed to find optimal designs which will be robust from stability and performance points of view.

References

1. Pinson, L.D., A.K. Amos and V.B. Venkayya (Eds.), "Modeling, Analysis and Optimization Issues for Large Space Structures," *Proceedings of the NASA-AFOSR Workshop,* Williamsburg, VA, May 13-14, 1982.

2. Khot, N.S., V.B. Venkayya and F.E. Eastep, "Structural Modifications to Reduce the LOS-Error in Large Space Structures," (84-0997-CP) AIAA/ASME/ASCE/AHS 25th Structures, Structural Dynamics and Materials Conference, Palm Springs, CA, May 14-16, 1984.

3. Venkayya, V.B. and V.A. Tischler, "Frequency Control and the Effect on the Dynamic Response of Flexible Structures," (84-1044-CP) AIAA Dynamics Specialists Conference, Palm Springs, CA, May 17-18, 1984.

4. Hale, A.L. and R.J. Lisowski, "Optimal Simultaneous Structural and Control Design of Maneuvering Flexible Spacecraft," *Proceeding of the Fourth VPI & SU/AIAA Symposium,* pp. 1-16, Virginia Polytechnic Institute, Blacksburg, VA, June 1983.

5. Salama, M., M. Hamidi and L. Demsetz, "Optimization of Controlled Structures," presented at the Jet Propulsion Workshop on Identification and Control of Flexible Space Structures, San Diego, CA, June 4-6, 1984.

6. Junkins, J.L., D.S. Bodden and J.D. Turner, "A Unified Approach to Structure and Control System Design Iterations," Presented at the Fourth International Conference on Applied Numerical Modelling, Tainan,

Taiwan, Dec. 27-29, 1984.

7. Khot, N.S., F.E. Eastep and V.B. Venkayya, "Optimal Structural Modifications to Enhance the Optimal Active Vibration Control of Large Flexible Structures," AIAA/ASME/ASCE/AHS 26th Structures, Structural Dynamics and Materials Conference, Orlando, FL, Paper No. 85-0627, April 15-17, 1985.

8. Miller, D.F. and J. Shim, "Combined Structural and Control Optimization for Flexible Systems Using Gradient Based Searches," 24th AIAA Aerospace Sciences Meeting, Reno, NV, January 6-8, 1986.

9. Miller, D.F., V.B. Venkayya and V.A. Tischler, "Integration of Structures and Controls - Some Computational Issues," *Proceedings of the 24th* Conference on Decision and Control, Ft. Lauderdale, FL, Dec. 1985.

10. Khot, N.S., H. Oz, F.E. Eastep and V.B. Venkayya, "Optimal Structural Designs to Modify the Vibration control Gain Norm of Flexible Structures," AIAA/ASME/ASCE/AHS 27th Structures, Structural Dynamics and Materials Conference, Paper No. 86-0840-CP, 1986.

11. Rao, S.S., "Game Theory Approach for Multiobjective Structural Optimization," *Computers and Structures,* 1986 (in press).

12. Crane, R.L., "Solution of the General Nonlinear Programming Problem with Subroutine VMCON," ANL-80-64, Argonne National Laboratory, Argonne, IL.

Table 1. Optimization Results of Two-Bar Truss

Quantity	Starting points, $\vec{z}_o$	Minimization of F_1 or F_3	Minimization of F_2 or F_4	Goal programming		Game Theory
				p=1	p=2	
Design variables†						
z_1	100.0	89.8788	2000.0	169.3107	179.0077	361.8625
z_2	100.0	10.0	221.6112	18.7606	19.8351	40.0963
Eigenvalues (open loop)						
ω_1^2	0.946	0.370	1.74	0.507	0.522	0.742
ω_2^2	1.89	1.45	6.83	1.99	2.04	2.90
Eigenvalues (closed loop)						
λ_1	-0.331±0.996j	-0.0373±0.371j	-0.175±1.75j	-0.0511±0.508j	-0.0525±0.522j	-0.0746±0.743j
λ_2	-0.904±1.81j	-0.958±1.42j	-4.52±6.69j	-1.32±1.95j	-1.35±2.00j	-1.92±2.84j
Damping ratios††						
ξ_1	0.3153	0.1002	0.1000	0.1000	0.1000	0.1000
ξ_2	0.4462	0.5599	0.5599	0.5599	0.5599	0.5599
Objectives†††						
F_1	100.0	49.9394	1110.8	94.0356	99.4214	200.9794
F_2	100.0	127.4131	1.2139	49.2823	45.3325	15.7726
F_3	100.0	73.0240	255.37	80.6034	81.5291	98.9855
F_4	100.0	114.5173	24.277	83.4402	81.1487	57.0750

† $z_i^{(\ell)} = 10.0, \quad z_i^{(u)} = 2000.0; \quad i = 1,2$

†† $\xi_i^{(o)} = 0.1$

††† $F_i = f_i c_i, \quad c_1 = 22.3607, \quad c_2 = 2.148\times10^{-4}, \quad c_3 = 17.4167, \quad c_4 = 48.0313$

Table 2. Optimization Results of 12-Bar Truss

Quantity	Initial Design	Minimization of F_1	Minimization of F_2	Goal programming solution		Game theory solution
				p=1	p=2	
F_1 †	100.0	78.8911	105.2115	86.7279	102.0096	78.2999
F_2	100.0	132.5619	25.5383	24.5643	30.0372	33.1521
F_3	100.0	48.9172	134.2457	112.7366	131.6969	99.6482
F_4	100.0	162.6467	23.6538	24.3611	26.8002	37.5417
ω_1^2 (open loop)	1.34	0.494	1.41	1.15	1.61	0.655
ω_2^2 (open loop)	12.9	11.8	17.1	16.5	16.2	13.2
λ_1 (closed loop)	-0.351±1.36j	-0.129±0.501j	-0.376±1.46j	-0.290±1.17j	-0.420±1.63j	-0.175±0.679j
λ_2 (closed loop)	-0.581±12.9j	-0.354±11.8j	-8.86±16.2j	-8.80±15.7j	-8.23±15.3j	-5.83±12.3j
ξ_1	0.2503	0.2503	0.2496	0.2398	0.2501	0.2501
ξ_{12}	0.0451	0.0301	0.4801	0.4882	0.4733	0.4289

† $F_i = c_i f_i$, $c_1 = 2.2885$, $c_2 = 0.4300$, $c_3 = 0.2886$, $c_4 = 68.3045$.

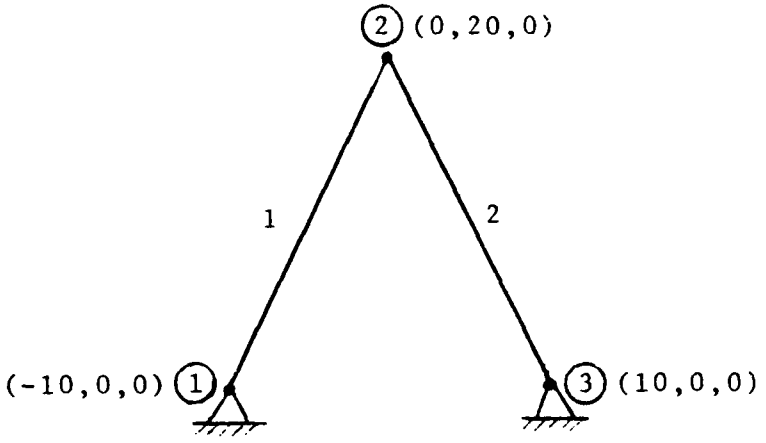


Figure 1 Two Bar Truss

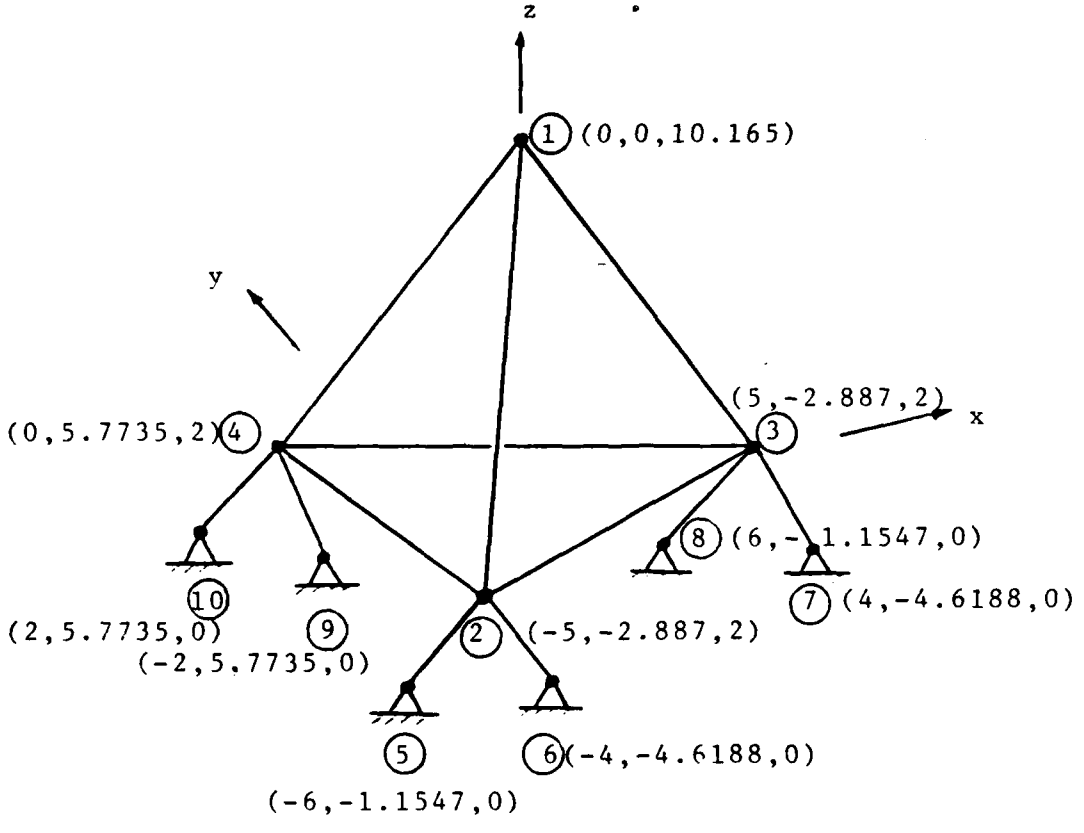


Figure 2 Twelve Bar Truss (ACOSS-FOUR)

1986 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT SUMMER SUPPORT PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

Development of a Rapid and Sensitive Assay Procedure for the Detection of the Protozoan Parasite Giardia lamblia in Drinking Water Supplies

Prepared by:	Ralph J. Rascati
Academic Rank:	Associate Professor
Department and University:	Department of Biology Kennesaw College
Research Location:	Occupational and Environmental Health Laboratory USAF/OEHL/ECQ, Brooks AFB, TX
USAF Researcher:	Major Robert D. Binovi
Date:	September 12, 1986
Contract No.:	F49620-85-C-0013

Development of a Rapid and Sensitive Assay Procedure for the Detection of the Protozoan Parasite Giardia lamblia in Drinking Water Supplies

by

Ralph J. Rascati

ABSTRACT

The development of procedures for concentration and detection of the protozoan parasite Giardia lamblia were initiated. Membrane filtration was used for concentration of the cysts from water supplies and an immunoassay is to be developed for the detection of cysts. Development of the immunoassay procedure will take longer than the ten-week period available but the preparation of antibodies against Giardia cysts was begun by injecting rabbits with cyst preparations. The injection schedule and part of the bleeding schedule were completed during the summer period. The bleeding schedule should be completed and the sera obtained used for development of the assay procedure. Testing of the membrane filtration procedure for concentration of cysts was undertaken using water samples deliberately seeded with known amounts of Giardia cysts. It was found that filters with a pore size of 3 um were more efficient (14-50% recovery) than filters with a pore size of 5 um ($\leq$ 17% recovery). Although the larger pore filters could process slightly larger samples before filter clogging occurred the difference in sample size did not fully compensate for the decreased recovery efficiency. These results are however, preliminary and further work is needed to establish the relationship between water quality, maximum sample size which can be processed and recovery efficiency for the two different pore size filters. A field trial of the concentration procedure demonstrated its ease of use. However, although Giardia-like objects were observed they could not be unequivocally identified. This serves to strengthen the need for the objective immunoassay procedures.

Acknowledgments

I would like to thank the Air Force Systems Command, The Air Force Office of Scientific Research and the Occupational and Environmental Health Laboratory (OEHL) for sponsorship of my research. I would especially like to thank several members of the OEHL staff, Major Robert Binovi, Colonel Marlin Sweigert, Colonel Bruce Poitrast, and Colonel James Rock, who made my stay intellectually stimulating, fruitful and enjoyable. I would also like to express my appreciation to those members of the OEHL staff who were instrumental in locating and providing many of the items needed to complete the project. It would not have been possible without them.

Finally, there are two people I would especially like to thank; my wife Daryleen for her encouragement and her willingness to relocate for the summer, and my niece Lisa Sproles for taking care of things at home in Georgia during our absence.

I. INTRODUCTION

I received my Ph.D. in Biochemistry from the University of Massachusetts studying the ability of isolated rat liver mitochondria to synthesize some of their own proteins. During that time I gained experience with the kind of immunological procedures used in the present project. I then spent 4 years at the Oak Ridge National Laboratory working on the regulation of eukaryotic gene expression and on genetic toxicology using viruses as a model system. After moving to Illinois State University in 1979 I continued to pursue the projects developed at the Oak Ridge National Laboratory. While there however, I became interested in water and wastewater quality standards with particular emphasis on standards pertaining to the presence of viruses and other potentially pathogenic microorganisms in water supplies. I have continued to pursue this interest since moving to Kennesaw College in Marietta Georgia in September of 1985. My goals are to develop and use procedures for detecting potential pathogenic organisms in order to assess the efficiency of currently available treatment procedures at removing and/or inactivating those organisms.

The OEHL is primarily a consultant division that deals with problems arising at various Air Force installations pertaining to occupational and environmental health and safety. One of the problems that they have been consulted about is the outbreak, on several Air Force installations, of a gastrointestinal disorder (giardiasis) caused by the protozoan parasite Giardia lamblia. Other outbreaks of this disease have been traced to contaminated water supplies. The OEHL needed therefore, an assay procedure for sampling water supplies and detecting the presence of Giardia lamblia. Since this problem was closely related to my own interest in water quality standards I was assigned the task of developing an assay procedure that could be used by OEHL personnel to detect the presence of Giardia in drinking water supplies. Once it is developed I will also use this assay procedure for an investigation of the

natural occurence of Giardia in water and wastewater and its removal and/or inactivation by various treatment procedures. Standards for removal of potentially pathogenic organisms such as Giardia can be developed rationally only when we have the means to rapidly and accurately assess the outcomes of procedures designed to remove such organisms.

II. OBJECTIVES OF THE RESEARCH EFFORT

The overall objective of this project is the development of an assay procedure for the detection of Giardia lamblia in water supplies that is relatively rapid, is sensitive, can be performed by laboratory personnel with a minimum of training, and which does not require subjective evaluation and interpretation of the results insofar as possible. Specific objectives are:

A. Analysis of the current literature and comparison of the reported efficiencies of the methods currently used for concentrating the organism from contaminated water supplies.

B. Analysis of the current literature and a comparison of the reported methods for identification of Giardia cysts in concentrated samples.

C. Development and field testing of a method for recovering cysts of the organism from large volumes of water which can be applied by relatively unskilled workers at the site of an outbreak.

D. Development of a method for identification of Giardia lamblia cysts in the field samples which is rapid, sensitive, easy to do, and does not rely heavily on subjective interpretation.

III. COMPARISON OF CURRENTLY USED CONCENTRATION METHODS

Giardia lamblia is a flagellated protozoan responsible for the gastrointestinal disorder known as giardiasis (see ref 1 for review). The organism has a two-stage life cycle. The cyst stage is a dormant stage. The

organism is encased in a relatively tough covering which gives to the organism resistance to environmental degradation and resistance to stomach acid. When ingested the organism survives in the stomach and is excysted in the small intestine where the conditions are no longer acidic. The organism then enters into the reproductive vegetative stage (trophozoite stage). Humans infected with Giardia can either be asymptomatic or may develop acute or chronic symptoms, including diarrhea, abdominal distention, flatulence, nausea, belching, anorexia, vomiting, fatigue, and/or abdominal cramps. As the organisms pass through the small intestine into the large intestine they are again encysted and the cysts are shed in the feces.

Waterborne transmission of giardiasis was suggested in 1946 when an outbreak in Tokyo was traced to sewage contamination of a water supply. Documentation of waterborne transmission was obtained in Aspen, Colorado in 1965 when cysts were detected in the water supply during an outbreak. Since the time of the Aspen outbreak over 80 outbreaks affecting over 25,000 persons have been reported. Most outbreaks are relatively small, affecting <200 persons, but several of the outbreaks have been quite large. Cysts were recovered from the water supplies in only some of the outbreaks while in others the waterborne etiology of the disease was suggested from epidemiological data only. In two of the largest outbreaks (Rome, NY, >4800 cases, >10% of the population; Berlin, NH, >7000 cases, >40% of the population) multiple cysts were detected in raw and treated water. In Berlin, NH contamination resulted from joint leakages and from a poor design which allowed filtered and unfiltered water to mix, thus bypassing the filter system. In fact, most of the outbreaks have been associated with the use of surface water without filtration or with inadequate filtration. Many have occurred in New England, the Rocky Mountain states and the Pacific states where the untreated water quality is often considered good enough that many water supplies are not filtered but are only chlorinated. This

may be inadequate for treatment of Giardia, even though it is adequate for the removal of many other pathogens, because Giardia is relatively resistant to chlorination. In the plains states there have been fewer outbreaks; possibly because the water quality is so poor that both filtration and chlorination must be used and therefore Giardia is effectively removed (2-4). It is clear that more efficient methods for detection of Giardia are needed in order to more fully document the waterborne transmission of this potential pathogen and also to more efficiently analyze the effectiveness of the various treatment procedures used in water treatment facilities.

For any detection method to be useful it must be able to process large samples of water and concentrate Giardia cysts so that they may be identified. It has been estimated that raw sewage should contain between 9.63×10^{3} and 2.4×10^{5} cysts per liter assuming average per capita water comsumption and an incidence of infection of 1-25%. Calculation of cyst levels in drinking water supplies will depend, therefore on the degree of raw sewage contamination of the drinking water supply but in any case should be considerably lower than the concentration in the raw sewage (5). Where contamination of the water supply is from some source other than raw sewage leakage the level of cysts in the water cannot be predicted but might be expected to be even lower. A number of methods have been used to concentrate cysts in order to detect possible contamination by Giardia lamblia. The most commonly used of these methods are summarized and compared below:

1) EPA Method (5) - The EPA developed a method for concentrating cysts from water samples which has been widely accepted as the standard method. In fact it is the method listed in the "Handbook of Standard Methods for the Analysis of Water and Wastewater". Briefly large samples of water can be filtered through a cartridge filter of wound orlon yarn. A water meter is

connected downstream from the filter and connections are made with standard garden hose fittings. The requisite volume of water is passed through the filter and the filter is then shipped on ice to the laboratory. The filter fibers are cut and homogenized in a blender with distilled water to extract the cysts. The extract is then subjected to flocculation and/or screening to further concentrate the cysts. The advantage of this method is its portability, ease of use in the field, and ability to process large water samples. The disadvantages are a complex series of manipulations once the sample has been filtered, an awkward processing procedure in extracting the cysts from the orlon yarn fibers and a relatively low efficiency of cyst recovery (Mean = 6.3%; Range = 3-15%) Analysis of the individual steps involved .ndicated that loss of cysts occurred at each stage in the process with drastic losses (>1 log) at the flocculation step if used.

2) **CDC Method** (6) - This method uses a swimming pool filter with sand as the filtering medium. Extremely large volumes can be filtered with this procedure. The filter backwash is collected in 55 gallon drums, coagulated with alum, and the sediment screened. Additional steps are then used to further concentrate the samples. In certain field trials with this method a few cysts were detected after passing 280,000 gallons through the filter. The procedure however, is difficult and complex and the efficiency has never been determined.

3) **Hygienic Laboratory Method** (7) - Hausler and coworkers developed a sampling and concentration procedure based on the MiniCapsule filter available from Gelman. This filter is an acrylic copolymer pleated membrane which is pre-sterilized by the manufacturer. Standard thread hose connections are used with a water meter connected downstream of the filter. Those authors conducted preliminary tests of the system by passing 100 gallon samples through the filters while incrementally introducing a cyst suspension such that the maximum concentration of cysts presented to the filter was 15 cysts per gallon.

Processing of the filter requires sawing open the capsule, removal and cutting of the filter to allow it to be flattened in a pan, and washing the filter with distilled water. The cysts were then concentrated by centrifugation. The system appears to have the advantages of the EPA and CDC methods in that relatively large samples (>100 gal) can be processed. Also the processing in the laboratory is simpler and seems to be better suited to routine use. However, processing is still somewhat complex and lengthy and the efficiency of the method has not been quantified.

4) Membrane Filtration Method (8) - Membrane filtration has been used unsuccessfully in the past (9,10) to attempt to identify cysts in water supplies suspected to be contaminated. The method as developed at that time however did not allow processing of large samples (>10 liters) and the efficiency of recovery had not been determined. More recently however, Wallis and Buchanan-Mappin (8) reported the use of 5 um Nucleopore membranes (110mm in diameter) to concentrate and detect cysts in untreated stream water which had been seeded with Giardia cysts in known amounts. After filtration the membrane was rinsed with small amounts of water (total = 80 ml). The cysts were then concentrated further by centrifugation and detected by microscopic examination. Using this method samples of 100-200 liters (25-50 gal) were processed. Cyst recovery in different trials indicated that the mean recovery was 71% with a range of 58-84% at optimum pressure. The lower limit of detection appeared to be a concentration of 0.5 cysts per liter if a 100 liter sample was processed. The infective dose of Giardia has been shown to be at least 10 cysts (11,12). Considering the average daily consumption this detection limit should be adequate to insure the safety of water supplies according to the authors. The disadvantage to the method is the relatively small sample size which can be processed and the unsuitability of the method for water with a high turbidity

level. However the small sample limitation is balanced by an extremely high efficiency of recovery. Furthermore, the processing of the filter in the laboratory is relatively simple and does not require any equipment not usually available in a reasonably equipped water laboratory.

IV. COMPARISON OF CURRENTLY USED DETECTION METHODS

Once water samples have been concentrated they must be accurately assessed for the presence of *Giardia* cysts. Several methods are available for detecting the presence of the cysts. Some methods are time-consuming and require highly skilled technicians while others are rapid and rely on an objective endpoint rather than the judgment of the technician. The choice of method may also depend on whether qualitative (presence/absence of cysts) or quantitative (level of cyst contamination) results are desired. Several of the most commonly used methods for detection of *Giardia* or other similar organisms are discussed below:

1) **Phase Contrast Microscopic Examination** (13,14) - The most common method for the detection of *Giardia* cysts in processed fecal or water samples is examination of stained preparations using a microscope equipped with phase contrast optics. After appropriate concentration procedures samples are transferred to a microscope slide, coverslip, or hemacytometer and stained with with dilute Lugol's Iodine. The slide is then scanned under the microscope for the presence of objects with the correct size and shape. Since many different objects that may be present in the sample can have the correct size and shape, objects meeting these criteria must be closely examined for the presence of characteristic identifying internal structures (two to four nuclei, axonemes, and/or median bodies). This method is time consuming and fatiguing both because of the low numbers of cysts found and because of the poor visual contrast between the *Giardia* cysts and contaminants of similar size and shape. The fewer

cysts present the longer it will take to scan enough of the sample to obtain accurate results. Negative samples could take as long as 20-134 hours for examination of the entire sample. Furthermore, this method requires a highly skilled and experienced technician to distinguish cysts from other similar objects.

2) **Immunofluorescence/Phase Contrast Microscopy** (15) - Immunofluorescence methods have been applied to the detection of aquatic microorganisms deposited on membrane filters (16). The filter is reacted with specific antibodies against the organism. These antibodies are coupled to a fluorescent dye. When incident Ultraviolet irradiation is used those microorganisms bound by the conjugated antibodies can be located by their bright green fluorescence in contrast to a dark background. Judith Sauch developed a similar method for the detection of Giardia cysts deposited on a membrane filter. Briefly, the sample containing Giardia cysts is deposited on a filter, reacted with specific antibody conjugated with a fluorescent dye and counterstained with Evans Blue. The treated membrane is then mounted on a microscope slide and cleared by reaction with glycerol containing propylgallate. The filter is then scanned with a microscope equipped with epifluorescence optics. The presence of Giardia cysts is then rapidly determined by the occurrence of organisms that flouresce a bright green and that have the correct morphological characteristics. Confirmation of the identity of these objects is obtained by switching to phase contrast optics and confirming the existence of the appropriate internal structures. In a comparison of this method with the simple phase contrast microscopy described above it was found that for comparable sample sizes the combination method was 28 times faster (15). Thus, if it took 20 min to scan a filter to find one immunofluorescent cyst it would take about 9 hours to examine a comparable portion of the sample by phase contrast microscopy alone. This procedure offers the advantage of being able to scan directly the filter

used to concentrate the water sample without any additional purification or concentration steps. The filter may be cut into appropriate size pieces, mounted on microscope slides, cleared, and reacted with the fluorescence tagged antibodies. The pieces may then be scanned directly for the presence of Giardia. The method can be complicated however, by the presence of non-specific fluorescence. Each fluorescent object must be scrutinized closely to determine if it is Giardia or something else. Furthermore, although faster than simple phase contrast microscopy it still takes some time to scan an entire filter. Although this would not be necessary with water samples contaminated with a high concentration of cysts, it would be necessary for low level contamination or for negative samples (13,14). Finally, while initial identification of fluorescing objects with the correct morphology would require only a minimal amount of training it is possible that confirmation by phase contrast examination might still require the trained eye of an experienced technician.

3) Enzyme Linked Immunosorbent Assay (17) - An ELISA assay has been developed for the detection of Giardia lamblia in fecal specimens. The basic reaction involves coating a plastic surface with Giardia, reacting the coated surface with specific antibodies against the organism, and subsequently reacting with antibodies against the immunoglobulin type represented by the anti-Giardia antibodies. This second antibody is conjugated to an enzyme such as alkaline phosphatase whose presence can then be measured by a simple and sensitive colorimetric assay. The basis for the assay lies in the presence or absence of Giardia in the original sample. If the organism is present then the specific antibody and subsequently the enzyme conjugated antibody will remain in the tube and their presence will be indicated by a positive colorimetric reaction. If the organism is not present in the sample then the specfic antibody and subsequently the conjugated antibody would be washed out of the

tube during the procedure and no colorimetric reaction would take place. Quantitative analysis of the procedure currently in use indicated that as few as 12 cysts could be detected. However, considering the small sample size that can be used in the procedure this actually represented a fairly high concentration of Giardia. It should be possible to develop a similar procedure for the analysis of concentrated water samples although certain steps would have to be taken to increase the sensitivity of the assay. The authors (17) used Giardia trophozoites as their antigen for the preparation of the specific antisera. Increased sensitivity might be obtained if cysts were used instead since this is the form of the organism to be detected in water samples. Additional steps can also be taken to increase the sensitivity and to increase the sample size that can be assayed. The major advantages to this procedure are that multiple samples can be processed simultaneously in a relatively short time and that quantitative results are obtained using a completely objective endpoint. Furthermore, no subjective judgment on the part of the laboratory personnel is required. The major disadvantage may lie in the small sample size which can be processed. It is not yet known if this will be a problem even for concentrated water samples.

V. DEVELOPMENT AND FIELD TESTING OF A CONCENTRATION METHOD

Of the concentration methods described in Section III the membrane filtration method was chosen for initial testing. Two major reasons for this choice were the ease of processing using this method and the high published recovery rates for the use of this method. Initial experiments involved seeding water samples with known quantities of Giardia cysts, pushing them through the filter, recovering them by washing the filter and concentrating them by sedimentation. Identification of the cysts and quantification of the recovery was done by counting wet mounted samples of known volume which had been stained

with dilute Lugol's Iodine. Cysts were identified by phase contrast microscopy since no other method is as yet available. In addition, the turbidity of the various water samples was measured in a nephelometer and the content of organic material was measured by determination of the chemical oxygen demand (COD). The maximum sample size which can be processed by this or similar methods has been demonstrated to correlate with turbidity (Wallis, personal communication) or COD (Binovi, personal communication). Therefore, these parameters are being measured in my experiments in order to more fully define the limits of application of this method of cyst concentration.

Recoveries were tested using two different filters. One had a nominal pore size of 3 um. In several experiments I obtained recoveries which ranged from 14 to 50% for Giardia cysts (8 x 16 um) and for smaller cysts of unknown origin (5 um dia.) which were also present in the preparations. None of the larger (Giardia) cysts were observed in the material which passed through the filter while 32% of the smaller cysts did pass through. This would suggest that the unrecovered Giardia cysts were either trapped on the filter or did not sediment after elution from the filter. Recommendations are made below to increase the recovery of these cysts. Alternatively, cysts might be damaged during the filtration process such that they are no longer recognizable. When a filter with a slightly larger pore size (5 um) was used, recovery of the Giardia cysts was reduced to $\leq$ 17% and recovery of the smaller cysts was reduced to <6%. Furthermore, 43% of the Giardia cysts and 91% of the small cysts passed through the filter and were detected in the filtrate. When turbid water was used as the sample it was found that the 5 um filters could process slightly larger volumes of sample than could the 3 um filter. However, the increase in volume did not completely compensate for the lowered recovery efficiency. These results are preliminary and need further investigation before firm recommendations are made but it would appear that the 3 um filter would be the best one to use, giving

the optimum compromise between between the maximum sample size which can be processed and the efficiency of recovery of Giardia cysts.

Near the end of the summer period one field trial of the concentration procedure was attempted at Makah AFS, WA. The water supply was a small reservoir fed by surface run-off. Furthermore, the area has much indigenous wildlife including a beaver population. Thus, conditions are conducive to the presence of Giardia although no outbreaks of giardiasis have been reported. This trial demonstrated the ease with which the apparatus can be shipped and set-up to process samples. Preliminary examination revealed several protozoa with the same approximate size and shape as Giardia. However, concrete identification was not possible without further testing. These results, although preliminary only serve to support the need for an objective identification procedure not strictly relying on the discriminatory skills of the technician. Finally, additional field trials should be attempted at sites where giardiasis has been diagnosed and where the water supply has been implicated on the basis of epidemiological data in order to fully validate the procedure.

VI. DEVELOPMENT OF A RAPID AND SENSITIVE ASSAY PROCEDURE FOR THE DETECTION OF GIARDIA CYSTS IN CONCENTRATED WATER SAMPLES.

It is clear from the discussion in Section IV that rapid sensitive assay procedures for the detection of Giardia will involve some type of immunoassay. It is not clear however, which of the two assays discussed, immunofluorescence-phase contrast microscopy (IFPC) or Enzyme Linked Immunosorbent Assay (ELISA), will be the best. It may even be found that both assays may be desirable and that specific circumstances may dictate the choice of assay to be used in a specific situation. In either case the first step in the development of the assay is the production of specific antibody directed against Giardia lamblia cysts. Cysts were obtained from Dr. Frank W. Schaeffer, III of the the Health

Effects Research Laboratory at the Environmental Protection Agency in Cincinnatti, Ohio. Three New Zealand White (NZW) rabbits were bled prior to immunization. The sera collected will be used as pre-immune control sera in subsequent assays. Each rabbit was then injected with cysts in the thigh muscle of the right hind leg with an emulsion containing 2.5×10^6 cysts suspended in 0.5 ml distilled water mixed with an equal volume of Freund's complete adjuvant. One month later the rabbits were hyperimmunized with the same dose of cysts in incomplete Freund's adjuvant. The rabbits are then bled 2, 4, and 6 weeks after the second injection. The sera collected are pooled and stored for later use in the immunoassays. As can be seen from the time schedule, under ideal conditions it would take the entire 10 week summer period to produce the needed antisera. Development of the actual immunoassays would take considerably longer and will be the major subject of a follow-on Mini-Grant proposal. In fact, because of delays in obtaining the cysts and in processing of a required animal use proposal, the initial injection of the rabbits did not take place until the fourth week of the project period. Consequently, there was only time to do the pre-bleed, the initial and booster injections, and the first (2-week) bleeding. Arrangements have been made for animal facility personnel to do the 4- and 6-week bleedings and to ship the antisera to me. If the Mini-Grant proposal is funded I will then develop and test the immunoassay procedures. Since it is not possible at this time to predict which of the two immunoassay procedures discussed (IFPC or ELISA) will be most advantageous under the circumstances normally encountered both procedures will be developed simultaneously and evaluated for their use potential.

VII. RECOMMENDATIONS

1) Analysis of the operational limits of the membrane filter procedure should be completed. Specifically, the maximum sample size that can be pushed

through the filter should be determined for water samples of various measured turbidity and COD levels. If either of these parameters can be used to predict the maximum sample size then these tests, which can be easily and quickly performed, should be run prior to filtration. This work has been started during this 10 week period but additional values for these parameters and additional replicates of samples with values already tested need to be performed.

2) Full comparisons of the two filter porosities (3 um and 5 um) should be obtained. Maximum sample volumes and recovery efficiencies should be compared under a variety of laboratory and field conditions so that a logical choice of filter can be made. Additional efforts to improve recovery efficiency should include more vigorous continuous agitation during the elution step and the use of centrifugation during the sedimentation step.

3) As a practical consideration, it is recommended that the membrane filtration apparatus (Gelman, $1,000) purchased for use in this project and the peristaltic pump used to push the sample through the filter be transferred to the Principal Investigator if the mini-grant is funded. This will avoid needless duplication of expenditure since the available funds are limited. The apparatus would be returned to OEHL upon termination of the funding period.

4) The immunoassays discussed in Section VI should be developed. Either the IFPC or the ELISA procedure should make the assay of water samples for the presence of Giardia cysts faster, more convenient, sensitive and objective. The need for this was strengthened by the reults of the field trial in which organisms were found which could have been Giardia cysts but which could not be identified unequivocally. The preparation of the antisera described in Section VI is the first step in the development of either assay procedure. In fact, it is my recommendation to develop both of these assays. The relative merits of the two procedures as routine assays for the presence of Giardia or as research tools can only be determined if both assays are available.

VIII. LITERATURE CITED

1. Erlandsen, L. & E.A. Meyer (Eds.) Giardia and Giarsiasis, Plenum Press, New York, 1984.

2. Craun, G.F. (1979) Waterborne Giardiasis in the United States: A Review, Amer. J. Public Health, 69:817-819.

3. _____ (1985) Giardiasis: The New Waterborne Disease, A.W.W.A., 73:33.

4. _____ (1985) Waterborne Giardia: It's Enough to Make You Sick, A.W.W.A. 77:14

5. Jakubowski, W. & T.H. Ericksen (1979) Methods for detection of Giardia cysts in water supplies, In: "Waterborne Transmission of Giardiasis" (W. Jakubowski & J.C. Hoff, eds.) EPA 600/9-79-001, pp.193-210.

6. Shaw, P.K., R.E. Brodsky, D.O. Lyman, B.T. Wood, C.P. Hibler, G.R. Healy, K.J.E. MacLeod, W. Stahl & M.G. Schultz (1977) A communitywide outbreak of giardiasis with evidence of transmission by a municipal water supply. Ann. Intern. Med. 87:426-432.

7. Hausler, W.J., Jr., W.E. Davis & N.P. Moyer (1984) Development and testing of a filter system for isolation of Giardia lamblia cysts from water, Appl. Environ. Microbiol. 47:1346-1347.

8. Wallis, P.M. & J.M. Buchanan-Mappin (1985) Detection of Giardia cysts at low concentrations in water using nucleopore membranes, Water Res. 19:331-334.

9. Chang, S.L. & P.W. Kabler (1956) Detection of cysts of Endamoeba histolytica in tap water by the use of a membrane filter, Am. J. Hyg. 64:170-180.

10. Moore, G.T., V.M. Cross, D. McGuire, C.S. Mollohan, N.N. Gleason, G.R. Healy & L.H. Newton (1969) Epidemic giardiasis at a ski resort, New Eng. J. Med. 281:402-407.

11. Rendtorff, R.C. (1954) The experimental transmission of human intestinal protozoan parasites. II. Giardia lamblia cysts given in capsules. Am. J. Hyg. 59:209-220.

12. Rendtorff, R.C. (1979) The experimental transmission of Giardia lamblia among volunteer subjects. In: "Waterborne Transmission of Giardiasis" (W. Jakubowski & J.C. Hoff, eds.) EPA 600/9-79-001, pp. 64-81.

13. Jakubowski, W. (1984) Detection of Giardia cysts in drinking water. In: "Giardia and Giardiasis" (S.L. Erlandsen & E.A. Meyer, eds.) Plenum Press, New York, pp.263-286.

14. Schaeffer, F.W. & E.W. Rice (1982) Giardia methodology for water supply analysis. In: "Proceedings of the American water Works Association Water Technology Conference. American Water Works Association, Seattle, pp.143-147.

15. Sauch, J.F. (1985) Use of immunofluorescence and phase-contrast microscopy for detection and identification of Giardia cysts in water samples. Appl. Environ. Microbiol. 50:1434-1438.

16. Schieman, D.A. (1981) Advances in membrane filter applications for microbiology. In: "Membrane Filtration" (B.J. Dutka, ed.) Marcel Dekker, Inc., New York, pp.537-566.

17. Ungar, B.L.P., R.H. Yolken, T.E. Nash & T.C. Quinn (1984) Enzyme-Linked Immunosorbent Assay for the detection of Giardia lamblia in fecal specimens, J. Infect. Dis. 149:90-97.

1986 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/
GRADUATE STUDENT SUMMER SUPPORT PROGRAM

Sponsored by the
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
Conducted by the
Universal Energy Systems, Inc.
FINAL REPORT

State Variable Model of the Cardiovascular System and a Controller Design for an Anti-G Suit

Prepared by:	Kuldip S. Rattan
Academic Rank:	Associate Professor
Department and University:	Department of Electrical Systems Engineering Wright State University
Research Location:	Armstrong Aerospace Medical Research Laboratory, Bioengineering and Biodynamics Division Acceleration Effects Branch Wright Patterson Air Force Base
USAF Researcher:	Dr. Daniel W. Repperger
Date:	August 22, 1986
Contract No:	F49620-85-C-0013

State Variable Model of the Cardiovascular System and a Controller Design for an Anti-G Suit

by

Kuldip S. Rattan

ABSTRACT

A state-variable model of the cardiovascular system under $+G_z$ stress was implemented. The model (which includes simulation of the arterial and venous systems, heart, baroreceptor control of the heart, and venous tone, and inputs for acceleration force and externally applied pressure) was used to study the impairment of cerebral function during $+G_z$ stress. It was found that even though eye level blood pressure decreases significantly during G_z stress, cerebral blood flow is maintained due to a compensatory mechanism which compares favorable with the experimental results found in the literature. This model will be used to investigate the effectiveness of anti-G suits. Finally, a preliminary design of a closed-loop control system for an anti-G suit was carried out. It was found that it is possible to control both the rise time (which is needed for the improvement of G-valve) and the overshoot of the suit pressure. More work needs to be done both in the simulation and design areas.

Acknowledgements

The author hereby expresses his gratitude to the Air Force Systems Command, the Air Force Office of Scientific Research and the Universal Energy Systems for providing him with the opportunity to spend a very worthwhile and interesting summer at the Armstrong Aerospace Medical Research Laboratory, Wright Patterson Air Force Base, Ohio. He appreciates the hospitality and excellent working conditions provided by Dr. Robert E. Van Patten and the Acceleration Effects Branch (AAMRL/BBS).

The author would like to thank Dr. Daniel W. Repperger for his conscientious coordination of this research effort, making sure that the author felt welcome. He would like to thank several members of the AAMRL/BBS branch, Mr. J.W. Frazier, Lt. Edward Hade, Mr. David Ratino and Mr. Lawrence Gould for their help during the course of this research. Last, but not least, he would like to express his thanks to Drs. Henning von Gierke and Ints Kaleps for many helpful discussions.

Finally, he would like to thank his graduate student, Mr. S. Aziz, for his help during the course of this research and Suzy McGovern of the College of Engineering and Computer Science, Wright State University for her careful and professional typing of this report.

I. Introduction

I received my Ph.D. from the University of Kentucky studying the problem of converting existing continuous control systems into digital control systems. I later was a post-doctoral fellow at the University of Kentucky in the department of Physiology and Biophysics.

The Harry G. Armstrong Aerospace Medical Research Laboratory (AAMRL) at Wright Patterson Air Force Base has been involved in studying the effects of high onset rate, high sustained $+G_z$ forces, and the development of new techniques for increasing $+G_z$ tolerance. The primary effect of high onset rate, high sustained $+G_z$ aceleration force is the pooling of blood in the lower extremeties and abdomen while blood pressure in the head is reduced and the brain is deprived of the blood. This results in the human going from a full clear visual field to loss of peripheral vision, loss of central vision, complete "black-out", and finally loss of consciousness. In recent years, G-induced loss of consciousness of the pilot has been implicated in several aircraft mishaps. Pilots of aircraft, such as the F-15 and F-16, that have very high sustained $+G_z$ and high onset rates capabilities are more likely to be involved in this type of mishap [8]. Tolerance to G_z forces can be extended to some extent by applying pressure to lower vasculatures by inflation of an anti-G suit and through straining maneuvers (M1 and L1).

Studies concerned with the effects of G_z forces and the anti-G suits on the cardiovascular system have been mostly carried out in the laboratory using the centrifuge, Dynamic Enviroment Simulator (DES). Well controlled and reproducible experiments are difficult to obtain using living subjects. Furthermore, it is difficult to obtain direct measurements of hemodynamic variables during acceleration. Modeling studies have produced some useful data. Cardiovascular models can be used to predict pressure and flow at the eye and leg level and therefore are important in evaluating techniques used to augment G tolerance. Because of my past experience in computer simulation, I was assigned to develop a simulation of the cardiovascular system for the laboratory.

The AAMRL has also been involved in the development of an anti-G valve [30]. A new concept for an anti-G valve was designed and built at the laboratory. The valve, called the Bang-Bang Servo Valve (BBSV), was designed to be a low cost, quick fix for the F-16 to protect its aircrew from the unique physiologial hazard of high onset rate, high sustained acceleration. The design is a hybridization of a conventional inertially operated High Flow Valve (HFOV) and uses an electronically controlled solenoid to drive the anti-G suit pressure to the maximum level when the level of acceleration exceeds both $+2G_z$ and the onset rate of $+2G_z$/sec. After a 1.5 sec. period, the valve reverts to the HFOV unless the trigger control is met again. The BBSV provides a 1G improvement over the standard valve and a 0.5G improvement over the HFOV during the period of greatest physiological vulnerability. Because the BBSV is designed to drive the suit pressure to the maximum level if the trigger condition is met, high suit pressure for an onset rate of $>2G_z$/sec and sustained $+G_z$ of less than $5G_z$ could be uncomfortable to the pilot. BBSV basically is an open-loop control system. Suit pressure for lower G_z could be reduced by closing the loop. In this study a closed-loop control strategy is proposed for an anti-G suit.

II. Objectives of the Research Effort

The overall objective of this research is to study existing methods of increasing $+G_z$ tolerance to acceleration forces and to propose a control strategy to improve the design of existing g-suits. We feel this can be done if the effects of the $+G_z$ forces on the cardiovascular system and the existing methods of improving $+G_z$ tolerance are better understood. Our objectives were:

1) To study existing mathematical models of the cardiovascular system under $+G_z$ forces. Develop a computer simulation of one of the models and study the effects of $+G_z$ forces on the cardiovascular responses such as eye and leg level pressures and flows.

2) Study the existing protective mechanism such as g-suits and straining maneuvers and develop a computer model of the g-suit.

3) Propose a closed-loop control strategy to improve the design of the existing g-suit.

III. Mathematical Model of Cardiovascular System Under +G_z Stress

The block diagram of the cardiovascular system for which the mathematical model was developed [Chu (4)] is shown in Fig. 1. A brief description of the model is given below:

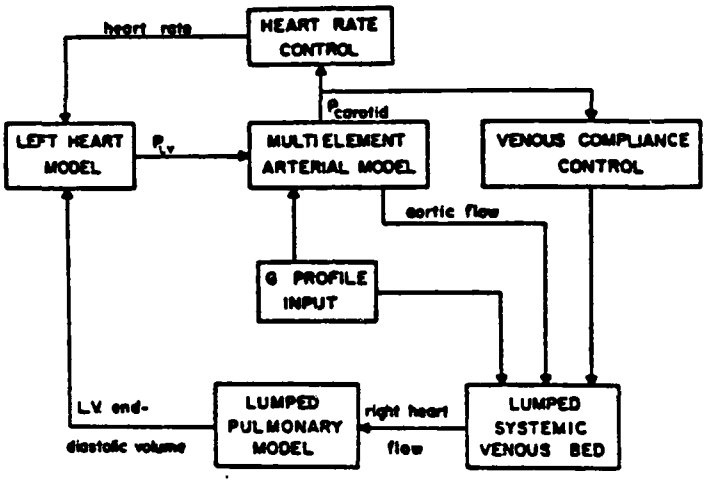


Fig. 1. Block diagram of the Cardiovascular System Model

3.1 Multilevel Arterial Model

The arterial model as shown in Fig. 2(a) consists of 24 distinct sections of aorta and major branches of arterial tree and is based on the Navier-Stokes equations. The model is nonlinear and is based on the the following assumptions: The blood is an incompressible Newtonian fluid; arterial wall thickness is small compared to vessel diameter; fluid flow is one dimensional; velocity profile is flat except near the vessel wall; wall stress-strain relationship in the physiological range is linear; wall motion in the radial direction is small and proportional to instantaneous volume stored in the elements.

The properties of each arterial segment are represented by circuit

elements shown in Fig. 2. The formulas to compute resistance, inductance and compliance are given by

$$R_{1n} = 8\mu\Delta z_n/8\pi r_n^4 \qquad L_n = 9\rho\Delta z_n/4\pi r_n^2$$

$$C_n = 3\pi r_n^3\Delta z_n/2Y_n h_n \qquad R_{2n} = .002/C_n$$

where

r_n and Δz_n = radius and length of nth arterial segment.

μ and ρ = blood viscosity and density.

Y_n ahd h_n = Young's modulus and wall thickness of the nth arterial segment

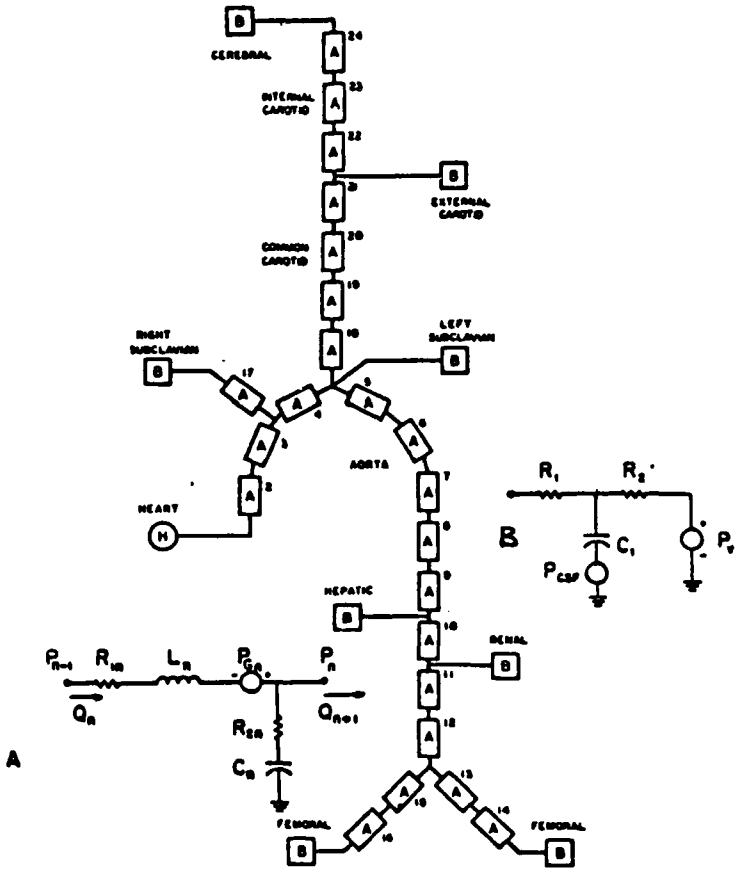


Fig. 2 Multielement Arterial Model

3.2 Lumped Systemic and Pulmonary Model

The lumped systemic and pulmonary model shown in the block diagram of Fig. 1 is given in the circuit form in Fig. 3

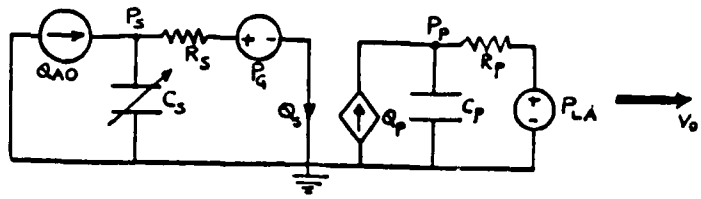


Fig. 3 Circuit Diagram of the Lumped Systemic and Pulmonary Model

Q_{AO} = Average flow generated by the multielement model

P_G = Voltage source due to hydrostatic pressure (Lumped Systemic Model is assumed to be separated by a vertical distance of 1.3 cm from heart (Green & Miller [13])

V_O = Left ventricular end diastolic volume
= Total blood volume - volume stored in C_s - volume stored in C_p.

As G_z acceleration increases, blood is pooled in the systemic venous compliance C_s and thus reduces the venous flow, Q_s, through the systemic venous resistance R_s. Therefore, V_O is calculated after every beat to adjust for the pooling of blood in C_s.

3.3 Left Ventricular and Coronary Model

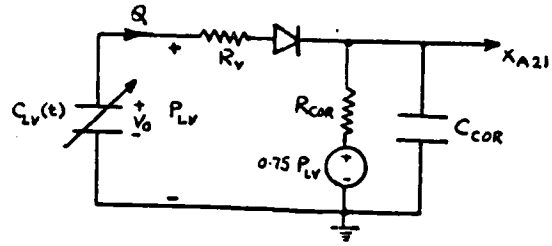


Fig. 4 Circuit Diagram for Left Ventricular and Coronary Model

$$P_{LV} = e(t)\left[V_O - \int_0^t Q(z)dz\right] \tag{3.3.1}$$

where e(t) represents the variation of left ventricular compliance $C_{LV}(t)$ during systole (Suga and Sagawa [27]) and is shown in Fig. 5

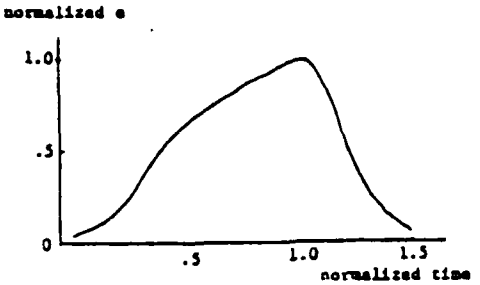


Fig. 5 Normalized e(t) curve

3.4 Reflex Compensation

The main response of the body to circulatory stress is the alteration of the heart rate and the change of vascular tone. These reflex compensations are attributed to the baroreceptors response to arterial blood pressure in the carotid sinus and aortic arch. Since the blood pressure in the aortic arch remains essentially constant, most of the stimuli responsible for the reflex compensation arise in the aortic sinus (Howard [15]).

3.4.1. Venous Tone Control:

The venous tone control followed a method developed by Green and Miller [13] and is shown in the block diagram form in Fig. 6

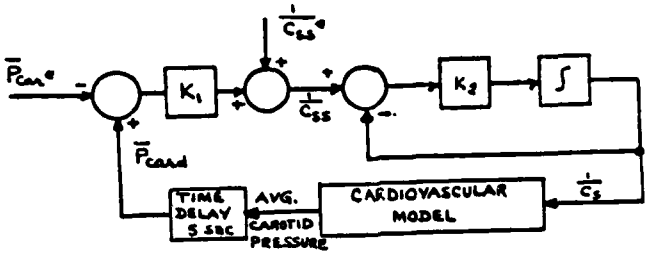


Fig. 6 Block diagram of Venous Tone Control

In Fig. 6, $\bar{P}^{*}_{car}$ and C^{*}_{ss} are the average carotid pressure and steady-state systemic compliance under no G_z stress.

3.4.2. Heart Rate Control:

The heart rate control follows a method developed by Katona [18] and modified by Snyder [26] and is shown in the block diagram form in Fig.7

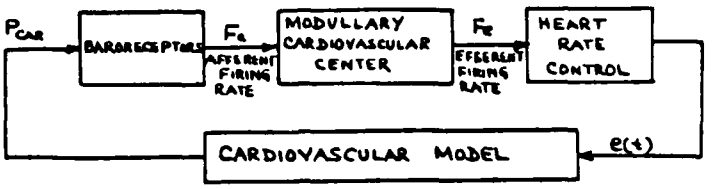


Fig. 7 Block diagram of Heart Rate Control

The baroreceptors are modelled as shown in Fig. 8

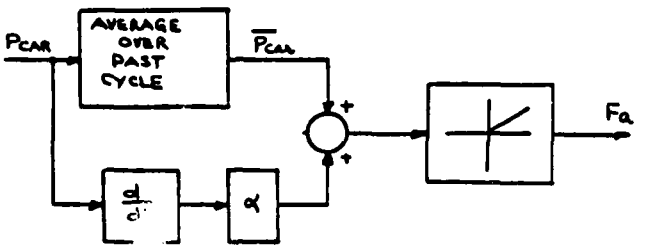


Fig. 8 Block diagram of the baroreceptors control

The efferent firing rate follows the afferent firing with a certain time constant as shown in Fig. 9

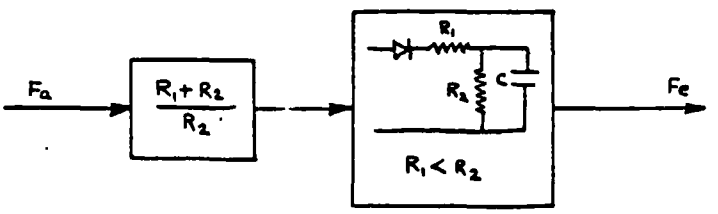


Fig. 9 Circuit diagram of efferent firing

If Fa > Fe, capacitor C charges through resistors R_1 and R_2 and Fe is given by

$$Fe((HB+)) = Fe(HB) + (Fa(HB) - Fe(HB))e^{-HP/\tau_1}$$

where HP is the heart period and $\tau_1 = (R_1R_2C)/(R_1+R_2)$.

If Fa < Fe, capacitor C discharges through resistor R_2 and Fe is given by

$$Fe((HB+)) = e^{-HP/\tau_2}$$

where $\tau_2 = R_2C$.

The value of Fe((HB+)) then determines the variation of the left ventricular compliance e(t) for the next cycle. This is done by determining the periods of systole (T_s) and diastole (T_d) as follows

$$T_d = \frac{1}{-AFe((HB+))+B} \qquad T_s = .11 + .099\ Td \qquad (3.4.1)$$

IV. State-Space Description of the Multielement Arterial Model

State-space description of the multielement arterial model shown in Fig. 2 was obtained by considering each of the 24 distinct elements. In this report, because of the page limitations, only the special cases of the model are described and the model is written in the matrix form. The state variables considered are

X_{An1} - current through the inductor of segment A (flow through segment n)

X_{An2} - voltage across capacitor of segment A ($1/C_n$ x volume of blood in element n)

X_{Bn} - voltage across capacitor of segment B

4.1 Segment A2

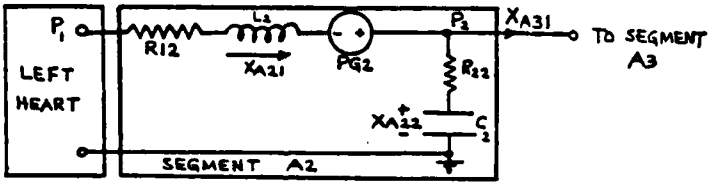


Fig. 10 Circuit Diagram of Segment A2 of Multielement Arterial Model

State and output equations of the segment A2 given in Fig. 10 are derived as

$$\begin{bmatrix} \dot{X}_{A21} \\ \dot{X}_{A22} \end{bmatrix} = \begin{bmatrix} A11 & A12 & A13 \\ A21 & A22 & 0 \end{bmatrix} X_{A2} + \begin{bmatrix} B11 & B12 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} P1 \\ PG2 \end{bmatrix}$$

$$P2 = [R22 \quad 1 - R22]\ X_{A2}$$

where $A11=-(R_{12} + R_{22})/L2$, $A12=-1/L2$, $A13=R22/L2$, $A21=1/C2=-A22$, $B11=B12=1/L2$ and $X_{A2}=[X_{A21}, X_{A22}, X_{A31}]^T$.

4.2 Segment A3

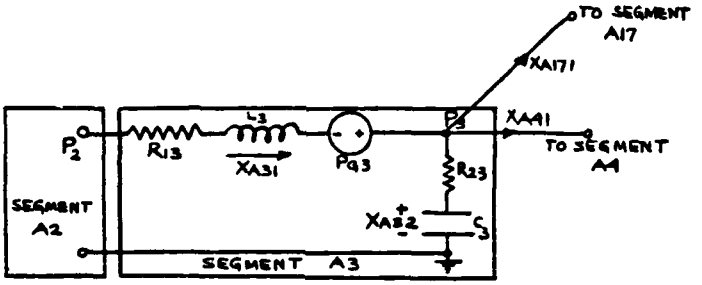


Fig. 11 Circuit Diagram of Segment A3

$$\begin{bmatrix}\dot{X}_{A31}\\ \dot{X}_{A32}\end{bmatrix} = \begin{bmatrix}A31 & A32 & A33 & A34 & A35 & A327\\ 0 & 0 & A43 & 0 & A45 & A427\end{bmatrix} X_{A3} + \begin{bmatrix}B33\\ 0\end{bmatrix} PG3$$

$$P3 = [0 \quad 0 \quad R_{23} \quad 1 \quad -R_{23} \quad -R_{23}]\, X_{A3}$$

where

A31=R22/L3, A32=1/L3, A33=-(R22 + R13 + R23)/L3, A34=-1/L3, A35=A327 =R23/L3, A43=1/C3=-A45=-A427 and $X_{A3}=[X_{A21}, X_{A22}, X_{A31}, X_{A32}, X_{A41}, X_{A171}]^T$.

4.3 Segment A4:

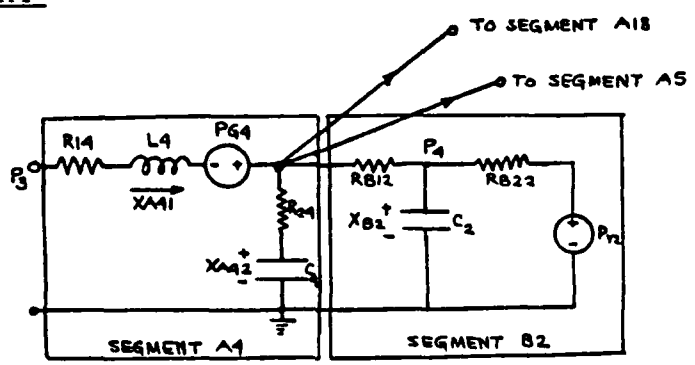


Fig. 12 Circuit Diagram of Segments A4 and B2

$$\begin{bmatrix}\dot{X}_{A41}\\ \dot{X}_{A42}\\ \dot{X}_{B2}\end{bmatrix} = \begin{bmatrix}A53 & A54 & A55 & A56 & A57 & A527 & A529 & A544\\ 0 & 0 & A65 & A66 & A67 & 0 & A629 & A644\\ 0 & 0 & A445 & A446 & A447 & 0 & A4429 & A4444\end{bmatrix} X_{A4} + \begin{bmatrix}B54 & 0\\ 0 & 0\\ 0 & B44V2\end{bmatrix} \begin{bmatrix}PG4\\ \\ PV2\end{bmatrix}$$

where A53=R23/L4, A54=1/L4=B54, F1=L4(RB12 + R24), F2=R24RB12, A55=-R14/L4-F2/F1-A53, A56=-RB12/F1, A57=F2/F1, A527=-A53, A529=A527, A544=-R24/F1, F2=C4(RB12+R24), A65=-R24/F2+(1/C4), A66=-1/F2, A67=R24/F1, A629=A67, A644=A65/RB12, F3=CB2(RB12+R24), A445=R24/F3, A447=1/F3, A447=-A445=A4429, A4444=(A445-(1/CB2))/RB12-1/RB22CB2, B44V2=1/RB22CB2, and $X_{A4}=[X_{A31}, X_{A32}, X_{A41}, X_{A42}, X_{A51}, X_{A171}, X_{A181}, X_{B2}]^T$.

4.4 State Variable Model

State-space representation of the other segment of the arterial segments can be obtained in a similar manner. Combining all the equations, the state-space model of the multielement arterial tree can be written in the matrix form as

$$\dot{\underline{X}} = A\underline{x} + B\underline{u}$$

$$\underline{Y} = C\underline{x} + D\underline{u}$$

where

$$\underline{x} = [X_{A21} \quad X_{A22} \cdots X_{A241} \quad X_{A242} \quad X_{B1} \cdots X_{B7}]^T$$

$$\underline{u} = [P_1 \quad G]^T$$

Because of the page limitations, A, B, C and D matrices are not given in this report, but can be obtained from the author.

V. Complete Simulation of the Model and Results under $+G_z$ stress

The state-variable model described in sections 3 and 4 was programmed on the VAX 750 computer using SYSTEM-BUILD of the $MATRIX_x$ computer aided package. The programming and the simulation results are given in Mr. S. Aziz's final report. These results show the pressures at left ventricular (LVP), aortic root (AOP), leg (LP) and eye level (EAP). The LVP, AOP, and EAP drop initally, then gradually return toward the control values. An overshoot appears when the G_z stress is removed. This is followed by a gradual return to pre-G values. However, LP does not follow this trend and has a higher pressure during $+G_z$ stess. This is because it is located below the heart level and experiences a postive hydrostatic pressure in addition to the blood pressure transmitted from the heart. It should be noted that although

initally, the systolic EAP pressure drops to a negative value, it reaches a value greater than 20mm Hg later. This result indicates that human subjects can recover central vision after a period of blackout or even unconsciousness.

VI. Methods of Increasing $+G_z$ Tolerance

The limiting factor in the cardiovascular system, from the viewpoint of $+G_z$ stress, is the reduction of eye level blood pressure and subsequent loss of vision. The pressure gradient changes caused by the G_z loading and blood pooling are, of course, the major contributing factors which reduce the pressure available at the eye and in more severe cases at the brain. Since World War II, many possible methods of raising the tolerance of pilots to acceleration including G-suit have been investigated. A block diagram of the cardiovascular system with the G-suit system is shown in Fig. 13

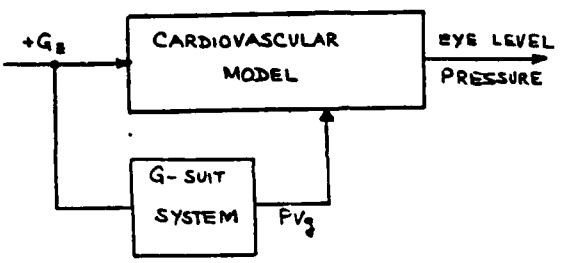


Fig. 13 Cardiovascular System with G-suit Protection

The protective G suit garment contains airtight bladders which are filled with pressurized air delivered from a G sensitive mechanical valve. The pressure delivered by the valve is a function of the current G level. The normal suit pressure G relationship curve is shown in Fig. 14.

The air bladders, the suit air feed hose, and the containing garment represents a dead space which introduces a time delay into the pressure system. The delayed suit filling is caused by both a transport delay and first order lag. The transport time is caused by the

connecting air feedline from valve to suit and is on the order of 5 msec. The bladder size and garment tightness govern the lag term which has a time constant on the order of 1 sec. These relationships can be shown in Fig. 15.

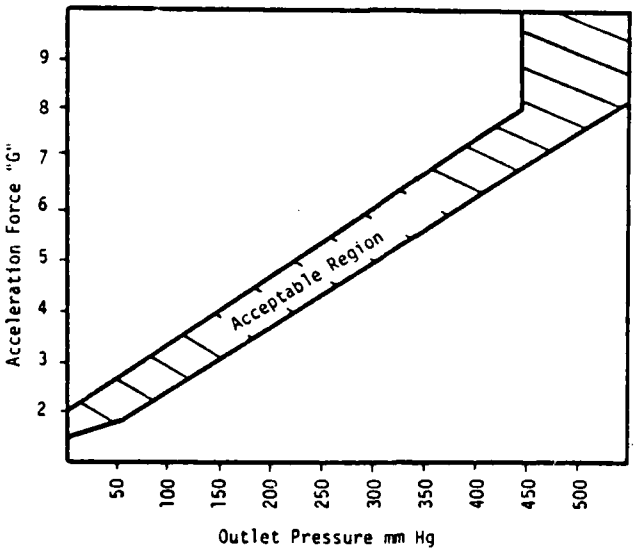


Fig. 14 Standard Pressure Suit-Valve Fill Schedule

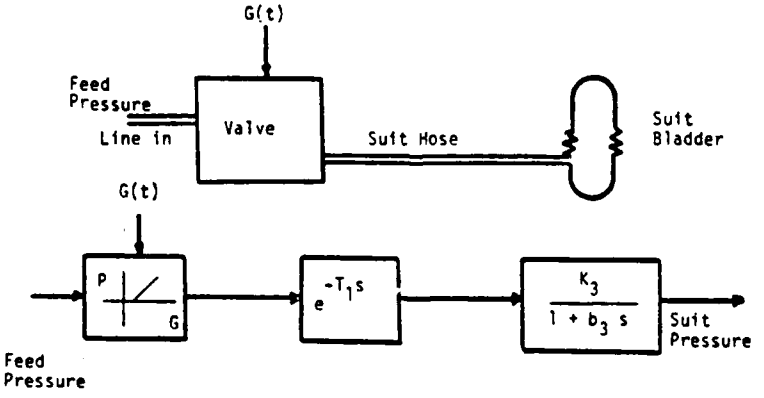


Fig. 15 Simplified G suit Pressure System Characteristic

6.1 G-suit with Bang-Bang Servo (BBSV) Anti-G Valve

The BBSV valve [Van Patten, et al.] is a modified version of the

high flow (HFO) G-valve. The modification to the valve consists of a new top housing containing an aerospace quality solenoid and associated bang-bang servo electronics. At $+G_z$ level below +2 G_z and at onset rates of less than 2G/sec., the valve acts as a HFO, Alar Corp. valve. This feature allows standard maneuvers without involving the rapid acting high G protection mode of the valve. Fig. 16 shows the block diagram of the modification circuit of the BBSV valve.

In high onset G maneuvering (onset rates > 2 G/sec.), the valve's high onset rate protection mode is activated when both trigger criteria, $+G_z$ greater than +2 and rate of onset greater than 2 G/sec , are met. When the two trigger criteria are met, the electronic circuit fires the solenoid for a period of 1.5 sec. After the suit fills to full pressure, the valve reverts to the standard inertial mode of operation.

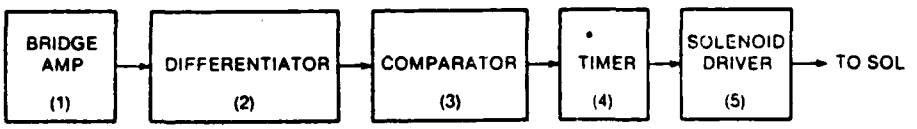


Fig. 16 Block diagram of BBSV Electronic Circuit

Fig. 17 shows the pressure profiles of the HFO and BBS valves for different sustained G_z levels with an onset rate of 3 G/sec. It can be seen from this figure that the BBSV has a very small rise time and possibly provides an additional 1 G_z protection to relaxed subjects (Van Patten, et al. [30]) during the period of greatest physiology vulnerability. Fig. 17 also shows that BBSV has a very large overshoot in pressure which could be uncomfortable to pilots for low sustained G_z levels. It is well know that analog differentiators are susceptable to noise and should be avioded. Even though Van Patten, et al. [30] claimed to have solved this problem with a novel comparator circuit, noise could still be a potential problem.

VII. Proposed Closed-Loop Servo Anti-G Valve

The block diagram of the proposed closed-loop control system for the anti-G valve is shown in Fig. 18

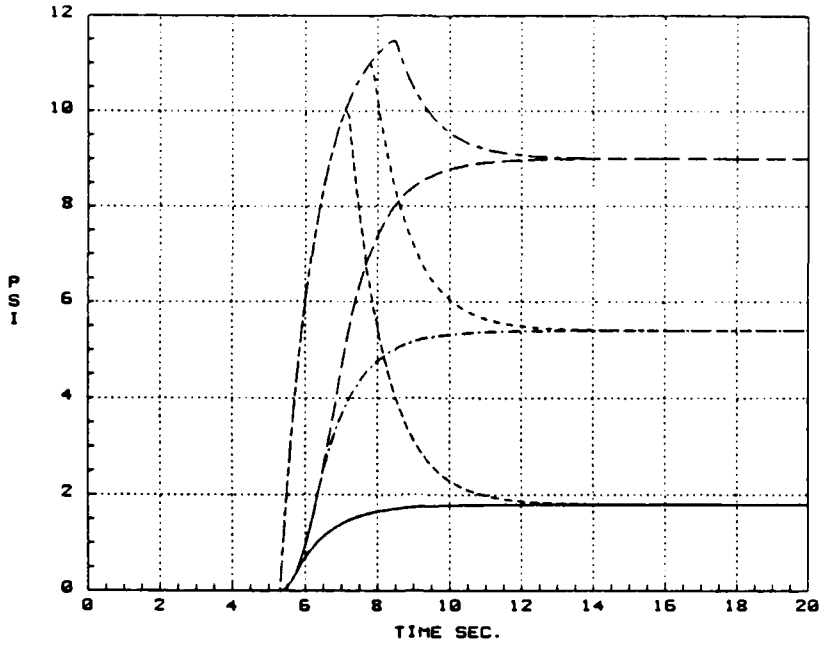


Fig. 17 Pressure Profiles of HFO and BBS valves for Different Sustained G_z Levels

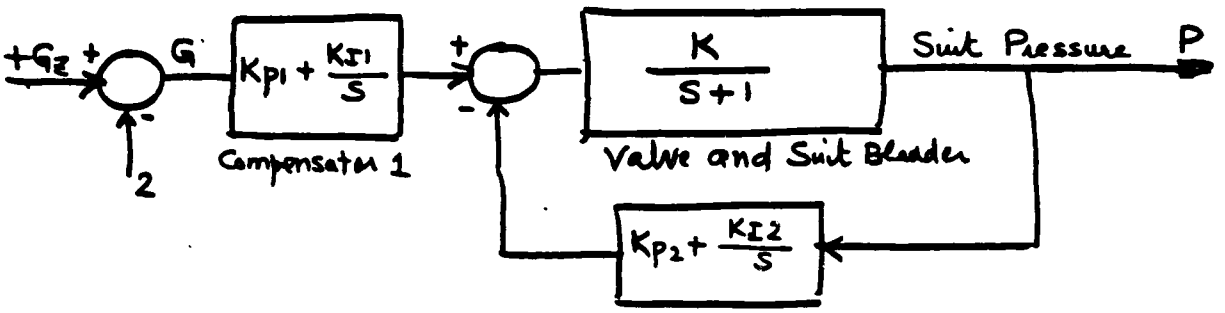


Fig. 18 Block diagram of the closed-loop servo Anti-G valve

The compensators 1 and 2 are designed to satisfy the following specifications:

1. Steady-state value of the suit pressure to unit step G input = 1.8.
2. Rise time to be as small as possible.

The transfer function of the suit pressure P to the acceleration input G can be written from Fig. 18 as

$$\frac{P(s)}{G(s)} = \frac{KK_{P1}(s+K_{I1}/K_{P1})}{s^2+(KK_{P2}+1)s+KK_{I2}} \tag{7.1}$$

Applying final value theorem to eqn. (6.1), we get

$$\lim_{s\to 0} sP(s) = \lim_{s\to 0} \frac{KK_{P1}s(s+K_{I1}/K_{P1})}{s^2+(KK_{P2}+1)s+KK_{I2}} \frac{1}{s} = 1.8 \tag{7.2}$$

$$\therefore \quad K_{I1}/K_{I2} = 1.8 \tag{7.3}$$

Selecting $K_{I1}/K_{P1}>1$, the root-locus plot of the open loop system shown in Fig. 18 can be drawn as shown in Fig. 19

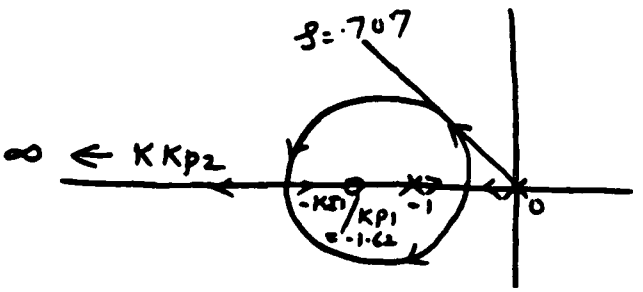


Fig. 19 Root-Locus Plot of the Open-Loop System shown in Fig. 18

Selecting $\zeta = 0.707$ and $\omega_n = 3$ rad/sec, the controller coefficients are obtained as

$$K_{I1}=1.8,\ K_{I2}=1,\ K=9,\ K_{P2}=.362 \text{ and } K_{P1}=1.11 \tag{7.4}$$

The transfer function of the closed-loop system given in eqn. (7.1) with coefficient given by eqn. (7.4) can now be written as

$$\frac{P(s)}{Q(s)} = \frac{10s+16.2}{s^2+4.242s+9} \tag{7.5}$$

Using the transfer function given by eqn. (7.5) and the G-profile shown in Fig. 20, the suit pressure is obtained as shown in Fig. 20. Suit pressure of HFO and BBS valves are also shown in Fig. 20. It can be seen from this figure that the rise time of the proposed system compares favorably to BBSV, but the overshoot is much smaller. Also note that no differentiator is used to obtain a faster pressure profile.

$$\frac{P(s)}{G(s)} = \frac{KK_{P1}(s+K_{I1}/K_{P1})}{s^2+(KK_{P2}+1)s+KK_{I2}} \tag{7.1}$$

Applying final value theorem to eqn. (6.1), we get

$$\lim_{s\to 0} sP(s) = \lim_{s\to 0} \frac{KK_{P1}s(s+K_{I1}/K_{P1})}{s^2+(KK_{P2}+1)s+KK_{I2}} \frac{1}{s} = 1.8 \tag{7.2}$$

$$\therefore \quad K_{I1}/K_{I2} = 1.8 \tag{7.3}$$

Selecting $K_{I1}/K_{P1}>1$, the root-locus plot of the open loop system shown in Fig. 18 can be drawn as shown in Fig. 19

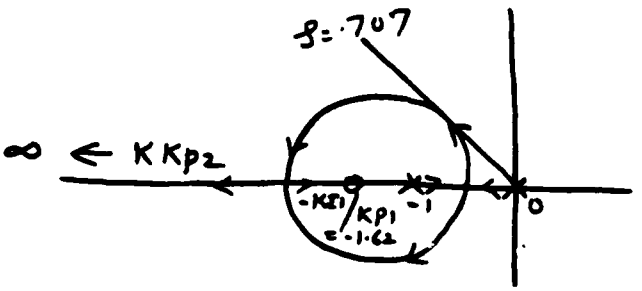


Fig. 19 Root-Locus Plot of the Open-Loop System shown in Fig. 18

Selecting $\zeta = 0.707$ and $\omega_n = 3$ rad/sec, the controller coefficients are obtained as

$$K_{I1}=1.8,\ K_{I2}=1,\ K=9,\ K_{P2}=.362 \text{ and } K_{P1}=1.11 \tag{7.4}$$

The transfer function of the closed-loop system given in eqn. (7.1) with coefficient given by eqn. (7.4) can now be written as

$$\frac{P(s)}{Q(s)} = \frac{10s+16.2}{s^2+4.242s+9} \tag{7.5}$$

Using the transfer function given by eqn. (7.5) and the G-profile shown in Fig. 20, the suit pressure is obtained as shown in Fig. 20. Suit pressure of HFO and BBS valves are also shown in Fig. 20. It can be seen from this figure that the rise time of the proposed system compares favorably to BBSV, but the overshoot is much smaller. Also note that no differentiator is used to obtain a faster pressure profile.

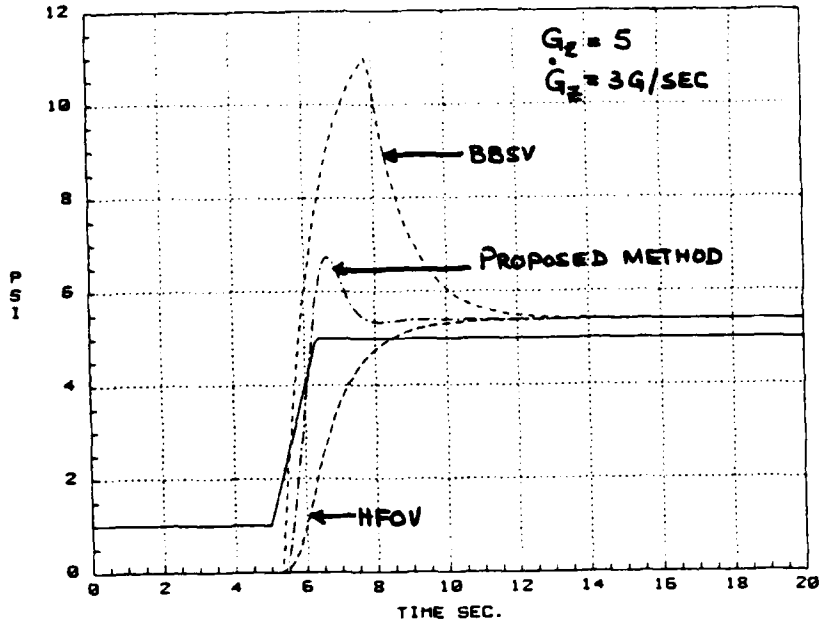
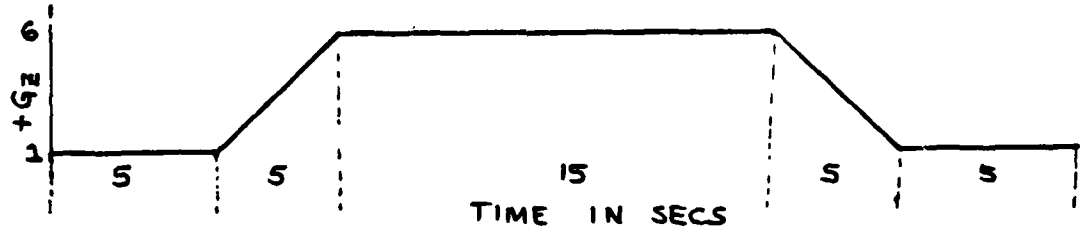


Fig. 20 Suit Pressure Profiles for HFOV, BBSV and the Proposed Closed-Loop System

VIII. Recommendation for Future Work

1. The computer simulation of the cardiovascular system we have developed takes approximately 24 hours of CPU time for the G-profile shown below



In order to effectively use this simulation, it is important that the simulation be speeded up. A new version of MATRIX$_x$ released recently has an option called HYPER-BUILD. Our initial study indicates that this option may speed up our simulation by a factor of approximately 10.

2. Present computer simulation assumes all the resistances, inductances, and capacitances of the arterial tree to be fixed. This simulation should be modified to include the nonlinear effects of these variables.

3. Modify the cardiovascular model and computer simulation to include G-suits and study their effects.
4. The present cardiovascular model assumes the leg as a peripheral element B. To study the effects of sequential G-suit, the multielement arterial tree should be modified to include the thigh and calf parts of the leg.
5. The lumped venous model needs to be replaced by a multielement model in order to get more accurate blood distribution under acceleration stress or external applied force.
6. To study the effects of various maneuvers, such as M1, L1, the model needs to include external pressure sources around the thoracic area.
7. Study the technique of pulsating G-suit synchronized with the heart.
8. Obtain a multivariable transfer function model using the state-space representation of the multielement arterial tree.
9. The proposed closed-loop servo controller designed is just a preliminary design. Time-domain and frequency-domain techniques should be used to design an optimal controller. Effort should be made to design a multivariable control system which should include eye level pressure and heart rate as inputs.

References

1. Burton, R.R., "Positive $+G_z$ Acceleraiton Tolerances of the Miniature Swine: Application as a Human Analog," Aerospace Med. 44(3): 294-298, 1973.
2. Burton, R.R., S.D. Leverett, and E.D. Michaelson, "Man at High Sustained $+G_z$ Acceleration: A Review," Aerospace Med. 45(10): 1115-1136, 1974.
3. Burton, R.R. and R.W. Krutz, Jr., "G-Tolerance and Protection with anti-G suit Concepts," Aviat. Space Environ. Med. 46(2): 119-1124, 1975.

4. Chu, C., "A Mathematical Model of the Cardiovascular System Under $+G_z$ Stress," Ph.D. Dissertation, Drexel University, 1984.
5. Cohen, M.M., "Combining Techniques to Enhance Protection Against High Sustained Acceleration Forces," Aviat. Space Environ. Med. 54(4): 338-342, 1983.
6. Erickson, H.H., H. Sandler, and H.L. Stone, "Cardiovascular Function During Sustaine $+G_z$ Stress," Aviat. Space Environ. Med. 47(7): 750-758, 1976.
7. Frazier, J.W., Robert E. Van Patten, Lloyd D. Tripp, Lora L. Howell and Charles D. Goodyear, "A Comparative Study of Two Anti-G Suit/Valve Combinations on High $+G_z$ Tolerance," AAMRL-TR-86-017, 1986.
8. Frazier, J.W., Robert E. Van Patten, Tom Jennings, Chuck Goodyear, David Ratino and William Albery, "Evaluation of an Electro-Pneumatic Anti-G Valve in a High Onset Rate Acceleraton Environment," AAMRL-TR-85-051, 1985.
9. Gauer, Zuidema, "Gravitational Stress in Aerospace Medicine," Boston: Little, Brown and Company, 1961.
10. Gillingham, K.K. and R.W. Krutz, Jr., "Effects of the Abnormal Acceleratory Environment of Flight," SAM-TR-74-57, pp. 92, 1974.
11. Gillingham, K.K. and J.J. Freeman, and R.C. McNee,"Transfer Functions for Eye-Level Blood Pressure During $+G_z$ Stress," Aviat. Space Environ. Med. 48(11):1026-1034, 1977.
12. Guyton, Jones, Coleman, "Circulatory Physiology, Cardiac Output and its Regulation", W.B. Saunders, 1973.'
13. Green, J.F., and N.C. Miller, "A Model Describing the Response of the Circulatory System to Acceleration Stress,"Annals of Biomed. Eng. 1: 455-467, 1973.
14. He, Ping, "Optimizaiton of Cardiac Assist Devices: Theoretical and Experimental Studies of Intra-Aortic Balloon Pumping," Thesis, Drexel University, 1984.

15. Howard, P., "The Physiology of Positive Acceleration," Chapter 23 in: A Textbook of Aviation Physiology, J.A. Gillies, ed. New York, Pergamon Press, 1965.

16. Jacob, Francone, Lussow, "Structure and Function in Man," W.B. Saunders, 1978.

17. Jaron, Moore, Chu, "A Cardiovascular Model for Studying Impairment of Cerebral Function During $+G_z$ Stress," Aviat. Space Environ. Med., 55(1) 24-31, 1984.

18. Katona, P.G., G.O. Barnett, and W.D. Jackson, "Computer Stimulation of the Blood Pressure Control of the Heart Period," P. Kezdi ed. Baroreceptors and Hypertension, Oxford: Pergamon Press; 1967: 191-199.

19. Laughlin, M.H., W.M. Witt, and R.N. Wittaker, "Regional Cerebral Blood Flow in Conscious Miniature Swine During High Sustained $+G_z$ Acceleration Stress," Aviat. Space Environ. Med., 51(11): 1197-1204, 1980.

20. Laughlin, M.H., J.T. Young, W.M. Witt, and P.P. Crump, "The Importance of Myocardial Perfusion in the Pathogenesis of the Cardiac Pathology Associated with $+G_z$ Exposure in Miniature Swine," Aviat. Space Environ. Med., 51(11): 1197-1204, 1980.

21. Ohley, W.J., C. Kao and D. Jaron, "Validity of an Arterial System Model: A quantitative evaluation," IEEE Trans. on Biomed. Eng., BME-27(4): 203-211, 1980.

22. Parkhurst, M.J., S.D. Leverett, Jr. and S.J. Shubrooks, Jr., "Human Tolerance to High, Sustained $+G_z$ Acceleration," Aerospace Med., 43(7): 708-712, 1972.

23. Rogers, D.B., "A Model for the Energetic Cost of Acceleration Stress Protection in the Human," AMRL-TR-79-58, 1979.

24. Rushmer, R.F., E.L. Beckman, and D. Lee, "Protection of the Cerebral Circulation by the Cerebrospinal Fluid Under the Influence of Radial Acceleration," Amer. J. Physiol., 151: 355-365, 1947.

25. Shurbrooks, S.J., Jr. and S.D. Leverett, Jr., "Effect of the Valsalva Maneuver on Tolerance to $+G_z$ Acceleration," J. App. Physiol., 34(4): 460-466, 1973.

26. Snyder, M.F., and V.C. Rideout, "Computer Simulation Studies of the Venous Circulation," IEEE Trans. on Biomed. Eng., BME-16(4): 325-334, 1969.

27. Suga, H., and K. Sagawa, "Instantaneous Pressure-Volume Relationships and their Ratio in the Excised, Supported Canine Left Ventricle," Circ. Res., 35: 117-126, 1974.

28. Westerhof, N., J. Bosman, C.J. DeVries, and A. Noordergraaf, "Analog Studies of the Human Systemic Arterial Tree," J. Biomechanics, 2:121-143, 1969.

29. Vander, Sherman, Luciano, "Human Physiology: The Mechanism of Body Function," McGraw-Hill, 1980.

30. Van Patten, R.E., T.J. Jennings, W.B. Albery, J.W. Frazier, C. Goodyear, B. Gruesbeck and D. McCollor, "Development of a Bang-Bang Servo Anti-G Valve for High Performance Fighter Aircraft: Final Report," AFAMRL-TR-85-024, 1985.

31. Vettes, B., H. Vielleford, and R. Auffret, "Cardiovascular Response of Man Exposed to $+G_z$ Acceleration in a Centrifuge, "Aviat. Space Environ. Med., 51(4): 375-378, 1980.

1986 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT SUMMER RESEARCH SUPPORT PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

A COMPARISON OF TWO MATHEMATICAL SYSTEMS FOR A STANDARD IMAGE ALGEBRA

Prepared by:	Barbara S. Rice
Academic Rank:	Associate Professor
Department and University:	Department of Mathematics Alabama A&M University
Research Location:	Air Force Armament Laboratory, Advanced Seeker Division, Electro-Optical Terminal Guidance Branch
USAF Researcher:	Neal Urquhart
Date:	August 4, 1986
Contract No.	F49620-85-C-0013

A COMPARISON OF TWO MATHEMATICAL SYSTEMS FOR A STANDARD IMAGE ALGEBRA

by

Barbara S. Rice

ABSTRACT

Advances in image sensor technology have provided large amounts of data for image processing. There are many techniques, but at this time there is not a unified method for image processing. The Air Force Armament Laboratory, Eglin AFB, Florida, wishes to obtain a standard for expressing image transformation algorithms. Two efforts, FO 8635-84-C-0296 of the Singer/Kearfott Division and FO 8635-84-C-0295 of the University of Florida, have been made in developing an Image Algebra.

The Singer effort defines basic operations which can be interpreted as certain manipulations of gray level functions. Special image processing techniques are effected through appropriate combinations of the basic operations.

The Florida group uses the template structure and defines operations on images and on templates. Specific image processing techniques are effected by utilizing appropriate templates and operators. Template representation and decomposition results are important.

This report provides a comparison of the two systems.

ACKNOWLEDGEMENTS

I would like to express my appreciation to the Air Force Systems Command and the Air Force Office of Scientific Research for sponsorship of this program. I appreciate very greatly this opportunity to participate in the Image Algebra Project at the Air Force Armament Laboratory.

Special thanks are due to Dr. Sam Lambert for his enthusiastic interest in the Summer Faculty/ Graduate Student Summer Research Program; and to Mr. Neal Urquhart for his unfailing interest and support, and for his guidance of the Image Algebra Project.

I am also appreciative of the encouragement I received from my home institution, Alabama A&M University, and from my family.

I. INTRODUCTION

Barbara S. Rice received her Ph. D. in mathematics from the University of Virginia in the area of analysis (stochastic integration). She was a member of the graduate faculty at Florida Institute of Technology prior to her present position of Associate Professor at Alabama A&M University.

This project is a study of two mathematical approaches in the development of an algebra for image processing algorithms. Advances in technology have improved the capabilities of imaging systems; this is particularly true for high quality infrared detectors. There is a need for fast, accurate processing of data recorded by sensors; and a desire to transfer image processing operations to an automatic machine system. Present image processing programs are lengthy and specific [4]. It is difficult to transfer one program to meet other needs.

In the general setting for image processing, a gray value (intensity) is recorded by a sensor for each pixel in an array. Due to physical and statistical constraints which lead to noise , blurring, and distortion, image processing

is needed for interpretation of the data.

The formulation of an image algebra is an effort to provide answers to these questions:

1. Is it possible to isolate a set of operators on gray level functions, so that image processing techniques could be described as operations on images?
2. Is it possible to machine encode the operators, so that an image processing transformation could be effected by specifying an algebraic expression of operators and operands?
3. Is it possible to create a system in which the choice of algorithmic process is selected by automated reasoning methods?

II. OBJECTIVES

At this time the Armaments Laboratory, Eglin AFB, Florida, has interest in obtaining a standard image processing algebra. Ongoing research is being conducted by two groups: the University of Florida, G. X. Ritter, principal investigator; and Singer/Kearfott Division , L. R. Giardina, principal investigator. Each group has proceeded independently to define and develop an image algebra [2, 7, 8].

As a Universal Energy System Summer Faculty Research

Program Fellow, my project has been part of this effort. My goals are these:

1. To compare and contrast the Florida and Singer algebras, and to determine whether any incorporation can be made of one within the other.
2. To note any points for clarification within each effort.
3. To investigate the versatility and applicability of image algebra software.

Section III of this report provides a background for the development of an image algebra; the Singer and Florida efforts are outlined in Sections IV and V, respectively; applications in each algebra are provided in Section VI; and recommendations arising from the comparison of the efforts are made in Section VII.

III. BACKGROUND

There is a need for an efficient uniform language for image processing transformations. The use of masks in basic transformations underscores the local nature of certain portions of image processing: the gray values at immediate neighbors of a pixel are important in considerations of the gray value at a point. The work of Serra in the morphological operations [9] also directs attention to a neighborhood structure. An earlier effort by Miller [5]

developed a Neighborhood Transform Algebra (NTA) for Logical (Boolean) Neighborhood Transforms.

The Florida and Singer efforts have assembled a collection of image processing techniques. From this they have defined basic operations, so that a large collection of image processing techniques could be expressed through the basic operations. This leads to a need to establish computational relationships and properties within the algebra of operators. There is a desire to find a machine implemented computational system for the algebra to make image algebra operations available as keyboard operations.

IV. THE SINGER ALGEBRA

The Singer group has defined ten basic operators which are described fully [2, pp. 46-67].

Four of the operations (addition $\oplus$, muliplication $\odot$, maximum $\heartsuit$, and division $\oslash$) are binary operations and produce a resultant image. For these, the indicated operation is produced from the corresponding pixel by pixel operation, suitably modified for pixels outside the common domain. Three operators (translation T , rotation N, and reflection D) are unary, taking an image into an image. The domain extractor K assigns to each image its domain; the parameter extractor G assigns to an image the gray value at

a particular pixel; and the existential operator E defines an image with constant gray value on a specified domain.

These operators provide a mechanical system for working with image processing data (gray value functions on a collection of pixels). The investigators have found that all strictly digital algorithms studied to date have an expression in terms of these operators. There is a larger collection of macro-operators [2, pp. 68-99] to facilitate particular image processing transformations. Each of these macros has a representation in the basic operators. With machine implementation, a user could input a gray level function, choose an appropriate finite sequence of processing operations, and output a processed image. For long range planning, artificial intelligence techniques could determine the most effective processing algorithm for the situation. In this description, the gray value f(i,j) corresponds to the pixel with horizontal - vertical location (i,j) with reference to a designated (0,0) pixel. The resultant image under transformation may have a domain different from that of the original image.

Macro expressions for the sine, cosine, and exponential are considered. Infinite powers series expressions are undefined, due to the lack of a limit process in the image algebra. Consequently, the Singer group considers only the specific truncations such as

$$\exp(f,n) = \sum_{k=0}^{n} a_k f^k \qquad (1)$$

using the coefficients from the powers series expansion of the real expontial, for images with gray values f(i,j) in the radius of convergence of the series.

For images on a rectangular domain, there are macros in the algebra to effect matrix multiplication and inversion. The processes are described in the algebra. The discrete picture transform P*f*Q is translated into the image algebra, for f a rectangular image, and matrices P and Q of appropriate size. One question arises: is it possible to utilize known computational processes for matrices to identify image algebra results?

Though computation of a convolution in the image algebra is demonstrated by example, there appears to be an omission in the definition of the convolution [2, p. 98].

Further classification and characterization of the operators is given, and subalgebras of the image algebra are noted. There is also a definition of domain induced and range induced operators, with a property of commutativity for such operators.

V. THE UNIVERSITY OF FLORIDA ALGEBRA

The University of Florida group designates an image by its domain and gray value function, $A = \{(x,a(x)) : x \in X\}$.

they have used the template to develop this image algebra. In this framework, to each pixel x in the domain, there is associated a collection of pixels $\mathcal{T}(x)$ (the neighborhood configuration), and a weight function t_x. The function $t_x(z)$ assigns weights to pixels in the configuration, and has value zero for pixels not in the configuration.

Example: The Moore neighborhood could be used as configuration $\mathcal{T}(x)$, and weights $t_x(z)$ as noted in Figure 1, with the location x noted by hatch marks.

Figure 1:

5	5	5
-3	0	-3
-3	-3	-3

A template T is a collection of weight functions, each with its corresponding configuration; $T = \{t_x, \mathcal{T}(x): \text{for } x \text{ in } X\}$. A generalization of this definition was made in [8]. The set of images on X is J_X. The collection of all templates on X is T_X. The operands in the algebra are images on X and templates on X. Scalars as operands are handled though the use of constant images. Rectangular images are indexed by row and column of pixel location.

For images in J_X, the binary operations are (pointwise) addition, A+B; multiplication, A*B; and the supremem, $A \vee B$; together with the dot product A•B. There are

also image-template binary operations $A \oplus T$, $A \circledcirc T$, and $A \boxtimes T$. In particular,

$$(A \oplus T)(x) = \sum_{y \in T_x(X)} a(y)\ t_x(y) \qquad (2)$$

so for each x, $A \oplus T$ produces the dot product of A with the image $\{(y, t_x(y)) : y \in X\}$. The operators $A \circledcirc T$ and $A \boxtimes T$ arise from the kinds of computations used in morphology. Finally, there are template-template operators, $S \oplus T$, $S \circledcirc T$, and $S \boxtimes T$; and $S+T$, $S*T$, and $S \vee T$. Explicit image algebra expressions are provided for a collection of image processing techniques. The problem is to find the image algebra expression and the templates involved in a particular transformation. A second problem is to find a template decomposition to yield an efficient and optimal method of performing the operations.

Example: H, V, and M are the templates of Figure 2.

Figure 2: H = | 1 | 1 | 1 |

V = | 1 |
| 1 |
| 1 |

M = | 1 | 1 | 1 |
| 1 | 1 | 1 |
| 1 | 1 | 1 |

Then for

$$A = \begin{matrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 2 & 3 & 2 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{matrix} \quad \text{and} \quad B = \begin{matrix} 1 & 3 & 4 & 3 & 1 \\ 3 & 8 & 11 & 8 & 3 \\ 4 & 10 & 15 & 11 & 4 \\ 3 & 8 & 11 & 8 & 3 \\ 1 & 3 & 4 & 3 & 1 \end{matrix}$$

we have $(A \oplus H) \oplus V = A \oplus M = B$, and it becomes clear that the computation $(A \oplus H) \oplus V$ is more efficient.

Lexicographic ordering is used to write an nxm rectangular image as an nxm vector. Since a template is a collection of images, with each weight function t being a gray level function, a second lexicographic ordering, on index x, provides a representation of the template as a matrix. Boundary pixels require special treatment in the computational framework. Often the neighborhood configurations and weights are given by a common mask. These considerations can be handled with the use of circulant templates; the templates can be identified with circulant matrices, which have a fairly rich development [1, 3].

In following the development of this image algebra, shifts have occurred. This is to be expected. However, certain portions of the theory seem fairly inaccessible. At this time the image algebra is presented in an intricate notational setting. It is hard to separate background material from the particular adaptations for the image algebra. There are few motivational examples, and the range in complexity of examples is limited. In particular, the result that the Fourier transform can be computed locally would benefit from further illustration. It would be helpful to have a more expository development of the theory.

VI. AN APPLICATION

The earlier Neighborhood Transform Algebra techniques were applied to the Hamadani FLIR algorithm [4]. The image enhancement, edge detection, and thresholding processes of that algorithm, though local transforms, could not be handled within that structure. These transformations have an expression in both the Singer and Florida image algebras.

The notational symbols are those used in the Florida algebra [7] and those of the Singer algebra [2], respectively.

Image and edge enhancements:

Both transformations of the Hamadani algorithm can be handled by the assignment of appropriate weights to pixels in the Moore neighborhood of a point. To be specific, we use the Kirsch operator:

$$\max\left(1, \max_{i=0}^{7} \left| 5(a_i + a_{i+1} + a_{i+2}) - 3(a_{i+3} + \ldots + a_{i+7}) \right| \right)$$

with subscript notation modulo 8, and the neighborhood reference system of Figure 3, for each x in domain X.

Figure 3:

x_4	x_3	x_2
x_5	x	x_1
x_6	x_7	x_0

For each i, use the mask M_i at x with weight 5 at x_i, x_{i+1}, and x_{i+2}; and weight -3 at x_{i+3}, . . . , x_{i+7}. Use template T_i which for each x has weight 1 at x_i, x_{i+1}, and x_{i+2}; else 0. Use template S_i which for each x has weight 1 at x_{i+3}, . . . , and x_{i+7}; else 0.

Florida: The enhancement is $I \vee (\bigvee_{i=0}^{7} |A + (5T_i - 3S_i)|)$.

Singer: The enhancement is $1_\chi \circledcirc N_\infty[\mathfrak{F}(f,M_0), \ldots, \mathfrak{F}(f,M_7)]$.

Linear scaling:

The range of gray values {a(x)} on X is scaled to the interval [0,1]; if M is the maximum gray value, and m is the minimum gray value on X, the scaled function is

$$(a(x) - m) / (M - m), \text{ for } M > m.$$

Florida: Let $A = \{(x,a(x)) : x \in X\}$. Define template $T = \{t_\chi\}$, with $t_\chi(x) = 1$; else 0. Then $M = (A \circledcirc T) \bullet I$, and $m = (A \circledcirc T) \bullet I$. The desired scaling is

$$1/(M-m) [A - mI].$$

Singer: For a=a(i,j), let $M = \bigvee_0 a(i,j)$ and $m = \bigwedge_0 a(i,j)$. The desired scaling is

$$1/(M-m) [a - E(X,m)].$$

Thresholding by mean and deviation:

The mean and deviation of the set of gray values is computed; a gray value is retained in the enhancement if it

is greater than the sum of the mean and deviation; otherwise the gray value 0 is assigned to that pixel.

Florida: For nxm image A, the mean μ is given by $1/(mn)\ [A \bullet I]$, where the domain is an mxn array; the deviation v is given by $[1/(mn)\ (A - \mu I)(A - \mu I)]^{1/2}$. The thresholded image is then $A * c_{>\mu+v}(A)$.

Singer: For nxm image a=a(i,j), the mean is $1/(nm) \Sigma_0 a(i,j)$ and the deviation v is given by $\{ 1/(nm)\ \mathcal{D}_0(a-E(x,\mu),\ a-E(x,\mu))\}$. The thresholded image is $CL(a, \mu+v)$.

Two-tiered threshold:

A non-zero gray value is retained in the edge thresholded image provided it is retained under a mean and deviation thresholding for edge value and also under a separate mean and deviation thresholding for intensity; otherwise value zero is assigned at that pixel.

Florida: Where A is the mean and deviation thresholded edge image, and B is the mean and deviation thresholded intensity image, then $c_{>0}(B)$ is the binary image with gray value 1 where B is non-zero. The desired image is $A * c_{>0}(B)$.

Singer: If a=a(i,j) is the mean-deviation thresholded edge image, and b=b(i,j) is the mean and deviation thresholded intensity image, then the binary image which is non-zero only where b is non-zero is given by $\mathcal{J}_0'(b)$. The desired

image is $a \odot \mathcal{T}_0'(b)$.

While these straightforward examples indicate a similarity of expression, there are differences. Though computations in the Singer algebra are very direct, the template of the Florida algebra is a powerful tool.

VII. RECOMMENDATIONS

Present image processing procedures require many computations involving data recorded pixel-by-pixel. The image algebra is an attempt to define computations at the image level.

The University of Florida algebra utilizes operations on images and templates, which are collections of weighting images. The operations are basic and few in number. Procedures are effected through appropriate operations and templates. Machine implementation of the operations is optimized through template decomposition.

The Singer algebra operations provide for the mechanics of manipulation, and are rooted in a geometrical and computational background. Processing procedures are effected through the macro operators, which are finite sequences of basic operations. The goal is machine implementation of the operators and macro operators on

images.

The following suggestions are made to promote the development and utility of these algebras.

Florida:

1.There is a need for a document in which the intuitive and motivational aspects of the algebra are developed. Present documents present the Florida algebra in a highly abstracted setting.

2. There is a need to highlight special processes which are used frequently and which yield direct implementation in the image algebra. Examples similar to those in the Quick Turnaround of [8] should be accompanied by text to provide linkage between the theory and practice of the algebra.

3. Since the identification and optimal decomposition of templates is a central consideration in this algebra, attention should be given to the choice of templates in applications.

Singer:

1. There is a need for a more efficient description of the convolution and the matrix macros of the algebra.

2. There is a need to highlight special processes which yield direct implementation, and provide computer software

to demonstrate these examples.

3. There is a need to proceed to machine implementation of the basic operators and the macro-operators. The basic operators, though straightforward, lie within a many typed algebra. This appears to offer a challenge in providing a setting for the more intricate macros.

The approach of developing an image algebra offers promise for achieving a standard for expression of image processing algorithms. At present, implementation of a processing algorithm involves a lengthy task-specific computational program. The design of an image algebra will provide a rigorous, efficient, and standard environment in which elemental operations can be combined to express any gray level image transformation; where the theorems of the mathematical structure provide for simplification and optimization of algorithms; and where the notational adaptability of the algebra to programming languages allows substitution of image algebraic expressions for equivalent blocks of code.

REFERENCES

1. Davis, P.J., Circulant Matrices, New York, John Wiley & Sons, Inc., 1979

2. Giardina, C. R. et al, Final Report for Phase I, FO 8635-84-C-0296, April, 1986

3. Graham, A., Kronecker Products and Matrix Calculus with Applications, Chichester, Ellis Harwood Limited, 1981

4. Hamadani, N.A., Automatic Target Cueing in IR Imagery, Thesis, Air Force Institute of Technology, AFIT/GEO/EE/81D-3

5. Miller, P.E., and Allan Corbeil, Development of Mathematical Structures for Image Processing, AFATL-TR-83-61, July, 1983

6. Ritter, G.X. A Synopsis of Current Image Transforms and Techniques, Technical Report 1, FO 8635-84-C-0295, September, 1985

7. Ritter, G.X. et al, Image Algebra Final Report for Phase I, FO 8635-84-C-0295, September, 1985

8. Ritter, G. X. et al, Image Algebra Project, Phase II Program Review, FO 8635-84-0295, 1986

9. Serra, J., Image Analysis and Mathematical Morphology, New York, Academic Press, 1982

1986 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT SUMMER SUPPORT PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

Multi-Echelon Inventory Models for EOQ Items

Prepared by:	Dan B. Ranks
Academic Rank:	Assistant Professor
Department and	Department of Quantitative Business Analysis
University:	Louisiana State University
Research Location:	AF Logistics Management Center Gunter AFS, Alabama
USAF Researchers:	Lt. Colonel Douglas Blazer/ Captain Craig Carter
Date:	September 18, 1986
Contract No:	F49620-85-C-0013

Multi-Echelon Inventory Models for EOQ Items

by

Dan B. Rinks

ABSTRACT

Two models of the Air Force's depot-base inventory system for EOQ (consumable) items have been constructed. The first model, a static model of the D062 wholesale (depot) requirements system, is useful for studying the effects, on the depot, of changes in depot stockage policies. The second model is a simulation model of the multi-echelon system. The simulation model can be used to study how changes in stockage policies at one level affect the stockage policies at the other level. Preliminary results from analyses using these models are reported.

Acknowledgements

I am grateful to the Air Force Office of Scientific Research for the sponsorship of this research. The Air Force Logistics Management Center provided an excellent environment to conduct the modelling of depot-base inventory stockage policies. The staff at AFLMC were very supportive of the project and were generous in sharing their knowledge with me. In particular, I would like to thank Lt. Colonel Douglas Blazer for the guidance he provided for this research. Many times, Doug kept me from straying in directions that would have proven unproductive. Captain Craig Carter was a valuable collaborator in the construction of these models; most of this research was a team effort with Craig. The assistance of Nick Godbey and Chuck Miller in resolving programming and computer related problems was a valuable resource. I appreciate the indulgence of the staff when our models dominated the computer and delayed the completion of their projects.

Finally, I would like to thank my wife, Gloria, for her willingness to handle my domestic responsibilities while I was away for the Summer.

I. Introduction

The Air Force Logistics Management Center is conducting research involving the analysis of depot-base stockage policies. How do changes in the stockage policies at the depot affect the ability of bases to service their customers? Conversely, how do changes in the stockage policies at the bases affect the stockage policy at the depot?

The Air Force stock fund employs standard reorder point and economic order quantity (EOQ) methods for inventory management as specified in Department of Defense Instructions 4140.39 and 4140.45. Policies, procedures, and models to be used for secondary-item management at the wholesale (depot) and the retail (base) levels are described in these instructions. Specific instructions for computing wholesale requirements are documented in AFLC Regulation 57-6 (May, 1982). This latter system is known as the EOQ Requirements System (D062).

The methodology for computing safety levels at the wholesale level in D062 is based on Presutti and Trepp (1970). In practice, unit fill rates at the depot, and subsequently at the bases, have been consistently lower than that predicted by the theory. Because of this, and other reasons, the method for computing safety levels has been changed several times by the Air Force Logistics Center (AFLC) as they attempted to "fine tune" D062. Despite these efforts, unit fill rates have steadily declined over the last several years. Changes have also been made in base stockage policies. It is not known what effect, if any, these changes have had on unit fill rates at the depots.

Obviously, the Air Force multi-echelon inventory system for EOQ items is a complex system. Policy changes at the depot have not necessarily led

to their anticipated results. Furthermore, interactions between stockage policies at the depot and the bases are not well understood.

II. Objectives of the Research Effort

The overall objective of this research effort is to construct models that facilitate the analysis of depot-base stockage policies for EOQ items. This work is divided into two parts.

1. A static model of the D062 wholesale (depot) requirements system. The primary purpose of this model is to study the effects, on the depot, of changes in depot stockage policies. For example, what is the expected cost (dollars invested in inventory) if a change is made in the method of computing safety stocks? And how will this change affect expected unit fill rates at the depot?
2. A simulation model of the Air Force multi-echelon inventory system for EOQ items. A multi-echelon model is necessary to study how changes in stockage policies at one level affect the stockage policies at the other level. Unfortunately, analytical multi-echelon inventory models that capture the essence of the Air Force system do not exist. Thus, a simulation model is required.

III. A Static Model of the D062 Wholesale (Depot) Requirements System

In order to construct an accurate model of the D062 system, the latest revision of the D062 documentation was acquired from the BDM Corporation. In addition, sample outputs were obtained so checks could be made that our model does, in fact, emulate the actual D062 system. Because of space limitations, this report focuses on the results obtained from an analysis using this model rather than on a description of the

model. Details on the D062 Emulator Model itself are contained in Rinks (1986a).

As stated previously, AFLC has changed the method for computing depot safety levels several times as they attempted to "fine tune" D062. One use of this model is to predict the consequences of such changes in depot stockage policy. Listed below are several methods for computing depot safety levels.

The following definitions apply; the subscript i refers to the ith item, where i = 1, . . . , n.

Q_i = order quantity (units)

ROP_i = reorder level (units)

k_i = safety factor

σ_i = standard deviation of demand during a leadtime

a_i = holding cost fraction

A_i = ordering cost (dollars per order)

c_i = item cost (dollars per unit)

μ_i = mean demand during a leadtime

Z_i = item essentiality

ARS_i = average requisition size

ISF = implied shortage factor

β = expected number of essentiality-weighted units in a backordered status at any point in time

Method 1 -- Original Presutti & Trepp Formula

$$k_i = -0.707 \ln \left[\frac{\sqrt{2}\, Q_i c_i a_i}{0.5\ ISF\ Z_i\ \sigma_i\ (1 - e^{-\sqrt{2}Q_i/\sigma_i})} \right] \quad (1)$$

$$ROP_i = \mu_i + k_i \sigma_i \quad (2)$$

Method 2 -- Square root of unit cost

$$k_i = -0.707 \ln \left[\frac{\sqrt{2}\, Q_i \sqrt{c_i}\, a_i}{0.5\ ISF\ Z_i\ \sigma_i\ (1 - e^{-\sqrt{2}Q_i/\sigma_i})} \right] \quad (3)$$

$$ROP_i = \mu_i + k_i\sigma_i$$

Method 3 -- Square root of average requisition size

$$k_i = -0.707 \ln \left[\frac{\sqrt{2}\, Q_i\, c_i\, a_i}{0.5\ ISF\ (1/\sqrt{ARS_i})\ \sigma_i\ (1 - e^{-\sqrt{2}Q_i/\sigma_i})} \right] \quad (4)$$

$$ROP_i = \mu_i + k_i\sigma_i$$

Method 4 -- Square root of unit cost and square root of average requisition size

$$k_i = -0.707 \ln \left[\frac{\sqrt{2}\, Q_i \sqrt{c_i}\, a_i}{0.5\ ISF\ (1/\sqrt{ARS_i})\ \sigma_i\ (1 - e^{-\sqrt{2}Q_i/\sigma_i})} \right] \quad (5)$$

$$ROP_i = \mu_i + k_i\sigma_i$$

Method 5 -- Fifty-five day time supply

$$ROP_i = \mu_i + SS_i \quad (6)$$

where the safety level (SS) is set equal to the average usage in fifty-five days.

Method 6 -- One standard deviation of demand during lead time

$$ROP_i = \mu_i + \sigma_i \quad (7)$$

Method 7 -- One standard deviation of demand during lead time + one average requisition lot size

$$ROP_i = \mu_i + \sigma_i + ARS_i \tag{8}$$

Method 8 -- Square root of unit cost and factoring of safety level depending on Mission Impact Essentiality Code

$$k_i = -0.707 \ln \left[\frac{\sqrt{2}\, Q_i \sqrt{c_i}\, a_i}{0.5\, ISF\, \sigma_i \left(1 - e^{-\sqrt{2}Q_i/\sigma_i}\right)} \right] \tag{9}$$

$$ROP_i = \mu_i + \frac{1}{Z_i} k_i \sigma_i \tag{10}$$

where Z_i = 1, 2, 3, or 4 depending on MIEC.

Method 1 is the formula for computing safety levels as originally developed in Presutti and Trepp (1970). The Presutti and Trepp model minimizes the total of variable ordering and holding cost subject to a constraint on time-weighted, essentiality-weighted, backorders.

$$\text{Minimize} \quad \sum \frac{A_i D_i}{Q_i} + \sum_{i=1}^{n} a_i c_i \left(\mu_i + k_i \sigma_i + \frac{Q_i}{2}\right) \tag{11}$$

$$\text{subject to} \quad \sum_{i=1}^{n} \frac{0.5}{2} \frac{Z_i \sigma_i^2}{Q_i} \left(1 - \exp\left(-\sqrt{2}\, \frac{Q_i}{\sigma_i}\right)\right) \exp\left(-\sqrt{2} k_i\right) \leq \beta. \tag{12}$$

The method of Lagrange yields Equation (1) as the solution for the safety factor for the items, where

$$ISF = \sum_{i=1}^{n} \frac{\sigma_i a_i c_i}{\sqrt{2}\beta} \tag{13}$$

and all other variables except Q_i are known or estimated quantities. Order quantities, Q_i, are also found as part of the solution methodology. However, D062 uses the Wilson lot size formula with certain modifications (e.g., order quantities must be at least a year's supply but not more than three years' supply).

In practice the implied shortage factor ISF is specified in such a manner that the aggregate safety levels are approximately equal to 55 days of demand (Hanks, 1985). Consequently, the problem is not really solved as a constrained optimization problem. Rather ISF is used as a "control knob." Larger values of ISF result in larger values for k_i, which, in turn, result in fewer backorders: smaller values for ISF have the opposite effect.

The average number of units in backorder status at a random point in time, B_T, for an item with stockage policy Q_i and k_i is

$$B_T = \frac{0.5}{2} \frac{\sigma_i^2}{Q_i}(1 - \exp(-\sqrt{2}Q_i/\sigma_i))\exp(-\sqrt{2}k_i). \qquad (14)$$

This expression assumes that demands during a leadtime are Laplace distributed.

Methods 2, 3, 4, and 8 are all variations of the original Presutti and Trepp formula. Each of these methods have either been actually used, or at least considered, by AFLC. One way of viewing the differences in these models is that they are distinguished by how the item essentiality, Z_i, is operationalized. The actual rationale behind these variations was probably much simpler. For instance, if the square root of unit cost rather than unit cost is used in the formula (Method 2), then the computed safety factor will be larger for the more expensive items. All other things being equal, Method 1 (original Presutti and Trepp) favors

investment in safety stock for lower cost items as the most efficient way of meeting the expected backorder constraint. Thus, Method 2 would tend to shift investment in inventory towards the more expensive items.

Method 5 is what some believe item managers actually use in implementing stockage policy for the items they manage. Each item has a safety level expressed as a fixed time supply of demand. Numerous studies have demonstrated that variable safety level policies dominate fixed time supply policies.

Methods 6 and 7 are the safety level policies used at the bases. We were interested in how simple variable safety level policies such as these would compare.

The "population" of items for the depot static analysis is derived from the SSD Emulator Data Base (Hanks, 1985). This data base (a 32-row by 10-column matrix) is essentially a 320 cell annual dollar usage array. Each cell represents a "typical item" at the wholesale level; and a typical item may, in turn, represent many actual items having similar characteristics. By taking these cell multiplicities into account, the data base represents 205,000 SSD items.

From the Hanks' template, information for a cell's annual demand, cost per unit, average requisition size, and supply management grouping code (SMGC) are available. Following Hanks, leadtimes were set in accordance with an item's SMGC: 323 days for 'T' items: 449 days for 'P' items: and 559 days for 'M' items. Next, means and standard deviations of demand during a leadtime were calculated based on a cell's annual demand and leadtime; for the purpose of computing the standard deviations, demand was assumed to follow a constant Poisson process.

To complete the parameterization, the following values were used in

the analysis. The implied shortage factor was set at $1100/unit/year and the holding cost fraction at 15%. The ordering cost was $620 if the annual dollar usage was less than $25000; otherwise, the ordering cost was $970. Item essentialities were all assumed equal to 1.0 as no information was available on which other values might be construed. Note that when Z_i=1, Method 2 and Method 8 are identical.

For every item, each method of computing safety levels was used to calculate safety factors. Order quantities were calculated according to the D062 EOQ machinery. Given these stockage policies (k_i, Q_i), expected number of backorders in place were computed by Equation (14). For the purpose of exposition, aggregation of results was done by SMGC. A comparison of the expected cost of the methods (dollars invested in inventory) versus expected number of backorders in place is given in Table 1.

Table 1.

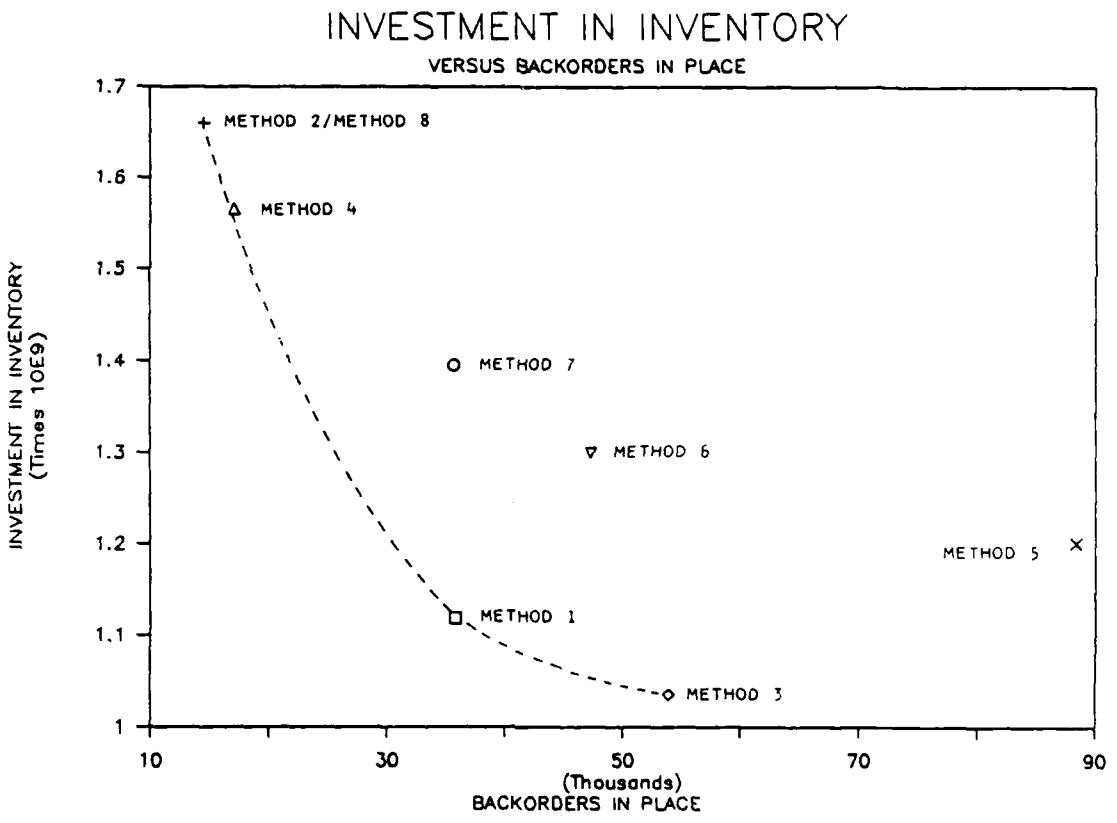
DOLLARS INVESTED IN INVENTORY VERSUS BACKORDERS IN PLACE

SAFETY FACTOR CALC. METHOD	SMGC = 'T'		SMGC = 'P'		SMGC = 'M'		TOTALS	
	B_T	$ INV (000)	B_T	$ INV (000)	B_T	$ INV (000)	B_T	$ INV (000)
1	17445	241863	12774	312621	5494	1118698	35713	1118689
2/8	12812	304489	1163	602199	513	7454033	14488	1667021
3	25285	206907	21105	273122	7561	554901	53951	1034930
4	13519	299741	2335	553233	1154	712840	17008	1565814
5	68891	184024	17808	317849	1630	699621	88329	1201494
6	31699	244861	13128	389975	2434	665774	47261	1300610
7	21946	278361	19185	436581	2190	677069	34321	1392011

A graph of total backorders in place versus total dollars invested in inventory for the various methods (Figure 1) is quite revealing. The first observation is that Methods 1, 2, 3, and 4 (Presutti and Trepp and

its variations) dominate the other policies. In fact, one can think of these as another type of "control knob for dialing in either expected backorders in place or, alternately, investment in inventory." These methods appear to be an effective way to do this, given the stated optimization problem. However, different formulations of the decision problem might prove even more attractive to the Air Force.

Figure 1



Method 5 -- a 55 day time supply of safety stock for each item -- is clearly dominated by the variable safety level (VSL) methods. This result is well known from inventory theory. Nevertheless, many practicing logisticians continue to use this method for safety stocks. From Figure 1, one can infer that for the same level of investment in inventory as required by Method 5, an efficient variable safety level policy would result in only about one-third as many backorders in place. Clearly, the Air Force does not want item managers to use Method 5 for establishing reorder levels. Method 6 is also inferior to VSL methods that consider pertinent information other than demand variability. Adding a requisition lot size to Method 6 (Method 7) reduces backorders, although the cost effectiveness in applying this to all items could be questioned. Adding a lot size for items having certain demand and cost characteristics might be desirable.

A single echelon model that focuses only on the depot has obvious limitations. Foremost, service levels at the bases are not addressed directly. Nevertheless, models such as this do allow aggregate analysis of depot stockage policies before they are implemented.

IV. A Multi-Echelon Inventory Simulation Model for EOQ Items

In order to study how changes in stockage policies at one level affect the stockage policies at the other level, a multi-echelon simulation model of the Air Force's depot-base inventory system for EOQ (consumable) items was constructed. Due to space limitations, this report documents results from a small scale study using this model rather than fully describing the model. See Rinks (1986b) for a complete description of the simulation model.

As far as we are aware, this is a first attempt at developing a multi-echelon simulation model of the Air Force's depot-base inventory system for consumable items. Consequently, there are many issues that warrant investigation. Some of these issues are related to stockage policies; others are related to the depot-base system; while others have to do with the simulation. A partial list of these issues include:

1. How do different methods for computing depot safety levels compare in the more realistic multi-echelon simulation model?
2. Are there item specific characteristics (demand level, cost/unit, SMGC. etc.) that influence the performance of different stockage policies?
3. How do the number of bases having demand for an item and the type of base (CONUS, PACAF, USAFE) influence the performance of stockage policies?
4. How many items as a sample population are required in order to generalize results to the entire population?
5. How long a simulation is required before the system reaches steady state? Are replications necessary?

Early experience with computer run times on the AFLMC's IBM 4331 system indicated that the scope of investigation would have to be limited. CPU times for a single simulation, in some cases, exceeded 20 minutes, and a full scale study might entail several thousand simulation runs. Given the constraints on computer time and the length of the research period, a pilot study was designed to address a selected number of issues.

The general framework for the pilot study was three dimensional: (1) methods for computing safety levels; (2) number of bases having demand for an item and type of base; and (3) number of items.

Based on the findings from the static analysis, we decided on five methods for computing safety factors: Methods 1, 2/8, 3, 5, and 7. The number of bases that demanded an item was studied at three levels: 5 bases (1 USAFE, 2 CONUS, and 1 PACAF), 12 bases (2 USAFE, 8 CONUS, and 2 PACAF), and 20 bases (3 USAFE, 14 CONUS, and 3 PACAF). Nine items, determined by item characteristics, were used in the study. Specifically, we selected items from the Hanks' template that had a cost/unit of "Low," "Medium," or "High" and had a demand of "Low," "Medium," or "High." Characteristics for the items used in the study are given in Table 2.

Table 2.

ITEM CHARACTERISTICS FOR THE SIMULATION STUDY[a]

DEMAND	COST: LOW	COST: MEDIUM	COST: HIGH
LOW	$0.26 25 T	$35.67 20 T	$359.45 10 T
MEDIUM	$2.52 2741 P	$72.86 375 P	$751.28 63 P
HIGH	$2.40 26458 M	$34.73 4599 M	$2500.00 2056 M

[a]Entries in the cell are: cost/unit
annual demand
SMGC

Customer arrivals at the bases were assumed Poisson distributed, with each customer demanding the same quantity (DSIZE). Thus, the demands at the base level follow a constant Poisson process. Demand levels at the bases were "backed out" from the overall demand at the depot for a given number of bases. Customer demand quantities at the base (DSIZE) were set in accordance with AFLMC past experience for base demand of that order of magnitude. The result of this procedure is the specification of a nominal value for the arrival rate at the bases. Next, in order to account for bases having different demand levels, arrival rates were randomly assigned to one of the three levels: Low (0.5 X), Medium (1.0 X), or High (1.5 X). For the cases of 5, 12, and 20 bases, the distribution of demand levels (Low, Medium, and High Demand) were (1, 3, 1), (3, 6, 3), and (5, 10, 5), respectively.

Order and ship times for both *Base-Depot-Base* and *Depot-Vendor-Depot* were assumed lognormal distributed. Table 3 gives the parameters used in the study for Base-Depot-Base OST. Note that these variates in the simulation vary by base location and order type (EOQ or Priority). The parameters for Depot-Vendor-Depot OST are listed in Table 4. Means for Depot-Vendor-Depot were the same as those used in the static analysis; standard deviations were specified as one-half the mean value. In addition, order and ship time variates for Depot-Vendor-Depot were restricted to be greater than or equal to one-half the mean OST and less than or equal to twice the mean OST.

For purposes of computing EOQs at the base level, the inventory holding cost fraction was 0.15 and the cost to place an order was $5.25. All other model parameters were the same as used in the static analysis.

Table 3.

PARAMETER FOR BASE-DEPOT-BASE ORDER AND SHIP TIMES

Base Location/ Order Type	Mean (Days)	Std. Dev. (Days)
CONUS		
EOQ	17	10
Priority	10	7
PACAF		
EOQ	42	28
Priority	18	7
USAFE		
EOQ	32	15
Priority	18	7

Table 4.

PARAMETERS FOR DEPOT-VENDOR-DEPOT ORDER AND SHIP TIMES

SMGC	Mean (days)	Std. Dev. (days)	Minimum (days)	Maximum (days)
T	323	162	162	646
P	449	225	225	900
M	559	280	280	1120

As outlined above, the pilot study consisted of 135 simulations (9 items x 5 safety factor methods x 3 number of bases). Relevant statistics were collected on performance at both the depot and the bases.

Analysis of the simulation results is continuing at AFLMC. An example of a comparison of simulation results with those from the static analysis for safety factor Method 2/8 is given in Table 5. Depot fill rates for the simulation are lower than from the static analysis, even though the simulation fill rates are considerably higher than what the depots are actually experiencing.

Table 5.

COMPARISON OF DEPOT FILL RATES

SAFETY FACTOR METHOD 2/8

ITEM[a]	5 BASES	12BASES	20BASES	STATIC ANALYSIS
T-LL	.9981	.9985	.9980	.999
T-LM	.9923	.9956	.9928	.999
T-LH	.9872	.9883	.9926	.998
P-ML	.9439	.9666	.9825	.999
P-MM	.9499	.9254	.9653	.999
P-MH	.9304	.9481	.9471	.996
M-HL	.9285	.9356	.8442	.999
M-HM	.9292	.9346	.9237	.999
M-HH	.8486	.8565	.8408	.993

[a]Item identification is X-YZ where X = SMGC, Y = demand level (L=Low, M=Medium, H=High), and Z = cost level (L=Low, M=Medium, H=High). See Table 2.

The static analysis assumes that the Depot-Vendor-Depot leadtime is constant. We were interested in what effect leadtime variability had on the simulation results. In order to investigate this we conducted three simulations in which the standard deviation of OST differed. Specifically, item P-MM was used with 12 bases having demand for the item. Method 2/8 was used for computing the safety factor. The OST was distributed as a lognormal with a mean of 449 days, and the standard deviation of leadtime was 225, 150, and 100 days for the three simulations. Performance measures at the depot are presented in Table 6.

Table 6.

EFFECT OF LEADTIME VARIABILITY ON DEPOT PERFORMANCE

OST STD DEV	AVERAGE FILL RATE	AVERAGE B/O DAYS	NUMBER OF B/O PER YR	AVG EOQ	AVG ROP	ONHAND INV
225	.9469	79.82	19.644	386.66	634.21	355.52
150	.9865	36.19	5.000	387.35	653.66	367.30
100	.9976	8.41	0.888	390.60	656.10	375.09
Static Analysis	.999	----	0.438	376.00	615.30	342.00

Note that the service performance at the depot deteriorates as the leadtime variability increases. However, there is not a pronounced effect on the depot stockage policy. Furthermore, the performance measures appear to be approaching those predicted by the static analysis as the standard deviation of leadtime approaches zero. This demonstrates with some credence that the system has been modelled correctly.

V. Recommendations

The construction of these models represents a significant first step in analyzing stockage policies for the Air Force's multi-echelon inventory system for EOQ items. However, much work remains to be done before their full potential can be realized.

1. There are a number of ways the models can be improved. For instance, the simulation model computes a "static" EOQ for each base at the beginning of the simulation that is then used for the remainder of the simulation. A "dynamic" stockage policy at the base level would allow investigation of how increasing or decreasing demand levels at the bases affect the stockage policy at the depot. Also, more detailed performance measures at the base level, such as time-weighted backorders, are needed.

2. In order to be able to generalize results, a representative sample of EOQ items is needed. Such a sample might involve as many as 1500 items selected on the basis of SMGC, annual demand, cost/unit, number of bases demanding the item, etc.

3. Service levels predicted by both the static model and the simulation model are considerably higher than those achieved by the system in practice. Identifying the reason(s) for this would represent as extremely important finding. These models may serve as a means of testing conjectures about the disparity between observed and predicted service levels.

4. The original intent of this research project was to construct a multi-echelon EOQ model for MICAP allocation. However, the construction and validation of the multi-echelon EOQ simulation model for consumables exhausted the entire research period. Extending this model to address the MICAP allocation problem will be the subject of my mini grant proposal.

REFERENCES

1. Department of Defense, "Procurement Cycles and Safety Levels of Supply for Secondary Items," DoD Instruction 4140.39, Washington, D.C.: Department of Defense, July 1970.
2. Hanks, C. H., The Influence of Systems Support Division Funding and Safety Levels on Aircraft Availability, Final Report, HQ USAF/LEXY, Logistics Management Institute, Bethesda, Maryland, October 1985.
3. Headquarters, Air Force Logistics Command, "Requirements Procedures for Economic Order Quantity Items," (EOQ Buy Budget Computation System - D062), AFLC Regulation 57-6, Wright-Patterson Air Force Base, Ohio: Headquarters, Air Force Logistics Command, 22 August 1984.
4. Presutti, V. J. and R. C. Trepp, "More Ado About Economic Order Quantity (EOQ)," Naval Research Logistics Quarterly, June 1970, pp. 243-251.
5. Rinks, D. B., "A Static Model of the D062 Wholesale (Depot) Requirements System," User/Programmer Guide, Air Force Logistics Management Center, Gunter AFS, Alabama, August 1986.
6. Rinks, D. B., "MECH: A SIMSCRIPT Simulation Program of a Multi-Echelon EOQ System for Consumables," User/Programmer Guide, Air Force Logistics Management Center, Gunter AFS, Alabama, August 1986.

1986 USAF-UES Summer Faculty Research Program/

Graduate Student Summer Support Program

Sponsored by the

Air Force Office of Scientific Research

Conducted by the

Universal Energy Systems, Inc.

Final Report

Preliminary Development of a Global Positioning System Package for use in Determining Exact Position of AFGL Research Balloons at Precise Times

Prepared by:	W. Paxton Robey
Academic Rank:	Assistant Professor
Department and	Electronics and Computer Technology
University:	Oklahoma State University
Research Location:	Air Force Geophysics Laboratory
	Hanscom AFB, Massachusetts
USAF Researcher:	Catherine Rice
Date:	August 14, 1986
Contract No:	F49620-85-C-0013

Preliminary Development of a Global Positioning System Package for use in Determining Exact Position of AFGL Research Balloons at Precise Times

by

W. Paxton Robey

ABSTRACT

The Air Force Geophysics Laboratory is involved in upper atmosphere research using high altitude balloons to carry scientific instrument packages. Two of the requirements are the tracking of each balloon for safety and recovery of the payload, and determining the precise position of the payload at the exact time each experiment is conducted. The Global Positioning System, a satellite radio navigation system, is now available and the adaptation of it to AFGL balloon flights is desirable. Moderate cost commercial GPS receiving equipment is now available which can be adapted to this use. This paper discusses the status of the hardware and software at this time and makes recommendations for proceeding with the incorporation of a balloon borne GPS receiver with telemetry for transmission to a ground station. Two areas of concern are the suitability of commercially available software for operation on an unmanned stratospheric balloon and the interfacing of the asynchronous RS-232 GPS output to S-band telemetry.

Acknowledgments

I would like to acknowledge and thank the Air Force Systems Command and the Air Force Office of Scientific Research for sponsorship of my research. I would like to express appreciation to Rodney Darrah, Sue Espy and Agnes Baine of Universal Energy Systems for the way in which they were able to administer the research program with a minimum of red tape and the least amount of confusion, at least from the point of view of the researcher.

I want to acknowledge the excellent facilities, especially the Research Library, which the Air Force Geophysics Laboratory made available to me. In particular, very helpful and pleasant assistance was afforded to me by Catherine Rice, Neal Stark, Art Korn, Ralph Cowie, Jack Griffin, Dolph Jursa and Alice McGinty.

I. Introduction

My background in electronics began in 1962 with military training in the repair and maintenance of communication and surveillance equipment. After my military obligation was completed, I received a BS degree in Engineering Physics from the University of Tulsa.

After ten years of engineering experience at American Airlines and seven years of operating my own engineering firm I opted for the less stressful world of teaching. Prior to joining the faculty at Oklahoma State University I was an instructor of electronics at Oklahoma Junior College.

At Oklahoma State I bacame familiar with the long standing association between the OSU electronics lab and AFGL. While the research carried out this summer was not connected with the OSU lab directly, it is hoped that this work will contribute to an expansion of OSU's involvement and contributions to the balloon research effort at AFGL.

The use of the Global Positioning System is an exciting development in todays technology. The potential for universal applications in navigation, location and precise positioning is vast and I would hope that I will have a long term connection with the Air Force in this area.

II Objectives of the Research Effort

The overall objective of this summer's effort was to formulate in specific terms a plan for incorporating the use of the Global Positioning System in research balloon flights carried out by AFGL. Specifically the objectives were to define any problem areas which need clarification or may require developmental work before the use of GPS for navigation and precise positioning can become a hassle-free, routine component available for all flights.

Some individual objectives were:

1. Determine the "state of the art" of low cost, commercially available GPS receiving/computing equipment and its applicability to the AFGL effort.
2. Create a feasible time table for incorporation of GPS equipment, including the effects of the limited number of satellites available (and the impact of the Space Shuttle explosion, since the shuttle was the primary launch vehicle for the Navstar/GPS satellites.)
3. Outline the steps required to overcome any electronic or logistical problems to the routine use of GPS in balloon flights.

III Background on Global Positioning System

It is not the purpose of this report to describe the operation of the Global Positioning System (GPS) in detail. However a brief overview of the system is presented here for

the benefit of readers who are not presently familiar with the concept. For more detailed explanations the reader is referred to the references listed, especially the compilation of articles on GPS published in two volumes by Navigation magazine.

GPS is not the first attempt at creating a world wide navigation system using satellite radio sources. As early as 1964 the Navy's Transit satellite system became operational. However GPS now offers the probability of a world wide, full time, system which provides 3 dimensional position information, velocity, heading and time with atomic clock accuracy. The system will be available to civilian users as well as military. Positioning accuracies for military users is purported to be possible with less than 1 meter spherical error probable (SEP). Civilian users may be able to get positioning data accurate to 5 to 10 meters SEP.

The Navstar/GPS is a space based radio positioning navigation and time transfer system. It will consist of 18 operating satellites in a constellation of 6 orbits with 3 satellites per orbit. Each satellite broadcasts on two L-band frequencies, L1 = 1575.42 MHz and L2 = 1227.60 MHz. Additional L-band frequencies are used by ground stations as up links for controlling the satellites and providing updated data to the satellite's almanac (ephemeris data, satellite health status, etc.). At the present time there are seven functioning Navstar satellites in two orbit planes. An additional eleven satellites are planned. The original plan was to begin launching these units in 1986 and have the constellation completed by the end of 1988. Present

speculation in light of the shuttle disaster would have future launches begin in 1988 with completion between 1990 and 1992.

The L1 signal is modulated with both the precision (P) code and the coarse acquisition (C/A) code. Both codes are pseudo random noise generated signals. The L2 signal is modulated with the P code only. Both the L1 and L2 signals must be received to use P code. The P code is available only to the military and authorized civilian users. While the P code provides greater precision, receivers for C/A code are much less complex and therefore much less expensive. The C/A code is sufficient for most AFGL operations and can be used without military clearances and authorizations.

In order for the GPS user's computer to be able to calculate position (X,Y,Z coordinates) and velocity (time), four equations in four unknowns must be solved. This obviously requires the presence above the horizon of four usable satellites. When all 18 satellites are in orbit the probability of having four usable sources at all times is greater than 95%. However, with the present seven satellites the system is capable of 3 dimensional positioning only about 25% of the time. For example, figure 1 is a print out of satellite visibility times for 15 July 1986 at Boston. Most AFGL flights, however, are launched from Holloman Air Force Base, New Mexico. The figure includes the altitude and azimuth of each satellite. (Courtesy of Rockwell Collins)

IV AFGL Mission

The USAF Geophysics Laboratory has been conducting upper

SATELLITE VISIBILITY ON 15-JUL-86 at BOSTON

DAY OF YEAR 196 GPS WEEK 340 LAT=N43 12'52.0" LONG=W72 43'11.9" ALT=0 meters

SV	3	6	8	9	11	12	13
REF WEEK	336	336	336	336	336	336	336
TIME (ZONE -4)	el az	el az	el az	el az	el az	el az	el az
00:00:00							38*146
00:20:00							29*151
00:40:00							20*155
01:00:00							11*158
01:20:00		6*104					2*161
01:40:00		13*98					
02:00:00		19*90		4*132			
02:20:00		25*81		12*127			
02:40:00		28*71		20*121			
03:00:00		30*60		28*114		0*160	
03:20:00		29*49		36*105		8*157	
03:40:00		26*39		41*94		17*153	
04:00:00	3*345	20*32		45*80		27*150	
04:20:00	9*340	14*26		45*65		37*146	
04:40:00	13*332	6*22		41*52		47*140	
05:00:00	16*324			35*42		56*130	
05:20:00	17*314			28*35		63*113	
05:40:00	15*305		3*351	20*31		66*89	
06:00:00	12*296		6*344	11*30		63*65	
06:20:00	8*287		8*336	3*30		57*49	
06:40:00	2*280		7*327			48*41	
07:00:00			5*319		1*5	40*38	
07:20:00			1*311		2*357	31*37	
07:40:00					2*349	22*39	
08:00:00					0*341	14*41	1*41
08:20:00						6*44	4*33
08:40:00							5*24
09:00:00							5*16
09:20:00							3*8
09:40:00							
10:00:00							
10:20:00							
10:40:00							
11:00:00							
11:20:00							
11:40:00							
12:00:00							
12:20:00	7*175						
12:40:00	16*174						
13:00:00	25*174						
13:20:00	35*174						
13:40:00	46*173						
14:00:00	56*173		8*196				
14:20:00	67*170		18*198				
14:40:00	78*157		27*201				
15:00:00	83*95	6*317	37*206				
15:20:00	77*44	14*321	47*211		1*216		
15:40:00	66*35	23*323	57*220		9*220		
16:00:00	56*34	31*324	66*232		17*225		
16:20:00	47*37	41*323	74*255		24*231		
16:40:00	38*40	50*318	78*300	6*329	32*239		
17:00:00	29*44	58*309	75*345	15*329	39*247		
17:20:00	21*49	64*291	68*9	23*327	46*258		
17:40:00	14*55	65*265	60*24	32*323	52*270		
18:00:00	7*61	61*242	53*34	39*315	58*285		
18:20:00		53*228	45*44	45*305	62*303		4*258
18:40:00		43*219	38*52	48*290	64*325	7*337	10*265
19:00:00		34*214	32*60	47*275	65*348	14*333	16*272
19:20:00		24*210	25*68	43*261	64*11	20*327	22*280
19:40:00		15*207	19*76	37*250	61*31	25*319	29*288
20:00:00		6*204	13*83	29*241	56*48	28*310	35*297
20:20:00			7*91	21*235	51*63	28*299	42*306
20:40:00			2*98	13*229	46*75	26*289	48*315
21:00:00				5*225	40*87	22*279	55*326
21:20:00					33*97	17*271	62*339
21:40:00					26*106	11*263	69*357
22:00:00					19*113	4*257	74*25
22:20:00					12*121		75*64
22:40:00					5*127		71*98
23:00:00							63*119
23:20:00							55*132
23:40:00							46*140
00:00:00							36*147

FIGURE 1

atmosphere research experiments since 1945. The Aerospace Instrumentation Division supports the other divisions of AFGL by providing the vehicle systems - balloons and sounding rockets - to carry the instruments that gather data for scientists in their study of the environment.

Typical altitudes attained by balloon flight are 100,000 to 130,000 feet with flights having gone as high as 170,000 feet. Depending on float altitude, payloads may exceed 4000 pounds. Typical duration of flights is 12 hours. Measurements of the vehicle's position and velocity are required for flight safety, recovering the payload, and also as a part of the atmospheric sampling data. The more precisely that the altitude and other parameters of a balloon's position are known when air samples or gravity measurements are made, for instance, the more valuable is the data. The use of GPS will free the Lab from dependence on the availability of precise radar equipment on test ranges, allow world wide operation, as well as increase precision.

The telemetry systems used for control and monitoring of the balloon and its payload are synchronous S-band multiplex systems. It is desired to have the asynchronous output of the GPS equipment be in a format which will conform to the existing telemetry systems.

V Status of Commercially Available GPS Hardware

Contacts were made with seven manufacturers of commercial C/A code GPS receivers: Rockwell-Collins, Magnavox, Trimble Navigation, Interstate Electronics, Texas Instruments, ISTAC,

and Motorola. Plant visits were made to Rockwell-Collins and Motorola. None of these manufacturers had operated their C/A code equipment in the stratosphere in exactly the manner which we desired. A Texas Instruments P code receiver has been flown at 100,000 feet on a balloon for an AFGL experiment in which a seperate telemetry channel was used.

Preliminary discussions and visits with manufacturers indicate a minimum of problems can be expected in adapting "off the shelf" C/A code commercial equipment to high altitude balloon flights. Particular areas of concern were:

1. Power consumption and voltage requirements
2. Heat dissipation or equipment cooling
3. Radiation hardness (cosmic rays only)
4. High altitude arcing
5. Shock g loading
6. Signal interfacing to telemetry
7. Reacquisition of tracking in case of momentary interuption

Although each of these areas will require flight testing (an altitude chamber is available at AFGL where environmental tests can be conducted), none of the first five items are anticipated to cause other than minor problems, no matter which manufacturer's equipment might be used. For instance, shock mounting and attaching a heat sink for heat dissipation is expected to be sufficient to handle the g loading and heat generation. The Rockwell-Collins and Motorola engineers anticipate no difficulty in automatic reacquisition of tracking.

There are at least two techniques which have been used succesfully in the past for interfacing an output like that of the GPS receiver to the telemetry transmitter and recovering the data on the ground. The output of commercial GPS receivers is typically either RS-232 or RS-422 serial data. In any case, it is anticipated that the output will be 5V nominal TTL data, possibly at a baud rate of 9600 bps.

One method is the use of a separate IRIG channel if frequency division (FM-FM) multiplexing is used. In this case a channel is chosen which has sufficient bandwidth to handle the 9600bps signal. The subcarrier voltage-controlled oscillator (VCO) is modulated directly by the RS-232 output. On the ground the signal is demodulated and supplied to the the computer's RS-232 input for display and processing.

If the preferred multiplexing method is pulse code modulation (PCM) then the asynchronous data can be piggy-backed on the bit stream. The effective frequency of the 9600 baud data will be either 9600 Hz or 4800 Hz depending on wheither "bi-phase" (BIϕ) or "non return to zero" (NRZ) bit representation is used. The standard PCM data may be on the order of 300 Kbs. By using an active mixer, the two frequencies may be mixed prior to entering the S-band transmitter. For protection from interference it is suggested that the 300 Kbs be put through a high pass filter with a cutoff frequency of approximately 165 KHz prior to mixing. Also the GPS data would be passed through a low pass filter with a cutoff frequency of around 50 Khz.

On the ground, the output of the S-band receiver is sent through a high pass filter to the standard bit synchronizer for signal processing in the usual manner. At the output of the receiver, the signal is also applied to a low pass filter and can then be fed to the computer's RS-232 input for display and processing. A block diagram is shown in Figure 2.

It is possible that in the future it may be desirable to develop a buffer memory to accept the standard RS-232 asynchronous output signal from the GPS computer and release the data in a synchronous manner for incorporation into the PCM bit stream for transmission to the ground.

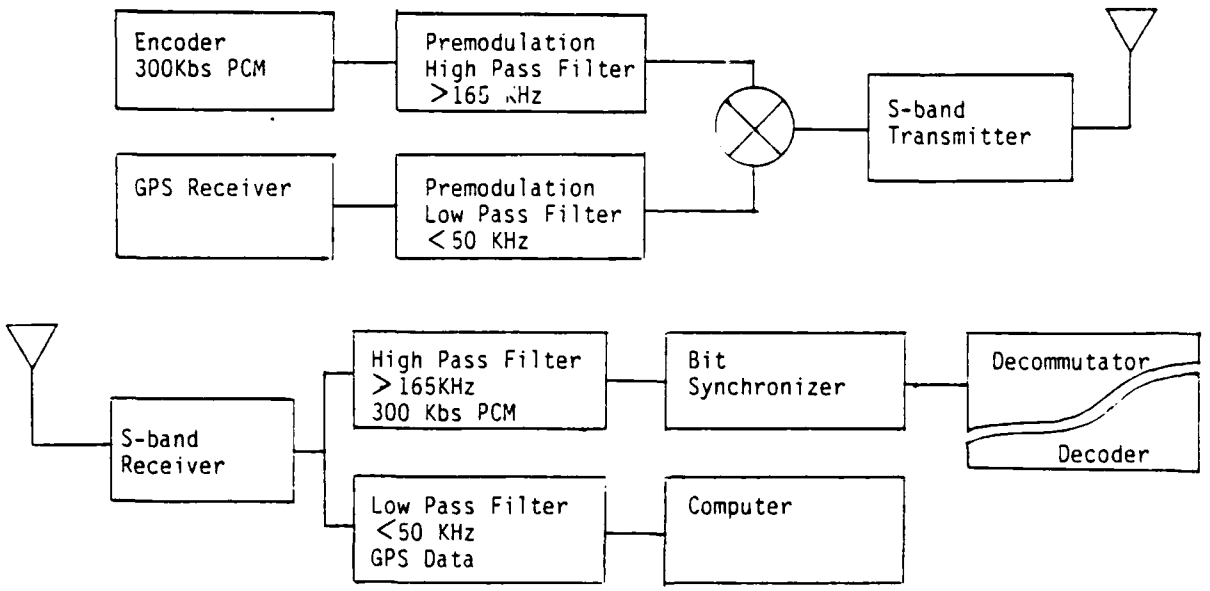


Figure 2

VI Status of commercially available GPS software

Although each manufacturer claimed to either have working software, or soon would have working software, which would allow the data to be displayed at the ground station by a portable computer, none was actually observed by myself. Typical claims were that the software would allow real time monitoring of position, velocity, heading, etc. and also could be used for post processing, i.e., smoothing of data or comparison with data from a second receiver at a known location (differential processing) to minimize errors.

Obviously the value of the equipment is dependent on the completeness, simplicity, user friendliness, and bug free operation of the software. This is definitely an area which needs thorough investigation before equipment purchases are made.

VII. Recommendations

As the work done this summer was preliminary and no equipment hardware nor software was used or tested, a complete hands-on investigation needs to be carried out. However, due to the apparent advanced state of low cost commercial grade GPS hardware, it is expected that within a year of the start of the investigation, a balloon GPS package can be on line and considered routinely available to all AFGL scientists. Specifically the following should be carried out:

1. After acquiring a GPS receiver /computer for test, the

following can be done on the bench:

a. directly connect the RS-232 output to a PC to verify the useability of the software

b. If the GPS data is to be inserted into the bit stream, develope the buffer for interfacing the asynchronous RS-232 output to the telemetry channel

c. In the altitude chamber test for heat dissipation, arcing or other altitude effects.

d. Operate and debug the complete GPS telemetry system.

2. Fly the equipment when precision radar tracking facilities are available for verification of the operation.

References

1. Jorgensen,P.S., "NAVSTAR/Global Positioning System 18-Satellite Constellations", The Aerospace Corporation, El Segundo, California
2. Milliken, R.J., and Zoller, C.J., "Principles of Operation of NAVSTAR and System Characteristics", The Institute of Navigation, Global Positioning System special issue, 1980
3. "NAVSTAR GLOBAL POSITIONING SYSTEM", The Analytic Sciences Corporation, Reading, Mass., Feb. 1981
4. Glazer, B. G., "GPS Receiver Operation", Magnivox, Torrance, CA
5. Aviation Week & Space Technology, numerous articles
6. "Navstar Forecast", Defense Electronics, May 1986

1986 USAF-UES Summer Faculty Research Program/

Graduate Student Summer Support Program

Sponsored by the

Air Force Office of Scientific Research

Conducted by the

Universal Energy Systems, Inc.

Final Report

KINETIC PROCESSES IN ADVANCED ALLOYS

Prepared by:	Kenneth C. Russell
Academic Rank:	Professor
Department and University:	Dept. of Materials Science & Engineering; Dept. of Nuclear Engineering; M.I.T.
Research Location:	Structural Metals Branch AFWAL/MLLS Wright Patterson Air Force Base
USAF Research:	Dr. C.F. Froes
Date:	September 30, 1986
Contract No.:	F49620-85-C-0013

KINETIC PROCESSES IN ADVANCED ALLOYS

by

Kenneth C. Russell

ABSTRACT

Various kinetic and thermodynamic criteria proposed to predict the occurrence of terminal solid solubility extension (TSSE) under rapid solidification are discussed. A new criterion based on Miedema's macroscopic atom model for solution thermodynamics is found to predict TSSE very well for Al-based alloys. The criterion is presently being extended to other binary alloys where there exists an adequate data base.

The theory of particle coarsening is extended to show that small dispersoid particles located on grain boundaries or dislocation lines may undergo preferential coarsening. A theory is developed which predicts that dispersoid particles may sometimes be dragged by migrating grain boundaries. The predictions of enhanced coarsening and particle dragging are in general accord with experimental observations on Ti and Ti_3Al based alloys.

Acknowledgements

I would like to thank the Air Force Systems Command and the Air Force Office of Scientific Research for sponsorship of my research. In order to be meaningful, scientific research must be conducted in an intellectually stimulating atmosphere: the Structural Metals Branch AFWAL/MLLS provided this environment. I am particularly grateful to Dr. C.F. Froes for accepting me into his group and for many stimulating discussions.

I. Introduction

My research interests in my 22 years on the MIT faculty have focused on the thermodynamics and kinetics of phase transformations in alloys. An excellent match was therefore obtained with the Structural Metals Branch, MLLS to work on two topics of great interest to them and myself: terminal solid solubility extension by rapid solidification and dispersoid formation and stability in advanced alloys.

Many desirable effects may be obtained from solute which is left in solid solution after solidification, including increases in strength and modulus, decreased alloy density, and avoidance of undesirable second phases. The amount of solute which can be retained in solid solution is thus a basic limiting factor in determining a number of important properties of alloys (1,2).

Modern high temperature materials are expected to serve at temperatures well over half the absolute melting point, where self diffusion in alloys is rapid, and non-equilibrium structures tend to quickly relax to near equilibrium. One microstructural modification which is often effective at high temperatures is the introduction of a fine distribution of insoluble dispersoid particles (3).

This report addresses terminal solid solubility extension under rapid solidification and in-situ dispersoid formation and subsequent dispersoid behavior during processing and service.

II. Objectives of the Research Effort

The overall objective of my research was to better understand the kinetic processes involved in the processing and service behavior of advanced alloys. The first sub-objective was to determine the circumstances which would produce the very desirable condition of

terminal solid solubility extension in rapid solidification. The second sub-objective was to develop a theory capable of predicting conditions giving a high number density of thermally stable dispersoid particles in alloys for high temperature applications.

III. Terminal Solid Solubility Extension by Rapid Solidification (TSSE by RS)

The importance of TSSE is so great that inevitably there have been several criteria proposed to predict the conditions of cooling and alloy composition which will give solubility extension. These criteria may be divided into three broad groups: kinetic, thermodynamic, and crystal chemistry.

Hornbogen (4) proposed that massive solidification (solidification without composition change) will occur when the advancing solid:liquid interface moves too rapidly for a concentration gradient to form in the liquid. Just below the melting point, T_m, the solidification interface moves very slowly, whereas diffusion in the liquid is very fast, so a concentration gradient forms very easily, preventing massive solidification. The diffusion coefficient in the liquid decreases rapidly with decreasing temperature, whereas the interface velocity at first increases due to an increasing thermodyanmic driving force. If, at some temperature, the interface moves faster than the concentration gradient, massive solidification is predicted to occur.

Buschow (5) related the temperature of crystallization of amorphous alloys to the enthalpy to form a hole in the glassy phase the size of the smaller atom. Vacancies in liquids or glassy alloys are thought to have a near-zero enthalpy of migration (6) so the diffusion coefficient is determined by the hole concentration. A

semi-empirical relation was in fact found to connect T_x, the temperature of crystallization and ΔH_h, the enthalpy of hole formation

$$T_x = 7.5\Delta H_h \qquad (1)$$

where T_x is in Kelvins and ΔH_h is kJ/mol holes.

A plot of T_x vs. ΔH_h for a number of amorphous alloys fits Eq. (1) very well, with some deviation toward slightly lower crystallization temperatures than predicted. The fit is good enough to conclude that the kinetic criterion is adequate to predict the temperature of crystallization of amorphous alloys.

Massive solidification above T_o requires that solutes <u>not</u> diffuse ahead of the solid liquid interface, so that ΔH_h and TSSE may somehow be related. This possibility is being studied.

Anantharaman, et al. (7) presented a review of rapidly solidified Al-based alloys in which they proposed that solutes with favorable Hume-Rothery size factors (within 15% of the solvent) should be amenable to TSSE. They noted, however, that 6 of the 17 such solutes (Au, Sn, Ti, V, Zn, and Zr had TSSE to 4 a/o or less. They sought a basis for this anomaly in thermodynamic considerations, as reflected in the phase diagram.

They noted for simple eutectic systems showing small deviations from ideal solution behavior, it should be easy to supercool the liquid below T_o and obtain massive solidification. For solids with large positive deviations from ideality, however, they noted that even at 0 K the free energy of the solid might be greater than that of the liquid, rendering massive solidification and TSSE thermodynamically impossible. Anantharaman et al. also argued that in some cases

TSSE will be limited by nucleation of an adjacent stable or metastable solid phase.

Several criteria to predict equilibrium solid solubility have been devised in terms of such elemental crystal properties as atomic radius, electronegativity, electron potential, or electron density at the boundary of the Wigner-Seitz cell. These criteria have some predictive power at equilibrium, and it is reasonable that they might be useful in predicting non-equilibrium solid solubility under rapid solidification.

Jones (8) analyzed equilibrium solid solubility in Al-based and Mg-based alloys in terms of the Wigner-Seitz atomic radius and the heat of solution of the solute in the matrix. In both cases Jones found a good division between soluble and insoluble atoms, with the more soluble clustered around the host atom in a plot of heat of mixing vs. atomic radius.

Jones then applied the coordinates to non-equilibrium solidification in Al-based alloys, and he found that elements with significant TSSE tended to be clustered around the host Al in the plot. The separation, however, was far from perfect. Some low solubility elements were located near the Al host and some high solubility elements were located well away.

This lack of either thermodynamic measurements or an accurate model has for years impeded theoretical studies of such kinetic processes as nucleation, growth, spinodal decomposition, and coarsening, as well as rapid solidification. A "macroscopic atom" model developed over the past decade by Miedema and his co-workers (9,10) provided for the first time an accurate model based on atomic properties only which may also be applied to solubility questions.

Miedema's approach gives mixing enthalpies of binary liquids or enthalpies of formation of binary compounds to an accuracy comparable to that of calorimetry. Mixing enthalpies of solid solutions are obtained to a lower accuracy.

Miedema's "macroscopic atom" approach depicts elemental A and B to be separated into Wigner-Seitz atomic cells, and then reassembled into the liquid or solid solution or intermetallic compound.

In the case of liquid solutions or intermetallic compounds the enthalpy of mixing is given by

$$\Delta H_{\infty} = P(\Delta\phi^*)^2 + Q(\Delta n_{ws}^{1/3})^2 - R + \Delta H^{trans} \qquad (2)$$

where

$\Delta\phi^*$ = difference in electron potentials of the elements

Δn_{ws} = difference between electron densities of the elements at the Wigner-Seitz cell boundary.

Unequal electron densities at the Wigner-Seitz boundary (n_{ws}) give a positive contribution to the enthalpy, and unequal electron potentials (ϕ^*) cause charge transfer and a negative enthalpy contribution. The constant R is from electron hybridization which occurs when d-valence electrons of transition metals hybridize with s or p valence electrons of polyvalent non-transition elements. The calculation works only for metallic elements. ΔH^{trans} is the enthalpy required to transform a non-metallic component (e.g. silicon) to a metal.

Chelikowsky (11) compared the effectiveness of Darken-Gurry and slightly modified Miedema plots as predictors of equilibrium solubility in divalent hosts. He found the Miedema coordinates (ϕ^*, $n_{ws}^{1/3}$) systematically superior, even for Mg, where the Darken-Gurry coordinates are most successful.

Chelikowsky also found that the Miedema coordinates were remarkably successful in predicting non-equilibrium site preference under ion implantation into beryllium. The injected elements divided without exception into three regions separated by smooth borders, depending on whether the injected ion occupied substitutional, tetrahedral, or octahedral sites.

The success of the Miedema plot in predicting equilibium solubility, crystallization temperature, and site selection in ion implanted alloys led me to extend it to TSSE in Al-base alloys. The ability of the Miedema coordinates to separate the results into regions of high, moderate, and near-zero TSSE is striking (12) and far superior to any other scheme known. The separation is perfect, except that cerium with an extended solubility of 1.9 a/o is included with solutes in the 2-10 a/o range. Furthermore the boundaries of the solubility regions are convex and smoothly curved, unlike the star-shaped boundaries found using other coordinates.

Why ϕ^* and $n_{WS}^{1/3}$ should have this predictive power for TSSE and what the functional relationship is, is not known. The predictive power of this criterion is, however, clearly great. Similar plots will be produced presently for TSSE in alloys based on titanium and magnesium, where extensive data bases are also available.

A more detailed analysis of criteria for TSSE under RS will be published separately.

IV. Dispersoid Formation and Thermal Stability

Let us consider the conditions for precipitation of a few volume percent of dispersoid AB in a solid alloy. The first requirement is that the molten alloy, M, dissolve A and B in amounts needed to form the desired amount of dispersoid. At the same time neither A nor B may form a compound with the constituents of M nor may they react with one another.

The tendency toward compound formation in the liquid is primarily determined by ΔH_{net}, the enthalpy change on forming the compound from elements in solution.

$$\Delta H_{net} = \Delta H_f - \Delta \bar{H}_A - \Delta \bar{H}_B , \qquad (3)$$

where ΔH_f is the enthalphy change on forming AB from elemental A and B, and where ΔH_A and ΔH_B are the partial molar heats of solution of A and B in M, respectively.

A large, negative ΔH_{net} will tend to cause AB formation in the liquid, producing a coarse, heterogeneously distributed dispersoid of no benefit in the alloy. The only way to prevent formation in the liquid of a dispersoid with large, negative ΔH_f is for $\Delta H_A + \Delta H_B$ in the liquid to be negative and fairly large. At the same time, neither ΔH_A nor ΔH_B may be so negative as to give a reaction of A or B with one or more of the constituents of M.

Formation of a fine, uniform distribution of dispersoids in the solid requires a high homogeneous nucleation rate, which in turn requires a high enthalpy of compound formation. The relevant

enthalpy is given by Eq. (3), except ΔH_A and ΔH_B now refer to the enthalpies of solution as A and B in solid M. We earlier required the enthalpies of solution of A and B in liquid M to be significantly negative, and now we want the same elements to have positive, or at least less negative enthalpies of mixing in the solid. If A or B, (or preferably both) do not fit well into the solid lattice, ΔH_A and ΔH_B in the solid may be substantially less negative than in the liquid. Then ΔH_{net} will then be considerably more negative for AB formation in the solid than in the liquid.

These restrictions on ΔH_f and ΔH_A and ΔH_B in the liquid and solid alloy make for a narrow thermodynamic window for producing a thermally stable in-situ dispersoid.

B. Thermal Stability

Dispersoids formed in-situ will usually be produced in rapidly solidified (RS) alloys which typically solidify in the form of fine powders, flakes, thin strip, or ribbon. The dispersoid particles thus must be stable against coarsening not only during service, but during hot isostatic pressing (HIP) and extrusion operations during fabrication.

Bhattacharyya and Russell (13) extended earlier analyses for elemental particles to describe coarsening of compound AB precipitates in a metallic matrix. Their expression is:

$$\bar{r}^3 - \bar{r}_o^3 = Kt \qquad (4)$$

where $K \approx 4 \times 10^{-15} \langle DC \rangle$ m^3/s and

$$\langle DC \rangle = \frac{D_A C_A^e D_B C_B^e}{D_A C_A^e + D_B C_B^e} \quad ; \qquad (5)$$

where D_A and D_B are the diffusivities of A and B atoms in the metallic matrix, and C_A^e and C_B^e are solubilities of A and B in the matrix in equilibrium with bulk AB phase. The quantity <DC> is thus equal to the smaller of $D_A C^e_A$ and $D_B C^e_B$, except for rare cases where the two are nearly equal.

The coarsening rate is minimized (at a given temperature) by adding a large excess of the faster diffusing element (say A). This excess A greatly depresses the concentration of the slower diffusing B atoms, and gives the lowest possible value of <DC> and of the coarsening rate.

Particles are sometimes located on grain boundaries, which are usually high diffusivity paths. The rate law for coarsening of particles on grain boundaries may be developed in a straightforward manner from the volume-diffusion case to give:

$$\bar{r}^4 - \bar{r}_0^4 \cong 4 \times 10^{15} \langle D_b \delta C_b \rangle \, t \; m^3/s \qquad (6)$$

where δ = grain boundary thickness and $\langle D_b C_b \rangle$ is given by a relationship equivalent to Eq. (5), except written in terms of boundary concentrations and diffusivities. Detailed calculations will be published separately. Ardell (14) derived an equivalent, but more complicated expression for grain boundary coarsening of elemental precipitates.

Comparison of equations (4,6) shows that boundary diffusion will

control the coarsening rate at smaller particle sizes and volume diffusion at larger. Boundary diffusion controls at small sizes because solute can add onto a larger fraction of the particle surface.

Coarsening may also be accelerated by enhanced diffusion along dislocation lines. Take a dislocation density of ρ_p lines/m^2 intersecting the particle surface, each line with a cross section of A_p m^2, through which diffusion may occur. The coarsening equation is then:

$$\bar{r}^3 - \bar{r}_0^3 = 4 \times 10^{-15} \langle D_p C_p \rangle A_p \rho_p t \ m^3/s \qquad (7)$$

where D_p = diffusivity of solute in the dislocation, C_p = solute concentration in the dislocations.

Eq. (7) is the same as Eq. (4) for volume diffusion controlled coarsening, except that $\langle DC \rangle$ is replaced by $\langle D_p C_p \rangle A_p \rho_p$.

C. Particle Dragging by Boundaries

Dispersoid particles are sometimes observed on only one side of grain boundaries after high temperature deformation. Such denuded zones suggest that particles were dragged along during deformation by the moving grain boundaries.

This analysis considered particle migration by dissolution of material from the rear of the particle, diffusion through the particle:matrix interface, and deposition onto the particle front.

When typical values of atomic volume and interfacial energy are inserted into the final equation it is found that the particles may move at a maximum velocity of:

$$V_{max} = \frac{27 \times 10^{-24} D_b C^e_b}{4\pi r^3} \quad m/s \qquad (8)$$

If the boundary moves at a velocity $V > V_{max}$, the particle cannot follow and is left behind, as is often observed. Eq. (8) shows that small particles are more easily dragged. Details of the calculation will be published separately.

D. Comparison of Theory with Experiment

We will restrict ourselves to dispersoid coarsening in titanium-based and Ti_3Al-based alloys because these systems are of particular current interest.

Whang (15) studied the coarsening of Ti_5Si_3 particles in a Ti-Al-Si matrix and of La_2Sn particles in a Ti-Sn-La matrix, and found that at 1073K, La_2Sn coarsened 300 times slower than did Ti_5Si_3. Whang, Lu, and Giessen (16) had estimated from coarsening data that the Si diffusion coefficient in α-Ti was about an order of magnitude greater than the self diffusion coefficient in α-Ti. Whang attributed the high coarsening rate of Ti_5Si_3, to a high Si diffusivity in α-Ti. However, at 1073K La_2Sn has a 300 times lower coarsening rate than does Ti_5Si_3. Unless La and Sn are anomalously slow diffusers in α-Ti, it is more likely that the decreased coarsening rate is due in large part to La_2Sn being less soluble than Ti_5Si_3 (see Eqs. 4,5).

Konitzer and Fraser (17) observed that a distribution of small Er_2O_3 particles in a Ti_3Al matrix did not coarsen in 10 hr at 1073 K or 1173 K. Appreciable coarsening did occur in 10 hr at 1273 K, with the particles going from about 10 nm to 50 nm diameter.

Sutliff and Rowe (18) studied alloys of Ti_3Al and 0.5 a/o Er to which additions of 0,5,6 and 7.5 a/o Nb had been made. Particles in the grain interior did not coarsen during aging of the 5 a/o Nb ribbons at 1173 K for one hour, an observation consistent with the results of Konitzer and Fraser. Particles at or near the grain boundaries coarsened, and a nearby particle-depleted zone was observed . This observation is consistent with enhanced coarsening by grain boundary diffusion. The denuded zones are probably due to larger, more stable particles on the grain boundary consuming smaller, less stable particles in the matrix.

HIP at 1173 K did not give particle coarsening. However, alloys HIP at 1123 K and extruded at 1143 K showed marked coarsening. Clearly something occurred during extrusion which promoted coarsening so powerfully as to overcome the higher temperatures and longer times available for coarsening during HIP. The presence of almost all the particles on grain boundaried indicated that the enhanced coarsening was due to a combination of grain boundary dragging and boundary diffusion enhanced coarsening.

Rowe, Sutliff, and Koch (19) studied the coarsening of dispersoids in Ti_3Al-based alloys to which 0,5,6,7.5 and 10.5 a/o Nb had been added. Erbium was added to some of the alloys to produce an Er_2O_3 dispersoid while in others cerium and sulfur additions produced a Ce_2S_3 dispersoid.

HIP of the alloys at 1123 K gave severe coarsening of the Er_2O_3 particles located on or near grain boundaries. Some grain boundaries showed dispersoid coarsening on one side but not the other, an effect attributed to boundary dragging. The Er_2O_3 particles in the study were less than 50 nm diameter, and according

to Eq. (8), could be dragged by moving grain boundaries.

Extrusion of the HIP alloys at temperatures near 1000 K caused severe coarsening to an average Er_2O_3 particle diameter approaching 0.5 μm. The authors noted that adiabatic heating may have heated the alloy above the estimated extrusion temperature. The coarsening constant increases rapidly with temperature, so a temperature excursion would give greatly increased coarsening.

The Ca_2B_5 dispersoid showed much less coarsening than did the Er_2O_3, under identical HIP and extrusion conditions. The lower coarsening rate was attributed to a lower solubility.

Denuded regions and large particles near grain boundaries were attributed to particle dragging and enhanced coarsening by boundary diffusion.

V. Conclusions

1. The Miedema coordinates successfully predict TSSE in RS Al-based alloys.
2. Conditions were derived for in-situ formation of thermally stable dispersoids.
3. Equations were derived for coarsening of dispersoids by boundary and dislocation pipe diffusion and for dragging by grain boundaries.

V. Recommendations

1. Computer programs should be written so that the Miedema model may be used further in modeling of TSSE under RS, and of dispersoid formation and thermal stability.
2. The Miedema criteria for TSSE under RS should be applied to other alloy systems.
3. Adiabatic solidification and absolute interface stability criteria for TSSE under RS should be tested.
4. The equations for particle coarsening and dragging should be tested quantitatively on α-Ti, Ti_3Al, Al, and Mg-based alloys.
5. A theory should be developed for predicting grain size in rapidly solidified alloys.

References

1. Jones, H., "Rapid Solidification of Metals and Alloys," The Institution of Metallurgists, London, 1982.
2. Froes, F.H., Y-W. Kim, and F.J. Hehmann, "Rapid Solidification of the Light Metals," Journal of Metals (1986, in press).
3. Ansell, G.S. (ed.), Proc. Second Bolton Landing Conference on Oxide Dispersion Strengthening, New York, Gordon & Breach, 1968.
4. Hornbogen, E., "The Origin of Microstructures of Rapidly Solidified Alloys," Z. Metallk. **77** (1986), 306-311.
5. Buschow, K.H.J., "Stability of Amorphous Alloys," Solid State Comm., **43** (1982), 171-174.
6. Faber, T.H., Introduction to the Theory of Liquid Metals, Cambridge, Cambridge University Press, 1972.
7. Arantharaman, T.R., P. Ramachanrardo, C. Suryanarayanan, S. Lele, and K. Chattopadhyay, "Structure and Constitution of Rapidly Solidified Aluminum Alloys, IIM Review No. 6," Trans. Indian Inst. of Metals **30**, no. 6 (Dec. 1977).
8. Jones, H., "Extent of Solid Solubility in Magnesium and Aluminum," Mater. Sci. & Eng. **57** (1983), L5-L8.
9. Miedema, A.R., P.F. de Chatel, and F.R. de Boer, "Cohesion in Alloys - Fundamentals of a Semi-Empirical Model," Physica **100B** (1980), 1-28.

10. Niessen, A.K., F.R. de Boer, R. Boom, P.F de Chatel, W.C.M. Mattens, and A.R. Miedema, "Model Predictions for the Enthalpy of Formation of Transition Metal Alloys II," Calphad **7** (1983), 1-70.

11. Chelikowsky, J.R., "Solid Solubilities in Divalent Alloys," Phys. Rev. B **19** (1979), 686-701.

12. Russell, K.C., Y-W. Kim and C.F. Froes, "Terminal Solid Solubility Extension Under Rapid Solidification," Journal of Metals (1987, in preparation).

13. Bhattacharyya, S.K. and K.C. Russell, "Activation Energies for the Coarsening of Compound Precipitates," Met. Trans. **3** (1972), 2195-2199.

14. Ardell, A.J. On the Coarsening of Grain Boundary Precipitates," Acta Met. **20** (1972) 601-609.

15. Whang, S.H., "Rapidly Solidified Ti Alloys for High Temperature Applications," J. Mat. Sci. **21** (1986) 2224-2238.

16. Whang, S.H., Y.Z. Lu, and B.C. Giessen, "Thermal Stability of Precipitates in a Rapidly Quenched Ti-Al-Si Alloy," in Rapidly Solidified Materials, ed. B.H. Kear and B.C. Giessen, New York, Elsevier Publishing (1984) 367-373.

17. Konitzer, D.G. and H.L. Fraser, "The Production and Thermal Stability of a Refined Dispersion of Er_2O_3 in Ti_3Al Using Rapid Solidification Processing," in Plasma Processing and Synthesis of Materials, ed. J.Szekely and D. Apelian, New

York, Elsevier Publishing (1985), 437-442.

18. Sutliff, J.A. and R.G. Rowe, "Rare Earth Oxide Dispersoid Stability and Microstructural Effects in Rapidly Solidified Ti_3Al and Ti_3Al-Nb" (1986, to be published).

19. Rowe, R.G., J.A. Sutliff, and E.F. Koch, "Comparison of Melt Spun and Consolidated Ti_3Al-Nb Alloys With and Without a Dispersoid," in Rapid Solidification Technology for Titanium Alloys, ed. F.H. Froes, et al., Warrendale, PA, TMS-AIME (1986, in press).

1986 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT SUMMER SUPPORT PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

COMPUTER MODELING OF INFRARED SIGNATURES

Prepared by:	Sally A. Sage
Academic Rank:	Assistant Professor
Department and University:	Mathematics and Computer Science Department West Georgia College
Research Location:	Air Force Armament Laboratory, Eglin AFB Division: AS Branch: ASE, IR Technology
USAF Research Contact:	Dr. Steve Butler
Date:	August 27, 1986
Contract No:	F49620-85-C-0013

COMPUTER MODELING OF INFRARED SIGNATURES

by

Sally A. Sage

ABSTRACT

This report describes a computer model for infrared radiation (IR) signatures. The goal of the system is to use a data file describing an air target to produce a computer-generated IR image that corresponds closely to an actual IR photograph of that target. The IR image provides information which can then be used to calculate the IR signature of the target. The target is modeled by a set of three-dimensional triangular facets which covers its outer surface. The target's plume is also based on a facet model, but the plume is generated interactively so that a user can control the details of the plume's construction. The three-dimensional coordinates of the target and plume are rotated using user-specified sensor angles. The coordinates are then projected into the YZ plane to produce a two-dimensional image. The image is composed of facets in the sensor's line-of-sight. These facets are assigned temperatures and then the area of each set of facets which corresponds to a particular temperature is calculated. The IR target signature can then be computed by the integration of the radiance over the projected area of the target along the line-of-sight of the sensor.

Acknowledgments

I would like to express my appreciation to the Air Force Systems Command, the Air Force Office of Scientific Research, and the Air Force Armament Laboratory for the opportunity to work in the Summer Faculty Research Program at Eglin AFB, Florida. I am especially grateful to Dr. Steve Butler and Captain Lawrence Jones for giving me the opportunity and the guidance necessary for my research.

I. Introduction

An important concern presently facing the aerospace industry is the design of aircraft which can avoid detection. Considerable effort has been directed to radar and visible signatures, but accurate models of infrared radiation (IR) signatures are still needed to provide a detailed analysis to aid designers in minimizing IR emissions. This research effort at the Eglin AFB Armament Laboratory is a continuation of a project developed last summer and includes the development of an integrated system of computer programs which model an aircraft's IR signature. The input to the system is based on a multifacet model which is available for current aircraft. Also included is a similar plume facet model that is generated geometrically.

Sally A. Sage received the B.A. degree in Mathematics from San Francisco State University in 1977 and the M.S. degree in Computer Science from the University of Pittsburgh in 1979. She has taught at West Georgia College in Carrollton, Georgia for six years and in the fall will begin teaching at Southern Technical Institute in Marietta, Georgia. Her research interests include image processing, computer graphics, and optimization techniques.

II. Objectives of the Research Effort

The goal of this research is to develop an integrated IR signature modeling system. The most significant IR emission of an aircraft is the plume which is formed behind the aircraft by the hot exhaust of the engines. Thus, a considerable amount of effort has been devoted to improving the plume model. My individual objectives were to:

(1) set up a user-friendly, interative plume generator

(2) implement a plume model that is more accurate than the model developed last summer

(3) assign temperatures and calculate corresponding areas so the IR signature can be computed

(4) integrate the separate parts of the system into a complete, interative package for the IR model

III. Overview of the Target Modeling System

Data files, representing a particular air target and its plume, are input to the system and the system generates an IR display image corresponding to the target as output. Each pixel in the image has a correspondence to a facet that is directly in the line-of-sight of the sensor. Temperatures are assigned to each facet and the area corresponding to each temperature is computed. Although the final calculations have not yet been implemented, this area will be used to calculate the infrared signature of the target by integrating the radiance over the projected area of the target along the line-of-sight of the sensor.

The target data is modeled by a set of triangular facets which completely cover the outside of the target. An interactive system generates a plume data file that can be attached to the target. These plume generation routines will be discussed in more detail in Section IV. In addition to the target data, parameters such as the IR sensor's azimuth and elevation are input to the system.

The IR model developed last summer consisted of three separate phases: (1) identify specific regions of the target,

(2) generate and attach a plume to the target (3) generate an image that corresponds to the sensor's line-of-sight. These separate routines have been integrated into an interactive package. (See Appendix A for a description of the menus.)

In addition, the data file format has been modified since last summer. The original data files contained a list of facets and coordinates but no information that related a particular facet to a target component (i.e. wing, nose, engine, etc.). This component information is available through other sources and is currently being added back into the data files by a contractor group. The codes are necessary to facilitate the assignment of temperatures to various components of the target.

The major data structures of the IR system include several arrays:

ICOMP -- A list of component identification codes and the corresponding facet indices which point into LLIST. The facet indices indicate the group of facets which correspond to each component.

TEMP -- A list of temperatures. TEMP is parallel to ICOMP; the identification code in ICOMP(1,i) has a corresponding temperature in TEMP(i).

LLIST -- A list of triangular facets describing the target. Each column corresponds to the three vertices of a facet and contains indices into array PNT. Note: a group of sequential columns describe a particular component.

PNT -- A list of three-dimensional coordinates. Each column contains the X,Y,Z coordinates of a facet vertex.

AREA -- A list of the areas corresponding to temperatures. The subscript corresponds to the temperature and the contents correspond to the calculated area.

FDIS -- A three-dimensional array that contains the generated image. Dimensions one and two correspond to the YZ plane. Dimension three corresponds to the transparent layers of the plume.

Figure 1 show a diagram of the arrays into which the input data

is read. The arrays contain both the target data and the plume data, but these data sets are kept separate (the target data is stored first, followed by the plume data). The data sets are stored externally in four files: the target data file, the target temperature file, the plume data file, and the plume temperature file. By convention, each temperature file has the same name as its corresponding data file but uses ".TMP" as a file extension.

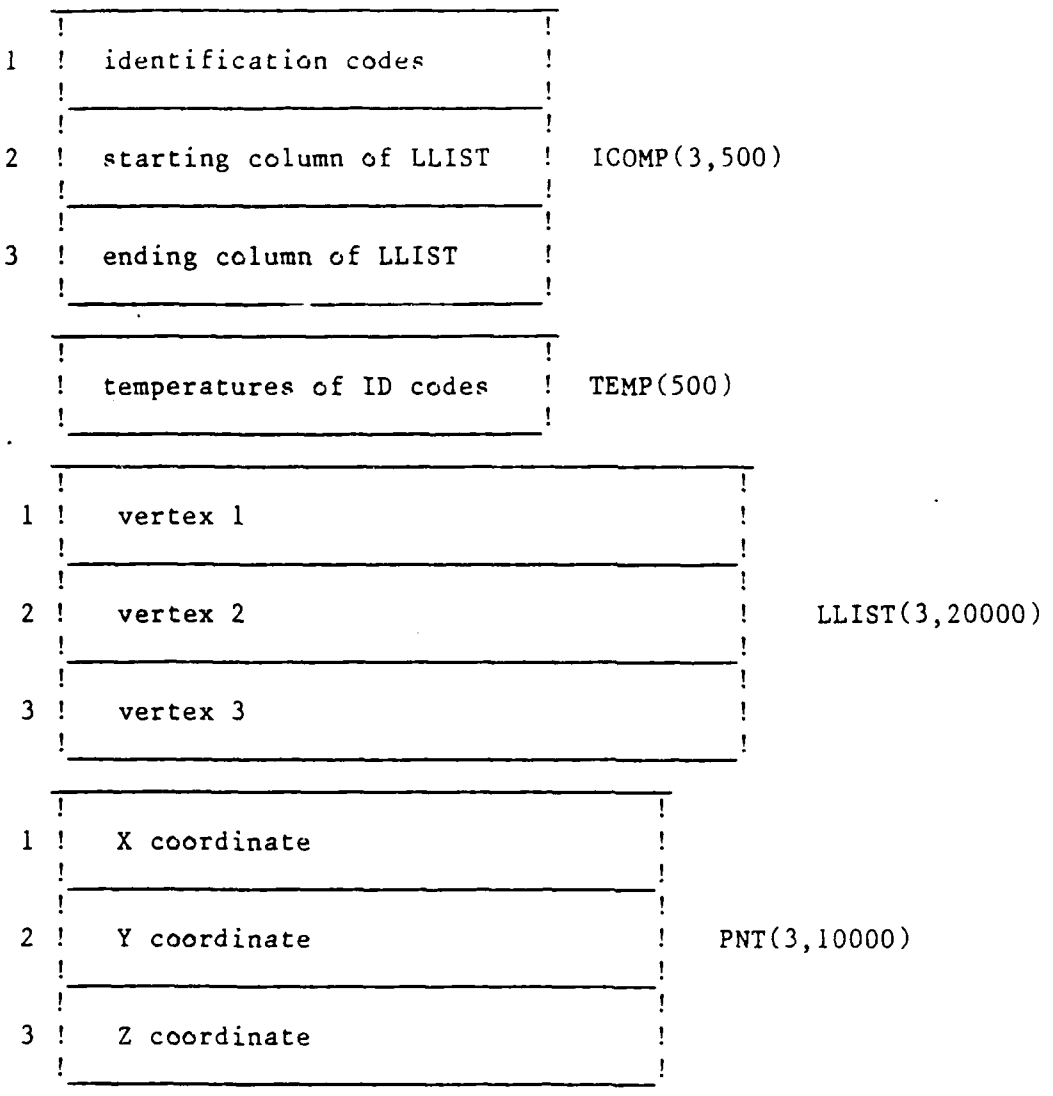


Figure 1. Data structures corresponding to data files

IV. Details of the Plume Implementation

The plume implemented last summer consisted of a single cone-like shape which was constructed geometrically by a sequence of disks and cones. The IR model, however, requires a plume model which is more sophisticated. A more accurate plume model consists of concentric layers of cone-like shapes; this corresponds more closely to the gaseous properties of the exhaust. The inner plume layers are assigned hotter temperatures while the outer layers are assigned cooler temperatures. Thus, the plume facets require an implementation which can simulate a "transparent" property. The image generated by the IR system developed last summer could display only those facets in the direct line-of-sight of the sensor. This is, only the outermost layer of the plume is visible, covered facets are not visible. These visible facets have cooler temperatures and do not contribute much to the plume signature, while the hotter facets, which are hidden, make a significant contribution to the signature. To keep a record of the hidden layers of the plume, the display image array (FDIS) has been changed from a two-dimensional array to a three-dimensional array. As the plume is generated from back to front, the currently visible facets, which are about to be covered, are saved. Later, when the IR signature is calculated, these transparent layers can be used to calculate areas which contribute to the overall radiance.

The plume is composed of cone-like shapes referred to as components. Each component consists of a group of truncated cones and disks. Each cone or disk is called a sub-component. The identification codes for the plume range from 6000 to 6999.

Each component is assigned to a unique 100's group and individual sub-components are numbered sequentially within their 100's group. For example, suppose a plume is composed of an inner plume and an outer plume. The inner plume is assigned the component group 6000 and the outer plume is assigned the component group 6100. Further, if the inner plume is constructed from two disks and one cone, then inner plume will have three sub-components: 6000, 6001, and 6002.

The plume generation routines are based extensively on a menu system so that plumes can be created and edited in an interactive environment. The routines include features such as adding, deleting and changing the disk and cone sub-components. The X, Y, Z coordinates of the sub-components are generated geometrically. To manipulate a disk, the user must know the X coordinate of the disk and the radius. To manipulate a cone, the user must know the starting X coordinate, the starting radius, the ending X coordinate and the ending radius.

The X, Y, Z coordinates of the plume are projected into the YZ plane to create a display image. Thus, the X coordinate is used for front to back positioning, while the Y and Z coordinates are related to the radius of a circle. The points generated to describe the disk and cone circles are centered around the Y and Z axes unless otherwise specified. If the user wishes to shift either the Y or Z coordinates, it is possible to enter the Y and Z offsets for individual components.

Further, the entire plume may be shifted along either the X, Y, or Z axis. This feature is currently used to attach the plume

to the target. The plume can be generated generically to fit several aircraft. By convention, the front X coordinate of the plume is zero and the X coordinates range from zero to some negative value. In most of the aircraft data files, the nose of the aircraft also has an X coordinate equal to zero and the tail is positioned at some negative value. When the generic plume is attached to a particular aircraft, it must be shifted so that it is positioned behind the engine.

V. Recommendations

1. A routine should be added to automatically scan the target data and determine the X,Y,Z offsets required to attach the plume behind the engine.

2. The routines used to assign temperatures to components are sufficient, but could be made more flexible.

3. The system has been extended so that temperatures can be assigned to the facets and the calculated area of each temperature is available. The next step is use this information to calculate the radiance to obtain the IR signature.

4. Additional parameters which can affect the IR sensor, such as speed and atmospheric conditions, should be added to the system.

5. The hidden-line algorithm used to sort the facets from back to front is not sufficiently accurate. A more sophisticated algorithm based on the concepts of partial ordering and topological sorting should be implemented.

Appendix A

```
IRSIG Main Menu
     E)dit
     D)isplay
     P)lume
     A)rea
     G)enerate Image
     Q)uit

PLUME Menu
     C)reate a plume component
          §Y,Z offset§
               D)isk
                    §X, radius§
               C)one
                    §X1, radius1, X2, radius2§
               L)ist
                    §ID code§
               E)xit
     L)ist a plume component
          §ID code§
     M)odify a plume component
          A)dd a sub-component
               §ID code, Y,Z offsets§
                    D)isk
                         §X, radius§
                    C)one
                         §X1, radius1, X2, radius2§
          D)elete a sub-component
               §ID code§
          C)hange a sub-component
               §ID code, X, radius§
          L)ist
               §ID code§
          E)xit
     T)emperature Assignment
          O)ne temperature for a range of ID codes
               §starting ID code, ending ID code, temperature§
          D)ifferent temperatures for a range of ID codes
               §starting ID code, ending ID code§
                    ID code => §temperature§
          I)ndividual ID codes, one temperature
               §ID code, temperature§
          R)andom list of ID codes, one temperature
               §temperature, list of ID codes§
          L)ist ID codes and temperatures
          E)xit
     S)hift
          §X,Y,Z offsets§
     W)rite a plume file
          §file name§
     E)xit Plume
```

The user may use the [BACKSPACE] key to return to a previous

menu. To select an option, the user types the first letter of the option sentence. This has been emphasized above by the placement of a right parenthesis (')') after the selection letter. Curly brackets ('§') are used to show the input that will be requested from the user.

1986 USAF-UES Summer Faculty Research Program/
Graduate Student Summer Support Program

Sponsored by the
Air Force Office of Scientific Research
Conducted by the
Universal Energy Systems, Inc.

FINAL REPORT

Swirling Flows in Dump Combustors

Prepared by:	Mo Samimy and Craig A. Langenfeld
Academic Rank:	Assistant Professor and Graduate Student
Department and University:	Mechanical Engineering The Ohio State University
USAF Researcher:	Dr. A. S. Nejad
Date:	September 24, 1986
Contract No:	F49620-85-C-0013

Swirling Flows in Dump Combustors

by

Mo Samimy and Craig A. Langenfeld
The Ohio State University
Department of Mechanical Engineering
206 W. 18th Street
Columbus, OH 43210

ABSTRACT

A series of experiments were conducted to study isothermal swirling flows in a dump combustor configuration. A two-component coincident LDV system was utilized for detail mean flow and turbulence measurements in the axial and tangential directions. To have optical access for two-component measurements and to minimize disturbances, a small flat window was used and the inlet to the combustor was moved relative to the combustor by using a novel traversing system. Two constant angle swirl generators with swirl numbers of 0.3 and 0.5 were tested. Only 0.5 swirler generated sufficient axial pressure gradient to produce central recirculation which extended approximately 4.5 times of the step height downstream of the expansion plane. The corner recirculating flows were present in both cases with a smaller recirculating region in the stronger swirler. Very large scale turbulence structure was measured in the central core of both flows; the structure was extremely large in the shear layer between the wake behind the hub and the main flow. The decay of large scale motion was very rapid in the stronger swirler flowfield. While 0.3 swirler flowfield became a single solid-body rotational flow after approximately two combustor diameters, the 0.5 swirler flowfield was combination of solid-body rotation at the central core and constant angle swirl flow outside of the central core even at four combustor diameters downstream.

I. INTRODUCTION

I received my Ph.D. from the University of Illinois at Urbana-Champaign studying applications of optical techniques in fluid mechanics especially flows with high mean velocity and turbulence level. Later as a Visiting Assistant Professor I studied flow-induced unsteady forces acting on the tube-launched missiles.

One of the research programs at the Air Force Wright Aeronautical Laboratories is focused on the experimental study of swirling flows in a dump combustor. The major thrust of the research effort is to obtain detailed mean flow and turbulence data using a two-component coincident laser Doppler velocimeter (LDV). Since I had an extensive experience with LDV, therefore, I was chosen and assigned to work on this project.

My student, Craig Langenfeld, did his undergraduate degree in Mechanical Engineering at the University of Dayton. He finished his degree April of 1986 and he started his Master of Science program at The Ohio State University this fall.

II. OBJECTIVES OF THE RESEARCH EFFORT

The overall objective of the first phase of the research was to design, fabricate, and set up a facility which could be used for a detailed experimental study of swirling flows in dump combustors. The objective for the second phase was to use the facility to study effects of influential parameters on the isothermal swirling flows in dump combustors and also to provide extensive data base for the numerical code development and testing.

The first phase of the research had been completed before the Summer Faculty Research Program. Therefore, our objective was to participate in the second phase of the program to perform the following task:

a. To conduct preliminary experiment and to check flow quality.

b. To use the facility to obtain detailed LDV data in two swirling flowfields.

c. To analyze and interpret obtained data.

III. GENERAL BACKGROUND

The significant effects of imposing a swirl component on a flow have been recognized for many years. Methods of generation, characterization, and applications of swirling flows have been discussed in detail by Beer and Chigier (1) and by Gupta, et al. (2). Swirling flows have significant impact on stability, mixing process, and control of combustors. As a result, swirling flows have been focus of many research efforts in the past few years (3,16). The design and characterization of swirlers with different swirl profile have been examined in detail by Buckley, et al. (3). The characteristics of swirlers have also been studied by Kilik (4).

Confined jet or dump combustor geometry has been studied in detail by Lilley and his co-workers (5). Extensive research effort has been focused on combustor geometry with co-annular swirling flows (6-14). Swirling free jet has also been the subject of some recent studies (15,16).

One can conclude from above referenced papers that even though our understanding of the swirling flows has been improved dramatically in the past few years, still a great deal of work is needed in this area. In Section V, the results of present experiments will be presented and compared with the results of earlier work.

IV. EXPERIMENTAL FACILITIES AND TECHNIQUES

The hot-wire anemometer and pressure probes have been utilized extensively in swirling flows (3,5,6,9,11). This type of flow is sensitive to any probing, there is a great ambiguity in flow direction, and the turbulence intensity is very high. Therefore, accuracy of conventional techniques is questionable. LDV technique could provide detail and accurate results if one

could overcome the problem of optical aberrations by the curved tube in a typical cylindrical combustor geometry (17).

A dump combustor configuration with a 102 mm inlet diameter and 152 mm combustor diameter was utilized in the present experiments. Two swirl generators of axial flow type with 12 curved inlet guide vanes were designed in house (3) and used in these experiments. The vanes of swirler are welded between a 102 mm i.d. outer ring and a 19 mm hub. The swirlers were constant angle swirl generators with 0.3 and 0.5 swirl numbers. The swirler was located approximately 50 mm upstream of the dump plane.

A unique feature of the experimental set up is that the inlet tube can be moved inside the combustor tube and as a result a small flat window of 38 x38 mm dimensions can be used for two-component coincident LDV measurements with a minimal disturbance to the flowfield. A two-component coincident LDV system was utilized for a detailed mean flow and turbulence measurements in the axial and tangential directions. Titanium dioxide particles of approximately 1 micron diameter were used to seed the flow. In order to reduce statistical uncertainty, 27300 samples were obtained in each location in the flowfield. The constant time sampling method and hold/sample method which uses particle inter-arrival time as a weighting factor were utilized to correct for velocity biasing. There was a good agreement between two methods, but because of larger sample size of the hold/sample method, the results produced by this method will be presented here.

The Reynolds number based on inlet diameter was approximately 120,000. The inlet center line velocity was monitored continuously and kept at 18.9 ± 0.4 m/s. In the following section, the incoming flow, the mean flow, and the turbulence field results will be presented and discussed.

V. EXPERIMENTAL RESULTS

Two-component LDV measurements have been carried out in one upstream location and 13 downstream locations up to 4 combustor diameters. Due to lack of time for appropriate presentation of the results and also due to limit of

number of pages of the report, only some selective results will be presented here.

a) Incoming Flow. The incoming flow data was obtained in a location approximately 75 mm upstream of the swirler. Fig. 1 shows the axial mean velocity, U and the tangential mean velocity, W, profiles. The mean velocities and also turbulence data are non-dimensionalized by the upstream center line velocity which is approximately 18.9 m/s. The axial mean velocity is a typical fully developed pipe flow, but the tangential mean velocity is not zero value perhaps due to the subsonic nature of the flow and the influence from swirler which is located only 75 mm downstream.

Fig. 2 shows turbulence intensities in the axial and tangential directions. The turbulence field is not isotropic, but the ratio of the axial turbulence intensity to the tangential turbulence intensity is approximately 1.25 in each radial location and the turbulence field is quite symmetric relative to the center line.

b) Mean Flow. Mean axial and tangential velocity profiles in selective streamwise locations are shown in Figs. 3-5. The measurements carried out from 0.06 combustor diameter to 4 combustor diameters downstream of the dump plane. The 0.3 swirler didn't generate sufficient axial pressure gradient to produce a central reversed flow. But, there was a central recirculation for 0.5 swirler which extended approximately 4.4 step heights downstream of the dump plane. This swirl induced reversed flow is reported by many researchers in cold flow dump combustor configuration (5), in multiple-stream swirling flows in both cold and hot flows (6-14), and in free jet flows (15,16). The corner recirculation zone extends approximately 4.3 step heights for 0.3 swirler and 3.2 step heights for 0.5 swirler. The shortening of corner recirculation with stronger swirler is consistent with other work (5). By two combustor diameters, the axial velocity for both swirling flows is almost uniform across the combustor diameter.

The tangential velocity shows a central core which behaves as a solid body rotation. While this region extends and covers the entire combustor

diameter by approximately 2 combustor diameters for 0.3 swirler, this region extends and then shrinks very gradually for 0.5 swirlers. At 4 combustor diameters downstream of step, still there is a strong central solid body rotation for 0.5 swirler. This result is consistent with findings of other researchers for stronger swirlers (5).

c) Turbulence Field. Axial and tangential turbulence intensity profiles are shown only at two streamwise locations in Figs. 6 and 7. As shown in Fig. 6, the axial turbulence identifies the extent of the central core and the locations of shear layers. The maximum axial turbulence intensity occurs in the shear layer between the central core and the main flow. Both the axial and tangenial turbulence intensities are higher for 0.5 swirler than 0.3 and the turbulence dissipation in streamwise direction is also higher for 0.5 swirler. Very high level of turbulence fluctuation in the central recirculation zone have been also reported for similar configurations (5) and for multi-stream flows (6-8).

At 12 step heights downstream of the dump plane, the axial turbulence intensity is almost uniform for both swirlers, but the turbulence level is higher for 0.3 swirler which is due to lower dissipation rate of turbulence. At the same location, the tangial turbulence intensity is approximately uniform across the combustor for 0.3 swirler, but shows large value for the central core of 0.5 swirler where the mean velocity shows a persistent solid body rotation region. The results at 24 step heights downstream still shows high levels of tangential turbulence intensity for 0.5 swirler. This behavior is also shown in other research work (5).

Shear stress and correlation coefficient profiles at two axial locations are shown in Figs. 8 and 9. Shear stress results show existence of very large scale motion in the shear layer between the core and the main flow in both 0.3 and 0.5 swirler flows and extremely fast decay of large scale structure within a short distance for 0.5 swirler. Experimental results show that it takes approximately 6 step heights for 0.5 swirler and approximately 10 step heights for 0.3 swirler to have almost uniform shear flow across the combustor. These distances correspond to the axial wall pressure relaxation which indicate a

strong correlation between pressure field and large scale structure. Lilley and his co-workers (5) used hot-wire for turbulence data in a similar type of configuration. Their results did show such a high level of shear stress for swirl number much higher than the present experiments but lower shear stress level for those comparable to the present swirl numbers.

Correlation coefficient results shown in Figs. 8 and 9 show very non-uniform turbulence structure in the flowfield. This non-uniformity is still strong at 24 step heights downstream of the dump plane.

Figs. 10 and 11 show distribution of turbulent triple products at 1 step height downstream of the dump plane. Turbulent triple products show the existence of large scale structure not only in the shear layer between the central core and the main flow, but also in the central core itself.

Fig. 10(a) shows that the axial component of turbulent kinetic energy is diffused toward the center line from both edges of the shear layer in the axial- tangential plane. This is similar to subsonic plane shear layers and shear layers formed by separation of boundary layer at a step, but it is opposite to supersonic shear layers (18,19). Fig. 10(b) indicates that the tangential component of turbulent kinetic energy is diffused in the direction opposite to mean tangential velocity direction.

Fig. 11 shows diffusion of turbulent kinetic energy in streamwise direction which is in the positive direction in the central core, but in the negative direction in the shear flows. The magnitude of turbulent kinetic energy diffusion is much smaller in the center line which causes very slow development of the center line mean velocity. Similar to shear stress profiles, the triple products decay much faster for 0.5 swirler than 0.3 swirler.

VI. CONCLUSIONS AND RECOMMENDATIONS

A two-component coincident LDV was used to obtain detailed mean flow and turbulence data in a dump combustor configuration with 0.3 and 0.5 swirl numbers constant angle swirl generator inlet. The data indicated domination

of large scale motion and enhanced mixing by the large scale structure specially for the stronger swirler. The obtained data is a good data base for computational code development and testing. The radial components mean and turbulence data would complement this data and would provide better understanding of the physics of this type of flowfield. From a practical view point, a similar experimental work with much higher Reynolds number would be very beneficial.

ACKNOWLEDGEMENTS

We would like to thank the Air Force Systems Command and the Air Force Office of Scientific Research for sponsoring this program. Any extensive experimental work can be accomplished only by a team of cooperative researchers. The swirl generators designed and fabricated by the experimental group in the Ramjet Technology Branch of APL/AFWAL. The LDV system was initially developed by Dr. Roger R. Craig. The experimental set up was designed and fabricated by Dr. A. S. Nejad. Mr. Kenneth G. Schwartzkopf was very helpful with data acquisition hardware and software. The support and guidance of Dr. A. S. Nejad during the research period are deeply appreciated.

REFERENCES

1. Beer, J. M. and Chigier, N. A., Combustion Aerodynamics, John Wiley and Sons, New York, 1972.

2. Gupta, A. K., Lilley, D. G., and Syred, N., Swirl Flows, Abacus Press, Turnbridge Wells, England, 1984.

3. Buckley, P. O., Craig, R. R., Davis, D. L., and Schwartzkopf, K. G., "The Design and Combustion Performance of Practical Swirlers for Integral Rocket/Ramjets," AIAA Journal Vol. 21, No. 5, May 1983, pp 733-740.

4. Kilik, E., "Better Swirl Generation Using Curved Vanes Swirlers," AIAA Paper-85-0187, 1985.

5. Lilley, D.G., "Swirling Flows in Typical Combustor Geometries," AIAA Paper-85-0184, 1985.

6. Vu, B. T. and Gouldin, F. C., "Flow Measurements in a Model Swirl Combustor," AIAA Journal , Vol. 20, No. 5, May 1982, pp. 642-651.

7. Gouldin, F. C. , Depsky, J. S., and Lee, S-OL., "Velocity Field Characteristics of a Swirling Flow Combustor," AIAA Journal, Vol. 23, No. 1, Jan. 1985, pp. 95-102.

8. Ramos, J. I. and Somer, H. T., "Swirling Flow in a Research Combustor," AIAA Journal, Vol. 23, No. 2, Feb. 1985, pp. 241-248.

9. Habib, M. A. and Whitelaw, J. H., "Velocity Characteristics of Confined Coaxial Jets With and Without Swirl," Journal of Fluids Engineering, Trans. of the ASME, Vol. 102, March 1980, pp. 47-53.

10. Brum, R. D., Seiler, E. T., LaRue, J. C., and Samuelsen, G. S., "Instantaneous Two-Component Laser Anemometry and Temperature Measurements in a Complex Flow Model Combustor," AIAA Paper-83-0334, 1983.

11. Mattingly, J. and Oates, G., "An Experimental Investigation of Co-Annular Swirling Flows," AIAA Paper-85-0186, 1985.

12. Brondum, D. C., Bennett, J. C., Weinberg, B. C., and McDonald, H., "Numerical and Experimental Investigation of Nonswirling and Swirling Confined Jets," AIAA Paper-86-0040, 1986.

13. Ramos, J. I., "Turbulent Nonreacting Swirling Flows," AIAA Journal, Vol. 22, No. 6, June 1984, pp. 846-848.

14. Kubo, I. and Gouldin, F. C., "Numerical Calculations of Turbulent Swirling Flow," Journal of Fluids Engineering, Trans. of the ASME, Sept. 1975, pp. 310-315.

15. Fujii, S., Eguchi, K., and Gomi, M., "Swirling Jet with and without Combustion," AIAA Journal, Vol. 19, No. 11, Nov. 1981, pp. 1438-1442.

16. Sislian, J. P. and Cusworth, R. A., "Measurements of Mean Velocity and Turbulent Intensities in a Free Isothermal Swirling Jet," AIAA Journal, Vol. 24, No. 2, Feb. 1986, pp. 303-309.

17. Durrett, R. P., Stevenson, W. H., and Thompson, H.D.,"Radial and Axial Turbulent Flow Measurements with an LDV in an Axisymmetric Sudden Expansion Air Flow," International Symposium on Laser Anemometry, ASME FED-Vol. 33, 1985.

18. Samimy, M., Petrie, H. L., and Addy, A. L., "A Study of Compressible Turbulent Reattaching Free Shear Layers, "AIAA Journal, Vol. 24, Feb. 1986, pp. 261-267.

19. Samimy, M. and Addy, A. L., "Interaction Between Two Compressible, Turbulent Free Shear Layers," to appear in Nov. 1986 issue of AIAA Journal.

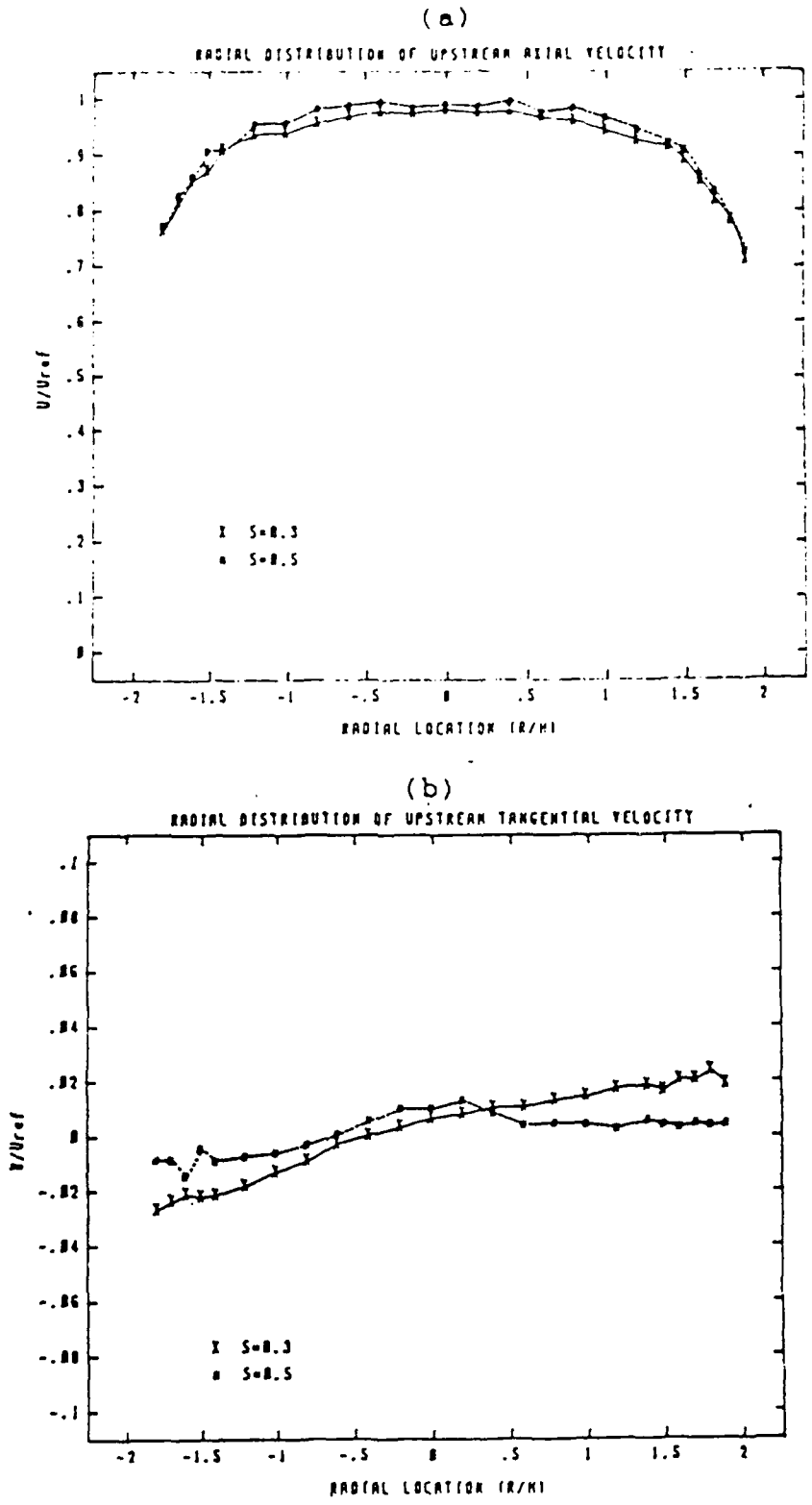


Fig. 1 Incoming axial and tangential mean velocities. (Approximately 75 mm upstream of swirler)

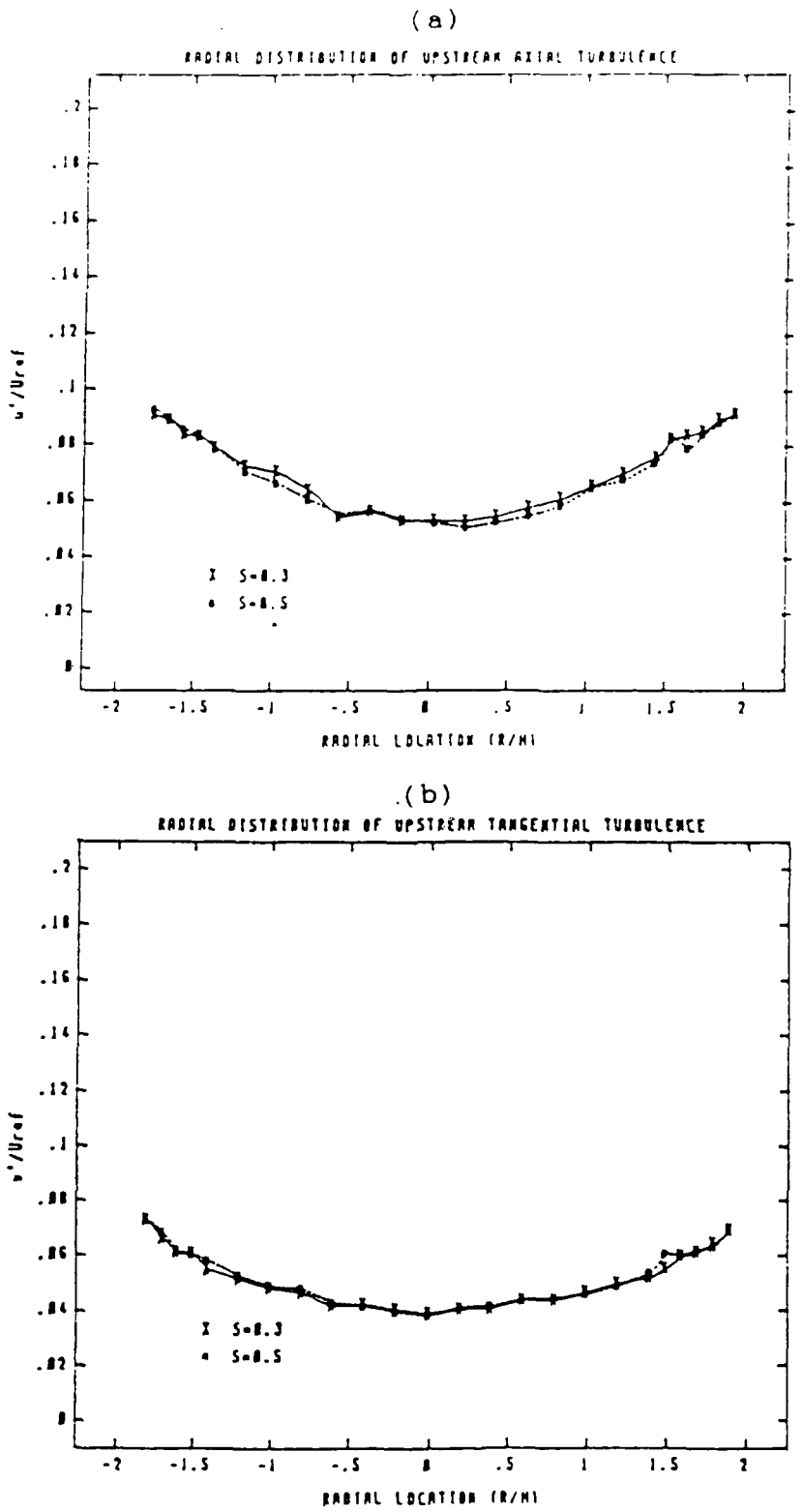


Fig. 2 Incoming axial and tangential turbulence intensities. (u' and w' in Figs. 2,6, and 7 are standard deviations of fluctuations)

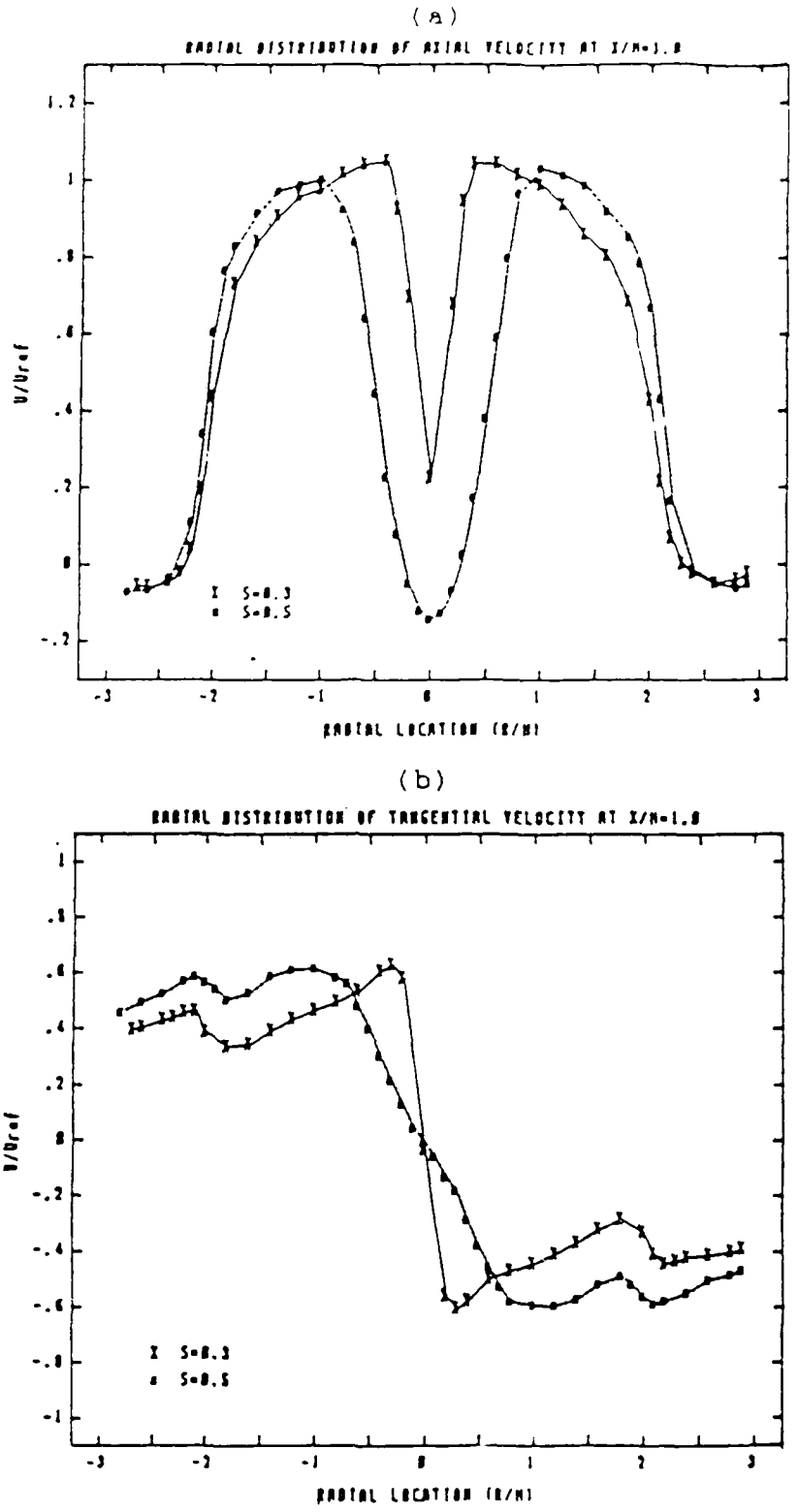


Fig. 3 Axial and tangential mean velocities at x/h=1. (h is step height; 25.4 mm)

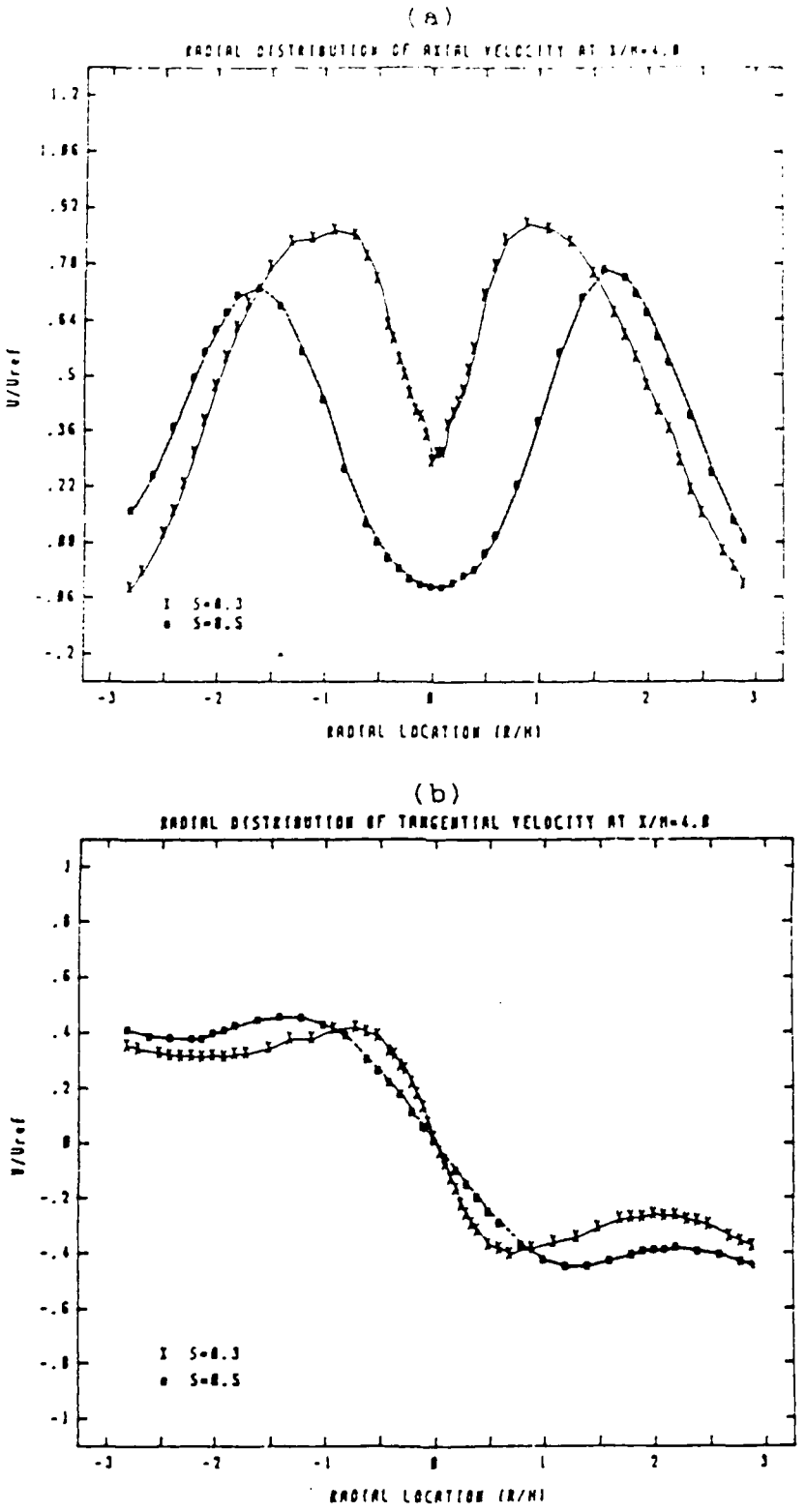


Fig. 4 Axial and tangential mean velocities at x/h=4.

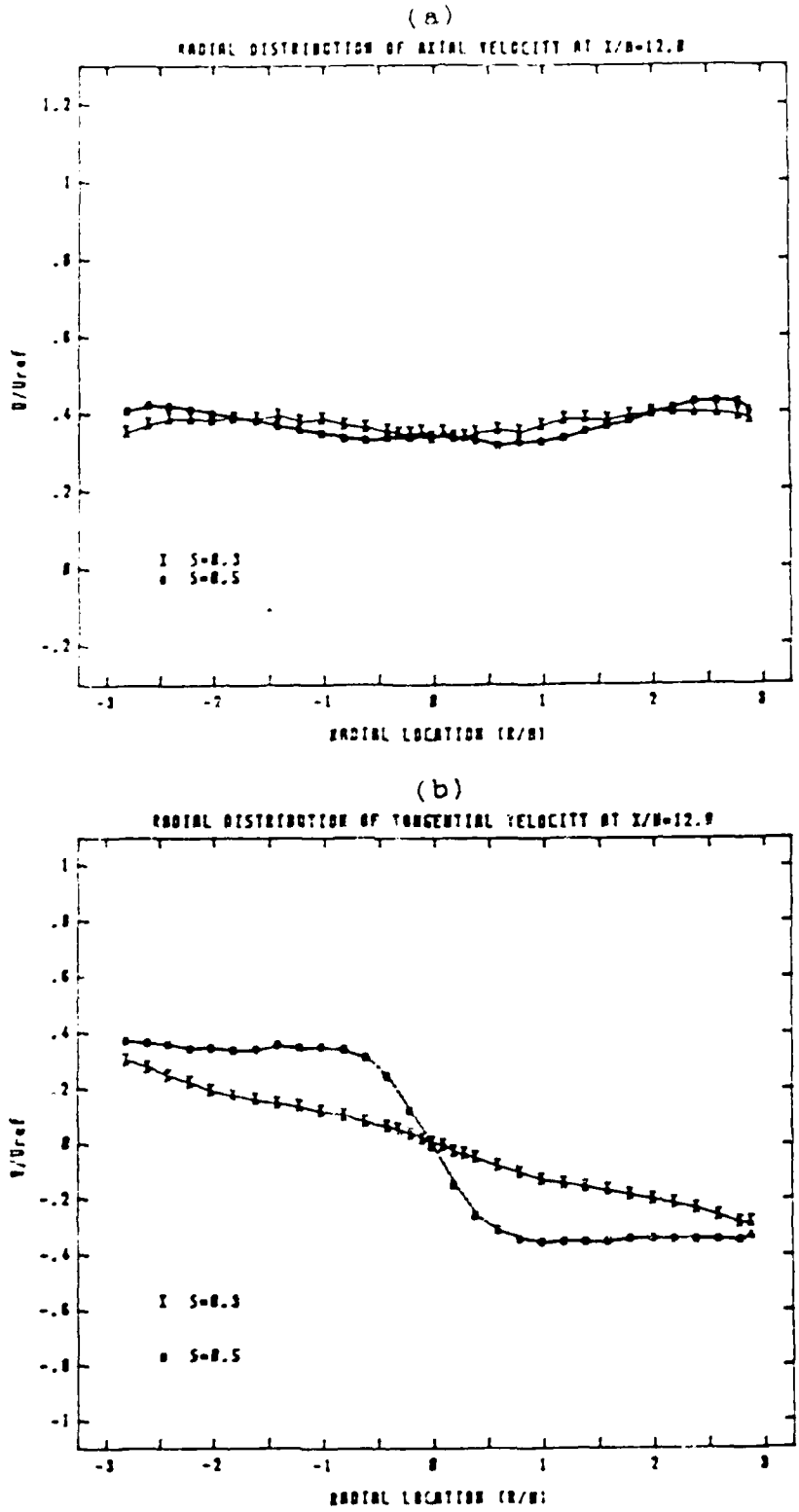


Fig. 5 Axial and tangential mean velocities at x/h=12.

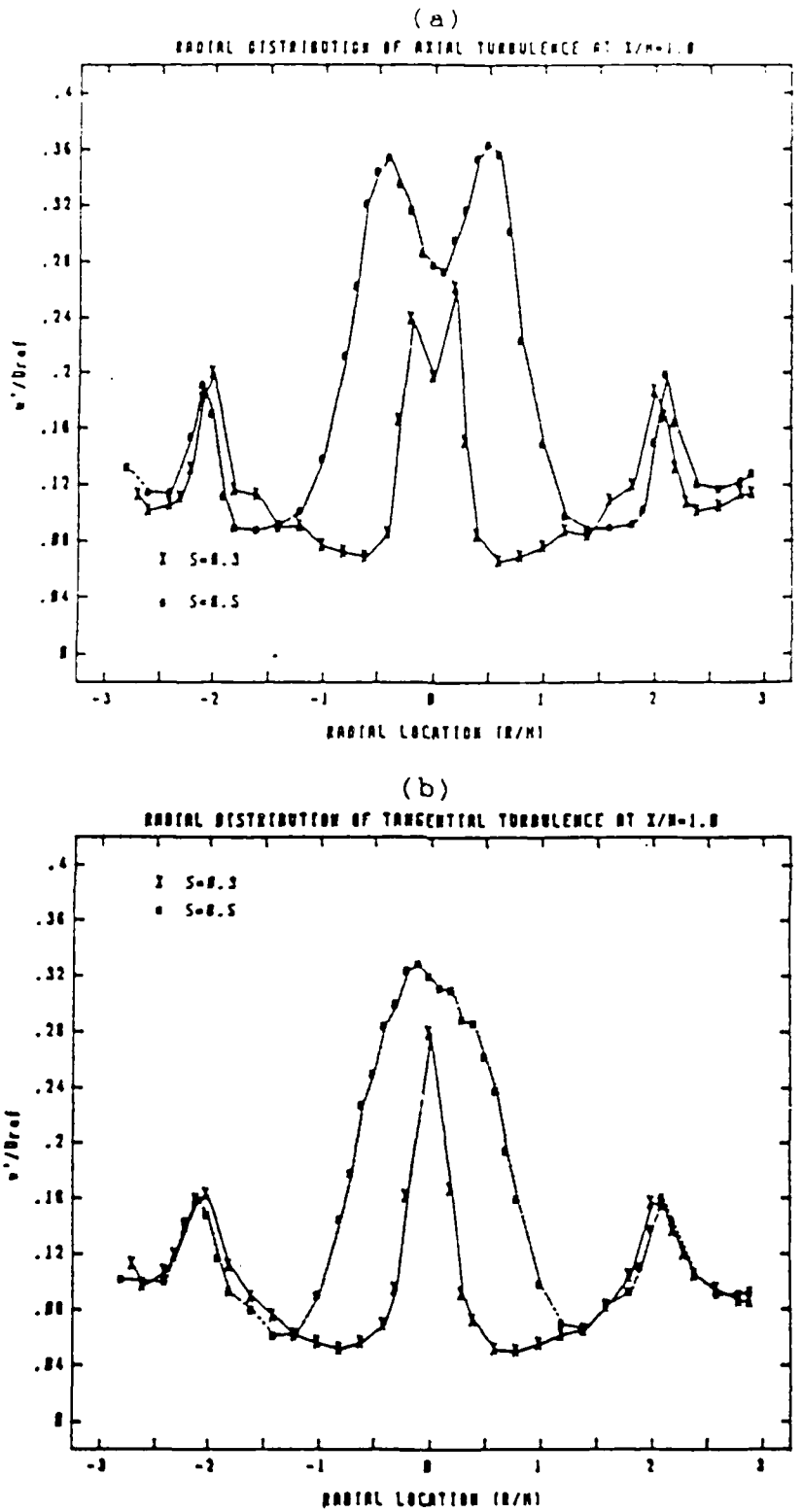


Fig. 6 Axial and tangential turbulence intensities at x/h=1.

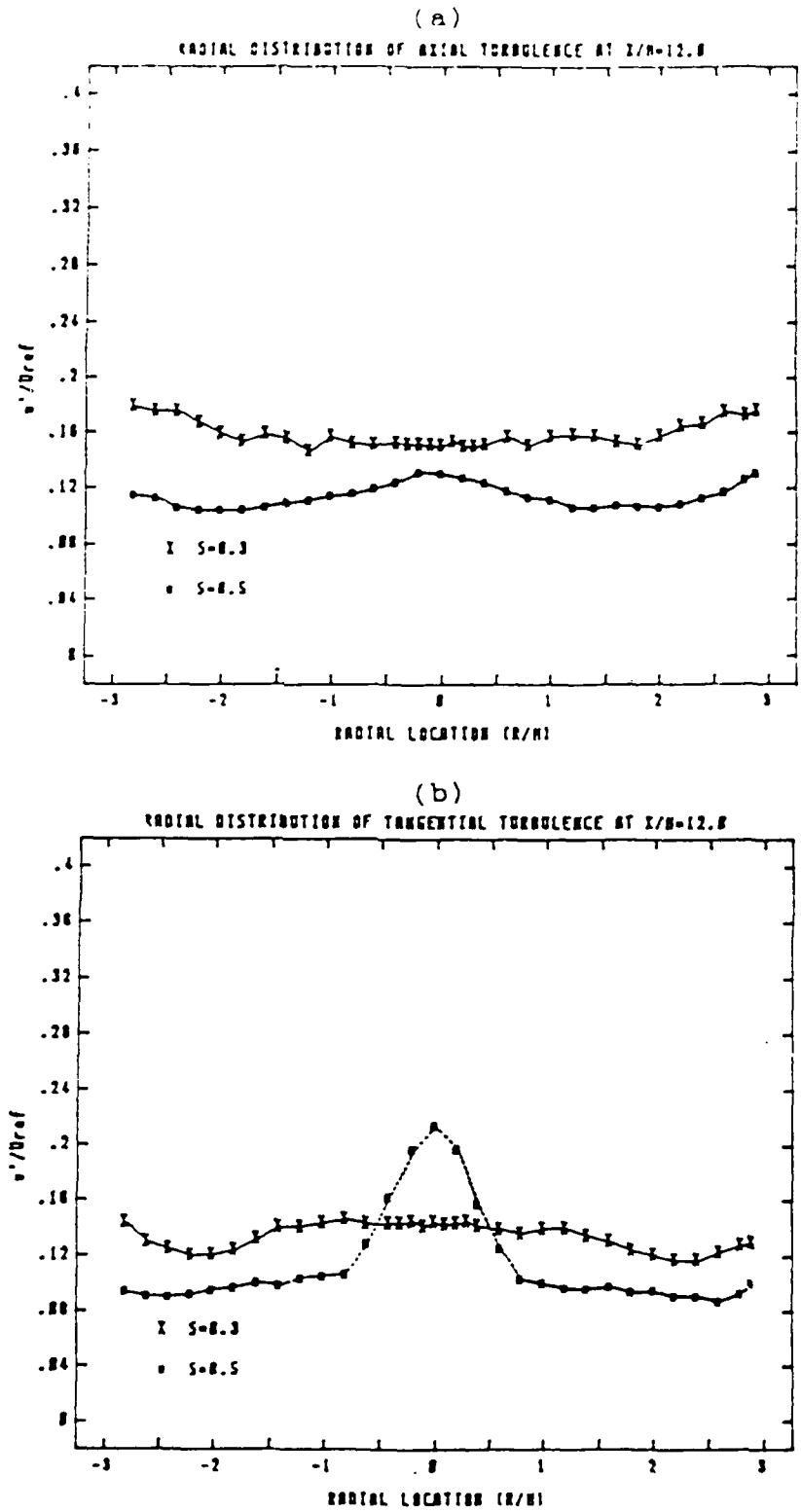


Fig. 7 Axial and tangential turbulence intensities at x/h=12.

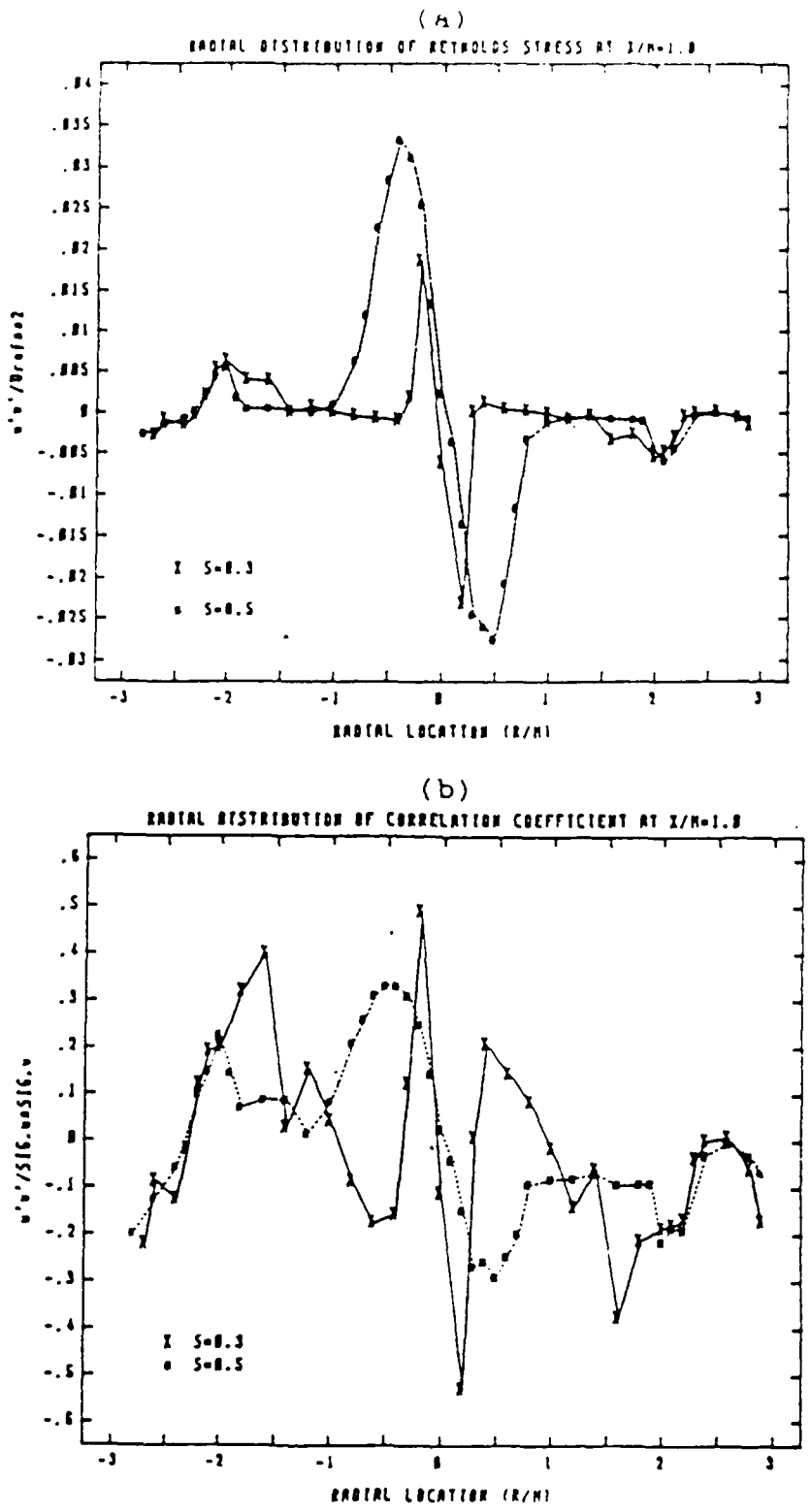


Fig. 8 Reynolds shear stress and correlation coefficient at x/h=1.

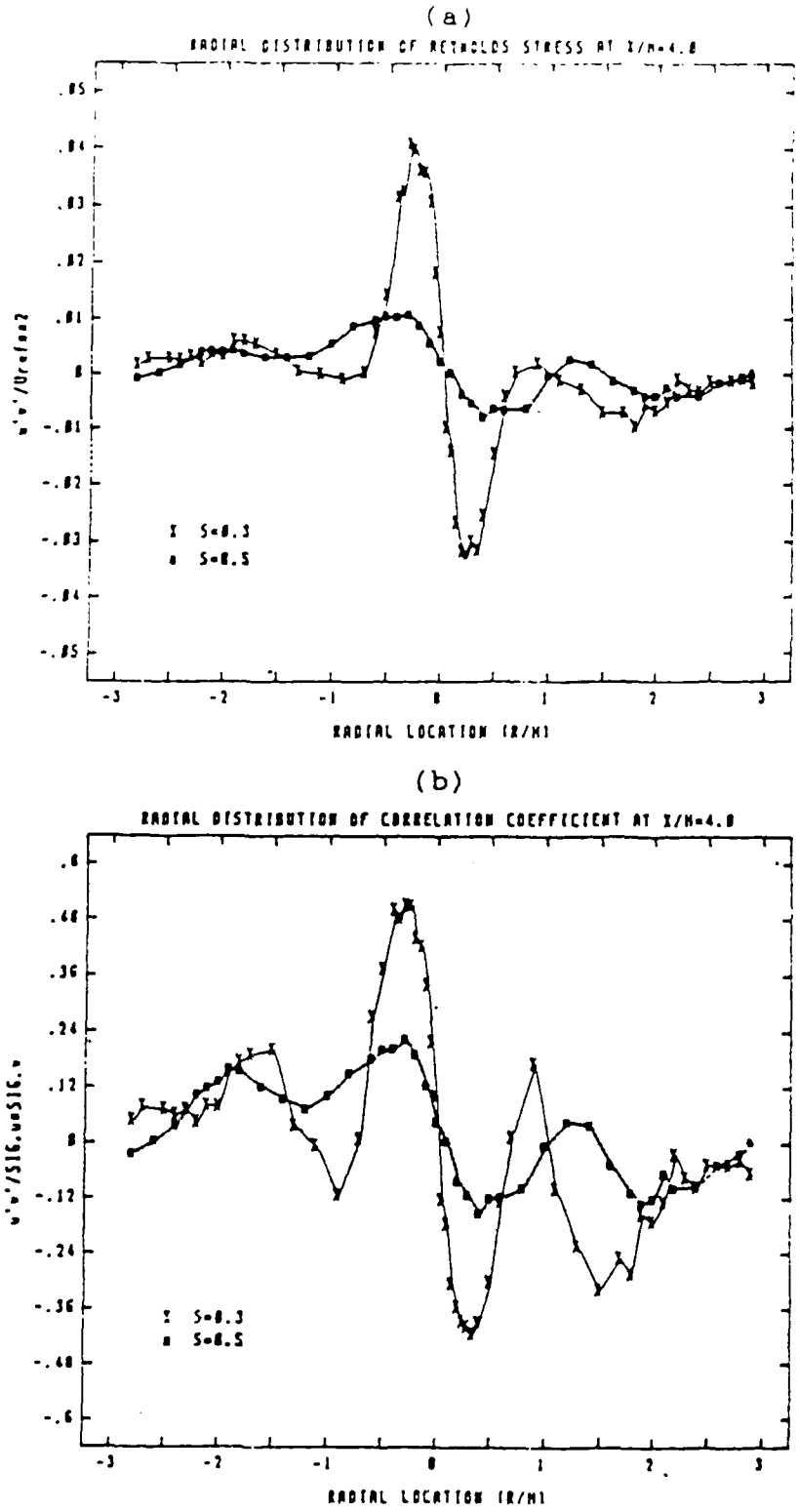


Fig. 9 Reynolds shear stress and correlation coefficient at x/h=4.

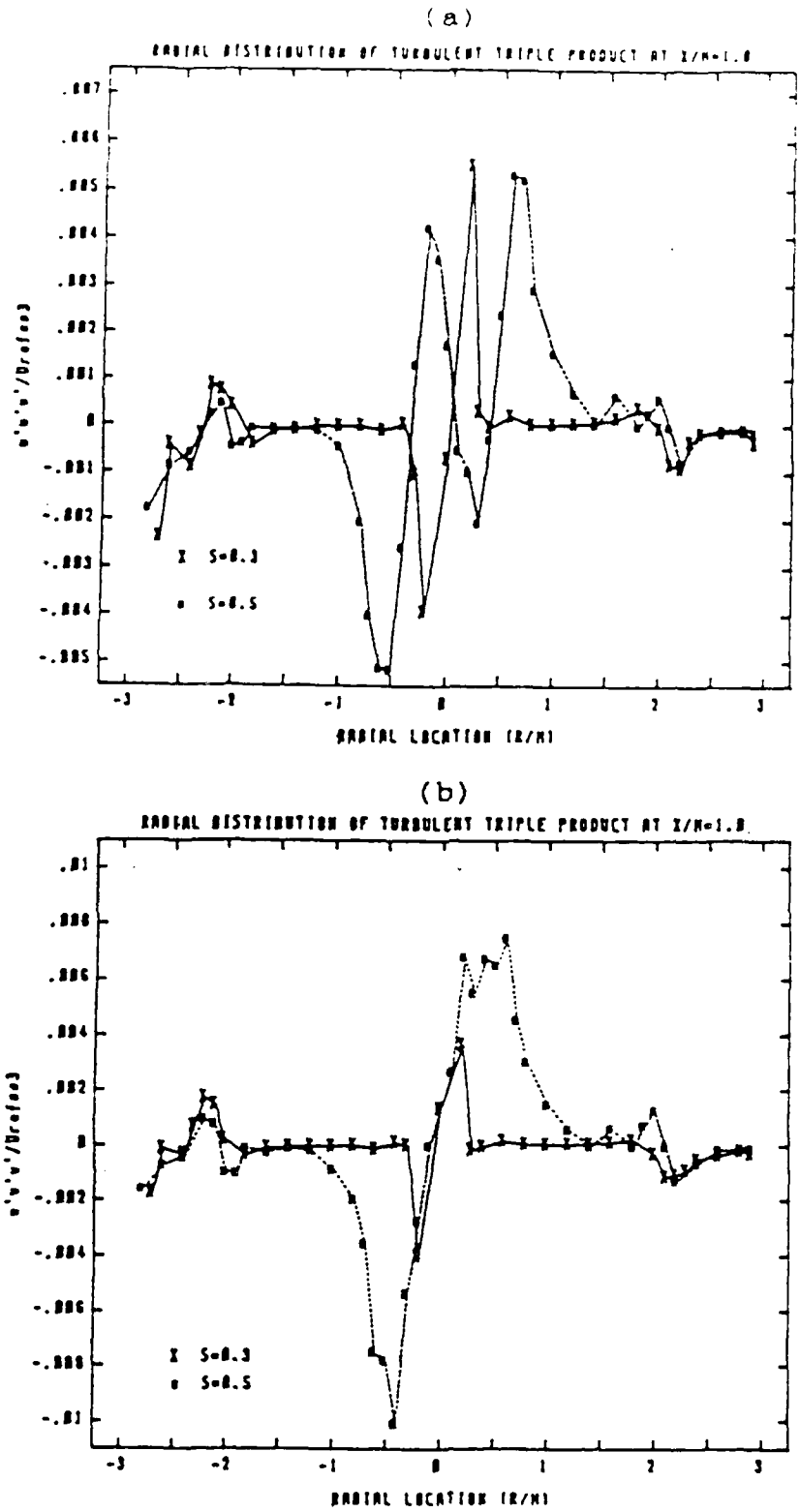


Fig. 10 u'u'w' and w'w'w' components of triple products at x/h=1.

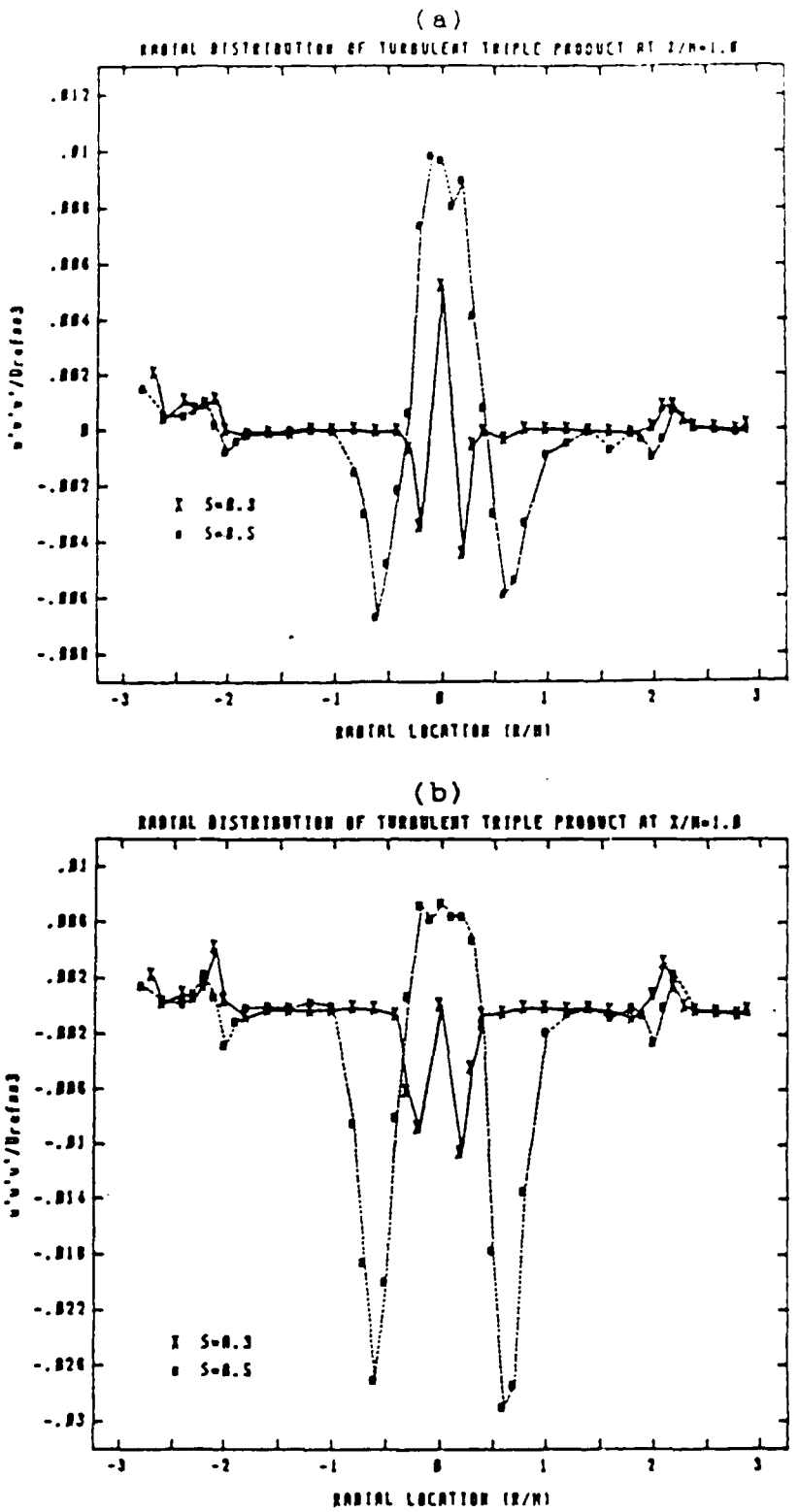


Fig. 11 u'w'w' and u'u'u' components of triple products.

1986 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT SUMMER PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

RESULTS OF A BRIEF INVESTIGATION OF TWO ALTERNATIVE ENERGY SYSTEMS

Prepared by:	John F. Schaefer
Academic Rank:	Associate Professor
Department and University:	Electrical Engineering The Citadel
Research Location:	Facility Systems and Analysis Branch Engineering Research Division Engineering and Services Laboratory Air Force Engineering and Services Center Tyndall AFB, Florida 32403-6001
USAF Researcher:	Thomas C. Hardy
Date:	15 September 1986
Contract No:	F 49620-85-C-0013

RESULTS OF A BRIEF INVESTIGATION OF TWO ALTERNATIVE ENERGY SYSTEMS

by

John F. Schaefer

ASTRACT

Analyses of two different alternative energy system concepts are treated. The first analysis is of certain characteristics of a generic class of photovoltaic energy systems. Specifically investigated was the fidelity of a commonly-used mathematical model of a generic photovoltaic array. The model was then used to study relative performance of two commonly-used techniques to extract near-maximum energy from photovoltaic arrays. Results indicate that the simpler, but infrequently employed, technique is superior to the more commonly implemented method.

The second analytical effort is a study of a novel wind energy machine, conceived by the author several years ago. The machine's potential advantage is performance/cost ratio. Analytical and computer simulation results confirm that the device will work; however, time constraints did not allow simulation confirmation of the analytically-predicted optimum choices for system parameters.

Acknowledgments

I am grateful to the Air Force Office of Scientific Research and my host organization, the Air Force Engineering and Services Center, for this opportunity to pursue several areas of long-abiding reseach interest. I particularly thank Tom Hardy for allowing the latitude to investigate my specific interests, Stan Strickland for numerous stimulating technical discussions of wind machines, and Joanne Brzowski for especially gracious assistance.

PART A.

INTRODUCTION

I hold a doctorate in electrical engineering from Stanford University. My specific expertise is in the field of automatic control theory; however, my technical interests have swung toward analysis and implementation of alternative energy systems over the past six years. I served as Manager of Solar Electric Projects at the New Mexico Solar Energy Institute prior to my recent move to The Citadel and full-time academia.

The overall research area at the Engineering and Services Laboratory involved near-term electrical energy systems for USAF bare-base, and other stand alone facility, needs. Thus I was particularly attracted to this opportunity to try to assist the USAF in improving its energy systems and, eventually, moving toward energy independence.

PART B.

RELATIVE PERFORMANCE OF TWO COMMON PHOTOVOLTAIC ARRAY MAXIMUM-POWER-POINT TRACKING SCHEMES

1. Introduction. This effort had a dual thrust: to analyze an important characteristic of a popular mathematical model of a photovoltaic (PV) array, and to use that model to gain insight into the relative performance of two commonly-used processes to maximize the output of PV energy systems. Neither thrust involved actual hardware or experimental data. Before discussing the approach and the results, a brief review of PV cell/array characteristics is in order. Figure 1 shows several current-voltage (I-V) curves of an hypothetical PV array.

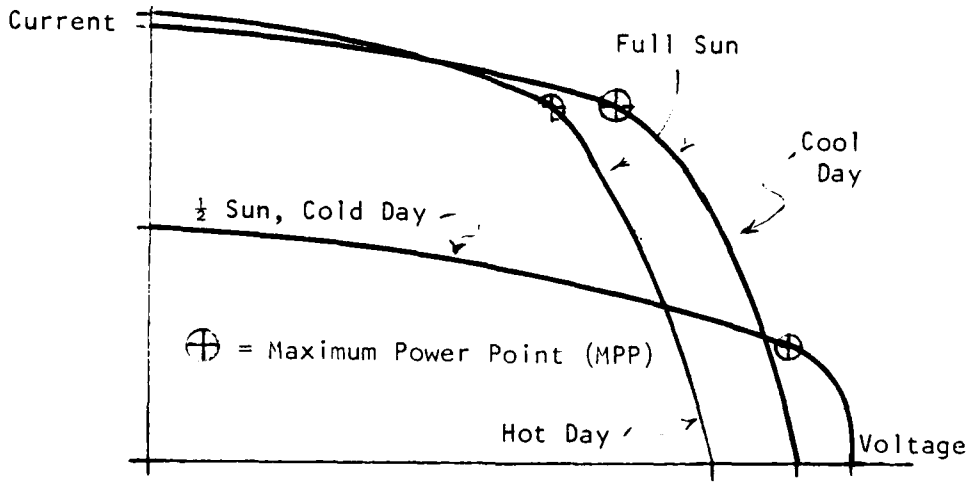


Figure 1. Generic PV Array I-V Curves, Various Environments

From it one concludes that array output is a function of amount of irradiance, PV cell temperature, and the load to which it is connected (power out is current times voltage; the "maximum-power-point" (MPP) is near the "knee" of the I-V curve. Current and voltage at the MPP are usually termed I_{mp} and V_{mp}). In practice, arrays are interfaced to loads either directly (in which case the operating point is determined by the intersection of the characteristic load line and the I-V curve) or by an inverter or converter which acts as an impedance transformer, forcing array operation at or near its instantaneous MPP. Since energy is lost when operating away from the MPP, all large PV arrays are of the latter form.

Motivation for the first thrust (analysis of the mathematical model) is as follows. One of the data points used to fit the model (curve) of interest, as well as all such models, is the MPP. However, in general the MPP of the model (the maximum of the ordinate-abscissa product) will not be at the specified I_{mp} - V_{mp} point. Thus, I felt it would be useful to PV researchers to gain some knowledge about the true versus specified MPP of the model as a function of the general shape (sharpness of the knee) of the I-V curve.

Motivation for the second thrust (comparison of MPP tracking schemes) was to gain insight into the expected relative performance of a pseudo MPP-tracking technique and a true, closed-feedback-loop technique. The pseudo approach is simple and reliable but open-loop and not widely used; the true trackers do indeed operate the array near the true MPP but characteristically lose energy by their (apparently necessary) continuous search for the true MPP. Briefly, operation of these units is as follows:

* In the open-loop approach, the power-producing array is caused to operate at a fixed fraction of the open-circuit voltage of a subelement (called a "pilot cell") of the array. The pilot cell produces no power, but is subjct to the same environment as the main array. The approach is reasonably valid because the true V_{mp} of most silicon PV elements remains at a near-constant fraction of the open-circuit voltage through a wide range of irradiance and temperature conditions [1]. My results, based on the array model and I-V curve translation techniques employed, confirm that validity.

* The closed-loop approach employes a perturb-observe-perturb technique: the inverter/converter picks a voltage, measures array power output, increments voltage, measures and compares power output, and by this means determines which way to increment voltage to increase output. The technique results in finding and eventually bracketing the true MPP. The wider the voltage increment bracket, the more power/energy is lost through operation either side of V_{mp}.

The issue: the author has observed the operation of numerous commercially-available PV inverters; a typical voltage bracket is 5% of array open-circuit voltage (V_{oc}) (the perturbation is +/-2.5%). While the resulting energy loss for a properly-operating unit is relatively small, the obvious question is, "Could it be smaller with a simpler pilot-cell tracker?" My results indicate the affirmative.

II. Array Model and I-V Curve Translation Routine. When designing, analyzing, or simulating PV energy systems, it is usually necessary to mathematically model the array I-V curve. Many models, of varying degree of complexity, have been used. One simple model, given as "Model 3" of [2], was used to generate the results reported here. The model is given as

$$I(V) = I_{sc}[1 - C_1\{EXP(V/[C_2V_{oc}]) - 1\}] \qquad (1)$$

where I_{sc} = Array short-circuit current, V_{oc} = Array open-circuit voltage, and C_1 & C_2 are curve-fitting constants. The curve is typically modeled at "Standard Rating Conditions" (SRC) of full sun (1 kw/m^2) and a cell temperature of 25°C. Constants C_1 and C_2 are chosen so that the curve passes through the two important points $I(V_{oc}) = 0$ and $I(V_{mp}) = I_{mp}$. The constants are given in closed form in [2] and several other references as

$$C_1 = (1 - I_{mp}/I_{sc})/(EXP[-V_{mp}/(C_2V_{oc})])$$
$$C_2 = (V_{mp}/V_{oc} - 1)/Log(1 - I_{mp}/I_{sc}) \qquad (2)$$

Unfortunately, I discovered in the course of this work that eqns.(2) yield erroneous results. I was unable to derive correct closed-form solutions. For all results reported here, I used an iterative technique to choose C_1 and C_2.

For realistic analysis of array performance under environmentally-dynamic conditions, it is necessary to shift the I-V curve of eqn.(1) as a function of sun and cell temperature. The routine used is from [3]:

$$I'(V,S,T) = I(V) + I_{sc}(S - 1) + C_1(T - 25)S$$
$$V'(V,S,T) = V + C_V(T - 25) - R_s[I'(0) - I_{sc}] \qquad (3)$$

Where I', V' are translated points
S is solar irradiance, kw/m^2
T is cell temperature, °C
C_I is the array current temperature coefficient, amperes/°C
C_V is the array voltage temperature coefficient, volts/°C
R_s is the array equivalent series resistance, ohms

III. Approach. To yield the results to be presented, I modeled 25 different "generic" PV modules and investigated voltage/current/power characteristics for five different sun/cell temperature conditions for each. The generic modules investigated had these characteristics:

V_{oc} = 21 volts C_V = -.05 volts/^oC

I_{sc} = 2.4 amperes V_{mp} = (.65,.7,.75,.8,.85)V_{oc} volts

R_s = 1.2 ohms I_{mp} = (.7,.75,.8,.85,.9)I_{sc} amperes

C_I = .002 amperes/^oC

Thus modules with significantly varying fill-factors were investigated (fill-factor is a measure of the squareness of the shape of a PV element I-V curve, and is thus a quality parameter. It is defined as $[V_{mp} \times V_{mp}]/[V_{oc} \times I_{sc}]$). The five sun/cell temperature employed were given by

$$S_i = .2 + .8\sin(\pi i/?) \quad kw/m^2$$
$$T_i = 20 + 32S_i \ ^oC \qquad \text{for } i = 1,2,...5 \qquad (4)$$

The cell temperature rise constant used, 32^{o}C/kw, is a realistic value based on past experience [4].

For each of the 25 models

* The constants C_1 and C_2 of eqn.(1) were evaluated
* 400 points of the eqn.(1) curve were filled at equal voltage intervals
* The filled curve's true MPP was found and the V_{mp}/V_{oc} ratio was found for later use.
* For each of the conditions of eqn.(4) the curve was translated according to eqn.(3). Each translated curve's
  - MPP I,V, and power were found
  - The pseudo MPP, as predicted by a pilot cell, was found.
  - The powers at voltges (true V_{mp}+/-2.5% of V_{oc}) were found.

Figures 2a and 2b are plots of the resulting I-V curves for two of the generic modules under SRC and the 5 realistic sun/temperature conditions. Figure 2a is for a very low

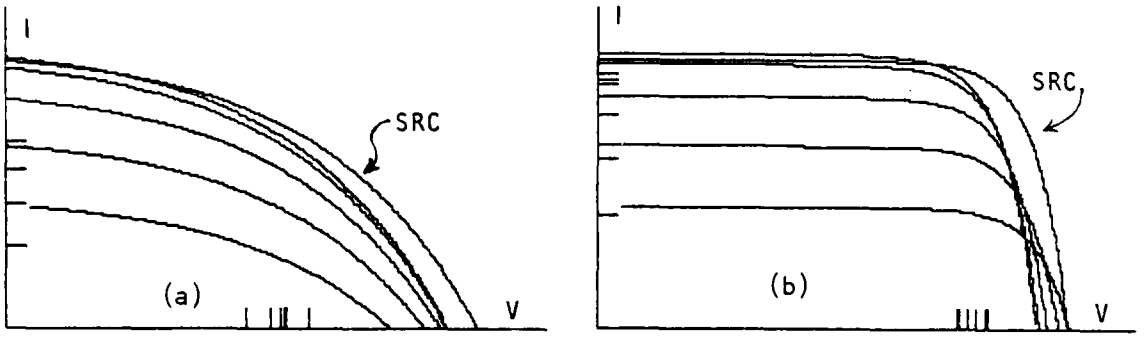


Figure 2. Model I-V Curves for Two Fill Factors, Six Conditions

(.455) fill-factor module, while Figure 2b is for a reasonably high (.765) fill-factor unit. These represent the extremes of the modules investigated.

IV. Results

A. Data relative to the actual versus specified MPP of the eqn.(1) model are presented in Figure 3. Figure 3a is a plot of the actual and specified V_{mp} versus the ratio (specified V_{mp})/(V_{oc}), with the ratio (specified I_{mp})/(I_{sc}) as a parameter. The data

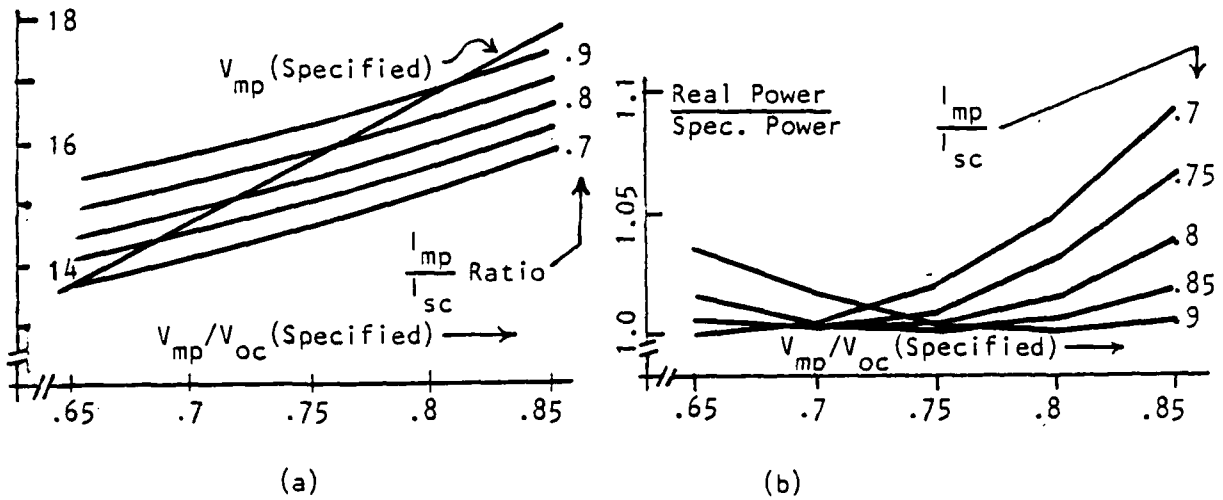


Figure 3. Actual and Specified V_{mp}, and Power Ratios, vrs. (Specified V_{mp})/(V_{oc})

show very clearly that the model of eqn.(1) does, in general, not have an ordinate-abscissa product maximum at the specified V_{mp}-I_{mp} point. The MPP's are close only for certain combinations of specified (V_{mp})/(V_{oc}) and (I_{mp})/(I_{sc}) ratios. The greatest error of those cases investigated is 11^{+}%.

Figure 3b is a plot of the ratio (actual maximum power of curve)/(specified $V_{mp} \times I_{mp}$ product). Clearly, substantial differences in power performance can occur. The greatest error is 9.1%, occuring for a high - V_{mp}/V_{oc}, low - I_{mp}/I_{sc} module.

The reader should not conclude that the model of eqn.(1) is not a "good" PV array model. The conclusion should, rather, be that if accurate results insofar as absolute power values are concerned, a more sophisticated model in which additional constraints (e.g., maximum ordinate-abscissa product at a specified point) are imposed should be employed.

B. On both an absolute and a relative basis, pilot-cell-controlled MPP trackers do a good job. The data can be summarized without resorting to a plot:

* For all of the 125 cases considered, the pilot-cell device extracts over 99% of maximum possible array power. The "worst" performance observed was the loss of a mere 0.2% of possible power.

* For all of the 125 cases, the pilot-cell device performance exceeds that of a true closed-loop MPP tracker whose voltage perturbations were +/-2.5% of V_{oc}. In no case was the performance advantage dramatic: the greatest superiority was for a high fill-factor module under medium irradiance conditions, where the performance advantge was 0.8%.

The pilot-cell performance results presented above confirm a long-held suspicion that the simple pilot-cell concept could be very competative with a true MPP tracking device if the latter, for whatever reason, makes significant voltage excursions in its maximum-seeking algorithm (theoretically, infinitesimally small perturbations will suffice. In practice, given a dynamic environment, measurement accuracy, repeatability, and precision, it is conceivable that voltage excursions less than a few percentage point are totally impractical, as current practice would seem to imply).

The reader should be aware that the pilot cell approach has potential drawbacks. For example, shadowing of the pilot cell itself (e.g., if small, by a fallen leaf) can result in misleading information: the shadowed cell could command the inverter to shut down, being unable to differentiate between a shadow and night, even though the remainder of the array is in full sun, ready and capable of providing full power. However, the reader should also note that the author has constructed and operated several units for periods of over one year without experiencing any problems.

V. Conclusions and Recommendations, Part B. Based on this exercise and the models and assumptions employed, one can conclude that

a. The popular I-V curve model given by eqns.(1) and (2) is invalid in the form given. The model does give a reasonable (in form and appearance) curve when the constants C_1 and C_2 are appropriately chosen. An iteration routine can easily yield proper constants. However, even with proper C_1 - C_2 values the model does in general not have a maximum ordinate-absissa product at the (desired) specified MPP. Power errors up to 9^+% were observed in the cases investigated.

b. A pilot cell-controlled MPP-tracker outperforms a true closed-loop MPP-tracker, if the latter perturbs array voltages by +/-2.5% of V_{oc} or greater. The USAF should consider use of the simpler concept for all future PV energy systems for both energy considerations and simplicity/reliability considerations. The work leading to conclusions here should be verified expermentally, particularly at low irradiance conditions, because of the known

decreasing accuracy of the curve translation eqn.(3) at increasing differences in irradiance and/or cell temperature[3].

PART C

ANALYSIS OF A SIMPLE CABLE/AIRFOIL

WIND ENERGY CONVERSION SYSTEM

1. Introduction. I considered the wind energy conversion system[5] depicted in Figure 4. It consists of a cable of length L under tension T, suspended horizontally normal to an

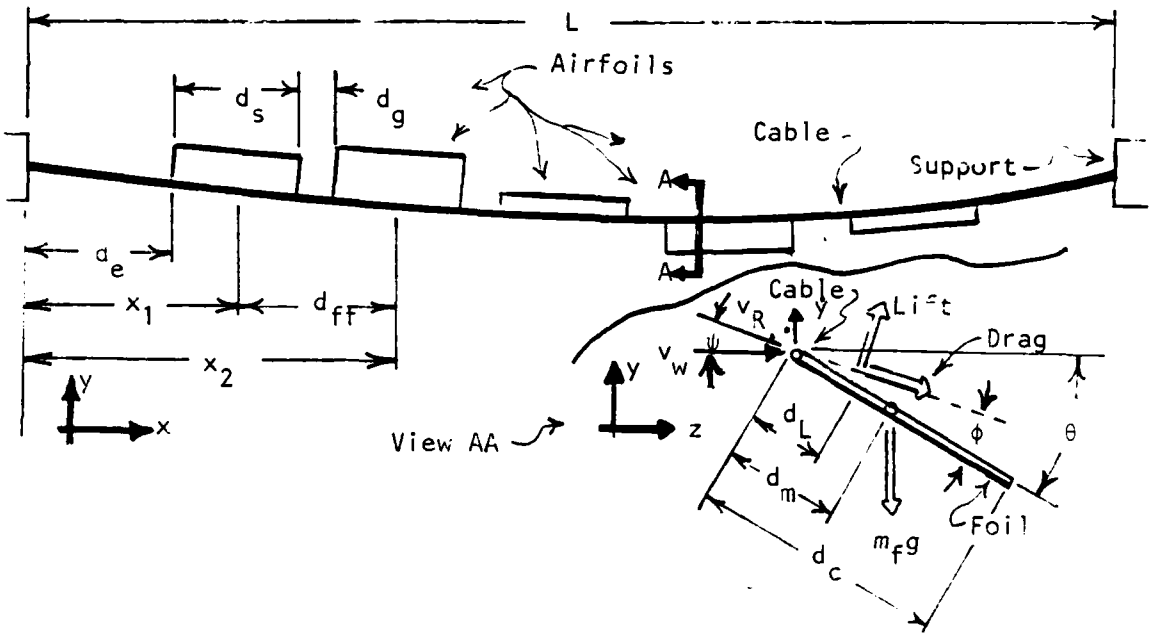


Figure 4. Cable/Airfoil Wind Energy System

horizontal wind of speed v_w. Suspended from the cable at, and free to pivot about, their leading edges are N airfoils, each of chord d_c, span d_s, mass m_f, moment of inertia about the center of mass J, and lift and drag coefficients C_L and C_D. The foils are assumed symmetric about a plane through the leading and trailing edges, and the center of lift is at a distance d_L aft of the leading edge (d_L is on the order of 1/4 of d_c for most low-speed airfoils, including flat plates). For this analysis, the foil center-of-mass distance d_m (positive aft of the leading edge) is assumed variable and adjustable, and its value does not affect J. The cable has a density ρ_c per unit length, and the foils are symmetrically spaced relative to the cable center, separated by a distance d_g. For analysis purposes, it is presumed that energy extraction mechanisms are tied to the cable at each airfoil mounting point; each opposes cable motion $\dot{y}_i$ (the dot signifies time differentiation) with a force $K'\dot{y}_i$ at the i^{th} foil position. My speculation was that for certain values of cable tension T and/or foil center-of-mass locations d_m, each presumably as a function of wind speed, energy could be extracted from the device. My results appear to confirm this speculation.

II. Equations of Motion. The system's equations of motion are derived in the coordinate system shown in Figure 4: y_i is the vertical displacement of the cable at the i^{th} foil position, relative to the fixed cable ends; up is positive. θ_i is the angular displacement of the i^{th} foil from the horizontal (the "pitch" angle); trailing edge down is positive. Foil yaw and roll and foil/cable upwind and downwind motions are neglected. Without showing details of the derivation, the 2N coupled, second-order, nonlinear differential equations which describe the system are

$$\begin{aligned} m_t\ddot{y}_i - m_f d_m \ddot{\theta}_i \cos\theta_i &= -m_t g + FC_i + FA_i - K'\dot{y}_i \\ -m_f d_m \ddot{y}_i \cos\theta_i + J'\ddot{\theta}_i &= m_f d_m \cos\theta_i g + M_i \end{aligned} \qquad i=1,2,\ldots N \qquad (1)$$

where

m_t = Total mass associated with i^{th} position = $m_f + \rho_c(d_s+d_g)$

g = Gravitational constant

J' = $J + m_f d_m^2$

FC_i = Cable force acting on the i^{th} cable/airfoil assembly

$$= -T\left[\frac{y_i - y_{i-1}}{\sqrt{d_{ff}^2+(y_i-y_{i-1})^2}} + \frac{y_i - y_{i+1}}{\sqrt{d_{ff}^2+(y_i-y_{i+1})^2}}\right]$$

[NOTE: FC_1 and FC_N are slightly different due to irregular spacing at end]

d_{ff} = Airfoil - to - airfoil spacing = $d_s + d_g$

FA_i = Aerodynamic force acting on the i^{th} airfoil

$= 0.5\rho_A A v_{Ri}^2 (C_{Li}\cos\psi_i - C_{Di}\sin\psi_i)$

v_{Ri} = Relative wind velocity at the i^{th} foil position = $\sqrt{v_w^2 + \dot{y}_i^2}$

ψ_i = Angle of relative wind velocity vector at i^{th} foil = arc $\tan(\dot{y}_i/v_w)$

ρ_A = Air density

A = Airfoil area = $(d_s)(d_c)$

M_i = Aerodynamic moment about leading edge of i^{th} foil

$= -.5\rho_A A d_L v_{Ri}^2 (C_{Li}\cos\phi_i + C_{Di}\sin\phi_i)$

ϕ_i = Angle of attack of i^{th} foil = $\theta_i - \psi_i$

In the above, second-order inertial forces are neglected. Additionally, any mass associated with the energy extraction devices is neglected: if significant it can be lumped in with cable mass.

The equations (1) were simulated on a microcomputer for various parameter values for values of N up to and including 8. Results of the simulations will be presented following discussion of the analysis of the linearized versions of the set (1).

II. Linearized Equations of Motion. For analysis (not simulation) purposes, the set of equations (1) are linearized to consider small motions ε_i and α_i about the equilibrium y_{ei} and θ_{ei}; that is, $\varepsilon_i = y_i - y_{ei}$ and $\alpha_i = \theta_i - \theta_{ei}$. Each of the airfoil's equilibrium position is found by evaluation of the second of the set (1) under zero motion/zero acceleration conditions; there results

$$m_f d_m g\cos\theta_{ei} = \tfrac{1}{2}\rho_A A v_w^2 d_L (C_{L_i}\cos\theta_{ei} + C_{D_i}\sin\theta_{ei}) \qquad (2)$$

a transcendental equation in θ_{ei} (each foil hangs at the same angle $\theta_{ei} = \theta_o$).

For the cable equilibrium position, zero motion/zero acceleration is assumed for the first N equations of the set (1) and those equations are summed; there results

$$0 = -Nm_t g + \sum_i FC_i + \sum_i FA_i . \qquad (3)$$

If each $|y_i - y_{i+/-1}|$ is small relative to the foil-foil distance d_{ff} then the sum over i of the $\{FC_i\}$ is the quantity $[-2Ty_1/d_e]$, where d_e is the distance from the cable end to the center of the first foil. Since, at equilibrium, $v_{Ri} = v_w$, the sum over i of the $\{FA_i\}$ is simply $(N/2)\rho_A A v_w^2 C_{Li}$. If a sinusoidal "first-mode" equilibrium shape of the cable is presumed, i. e., $y_i = H_y \sin(\pi x_i/L)$, i=1,2,....N, then $y_1 = H_y \sin(\pi d_e/L)$ and the cable equilibrium shape is sinusoidal with a peak displacement (at the center) found by evaluating eqn.(3) for H_y:

$$H_y = N d_e (\tfrac{1}{2}\rho_A A v_w^2 C_L - m_t g)/\{2T\sin(\pi d_e/L)\} \qquad (4)$$

where C_L is a function of θ_o, the equilibrium foil position given by solution of equation (2).

Before derivation of the linearized equations, it is necessary to model the lift and drag coefficients C_L and C_D: for all simulation work and all subsequent analyses these parameters were taken as

$$C_{L_i} = C_L'\phi \quad \text{if} \quad |\phi| < .4 \quad \text{or} \quad C_{L_i} = 0 \quad \text{if} \quad |\phi| > .4$$
$$C_D = C_{DO} + K_O \phi_i^2 \qquad (5)$$

where C_L' (the lift coeficient derivative), C_{DO}, and K_D are constants. Thus C_L is, for small angles, a linear function of angle of attack.

If, now, the results (3) and (4) with models (5) are used in equations (1) and the set (1) is summed over N there results, after second-order terms are dropped,

$$m_t\Sigma\ddot{\varepsilon}_i - m_f d_m \cos\theta_o \Sigma\ddot{\alpha}_i = -2T\varepsilon_1/d_e + qC_L'\Sigma\alpha_i - \{K' + q(C_L' + C_{DO})/v_w\}\Sigma\dot{\varepsilon}_i$$

$$-m_f d_m \cos\theta_o \Sigma\ddot{\varepsilon}_i + J'\Sigma\ddot{\alpha}_i = -qd_L f_o \Sigma\alpha_i + qd_L f_o \Sigma\dot{\varepsilon}_i/v_w \qquad (6)$$

where all summations are from 1 to N and where

$$q \equiv \tfrac{1}{2}\rho_A A v_w^2 \quad \text{and} \quad f_o \equiv (C_L' + C_{DO})\cos\theta_o - C_L'\theta_o \sin\theta_o.$$

The set (6) form a set of two linear, constant-coefficient ordinary second-order differential equations in the variables ε_1, $\Sigma\varepsilon_i$, and $\Sigma\alpha_i$ where all ε_i and α_i are small if (6) is to be valid.

IV. Analysis of the Linearized Equations. The goal is to determine if any one or more combinations of controllable parameters of the set (1) can be found such that, given a wind v_w, a steady oscillation of the variables ε_i and α_i can exist. If so, power (in the amount $K'\Sigma i\ \dot{\varepsilon}_i^2$!) will be generated. To this end, the solutions

$$\varepsilon_i(t) = \{B_1 \cos\omega t\} \sin(\pi x_i/L)$$

$$\alpha_i(t) = \{B_2 \cos\omega t + B_3 \sin\omega t\} \sin(\pi x_i/L) \qquad (7)$$

are substituted in (6) [again, "first-mode" amplitude of the sought oscillations is assumed. Since time = 0 is arbitrary, no (sin ωt) term is required for the ε_i in (7)]. The result of this substitution is two sets of simultaneous equations in the constants B_1, B_2, and B_3. The requirement for a nontrival solution yields not only the "characteristic" equation

$$\left[1 - \frac{(m_f d_m \cos\theta_o)^2}{m_t J'}\right]\omega^4 - \left[\omega_c^2 + \omega_f^2\left(1 - \frac{m_f d_m C_L' \cos\theta_o}{m_t d_L f_o}\right)\right]\omega^2 + \omega_c^2\omega_f^2 = 0 \qquad (8)$$

but also a constraint which yields the "damping constant" (power extraction rate coefficient) K':

$$K' = \frac{\omega^2\left[\frac{m_f d_m \cos\theta_o}{v_w}\omega_f^2 - \frac{q}{v_w}(C_L + C_{DO})\right] + \frac{qC_{DO}}{v_w}\omega_f^2}{\omega^2 - \omega_f^2} \qquad (9)$$

where $\omega_f^2 = qf_o d_L/J'$, the squared natural frequency of one of the airfoils when the cable is infinitely taut, and where ω_c^2, the squared natural frequency (fundamental mode) of the cable/foil assembly when foil rotation is disallowed, is

$$\omega_c^2 = \frac{2T}{m_t} \frac{\sin(\pi d_e/L)}{\sin(\pi x_i/L)}$$

One concludes that if the power extraction rate coefficient K' is chosen according to eqn.(9), where ω^2 is a solution of eqn.(8), then indeed the cable/foil system can oscillate in the mode given by eqn.(7) and, if K' is positive, can deliver energy as hoped. The next issue is to analyze the relations (8) and (9) to attempt to gain insight into how to choose parameters so as to maximize power output.

Equation (9) suggests that if one could choose parameters such that $\omega^2 = \omega_f^2$ (implying a "resonance" condition), K' becomes undefined. To see if this can occur, the characteristic equation (8) is rewritten as

$$\omega_c^2 = \frac{\omega^2 \left(1 - \frac{(m_f d_m \cos\theta_o)^2}{m_t J'}\right)\left(\omega^2 - \omega_f^2 \left[\frac{1 - \frac{m_f d_m \cos\theta_o C_L'}{m_t d_L f_o}}{1 - \frac{(m_t d_m \cos\theta_o)^2}{m_t J'}}\right]\right)}{\omega^2 - \omega_f^2} \qquad (10)$$

Figures 5 show in the complex ω^2 plane the three possible forms of the loci of ω^2 as the parameter ω_c^2 (which varies linearly with cable tension) is varied from zero to large values[6].

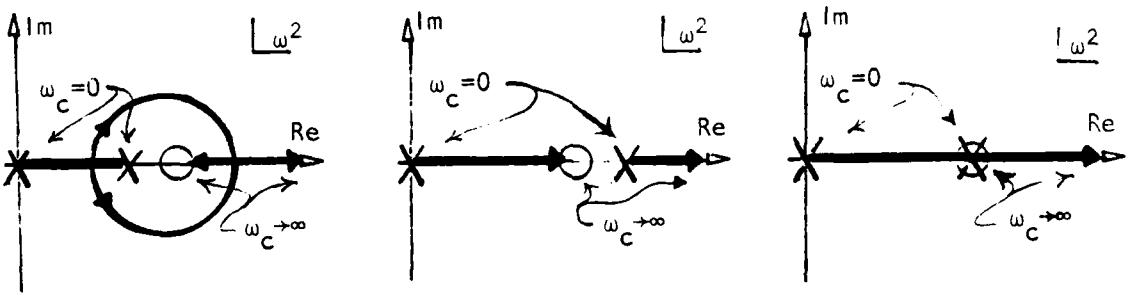


Figure 5. Possible loci of ω^2 as a solution of eqn. (8)

Figure 5c is of most interest: from it one concludes that the condition $\omega^2 = \omega_f^2$ can occur for finite values of ω_c^2, but only if the bracketed coefficient of ω_f^2 in the numerator of (10) is unity. The only reasonably-adjustable parameter therein is d_m, the location of the foil center-of-mass. But these results are valid only if eqn.(2) is satisfied: it, of course, shows that θ_o is a function of d_m. Simultaneously satisfying

eqn.(2) and setting the bracketed numerator coefficient of ω_f^2 in (10) to unity, there results the condition

$$\frac{qd_L(C_L\cos\theta_o+C_D\sin\theta_o)}{m_f g\cos\theta_o} = d_m = \frac{1}{2}\left[\frac{d_L f_o\cos\theta_o}{C_L'} \pm \sqrt{\left(\frac{d_L f_o\cos\theta_o}{C_L'}\right)^2 \frac{4J}{m_f}}\right] \quad (11)$$

Equation (11) will have one or more solutions for θ_o and/or d_m only if $J < m_f d_L^2(C_L+C_{DO})^2/4C_L'^2$. This is probably an achievable constraint, depending on airfoil design. If that constraint is met, then only solutions for which $|\theta_o| <= 0.4$ are of interest, to insure that C_L is in the linear-with-angle-of-attack regime. Finally: if d_m is adjusted so that (11) is satisfied, and (from eqn.(10)) if cable tension T is adjusted such that

$$\omega_c^2 = \left[1 - \frac{(m_f d_m\cos\theta_o)^2}{m_t J'}\right]\omega_f^2 \quad (12)$$

these analyses indicate that ... for want of a more rigorous expression ... interesting behavior may well occur. These results seem to make reasonable physical sense: when parameters are properly adjusted, a double solution $\omega^2 = \omega_f^2$ of eqn.(8) can occur, implying a motion of the form $y(t) = (YM)t\sin(\omega_f t)$. Loosely interpreted, eqn. (9) implies that only energy extraction at an extremly high rate can prevent a growing solution of that form. In reality, the nonlinearities of the realizable system will result in either a bounded solution or ... as happens in many unwanted resonance conditions ... system destruction.

V. Simulation Results. The set of equations (1) for an N-airfoil system, where N's of 2,3,4,5, and 8 were investigated, was simulated on an Apple microcomputer. For all simulations, the following parameters were held constant:

Total airfoil area was 1600 ft^2
Cable length L was 200 ft
Airfoil chord d_c was 20 ft.
Interfoil gap d_g was 2 ft.
Windspeed v_w was 25 ft/sec
Gravitational constant g was 32.2 ft/sec^2
Airfoil mass/area was .033 $slug/ft^2$
Moment of inertia J was $m_f d_c^2/50$
Center of lift dimension d_L was 5 ft.
Cable density ρ_c was .06 $slug/ft.^3$
Air density ρ_A was .0027 $slug/ft^3$

A fourth-order Runge-Kutta integration scheme was used; integraton step sizes varied but were in the vicinity of .02 sec for most runs. Additionally, a code was written and exercised which solved eqns. (2), (4), (8), (9), (11), and (12) for the various equilibrium values and "optimum" system parameters and which also solved for the constants B_1, B_2, and B_3 (to give mode shapes; see equation (7)).

The simulation code showed that indeed "fundamental mode" motions could occur for wide ranges of cable tensions and power extraction rates. It also confirmed that nonfundamental mode motions can exist (as suggested by Fig. 5a, which indicates complex values of ω^2 are possible for a range of cable tensions). Figure 6a is a depiction of cable-foil position for an 4-airfoil system at several instants in time; a clear fundamental-mode motion exists. Figure 6b is the same system for a different cable tension, clearly in an harmonic mode.

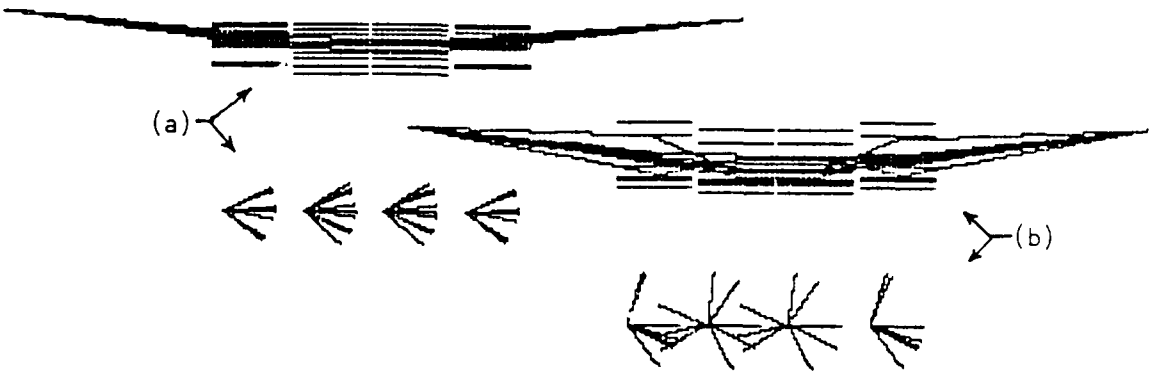


Figure 6. Four-Airfoil System. 6a: Fundamental Mode. 6b: Harmonic

To give example of device performance: for a 4-airfoil system with the parameters given above the code which calculates the "optimum" center-of-mass location from eqn. (11) gives d_m as 4.87 ft. aft of the leading edge. This yields an equilibrium foil angle of .257 (14.7°), a foil natural frequency ω_f of 4.34 rad/sec, and an "optimal" cable tension of 6965 lbs. A 20-second simulation run (which required, in real time, about 1.5 hours) gives an average power output of slightly over 48 hp for a K' (power extraction rate constant) of 60 lb-sec/ft.

Unfortunately, time constraints did not allow a thorough verification of the analytical results by the simulation code. The author intends to continue pursuit of a more substantial verification of the analysis.

VI. Application. It has been shown that a simple system of one cable and one or more airfoils (which can be simple flat plates) can in theory be attached to two fixed supports (possibly existing, e.g., mountains) in a wind regime and substantial motion can result. The linear mathematical model implies an extremely precise rate of power extraction may be required, but simulation of the realistic (nonlinear) model indicates otherwise. Thus, if simple power extraction mechanisms can be implemented, the device may well prove to have significant performance-to-cost ratio advantages over conventional wind machines. The author has devoted little effort toward details of a power-extracting mechanism, but

suggests that one or more rods connecting the cable to linear alternators or linear homopolar generators may well be a satisfactory solution (power extraction schemes to match the model used herein would have to be multiple; one rod alternator or generator per airfoil). The ideal single extraction point would be at one end of the cable, where large forces but small displacements would be available. However, I know of no simple mechanism which could make use of such forces and displacements.

VIII. Conclusions and Recommendations, Part C. Analysis of a mathematical model has shown that a simple cable/airfoil system can, if properly adjusted, extract significant energy from a wind field. Extremely inexpensive components can perform the power extracting function; the problem of conversion to electrical or other useable energy forms was not addressed. The author recommends that the USAF pursue this concept. The suggested approach is to construct a small-scale model to verify these analytical results, followed by a study of power conversion mechanisms, followed by a medium-scale power-producing demonstration. Patent protection should be considered.

REFERENCES

1. Pickerell, R. L.; O'Sullivan, G. ; and Merrill, W. G. "An Inverter/Controller Subsystem Optimized for Photovoltaic Applications." In Proceedings of the 13th IEEE Photovoltaic Specialists' Conference, Washington, D. C., 1978.

2. Rauschenback, H. S. Solar Cell Array Design Handbook, New York, Van Nostrand Reinhold Co., 1980.

3. Russell, M. C. Residential Photovoltaic System Design Handbook. DOE/ET/20279-255, April 1985, MIT Energy Laboratory.

4. New Mexico Solar Energy Institute. "Establishment and Operation of the Southwest Residential Experiment Station." DOE/ET/20279-213, 1982.

5. Schaefer, J. F., and Hastings, R. M. "A Potpourri of Innovative Wind Machines." In Proceeding of the 1984 Annual Meeting of the American Solar Energy Society, Anaheim, CA, June 6-9,1984.

6. D'Azzo, J. J., and Houpis, C. H. Linear Control System Analysis and Design, New York, McGraw-Hill Book Co., 1981.

1986 USAF-UES SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems. Inc.

FINAL REPORT

RAMAN SPECTROSCOPIC INVESTIGATIONS ON GROUP IIIA DOPED SILICON CRYSTALS

Prepared by:	Dr. James Schneider
Academic Rank:	Professor
Department and University:	Physics Department University of Dayton
Research Location:	Air Force Wright Aeronautical Laboratories Air Force Materials Laboratory Electromagnetic Materials Division Laser and Optical Materials Branch
USAF Researer:	Dr. Patrick M. Hemenger AFWAL/MLPO
Date:	August 1, 1986
Contract No:	F49620-85-C-0013

RAMAN SPECTROSCOPIC INVESTIGATIONS ON GROUP IIIA DOPED SILICON CRYSTALS

by

James Schneider

ABSTRACT

Raman spectroscopic studies of extrinsic p-type silicon crystals conventionally doped with group IIIA elements were undertaken with samples near 77K and 4K cryogenic temperatures. Incident light pulses from a Q-swsitched Nd:YAG laser with near IR wavelengths available at 1064 nm, 1074 nm, or 1123 nm among others were used with rectangular shaped doped silicon crystals to provide a 90 degree scattering geometry. Boron doped samples definitly show an electronic Raman scattering interaction at 184 cm^{-1} (23 meV) from the incident laser energy. With gallium doped samples a background fluorescence is evident even with samples held near 4K and when using 1074 nm light. At that temperature this laser energy should be below the band-gap energy for high quality silicon, so the fluorescence was not expected. The existence of a peak due to a Raman interaction is not observed above the level of this fluorescence in gallium doped silicon.

Acknowledgement

The author thanks the Materials Laboratory of the Air Force Wright Aeronautical Laboratories, the Air Force Systems Command, the Air Force Office of Scientific Research, and the Universal Energy Systems, Inc. for sponsorship of this research effort. I especially appreciate the opportunity to interact with the many stimulating Air Force scientists at Wright Patterson Air Force Base, Ohio. In particular, I acknowledge the collaboration of Dr. Patrick Hemenger, who suggested this Raman scattering problem associated with silicon doped with group IIIA elements, and continues to show concern and interest in its solution. The helpful assistance of Tim Peterson and other staff at the Materials Laboratory in searching their inventory for appropriate samples of high quality silicon crystals is acknowledged with sincere gratitude.

I. INTRODUCTION

The research and development of high quality extrinsic silicon material for fabrication of infrared detectors and use in operational systems continues to be important. The optical and electrical characterization of these materials is necessary for the scientific understanding of the mechanisms and properties that are to be employed in useful devices. Hall effect studies are in the mainstream of techniques that yield interesting and useful details concerning the electrical transport properties of semiconducting materials. Visible and infrared absorption and emission spectroscopies are among the spectral characterization techniques that yield a great deal of information about the electronic excitations in semiconductors. Raman scattering spectroscopy is a useful tool for studying electronic transitions which may be forbidden by selection rules or which occur in experimentally inconvenient spectral regions for infrared spectroscopy.

In this study we consider the application of Raman scattering spectroscopic techniques to silicon samples conventionally doped with one of the group IIIA elements such as boron, aluminum, gallium, indium, or thallium. Silicon is a much studied, technologically important, material. The quality and purity of these crystals has been driven by their frequent use as the base materials for sophisticated device fabrication. Among the open questions in silicon that could yield under the continued improvement in state-of-the-art experimental techniques and the availability of higher quality crystals is a better understanding of shallow p-type acceptor impurities such as the group IIIA elements. Raman spectroscopy has been used extensively to study phonon spectra and phonon interactions in pure and doped silicon,[1-7] but comparatively few researchers[8-11] have reported studies of electronic transitions in this semiconductor. In their now classic paper, Wright and Mooradian[8,9] reported on Raman scattering from phosphorus donor and from boron acceptor impurities in silicon. They observed a rather weak, sharp line at 23.4 meV which was attributed to the boron acceptors. Jain, Lai and Klein[11] produced an extensive Raman scattering study of phosphorus-doped silicon which included some observations concerning

antimony and arsenic donor impurities in silicon. Their paper also contained a section concerning their observations on p-type acceptors in silicon. They reported a sharp line at 184 cm^{-1} (22.8 meV) in the Raman spectra of boron-doped silicon which was attributed to an electronic transition associated with the boron acceptor state. In addition, they reported unsuccessful attempts to observe similar transitions in aluminum-doped and gallium-doped silicon samples, and commented on not understanding this absence in their spectra.

Due to the similarity of Group IIIA acceptors, it would seem reasonable to expect transitions similar to the reported boron electronic transition to occur in aluminum, gallium, indium, and thallium-doped silicon. To the best of our knowledge, attempts have only been made in aluminum and gallium-doped samples and these previous negative results were reported some years ago when silicon sample quality was less developed. As was pointed out by Klein[12] in a review on electronic scattering in semiconductors, it is generally felt that additional theoretical and experimental work is necessary for better understanding of acceptor states in silicon.

It is, however, not an easy experiment to undertake due to a combination of facts. Firstly, while the ratio of Raman to Raleigh scattering intensities is generally in the order of 10^{-8} to 10^{-12}, with incident high power lasers a Raman scattered intensity is, in general, readily detectable. However, Raman interactions in silicon, other than with the LO phonon, are known to be very weak. Signal levels are therefore very low and even small electronic noise such as dark current in the photomultipler detector is a big problem. Secondly, since silicon is not transparent to visible light, lasers operating at near infrared wavelengths must be employed in order to probe the volume of the silicon samples. Detector sensitivity becomes a problem at these IR wavelengths because the only photocathode that even has any sensitivity in that region is the S-1 cathode and its quantum efficiency for the production of photoelectrons is very poor.

II. OBJECTIVES

The research was directed at the spectral properties of acceptor impurities in silicon crystals doped with various group IIIA elements (boron, aluminum, gallium, indium, or thallium). The general goal is to make some contribution to the enhanced understanding of the role of these dopants in silicon. Raman spectroscopy, which has provided a great deal of information about electronic transitions in other semiconductors, was the principal experimental technique that was used in an attempt to improve and expand the current understanding of the role of the acceptor levels resulting from these dopants in high quality silicon semiconductor samples that are now available. In particular, experimental Raman data was sought on doped silicon crystals at cryogenic temperatures using liquid nitrogen and liquid helium cooling fluids. Information obtained from the Raman scattering data should compliment the ongoing experimental work in infrared absorption in these materials at the WPAFB Materials Laboratory.

In a previous study a rather complete and convincing Raman scattering study of boron impurity dopant in silicon was undertaken. Prior to that work, there was some dispute about whether an experimentally observed interaction was due to an electronic Raman transition involving boron impurities, since expected Raman lines for similar electronic transitions in samples of Si:Al and Si:Ga had not been observed. Using the 1.064 micron, the 1.074 micron, and other similar laser lines from a pulsed Nd:YAG laser and using gated detector electronics with an extended IR-sensitive supercooled photomultiplier, an experimental facility and technique was established which allowed us to reproduce and confirm the existence of a Raman interaction at 23.4 meV (184 cm^{-1}) in boron doped silicon. The signal attributed to boron was very, very weak compared to the phonon interaction, but it was repeatedly observed with this IR-Raman scattering experimental setup. The study used samples of various known concentrations of boron dopant to observe signals that were dependent on concentration. The signal attributed to boron was observed at the same energy-shift from several

different incident laser wavelengths in each sample.

With the experience gained in that work wherein an improved signal-to-noise ratio was achieved, it was the objective of this work under the Summer Faculty Research Fellowship to investigate carefully the expected similar interactions using aluminum, indium, and/or gallium doped samples. If well known and documented doped samples of high quality are obtained, this study attempts to settle the issue of the existence or non-existence of electronic interactions in group IIIA (other than boron) doped silicon.

III. RAMAN SCATTERING SYSTEM; EXPERIMENTAL TECHNIQUES

The laser Raman experimental set-up is based on a pulsed Nd:YAG laser, a SPEX model 1401 (0.85 m) double monochrometer, photon counting detection, and a gated digital signal processing system which is computer controlled by an IBM/AT microcomputer. Figure 1 shows a schematic representation of this pulsed laser Raman spectroscopy system[13]. The laser is capable of operating in the burst, Q-switched mode which generates about 10 output pulses of approximately 200 nanosecond duration during each 150 to 200 microsecond long bursts. Burst rates are normally set at about 50 per second and the average power delivered is in the order of hundreds of milliwatts with peak power during each of the Q-switched pulses in the kilowatt range.

The basic idea is to reduce unwanted detection of background signals due to certain fluorescences and other stray light as well as photomultiplier dark current by gating the digital signal detection electronics "ON" only when there is temporal coincidence with an input laser pulse. A small portion of the incident laser light pulse is reflected onto a PIN diode. The output of this diode unit serves both as a timing pulse and as a means to sample, monitor,and record the incident laser power. A Time Pickoff Unit (Ortec 260) feeds this pulse from the diode to an integrator and sample/hold unit that yields an analog signal representative of input laser power. This analog output is digitized for recording. This same output from the PIN diode provides a

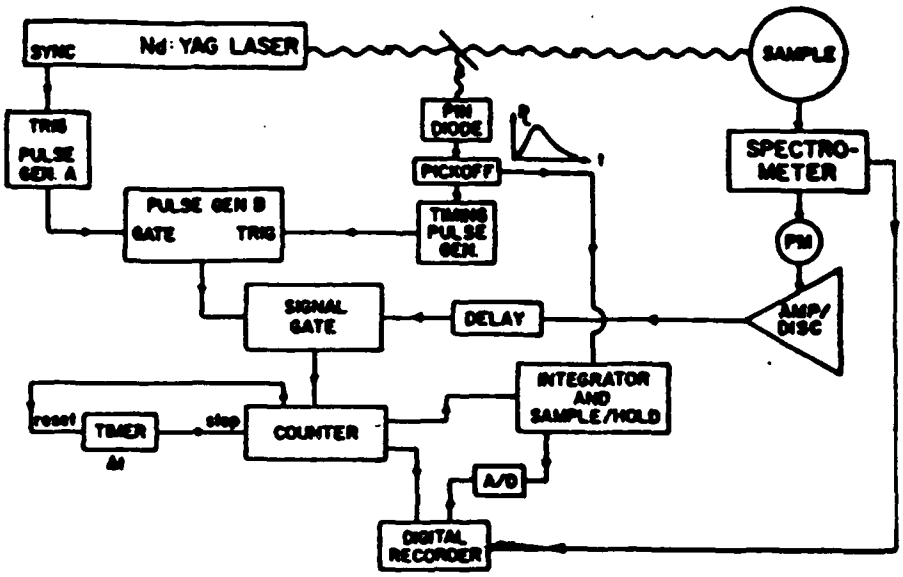


Figure 1. Schematic of Pulsed Laser Raman spectroscopy System

model 100A Pulse Generator). Additionally, the synchronized output from the Chromatix laser lamp flash circuit is used to trigger another pulse generator (pulse generator A in figure 1; portion of PAR CW-1 Boxcar Integrator) that is used to feed the enable gate of pulse generator B. In this configuration, pulse generator B provides an output pulse to open a gate to the digital counter (SSR 1120 Photon Counter) only during the time duration of each pulse of the laser, and spurious firing of pulse generator B is virtually eliminated. Since scattered photons detected by the photomultiplier on the exit slit of the spectrometer are counted by the digital counter only during the short time (100-200 ns) that the laser is active, background signal from photomultiplier dark current, fluorescence, ambient light, etc., are essentially reduced to near zero and weak signals are in principle observable. Typical operation is to hold at a spectrometer setting for a fixed number of input laser pulses, with signal to noise ratio improved at larger numbers of input pulses.

Three digital signals are recorded for each setting of the spectrometer, the scattered photon count from the counter, the digitized

output from the integrator and sample/hold circuit that is proportional to the incident laser power, and a digitized signal from the spectrometer setting that gives the wavenumber shift measured from the incident laser wavelength. These three signals are processed through a PAR Series 260 multiplexer unit and recorded on the IBM/AT microcomputer.

To investigate the Raman scattering from samples of silicon doped with various group IIIA elements, the optical elements required to effectively get the incident laser light to the sample and to direct the scattered light on the entrance slit of the double spectrometer had to be assembled to be appropriate for this infrared experiment. Gratings in the spectrometer had to be changed and the photomultiplier had to be installed for operation in the near infrared region of the spectrum as required with 1064 to 1123 nm laser light. A sample holder, capable of temperature dependent studies near 4K was adapted to the experimental arrangement. The signal detection electronics were slightly modified for operation in the gated digital signal processing mode, and programs were modified to allow computer control of the data recording and storage process with the newly installed IBM/AT minicomputer and ASYST signal processing software.

IV. RAMAN SCATTERING SPECTRA OF Si;B

A sample of boron doped silicon (#1742, 1.3E16) that had been previously investigated and whose Raman spectra with infrared laser pulses was earlier recorded at both 4K and 77K was used to check out the Raman scattering system as configured for this current experimental series. Figures 2 and 3 for boron doped samples bring out some of the features of the Raman spectra that might be useful in the search for similar data in silicon with other IIIA dopants. In figure 2a it is observed that the peak at 523 cm^{-1} shift from the 1074 nm laser energy is rather strong, typically about 1 detected count per incident laser pulse. Incident laser light is at zero on this Raman shift or energy shift scale. This peak is attributed to the LO phonon interaction in silicon and is a property of the bulk crystal. Note that it is common

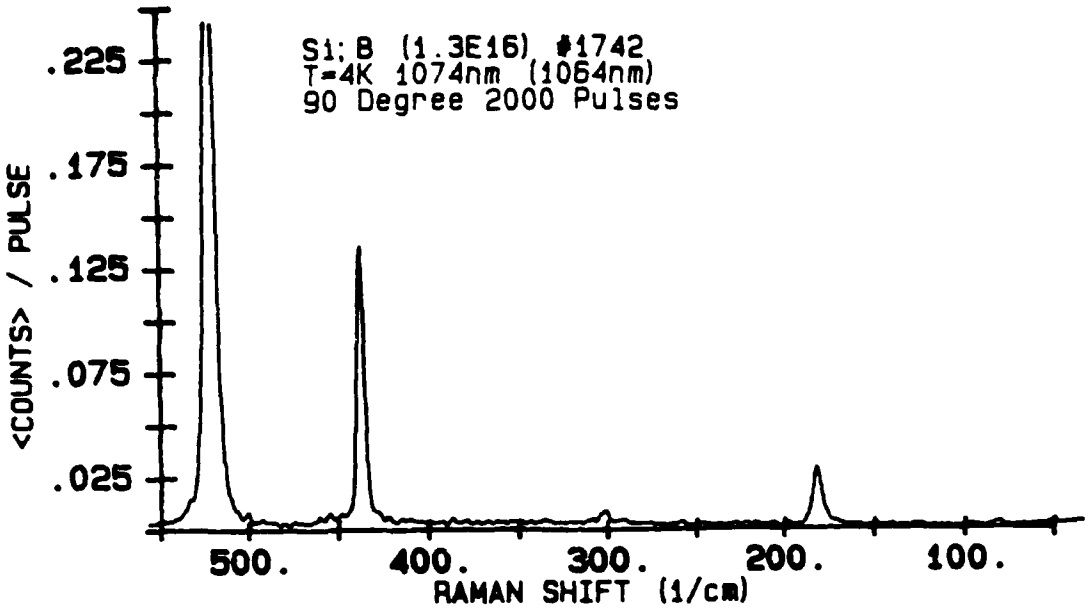


Figure 2a. Raman Spectra of Si;B (1.3E16); 4K ; 2000 Pulses

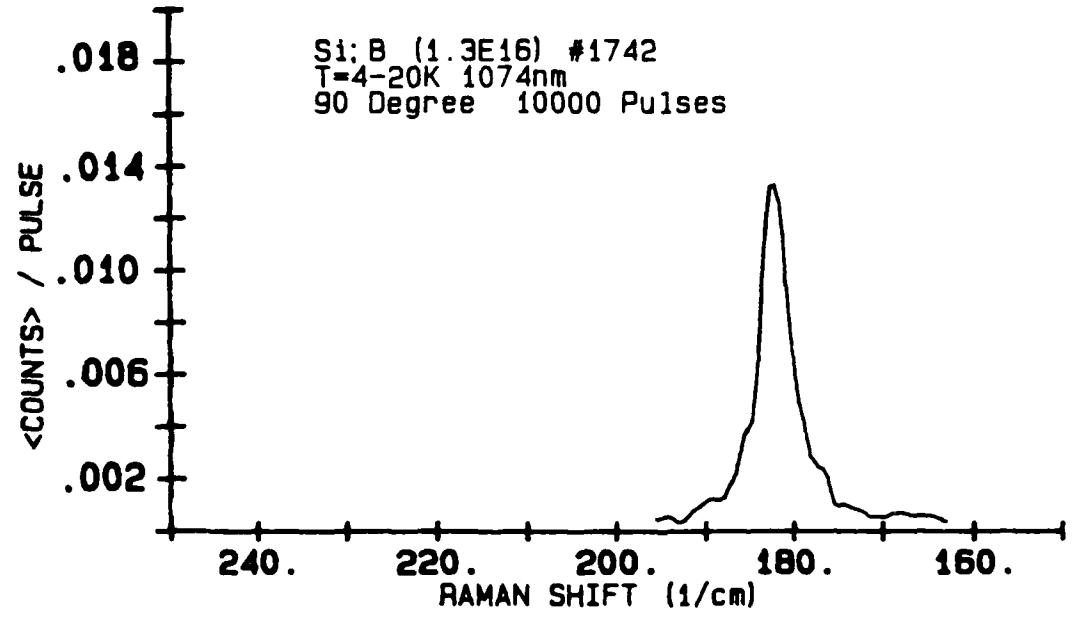


Figure 2b. Raman Spectra of Si;B (1.3E16); 4K ; 10,000 Pulses

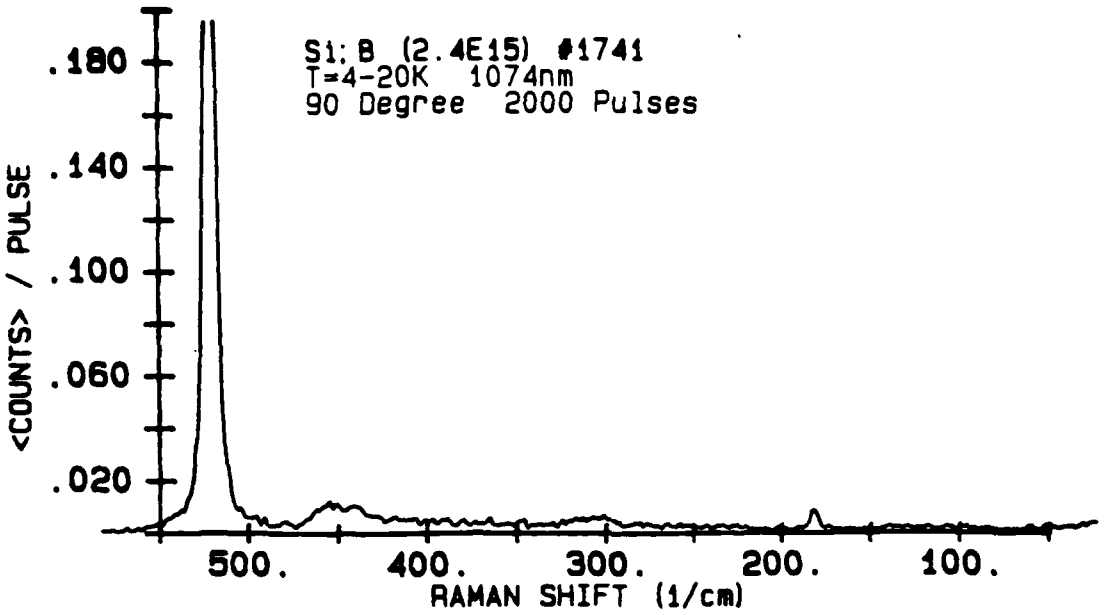


Figure 3. Raman Spectra of Si;B (2.4E15); 4K ; 2000 Pulses

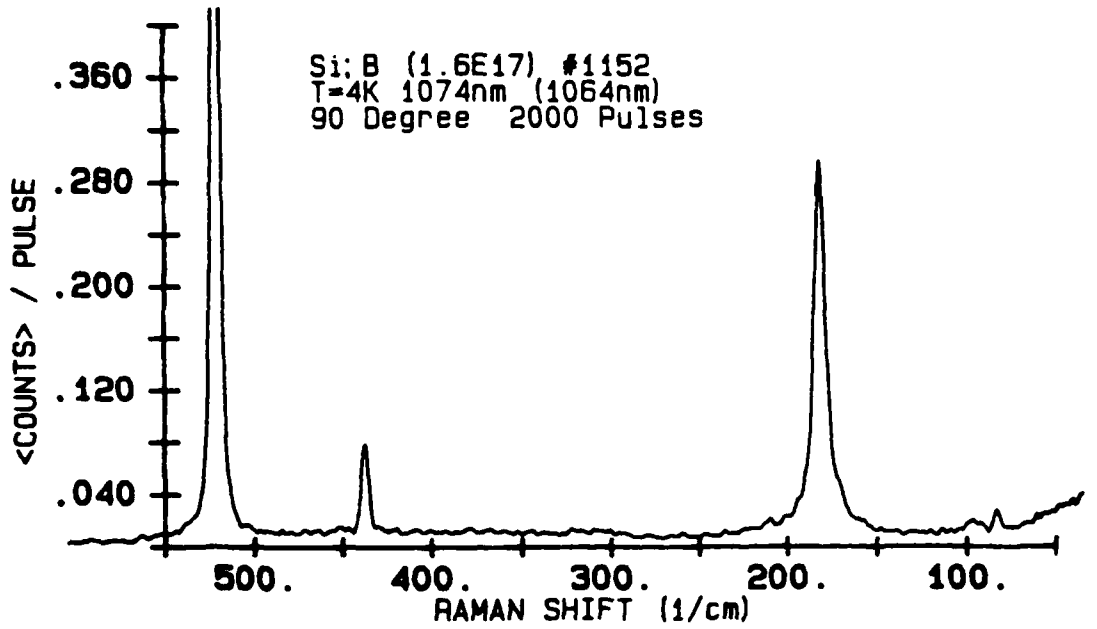


Figure 4. Raman Spectra of Si;B (1.6E17), 4K ; 2000 Pulses

with this Nd;YAG laser (prism tuneable to various laser wavelengths) for there to exist some pulses at wavelengths other than that for which it is intended to be tuned. In figure 2a the peak that shows at 436 cm^{-1} on the plotted "energy shift from the energy of the 1074 nm laser light" scale is actually a superimposed peak at 523 cm^{-1} from the also present 1064 nm laser energy. Note that 1074 nm is at 9311 cm^{-1} and 1064 nm is 9398 cm^{-1} on actual energy scales; a difference of 87 cm^{-1}. Thus it is obvious that this 436 cm^{-1} peak is a "false" or "superimposed" peak on the scale plotted and is actually another manifistation of the LO phonon interaction in silicon crystals.

The electronic Raman interaction at 184 cm^{-1} associated with the boron impurity is very, very weak. About 0.02 detected counts per incident laser pulse or less is observed for this sample with 1.3E16 boron impurities. Note that the signal to noise ratio can be improved by using 10,000 incident pulses as in figure 2b compared to the 2000 used in figure 2a. Note also that a "background count" in the order of only 0.001 counts per pulse is observed at helium temperatures.

Figures 3 and 4 show two other boron doped samples with different dopat levels. Figure 3 shows a sample with 2.4E15 boron impurities in which the boron peak at 184 cm^{-1} is much reduced on the scan of energy shift from under 50 cm^{-1} to beyond the LO phonon energy. Note that in this run the laser was well tuned at 1074 nm since no superimposed LO phonon interaction with 1064 nm light shows in the spectra. Knowing that the location of the electronic Raman interaction is near 184 cm^{-1} allows one to scan only a short region of the spectra while increasing significantly the number of incident laser pulses at each specrometer setting. While time consuming, the signal to noise ratio is improved such that the 184 cm^{-1} peak is clearly observed at about ten times the 0.001 counts per pulse background at this temperature.

Figure 4 shows an observation from a sample with 1.6E17 boron impurities. The 184 cm^{-1} peak attributed to the electronic interaction with the boron is readily apparent in this case. Note that the values of counts per pulse on these plots cannot strictly be compared since from run to run it depends on laser pulse power, focus, collection efficiency, etc., but a rough comparison is certainly reasonable.

V. RAMAN SCATTERING SPECTRA OF Si;Ga

Three samples of gallium doped silicon were prepared for a series of Raman scattering experiments. A very heavily doped sample #2042 with 2E19 Ga impurities was examined at 4K using laser pulses at 1074 nm. Figure 5 show that the spectra is relativily noisy, appears to exibit a "tail" of the Raleigh scattered incident laser peak which extends to near 200 wavenumbers before the spectra shows the "flat" background response in the order of .001 or .002 counts per input laser pulse that has been typical of silicon samples at 4K with this experimental apparatus. The "LO phonon peak" near 524 cm^{-1} shift is very weak in this sample, down about two orders of magnitude from typical peak response at this phonon energy. Such results are perhaps consistent with the expected scattering from a poor quality crystal that has its long range order perturbed by a high doping level of the gallium impurities.

Sample #2040 with 2E18 Ga impurities plotted in figure 6 likewise shows a "tail" extending from the incident laser energy to beyond 200 cm^{-1} shift before the background is "flat" at about .002 counts per input laser pulse. The phonon peak at 524 cm^{-1} in this sample was stronger than the corresponding peak observed from the sample with 2E19 Ga impurities, but is still weak by comparison with previously observed LO phonon spectra in silicon which typically gives about one count per incident laser pulse. Again, this observation would be consistent with a poor quality single crystal that has been disturbed by the high impurity density.

Figures 7a and 7b show the Raman spectra from sample #2041, a silicon crystal with 1.3E17 Ga impurities. This sample appears to have better long range order or higher crystal quality than the higher gallium doped silicon crystals as judged from the strong Raman peak associated with the phonon interaction at a 523 cm^{-1} shift from the incident laser energy. This phonon peak is more than an order of magnitude stronger than the 2E18 Ga doped sample (figure 6) and almost two orders of magnitude stronger than the sample with 2E19 Ga impurities (figure 5). Some fluorescence is observed from this sample near liquid

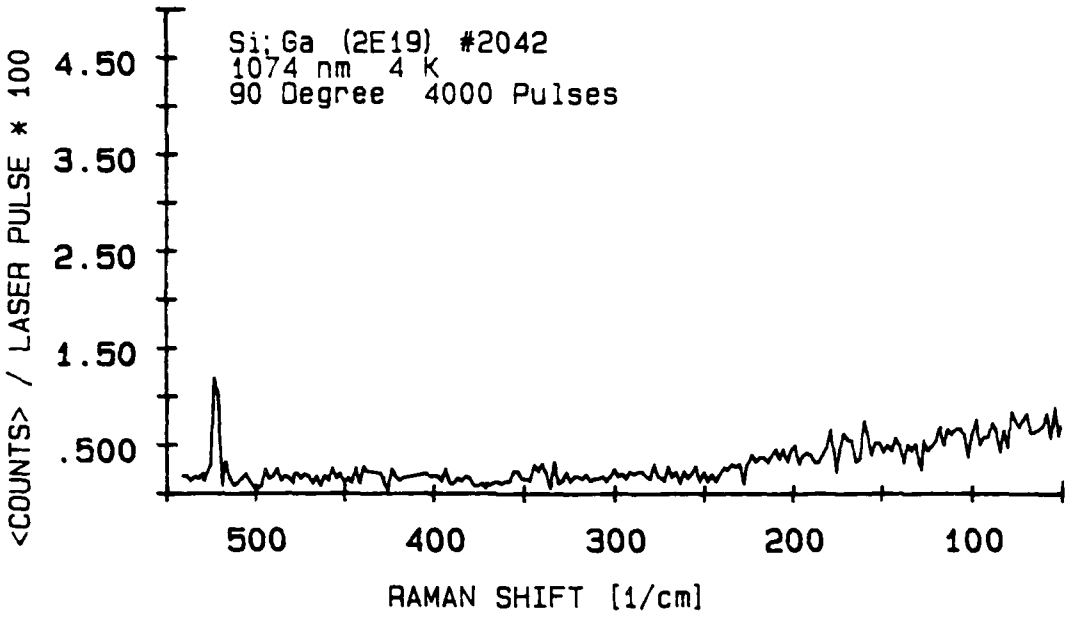


Figure 5. Raman Spectra of Si;Ga (2E19); 4K ; 4000 Pulses

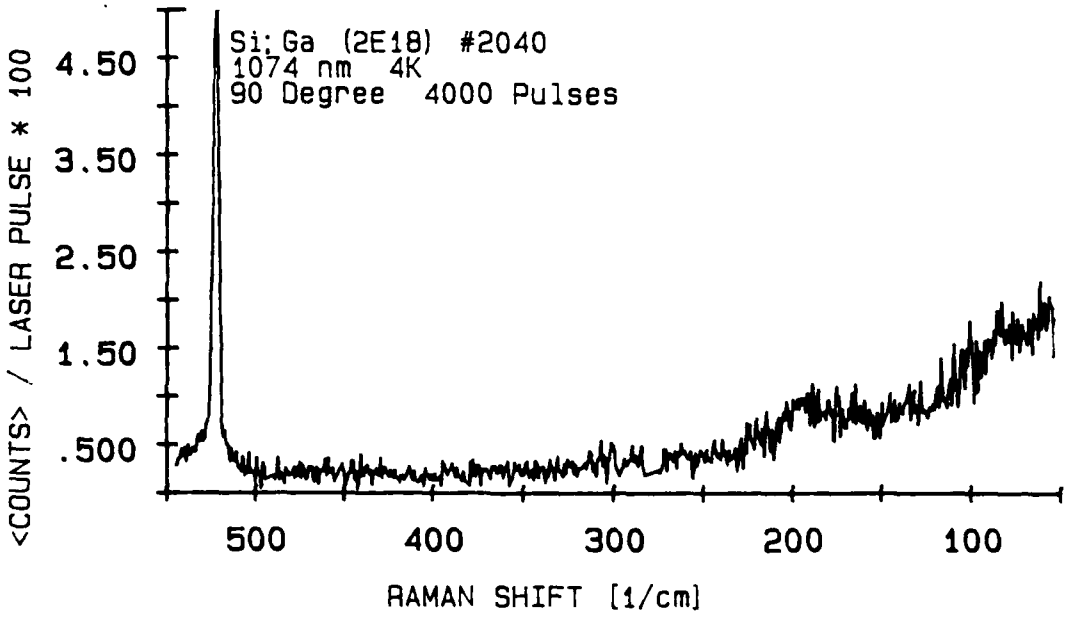


Figure 6. Raman Spectra of Si;Ga (2E18); 4K ; 4000 Pulses

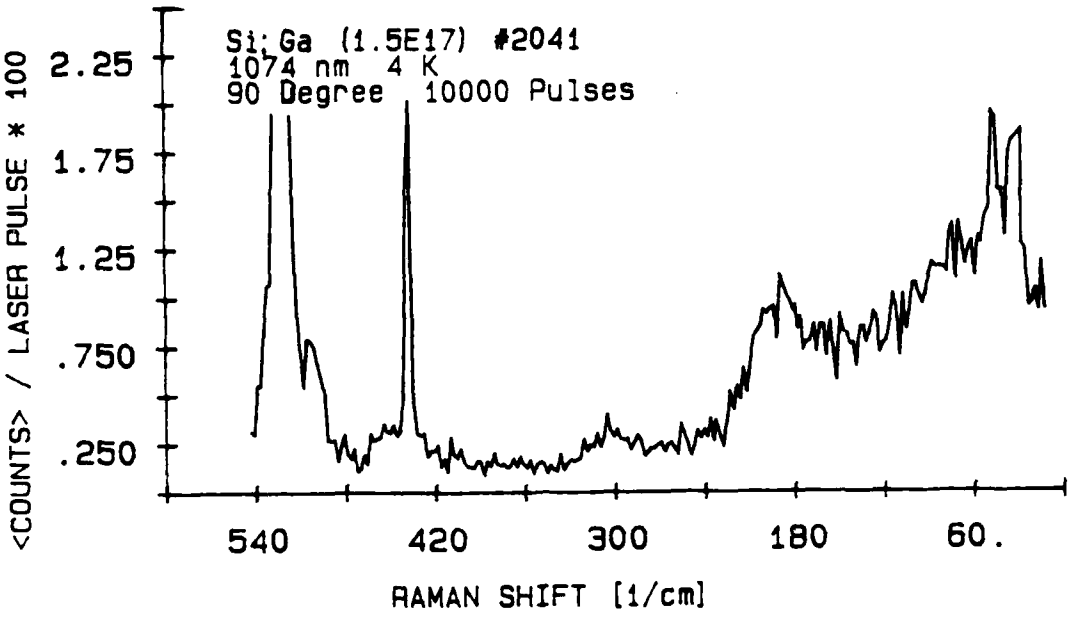


Figure 7a. Raman Spectra of Si;Ga (1.5E17); 4K ; 10000 Pulses, 1074 nm.

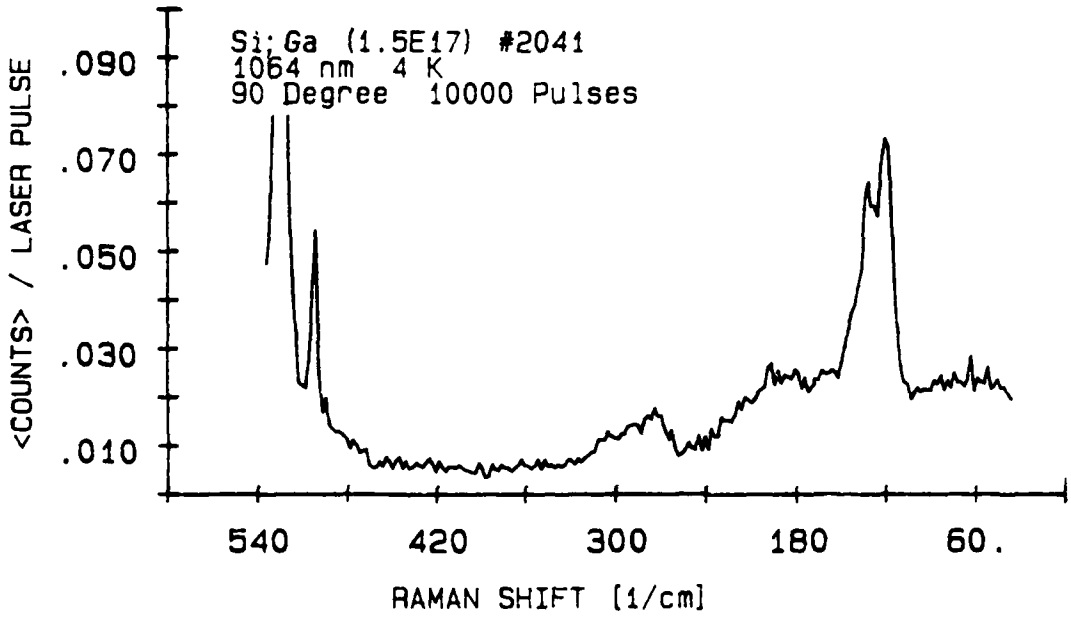


Figure 7b. Raman Spectra of Si;Ga (1.5E17); 4K ; 10000 Pulses, 1064 nm.

helium temperature when irradiated with incident laser light at either 1064 nm (figure 7a) and at 1074 nm (figure 7b). This is somewhat surprising. While 1064 nm light (1.165 eV) is just above the band-gap energy, the 1074 nm laser light (1.154 eV) was expected to be below band gap energy of about 1.16 eV at helium temperatures for high quality silicon as was the case with the boron doped samples. When the experimentally detected light from the crystal is plotted as a Raman shift, a sharp peak shows at 120 cm^{-1} shifted from the 1064 nm laser energy and a similar peak shows at about 30 cm^{-1} shifted from the 1074 nm laser energy when that laser is used. However, both peaks calculate to be near 9280 cm^{-1} or 1.150 eV when on an absolute scale and are thus believed to be fluorescence emissions superimposed on the detected scattered light. Another sharp peak shows on the Raman shift plots at 135 cm^{-1} from the 1064 nm laser and at about 45 cm^{-1} from the 1074 nm laser energy. These peaks likewise calculate to both exist at about 9265 cm^{-1} or 1.148 eV on an absolute energy scale and are judged to be fluorescence emission. Other broader emission exists with one such broad peak around 9123 cm^{-1} or 1.131 eV.

It's hard to say whether there is any true electronic Raman scattered signal among these emissions that might be associated with the gallium impurity. Fluorescence levels were high enough in the samples used to possibly mask a weak electronic Raman interaction peak. A rather obvious sharp peak shifted a consistent 503 cm^{-1} from the 1064 nm and also from the 1074 nm laser energies should be further investigated. It looks like a typical Raman peak but on the other hand its existence near the strong LO phonon line at 523 cm^{-1} suggests some caution. With strong light incident onto the spectrometer slit it is possible, on occasion, to observe spurious "peaks" associated with grating imperfections or geometrical reflections inside the double spectrometer. A slight change of scattering geometry and/or masking the gratings in subsequent experiments should allow us to determine if this apparently observed Raman peak at 503 cm^{-1} is a real Raman signal or a spurious peak. Additional Ga doped samples should also be helpful in sorting out those effects which are unique to a particular sample or experimental run and those that are characteristic of the doped silicon material.

VI. RAMAN SCATTERING SPECTRA OF Si;Al

One sample of aluminum doped silicon was available for investigation, but its level of doping was only estimated to be about 5E15. The Raman spectra of this sample is shown in figure 8. The only features are the strong LO phonon peak at 523 cm^{-1} from the 1074 nm incident laser light, and a much weaker superimposed LO phonon peak showing at 436 cm^{-1} on this scale but at 523 cm^{-1} from some 1064 nm pulses also present in the incident beam. No peak attributable to an electronic impurity interaction with the aluminum is obvious, at least not at this dopant level.

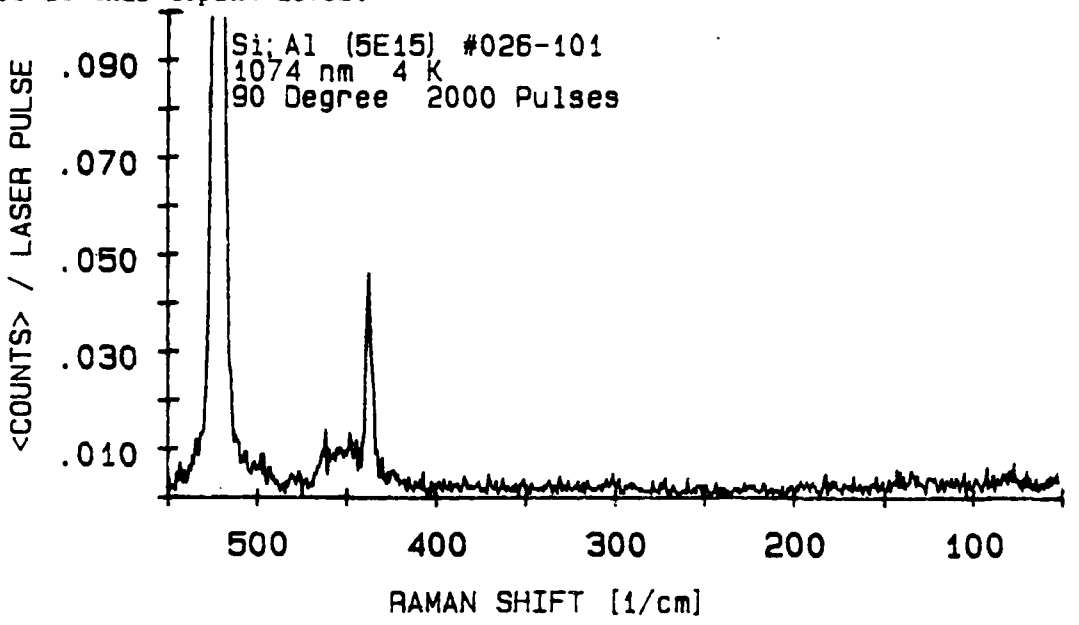


Figure 8. Raman Spectra of Si;Al (5E15); 4K; 2000 Pulses

VII. RECOMMENDATIONS

The observation of an electronic Raman interaction in silicon doped with the group IIIA elements has been an elusive and frustrating task. The well verified observation in boron doped silicon of a Raman active peak at 184 cm^{-1} or 22.8 meV was a challenge, but a satisfying one. Similar observations in the other group IIIA doped silicon crystals does not look promising at this point based on the evidence so

far. But in each case examined there is a possible reason for not observing a Raman interaction, such as a weak dopant level, poor crystal quality or the presence of a masking fluorescence. The experiments should continue when and if appropriate crystals are available.

REFERENCES

1. J. P. Russel, Appl. Phys. Lett. 6, 223 (1965).

2. J. H. Parker,Jr., D. W. Feldman and M. Ashkin, Phys. Rev.,155 (3) 712 (1967)

3. R. K. Chang, J. M. Ralston and D. E. Keating, Light Scattering Spectra of Solids, Edited by G. B. Wright (Springer-Verlag 1968)

4. Paul A. Temple and C. E. Hathaway, Phys. Rev. B 7 (8), 3685 (1973)

5. B. A. Weinstein and Manuel Cardona, Solid State Com., 10, 961 (1972)

6. K. Uchinokura, T Sekine and E. Matsuura, Solid State Com., 11, 47 (1972)

7. K. Uchinokura, T Sekine and E. Matsuura, J. Phys. Chem. Solids, 35, 171 (1974)

8. G. B. Wright and A. Mooradian, Phys. Rev. Lett., 18 (15), 608 (1967)

9. G. B. Wright and A. Mooradian, Proc. 9th Intern. Conf. Physics of Semiconductors, Moscow, 1067 (1968)

10. J. M. Cherlow, R. L. Aggarwal and B. Lax, Phys. Rev. B, 7 (10), 4547 (1973)

11. Kanti Jain, Shui Lai and M.V. Klein, Phys. Rev. B, 13 (12), 5448 (1976)

12. M. V. Klein, Light Scattering in Solids, Ed. M. Cardona, (Springer-Verlag, 1975, 1983-2nd)

13. P.P. Yaney, J. Raman Spectry., 5, 219, (1976)

1986 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT SUMMER SUPPORT PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

AN INTENTIONAL TUTOR

Prepared by:	Richard M. Schori
Academic Rank:	Professor
Department and University:	Department of Mathematics Oregon State University
Research Location:	USAF School of Aerospace Medicine Brooks AFB, Texas and Southwest Research Institute San Antonio, Texas
USAF Researcher:	Dr. Bryce Hartman and Mr. Glenn Humphress
Date:	September 16, 1986
Contract No:	F49620-85-0013

AN INTENTIONAL TUTOR

by

Richard M. Schori

ABSTRACT

The Air Force is interested in the development of a tutor for helping with the ground-based training of tactical fighter pilots. Such a training device would have applications in any teaching/learning environment. Contributions were made in 1) the collection and mathematical modeling for learning strategies for the tutor and 2) the method of knowledge representation for the tutor. Fuzzy set theory was used in the mathematical modeling project. Fuzzy sets give a convenient mathematical method of describing subjective concepts such as a "moderately high" level of knowledge. For knowledge representation, the notion of a genetic graph is used to develop what is called a "graph of contents" of a given text file. In a genetic graph the nodes represent concepts, facts, or rules and the links represent relationships between the nodes such as prerequisite, specialization, or analogy.

ACKNOWLEDGEMENTS

I thank the Air Force Systems Command and the Air Force Office of Scientific Research for sponsoring this research. I thank Dr. Bryce Hartman of the School of Aerospace Medicine, Brooks Air Force Base, Texas, for introducing me to this project, for his continual support, and for his efficient management and active interest in the project. I thank Southwest Research Institute for hosting me for the summer and in particular I thank Glenn Humphress for his congenial and imaginative technical leadership on the project. I thank Karen Westine of SwRI and Stanford University for her efficient support on the project and I thank Dr. James Schroeder of SwRI for his insightful direction in matters of learning theory.

I. INTRODUCTION

I received my Ph.D. in mathematics from the University of Iowa in 1964 and have been actively publishing in topology since that time. One of my specialties is infinite-dimensional topology and through this I have gained considerable experience in abstract geometric thinking and reasoning. On the other hand I have had a considerable practical experience in applied research, administration, teaching and pedagogy, athletics and coaching, psychology, and computers. I had my own research projects for four summers during college in a research department at Bell and Howell Company, I have held full time academic positions at major universities continually since my Ph.D., was promoted to (full) professor in 1973, and during the five year period 1978-83 I was Chairman of the Department of Mathematics at Oregon State University. I have had an interest in the various psychological types of people and the corresponding pedagogical implications. I have been using microcomputers extensively for the past four years and own two of them.

The USAF research project to which I was assigned can be called an Intentional Tutor, Personal Aid, or Pilot's Associate. It is ambitious, long term, and from a developmental point of view highly interdisciplinary. My versatile background which includes working well with other people was a deciding factor in me being selected for this project. Additionally, the existing research group was predominately from the areas of psychology and biology and whereas I could productively interact with them, I could also provide a bridge from them to the mathematical modeling, computer science, and AI communities.

II. OBJECTIVES OF THE RESEARCH EFFORT

The overall objective of the Intentional Tutor project is to develop a training device that would be deployed with tactical fighter squadrons and used by pilots for ground-based training in tactics, aircraft systems, etc. Such a training device would have applications in any teaching environment.

My specific objectives were:

A. Be an sounding board and advisor/consultant for the project.

B. Identify and mathematically model appropriate pedagogical principles and operational strategies for the Tutor.

C. Study the problem of knowledge representation as it pertains to the Tutor and make recommendations.

III. ADVISORY/CONSULTING CAPACITY

I consulted on a regular basis with Dr. Bryce Hartman, the managing director of the project, and I met on essentially a daily basis with Glenn Humphress, the technical director of the project, and with other members of the research team. Contributions mentioned here will be limited to problems different from those mentioned in B and C above. Individual contributions to group discussions and decisions can be difficult to identify particularly in a smoothly working group where synergism is prevalent. Nevertheless, I mention contributions to the overall structure of the Tutor including scenarios and flow diagrams, and to the graphical representation of knowledge and interest.

IV. MATHEMATICAL MODELING of PEDAGOGICAL PRINCIPLES

In this section I will discuss some important pedagogical principles that are expected to be relevant to

the Intentional Tutor. I will then discuss the mathematical modeling using fuzzy set theory of a small number of what appear to be the most important parameters for a first generation Tutor. An important research and development component of this project is to determine which models of which parameters yield the best results. The main objective is to maximize the subject's performance in learning the material presented by the Tutor. It is felt that an important intermediate objective is to help the subject achieve a flow state in the learning process. A flow state is a hyperattentive state where the person is so intent on the task at hand that he/she naturally disregards irrelevant stimuli such as hunger or the passage of time. The following is a brief summary of some important pedagogical principles.

1. During a given lecture or session at a tutoring device it is known that the first and last parts of the session have the most influence on the subject (Ellis, 1978). Consequently the Tutor will attempt to start and terminate a lesson with a topic of high interest to the subject. The terms primacy and recency are used in reference to the importance of the first and last parts of serial presentation.

2. Use of the Premack Principle (Ellis, 1978) in the choice of topics is good pedagogy. That is, the scheduling of topics of high interest to the subject immediately after topics of low interest is an example of the Premack Principle.

3. The theoretical learning curve (knowledge plotted against time) shows a sharp increase at first followed by a gradual leveling off as the subject reaches his/her

capacity for learning a given topic. The curve is called a negatively accelerated curve. In practice, there is sometimes an actual decline in performance when the process of study first starts. This can happen when the subject originally attempts to pursue inaccurate strategies and feels a sense of confusion before the pieces start coming together.

4. A high level of interest in a given topic is important for achieving a flow state. We will equate, for this discussion, interest with intrinsic motivation. We know that in general motivation includes extrinsic motivation which includes incentives such as fame and money. Motivation can be too high for efficient learning. There appears to be an optimal level of motivation for learning of any task, the optimum depending on the difficulty of the task. In general, the more difficult the task, the lower the level of motivation that will promote efficient learning (Logan,1976).

5. The notion of irrelevant drive plays a role in learning theory. Typical irrelevant drives are fear and anxiety. It is known that the right amount of irrelevant drive increases performance but that too much reduces performance (Logan, 1976).

Other factors to be considered as a strategy for a tutor are as follows:

1. In the first few lessons the Tutor will structure the choice and presentation of lessons to maximize the subject's opportunities for achieving a flow state in the learning process. One hypothesis is that the topics in the library for which the subject is most likely to achieve a

flow state are those for which the subject has a moderate knowledge and a high interest level.

2. After the first few sessions, an emphasis compatible with the Premack Principle will be given to those topics for which the distance from the current level of knowledge to the goal is greatest.

3. Some preference may be given to the topic the subject was studying when he finished his last session at the Challenge Work Station in the decision as to which topic to start with in the next session.

4. Considerations will be given to the relative importance of one topic as a prerequisite for another topic and other relationships between topics as discussed in the next section that is entitled Knowledge Representation by the Tutor.

A key feature of the Tutor that makes it possible to take into account many of the above mentioned parameters is its ability to measure or model the arousal/interest of the subject in the various topics. This model of interest will be used to help the subject achieve a flow state in the learning process. With this in mind, we will use fuzzy set theory techniques for selecting those topics for which the subject has a moderate knowledge and a high interest level as suggested in item 1 above. We now show how to model this concept in such a way that it can be easily programmed on a computer. We define functions "k" and "a" from L into appropriate evaluation sets to model the subject's "knowledge" of and "arousal" in the library

$$L = \{t_1, t_2, \ldots, t_n\}$$

consisting of various topics t_i. Let

$$k:L \dashrightarrow \{0,1,2,3,4,5,6,7,8,9,10\}$$

be a knowledge function from the "library" L into the given evaluation set where 0 corresponds to essentially no knowledge and 10 corresponds to mastery of the topic as displayed by an expert. Correspondingly, let

$$a:L \dashrightarrow \{0,1,2,3,4,5,6,7,8,9,10\}$$

be an arousal or interest function which models the subject's interest in the various topics in the library at a given time. Zero indicates no interest, while 10 is great interest. The function a can also be considered to be a mapping from L to a set of 11 shades of blue, as in the case of our graphic representation of interest. In this case, white indicates no interest and dark blue shows a great interest.

We start with the assumption that we have a given library L and models k and a of the subject's knowledge and interest in the library at the given time. We will mathematically display the hypothesis that moderate knowledge and high interest in a topic yields the highest payoff for studying that topic.

We now introduce a fuzzy set or approximate reasoning version of the linguistic descriptions "moderate" and "high". Fuzzy set theory, as introduced by (Zadeh, 1965) is an extension of conventional set theory, with the degree of membership for an element in a set taking a value anywhere in the interval [0,1], instead of 0 or 1 only. It was developed as a means of describing subjective or ill-defined concepts. Approximate reasoning appears to be a valid way to externalize peoples' perceptions of linguistic (qualitative) variables and their interactions. For example, a person's knowledge level in a given subject may

be described as "moderate". In distinguishing this from "high" we would infer that the level of knowledge is probably not as high as 8 or 9 and certainly not 10. We will now give a fuzzy set description of moderate. For convenience of notation let U = {0,1,2,3,4,5,6,7,8,9,10} be our universe of discourse. We may want the fuzzy set M = "moderate" to include 4,5 and 6, i.e. these integers should be (almost) completely in the fuzzy set, and to include only fractions of 2,3,7, and 8. A mathematical definition of the fuzzy set M is to define M as a function m from U into [0,1],

$$m:U \longrightarrow [0,1]$$

where for each t in U, m(t) gives the degree of membership of t in M. The following is a sample definition of M. We include corresponding definitions of the linguistic variables high, moderately high, and low.

Levels of Knowledge

U	m(t) = moderate	h(t) = high	moderately high	low
10	0	1	.1	0
9	.1	1	.5	0
8	.3	.9	1	0
7	.7	.7	1	0
6	.9	.3	.7	0
5	1	.1	.3	.1
4	.9	0	.1	.3
3	.7	0	0	.7
2	.3	0	0	.9
1	.1	0	0	1
0	0	0	0	1

Fig. 1

If we compose m, the function defining "moderate", with the knowledge function k as follows,

$$L \xrightarrow{k} U \xrightarrow{m} [0,1]$$

then we have a fuzzy subset of the library L, denoted

$$f_{mk}: L \longrightarrow [0,1]$$

consisting of those topics for which the subject has moderate knowledge (mk). Specifically, we have a measure $f_{mk}(t) = m(k(t))$ on a scale [0,1] of how close to a "moderate" level is the knowledge of the subject for the topic t.

Correspondingly, a fuzzy set description of a "high" level of interest is a function (*see Fig. 1*)

$$h: U \longrightarrow [0,1]$$

where for t in U, h(t) gives the degree of membership of t in "high". If we compose the function h that defines "high" with the interest function a as follows

$$L \xrightarrow{a} V \xrightarrow{h} [0,1]$$

then we have a fuzzy subset of the library L, denoted

$$f_{hi}: L \longrightarrow [0,1]$$

consisting of topics for which the subject has a high interest (hi). Specifically we have a measure $f_{hi}(t) = h(a(t))$ on a scale [0,1] of how close to a "high" level is the interest of the subject for the topic t.

Recall that we are trying to pick the topics for which the subject has moderate knowledge and high interest. This corresponds to taking the intersection of the fuzzy sets "moderate knowledge" (mk) and "high interest" (hi). In classical set theory, the <u>intersection</u> of two sets contains

those elements of the sets that are in one and in the other. But, by definition, an element of a fuzzy subset may be only partly in the set. Thus, by accepted conventions in fuzzy set theory, the degree to which an element is in the intersection of the two fuzzy sets "mk" and "hi" is equal to the minimum of the degree that it is in "mk" and the degree that it is in "hi". Because a fuzzy subset is a function, one can write

$$f_{mk \text{ AND } hi} = \min (f_{mk}, f_{hi})$$

The following chart gives the resulting values for

Moderate Knowledge AND High Interest

			Moderate Knowledge										
			0	.1	.3	.7	.9	1	.9	.7	.3	.1	0
			0	1	2	3	4	5	6	7	8	9	10
	1	10	0	.1	.3	.7	.9	1	.9	.7	.3	.1	0
	1	9	0	.1	.3	.7	.9	1	.9	.7	.3	.1	0
	.9	8	0	.1	.3	.7	.9	.9	.9	.7	.3	.1	0
	.7	7	0	.1	.3	.7	.7	.7	.7	.7	.3	.1	0
	.3	6	0	.1	.3	.3	.3	.3	.3	.3	.3	.1	0
High	.1	5	0	.1	.1	.1	.1	.1	.1	.1	.1	.1	0
Interest	0	4	0	0	0	0	0	0	0	0	0	0	0
	0	3	0	0	0	0	0	0	0	0	0	0	0
	0	2	0	0	0	0	0	0	0	0	0	0	0
	0	1	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0

This is a model for the possibility of the subject achieving a flow state while studying the topic in question.

Fig. 2

A way of interpreting the chart in Fig. 2 is that on a scale [0,1] the values in the chart tell how close you are to "moderate" knowledge and "high" interest. For example, for a given topic t, if the level of knowledge of the subject is 6 and that of interest is 7, then reading from the chart we find the value .7 which is the degree to which k(t)=6 and a(t)=7 is in the fuzzy set "mk AND hi".

Modeling Learning in the absence of a flow state.

It is known (Logan, 1976) that one can learn material even in the total absence of interest in the topic. We learn through exposure, but motivation does indirectly affect our learning. A potential model for maximizing short term payoff for studying a given topic is the criterion of inverse proportionality of the current level of knowledge and direct proportionality of the interest. A rationale for the statement about knowledge is that if we don't know much about a topic, then there is much to learn, and if we know a great deal about it, then there is not much more to learn (we are assuming a fixed domain of knowledge). In the case of interest level, we are assuming that with no interest, some learning will occur and that increased interest leads to increased learning.

Using fuzzy sets as illustrated in the previous example we can easily create a mathematical model of a potential criterion for maximizing short term payoff for studying a given topic. These mathematical models illustrate the use of fuzzy sets in the control problem for the Intentional Tutor.

V. KNOWLEDGE REPRESENTATION BY THE TUTOR.

Knowledge representation perhaps is the central issue in the design of expert systems, in particular tutors, and possibly the most important issue within AI itself. Since the Intentional Tutor will be using many of the existing tools and techniques of AI, but put together in quite different ways than with existing tutors, we must take a hard and creative look at the way knowledge is represented. Since the Tutor will be a multi-media device, it will use combinations of video disks, computer graphics, and text material (text files).

The question is: How does one organize this material so that it is readily accessible and so that topics or concepts can be selected on a basis of both interest and important learning concepts such as prerequisites, specializations, and analogy? A proposed solution is to use a genetic graph, introduced by Goldstein (1982) and explained below, and then properly add to it so that it becomes a glorified "table of contents" for the multi-media tutoring material. I suggest the name "graph of contents" for this augmented genetic graph.

The issue of "choice of topic" has been a major problem for all tutors. Traditional CAI gives no choice to the student and more sophisticated systems described as AICAI (e.g., Goldstein (1982)) deal will the problem by taking into account both the topic material and the current level of knowledge of the student. The Tutor will do a better job of this by also paying attention to the interest level of the student in the various topics. We will now define the notion of genetic graph which has built into it the capacities for utilizing the learning theories of Piaget (1977).

(a) Genetic Graphs

A genetic graph is a descendant of a semantic network. A semantic network is a knowledge representation model consisting of a network of nodes, standing for objects, concepts, and events, and links (or arcs) between the nodes. These links represents the relationships between the nodes. For our purposes a genetic graph will be what Goldstein (1982) refers to as an extended genetic graph. In particular the nodes can be described as knowledge nodes and the relationships are as follow:

prerequisite/postrequisite,
generalization/specialization,
simplification/refinement,
deviation/correction, and
analogy.

We can also group together as islands the knowledge nodes that are closely related and generally learned together. An example criterion for forming islands is to group together rules that have the same goal. These islands can then be related with links in a way similar to which nodes were related. In this way we can think of the islands as being the nodes of a higher order genetic graph and consequently we can construct a hierarchy of genetic graphs. Genetic graphs can be used for modeling the given library of topics as well as the student's knowledge of the given library.

The genetic graph can be viewed on command as a window on the computer monitor where a node is represented as a circle with a few key words in it identifying the concept and with arrows or arcs going to related nodes. When a given node has been picked, there will be a pointer method

that will direct the student to the appropriate media and location for study of the concept. If the selected concept appears in several locations, then the probability ranking principle will be used. This idea, discussed in Maron et al. (1986) is a document retrieval system, which ranks its documents, relative to each input query, in descending order of computed values for probability of relevance.

(i) The genetic graph as a basis for tutoring.

First, the genetic graph can be used as a tool in the "choice of topic" problem. (Another important factor is the student's interest in the various topics.) There are two methods for using the genetic graph for topic selection: the frontier and the web methods. The first refers to picking topics, concepts, facts, or rules that are on the frontier of a student's genetic graph. The web method is where the teacher seeks to explain the syllabus as a whole to give the student perspective and then later the teacher returns to various subtopics and presents those at deeper levels of understanding. The web method is preferred for a syllabus of facts and the frontier method preferred for procedural skills that have prerequisite relations.

Second, once a topic or more particularly a concept is selected the genetic graph supplies guidance for explaining the concept in more than one way to the student by means of relating it to its evolutionary predecessors. For example, students will have their choice of seeing an analogous concept, or a special case, or a generalization of the concept. The ability to explain a concept in more than one way is an important tutoring technique.

Finally, we give a discussion of the role of using a

genetic graph as a "graph of contents" in the large and complex field of knowledge representation. In AI, semantic nets, frames, and scripts are three declarative mechanisms for representing knowledge. Probably the most ambitious project to date in knowledge representation is taking place at MCC with their CYC project (Lenat et al, 1986). This project is the building, over the coming decade, of a large knowledge base of real world facts and heuristics and methods for efficiently reasoning. These ideas and techniques are very sophisticated technically and conceptually and they take a great deal of time and effort to accomplish. On the other hand, the "graph of contents" idea is technically much simpler, and although it has no built-in reasoning capacity it does have built-in pedagogical concepts. Furthermore, it could be implemented relatively quickly and easily and this apparently is a prerequisite for many of the intended uses of the Intentional Tutor by the Air Force.

VI. RECOMMENDATIONS

A. Mathematical Modeling of Pedagogical Principles.

Fuzzy set theory appears to be a reasonable tool for mathematically modeling the pedagogical principles and operational strategies for the Tutor. Since these represent the main operational controls for the Tutor it is essential that this work be continued. The work to date suggests only the feasibility of these ideas and now they need to be implemented. Dr. James Schroeder of Southwest Research Institute is a learning theory specialist, was an important resource for that aspect of this report, and is an active member of the research project. Consequently, it is recommended that Dr. Schroeder and I continue our work on identifying and modeling the pedagogical and control strategies for the Tutor.

The next stage of the process will be to choose a computer language and to program, probably as production rules as used in expert systems, the mathematical modeling as described above.

B. Knowledge Representation by the Tutor.

There are several steps needed to validate the use of a genetic graph as a "table of contents" for a given Library and solve the knowledge representation problem for the Tutor. It must be tried out, that is, an appropriate manual selected, a genetic graph built, and appropriate software developed to display the graph on a monitor. This graph would have built in pointers, activated by a mouse, that would either put on the monitor a specified section of the Library or list several references with associated values rating their relevance.

Building a genetic graph is a research field in itself but a procedure would be to start with a relatively simple document and create a genetic graph with the coordinated efforts of a team of experts identifying key concepts and interrelationships and with computer search techniques.

The next step would be to develop a written procedure along with appropriate software which would enable a group of experts in a given discipline, without the aid of a knowledge engineer or computer programmer, to take a document, manual, or textbook and relatively quickly develop a high quality "graph of contents".

F. REFERENCES

Ellis, H.C. (1978) Fundamentals of Human Learning, Memory, and Cognition, Wm. C. Brown Co., Dubuque, Iowa

Goldstein, I. (1982). The Genetic Graph: a representation for the evolution of procedural knowledge. D. Sleeman and J.S. Brown (Eds.), Intelligent Tutoring Systems. London: Academic Press.

Lenat, D., Prakash, M., & Shepherd, M. (1986). CYC: Using Common Sense Knowledge to Overcome Brittleness and Knowledge Acquisition Bottlenecks, The AI Magazine.

Logan, F.S. (1976). Fundamentals of Learning and Motivation, Wm. C. Brown Co., Dubuque, Iowa

Maron, M.E., Curry, S., and Thompson, P., (1986) An Inductive Search System: Theory, Design, and Implementation, IEEE Trans. on Systems, Man, an Cybernetics, Vol. SMC-16, No.1, Jan/Feb 1986.

Piaget, J. (1977) The Development of Thought. Equilibration of Cognitive Structures (A. Rosin, tran.). New York: Viking Press.

Zadeh, L.A. (1965). Fuzzy Sets, Information and Control, Vol. 8, 338-353.

1986 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT SUMMER SUPPORT PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

UNIVERSAL ENERGY SYSTEMS, INC.

FINAL REPORT

POLYNUCLEAR AROMATIC HYDROCARBONS IN PARTICULATE TURBINE ENGINE EXHAUST AND FROM COMBUSTION OF SINGLE COMPOUND FUELS

Prepared by:	William D. Schulz and Mary R. McGill
Academic Rank:	Professor of Chemistry; M.S. degree Candidate
Department and University:	Department of Chemistry Eastern Kentucky University
Research Location:	Headquarters Air Force Engineering Services Center, Engineering Research Directorate, Environics Division, Tyndall Air Force Base, Florida
USAF Research:	Mr. Surendra B. Joshi, Captain, Paul E. Kerch
Date:	September 30, 1986
Contract No.:	F49620-85-C-0013

Polynuclear Aromatic Hydrocarbons in Particulate Turbine Engine Exhaust and From Combustion of Single Compound Fuels

By

Dr. William D. Schulz and Miss Mary R. McGill

ABSTRACT

Particulate exhaust samples of turbine engines and a combustor rig were sequentially extracted with methylene chloride, benzene-methanol and toluene. No solvent was found superior to methylene chloride for the entire range of polynuclear aromatic hydrocarbons (PNAs) in these samples.

Turbine engine exhaust samples from three engines at four different power settings were extracted, concentrated and analyzed. Analysis was by G.C. (F.I.D.) and GC/MS. Blanks and samples contained large amounts of plasticizers and filter loadings were too low for reliable quantitation or identification by GC/MS. G.C. (F.I.D.) analysis and identification by retention times indicated that PNA concentration was highest at the "idle" power setting and that it was also higher for the one afterburning sample.

Samples of exhaust particulates from a bluff body combustor, burning pure acetylene, propane, propene, butane, 1-butene and isobutene were extracted, concentrated and analyzed. Soot yield of the fuels was not determined. Higher concentration of higher molecular weight PNAs were found for the aliphatic compounds with unsaturation and for isobutene, with unsaturation and branching.

Acknowledgments

We wish to thank the Air Force Systems Command and the Air Force Office of Scientific Research for sponsorship of this research. We also thank the entire Environics Division of the Air Force Engineering Services Center for their friendliness, helpfulness, and for allowing total access to their fine laboratory equipment. The civilian and military laboratory personnel worked beyond the call of duty to make room for, and cooperate with, a number of visiting scientists equal to their own number, in limited space.

We thank Lt. Cols. Brocato, Bramlett and Olfenbeutel for doing everything that made the summer very convenient and enjoyable. Doctoral candidate Abdulraluman Touati and Capt. Paul Kerch supplied total cooperation in supplying combustor laboratory samples, information and reference material. We thank Mr. Surrendra Joshi, Dr. Howard Mayfield, Lieutenant Glen Seitchek, Mr. Mike Henly and M/Sgt. Dan Stork for help and guidance that extended even to nights and weekends if needed.

I. Introduction

One area of interest to the Environics Division at Tyndall Air Force Base is the mechanism of soot formation. This interest is evidenced by the Combustor Laboratory, with equipment for precise conditions of combustion in a centerbody burner (or "bluff body combustor") and laser light scattering equipment for particle size and density measurements. Information gathered in soot formation studies is of fundamental importance in a number of areas: Air Quality (as measured by visibility), Health Effects (the concentration of hazardous materials adsorbed on air particulates formed by combustion) and Tactical Concern (possible abatement of turbine engine soot trails). Applications of such studies can range from waste incineration particulate formation to particulates emitted from tactical fighter aircraft turbine engines.

Another area of interest to the A.F.E.S.C. Environics Division is the total characterization (and Photoreactivity) of turbine engine exhaust. This area is primarily with concern for environmental and health effects of jet engine exhaust. Study of this area of concern has been contracted to Dynamac Corporation[1] for literature review and to Battelle Corporation[2-4] (Columbus) for actual implementation. The area not covered by the Battelle contract is the concentration of hazardous substances on turbine emitted particulates. Although this area is covered, at the pr ent, in a recent contract with Battelle Corporation, particulate samples would become available from May or early June 1986 sampling that would not be analyzed by Battelle, and these samples could not only be analyzed for hazardous materials (principally polynuclear aromatic hydrocarbons-"PNAs"), but could also be used for experiments to determine the optimum methods and conditions for extraction, concentra-

tion and analysis of particulate turbine exhaust samples.

I have had intensive experience in sampling, extraction and high resolution gas chromatography-mass spectrometry analysis of air particulate samples over an academic year and two summers under NSF grant ATM-801 4191, with Kent J. Voorhees, at the Colorado School of Mines in 1981-82. Because of this experience, Mr. S. B. Joshi of the Environics Division and Dr. C. W. Spicer of the Battelle Institute thought it desirable for me to experiment with extraction and concentration techniques and to attempt to analyze turbine exhaust samples preliminary to a contract effort by Battelle to analyze particulate samples collected in the late summer or fall. It was thought that this effort, at a minimum, would provide data necessary for successful turbine exhaust particulate collection and analysis in the Battelle contract work.

Miss Mary R. McGill is a Masters Degree candidate at Eastern Kentucky University who had recently completed my Analytical Separation Methods course. She expressed an interest in working on this problem as a part of her thesis research, under my direction, and was awarded the A.F.O.S.R. Graduate Fellowship by Universal Energy Systems upon approval by the Environics Division at Tyndall Air Force Base.

II. Objectives of the Research Effort

The initial objective of this research effort was to be evaluation of extraction techniques, solvents and extract concentration methods for analysis of samples from aircraft turbine engine exhaust particulates. Samples collected at Tinker A.F.B., Oklahoma, by the Battelle Corporation in May or early June were to be analyzed for polynuclear aromatic hydrocarbons (PNA). The analyses and method development both require gas chromatography-mass spectrometry for positive identification of analytes.

Upon arrival at Tyndall A.F.B., a serendipitous decision was made to begin extraction experiments with soot from the combustor laboratory, where pure aliphatic hydrocarbons (acetylene, propane, propene, 1-butene and isobutylene) were being burned in soot formation studies. Analysis of extracts of soot formed by these fuels indicated that the nature of the PNA content of the soot varied with the nature of the fuel.

A literature search indicated that, while soot formation had been studied for several pure compounds, essentially no research had ever been done involving qualitative and quantitative analysis of the PNAs formed from combustion of pure, simple aliphatic compounds. With this finding, extraction and analysis of soot samples from the combustion of pure fuels from the combustion laboratory because of at least equal importance to analysis of the turbine engine exhaust particulates.

The major objectives of this effort then became:

1. Method development and analysis of authentic particulate exhaust samples from aircraft turbine engines with the goal of at least identifying problem areas and advising Battelle personnel on optimum conditions for their own subsequent sampling and analysis.

2. Extraction and analysis of particulates formed by the combustion of pure aliphatic hydrocarbons in the combustor laboratory.
3. Establishing a pattern from the PNA content of pure fuel particulates and attempting to understand the significance of the pattern.
4. Attempting to devise a sampling procedure for the combustor laboratory that would give samples with "real world" combustion product implications.

III. Background

A. Extraction and Analysis

Particulate matter associated with incomplete combustion of organic materials (soot) is of environmental concern for several reasons, including esthetic, but the paramount concern is for the health-related aspects. Chief among the health related concerns is the association of polynuclear aromatic hydrocarbons ("PNAs") with soot. Many of the PNAs (also known as "PAHs" or "PACs", for polycyclic aromatic compounds) are proven carcinogens. The literature of analysis of PNAs in combustion products is voluminous. A 1981 review of chemical analysis of particulate carbon by Lee and Bartle[5] contains 127 references. The 1986 literature survey by Dynamac Corporation[1] for the A.F.E.S.C. Evironics Division contains 282 references.

Examination of the literature shows that there is no general agreement on the method of extraction of PNAs from particulate carbon. In the reviews cited above, there are references citing the comparison of soxhlet extraction vs. ultrasonication and recently supercritical fluid extraction[6] has been claimed to be more efficient for extraction. Solvents for Soxhlet extraction include cyclohexane, dichloromethane,

benzene-methanol, toluene, di and trichlorobenzenes and even napthalene at reduced pressure. All of these have been preferred by various groups. In a very recent study by Junk and Richard[7], pyridine, benzene, dichloromethane, dimethylsulfoxide, dimethylformamide and n-methylpyrrolidone were compared, both by Soxhlet extraction and ultrasonication, for extraction of PNAs from urban air, diesel and stack particulate samples. The results were ambiguous. The researchers conclude that "there is neither a single extraction technique nor a solvent that is effective for the efficient extraction of all organic compounds from all types of solid matrices". This is not surprising if one considers that the extractable compounds are formed with, and in a matrix of, the graphitic, insoluble portion of the particulate matter. The nature of a soot sample then, depends upon a number of combustion parameters and extraction efficiency can not be absolutely determined. This is in contrast to an adsorbtion model, in which extraction efficiency could be simply determined by "spiking" a known amount of deutero, or other "tagged" compounds into a known amount of sample. The conclusion would seem to be that, for an uninvestigated matrix, the analyst should, at the very least, sequentially extract his samples with known efficient solvents until satisfied with the recovery is maximized.

Analysis of extracts of exhaust particulates is indeed the analysis of complex mixtures. Capillary gas chromatography of such extracts typically gives dozens to a few hundred peaks. Samples are usually expensive and difficult to collect, so it is usually undesirable to apply clean up procedures that would remove all but one class of compounds. Analysis of extracts under these conditions requires high resolution gas chromatography-mass spectrometry. The mass spectra of isomeric PNAs are

essentially identical and identification must be made on the basis of retention time but the mass spectrum is necessary to assure a peak is due to a PNA for example, and not a nitro-PNA. In spite of recent publication of relative retention indices[8,9] it is essential to have authentic compounds for identification standards since relative retention can change with cross-linking of the stationary phase or with nominally identical stationary phases from different suppliers. In spite of earlier problems with the separation of isomeric PNAs (such as triphenylene and chrysene or benzo [a] and [e] pyrene) and even synthesis of stationary phases as exotic as bonded phase liquid crystals[10], common modern, bonded phase capillary columns essentially separate PNAs, up to a molecular weight of about 400 Daltons. Quantitation can only be done at high levels of precision by the use of standards due to the vastly different detector response for different PNAs. Ideally, quantitation would be with per-deutero internal standards for each compound of interest, by GC/MS.

B. Soot Formation and PNA Identity

The mechanism(s) of particulate carbon and PNA formation by combustion are certainly complex and are certainly far from being well elucidated. It is, however, generally accepted that soot formation involves two distinctly different mechanism steps; pyrolysis and pyrosynthesis. Organic molecules are partially cracked (pyrolysis) to give smaller, highly reactive species. These species, mostly radicals, recombine to give larger relatively stable aromatic compounds (pyrosynthesis). The literature of formation of PNA in combustion has been reviewed by Crittenden and Long[11]. It is generally accepted that aromatic rings in fuels accelerate soot formation and that both chain branching and un-

saturation result in increased soot and PNA formation. It is known that soot and PNA yields, as well as the identity of the PNA formed, can vary widely depending on temperature, residence time in a combustion zone and the nature of the organic fuel being burned. It is assumed that chemical characterization of PNA products of combustion can lead to valuable information for the understanding of the mechanisms of formation in combustion. Whether or not this is true, it is rapidly becoming apparent that there may be immediate and practical applications for the identification and quantification of PNA's from combustion. In 1982 Spitzer and Dannecher[12] found substantial amounts of three PNAs in the exhaust of Pratt and Whitney JT3D3 turbine engines on a Boeing 707 aircraft that are absent, or present in trace quantities, in urban air samples. Such findings obviously have health implications for personnel such as flight line and test cell mechanics.

A very recent report by National Bureau of Standards personnel suggests that the identities of PNAs in soot from arson can be used to identify the accelerant[13] or even to prove arson through the detection of an accelerant.

Findings such as these indicate that, as more information on the identity and relative quantity of different PNAs from different fuels and flame conditions become known, there are immediate and practical uses, as well as the theoretical possibilities for such information. One very exciting prospect is that of source apportionment of particulate pollution and another is that of proving that all soot producing process are not equally health-threatening per gram of soot produced.

IV. Experimental Materials and Methods

A. Reagents

High purity analytical standards were obtained from a variety of

sources, chiefly Chem Service, Inc., and the U.S.E.P.A. and were checked by G.C./M.S. for 70 eV spectra and retention indices. Solvents were HPLC grade (Fischer Chemical and Burdick and Jackson) and were blanked by FID-G.C. analyses.

B. Sample Extraction and Concentration

Samples were Soxhlet extracted employing Whatman 33x80 mm single thickness cellulose thimbles with a light glass wool plug at the top. The glass wool and thimbles were always pre-extracted with the extraction solvent, dried at 50°C/3 torr and tared. Solvent volume was approximately 250 mL and the cycle time for the extraction was approximately 6 minutes. Samples were extracted for 24 hours, replenishing solvent if necessary. After extraction, samples were concentrated to 5.0 mL in a Kudena-Danish apparatus, at 98°C in a water bath. Samples for GC/MS analysis were further concentrated by "blowing down" with dry nitrogen to a final volume of 0.20 mL.

"Free" soot samples, scraped from the combustor rig burner face, of between 0.090 and 0.10 grams were placed directly in the thimbles, weighed to 0.1 mg. and extracted. Variable weights of soot collected on various filter media were extracted by placing the folded filter into the soxhlet thimble. In all cases the samples were contained by a preweighed light plug of glass fiber. Solvent from pre-extractions of the thimbles and glass fiber plugs was routinely concentrated and analyzed by FID gas chromatography.

C. Sample Collection

Soot samples were collected from the combustor rig (center body combustor) located at the Tyndall A.F.B. Combustion Laboratory. The combustor has been described elsewhere[14]. In all cases, the soot col-

lection was secondary to light scattering studies of particle size and rate of formation. Samples were simply scraped off the face of the burner after any run in which a pure hydrocarbon fuel was burned. The samples were scraped onto a glassine weighing paper, weighed, the weighing paper folded around the sample, then placed in a small paper envelope and stored in at $0^{\circ}C$ in a glass container in a freezer until they were extracted.

Other samples were obtained on 47 mm. teflon coated "Millipore" filters, using a pumped probe set at fixed distances behind the burner head. Later, a filter holder was designed which fit into the stainless steel exhaust system, in order to obtain soot samples at room temperature. The filter holder was a 12 inch diameter aluminum plate with a centered 4 1/2"x 7 1/2" rectangular aperture, fitted with a frame and wing nuts to hold one half of a Schleicher and Schull binder-free quartz micro fiber 8x10 inch filter. The filters were conditioned in an oven at $400^{\circ}C$ for four hours, cooled, weighed and stored in a dessicator prior to use. Samples were collected, using this system, for burns of each pure hydrocarbon fuel, weighed, protected by $400^{\circ}C$ fired aluminum foil and stored with samples collected from the burner head, from the same burn, in a covered cake pan in a freezer at $0^{\circ}C$.

Exhaust particulate samples from 3 different engines at 4 different power settings were collected by personnel from Battelle Institute, Columbus, Ohio at Tinker AFB on 6/6 to 6/10, 1986. The filters were Pallflex Corporation 2500 QAST 6 inch diameter quartz fiber filters. The sampling rake and collection procedure are described elsewhere[2]. The filters were preweighed in a controlled humidity chamber and held at $150^{\circ}C$ for the duration of sampling. They were again equilibrated in

the humidity chamber and reweighed after sampling. The filters were stored in glassine envelopes enclosed by "zip lock" plastic bags. The samples were delivered to Tyndall A.F.E.S.C. in ordinary packaging by common carrier.

D. Analytical Conditions and Instruments

The gas chromatographic analysis were carried out using a J and W Scientific 30 M. x 0.25mM x 0.1 μM bonded film of DB-5 (J and W Scientific) in a Perkin-Elmer Sigma 2000 instrument using the flame ionization detector (FID). The injection mode was a 30:1 split ratio (because of lack of a splitless injection port liner) and the linear velocity of helium through the column was 35 cm/sec at 100°C. The injection port and detector were maintained at 300°C. The oven temperature program was 2 minutes at initial temperature of 50°C, then increased to 125°C at 5°C/min. and then increased to 300°C at 8°C/min., with a 10 minute final temperature hold. One or 2 μL samples were injected. The initial gas chromatography-mass spectrometry (GC/MS) analyses were performed on a Hewlett-Packard (H-P) 5987 mass spectrometer with capillary-direct sample introduction from an H-P 5880 gas chromatograph with control and data collection by an H-P RTE-6 data system. The chromatography was with the splitless injection mode, system purge on at 0.5 minutes with a 50M by 0.20 mM by 0.10 μM bonded phase "HP-5Ultra" column (Hewlett Packard Corporation). The linear He velocity was 2.5 cm/sec., at 200°C. The same temperature programs used in the G.C. FID analysis were used. The mass spectrometer was programmed to scan from 50 to 400 amu and operated in the 70 eV electron impact mode.

Later analysis were carried out using an identical capillary column with an H-P 5890 GC/5970 "MSD" system employing essentially identical G.C. and M.S. conditions. Except that the 5970 analyzer was operated at above five times the source pressure ($2\text{-}3\times10^{-5}$ torr) as the 5987 instrument ($4\text{-}5\times10^{-6}$ torr).

V. Results and Discussion

A. Combustor Soot Formation: Initial data from the analysis of PNAs formed in the combustion of the different pure compounds is summarized in Table I. Representative chromatograms from these analyses are shown in Figure I. The elution order in these chromatograms is essentially the same order as increase in boiling point. Many unidentified peaks are due to methylated PNAs for which standards were not available. The difference in PNA composition patterns is quite dramatic in the chromatograms. The 1-butene and isobutene soot extracts, particularly, are much richer in five and six ring components. The flame temperatures and fuel/air ratios are, unfortunately, not known for these samples.

The lower set of two chromatograms in Figure I is a comparison of sequential (1. methylene chloride, and 2. benzene/methanol) extracts of the same sample. The results are representative; in no case was more than 10% of a given PNA found in a second extraction, when solvents were benzene/methanol, toluene or dichloro benzenes. This observation, coupled with increased difficulty in solvent removal made methylene chloride the obvious choice for sample extraction with these samples. The same results were observed for the same experiments with turbine engine exhaust particulates. (Lack of space prevents presentation of more chromatograms).

Table II is a representation of the PNA content of soot collected for three fuels, (a) from the burner head, and (b) from a filter at essentially ambient temperature about 10 ft. down the 75 ft.^{3} min^{-1} exhaust stream. Soot samples from the other three fuels are currently being analyzed. Representative chromatograms are not presented due to the

page constraint. The quantitation is relative. Quantitation was to have been done with internal standards by SIM mode GC/MS. Instrument problems made this impossible in the 10 week period.

These soot extracts do not show the dramatically higher concentration of higher molecular weight PNAs for 1-butene. The data are consistent in the absence of lower PNAs in soot collected at low temperature, with high air dilution. This could either be attributed to loss of volitiles from the sample or to quenching of the pyrosynthesis mechanisms at the burner face. We believe that it is the latter. We believe very strongly that this work needs to be repeated with known flame conditions, collection temperatures and for a statistically significant number of samples.

B. Turbine Engine Exhaust Particulates: Table III represents FID data from turbine engine particulate exhaust analysis. Quantatative data is given only for benzo (a) and benzo (e) pyrene. These data were generated from external standards run several days before the samples. Detector response could not remain constant over the time interval and the data should be considered semi-quantitative at best. Samples of these extracts were run by GC/MS but due to a very high noise level, with numerous large noise spikes the PNAs were essentially below the limits of detection. Filter blanks of these samples were not subjected to the 150^{o}C collection temperature or airflow, but were otherwise handled the same, including installation in filter holders. The "blanks" all contained two different phthalates at more than 100 times the concentration of any analyte. Only comparative amounts of phthalate were found in actual samples. Several of these extracts were run on an HP 5970B instrument about six weeks after extraction. Compounds were well above limits of detection but there was some indication of sample

degradation/oxidation. This, coupled with the initial gross imprecision of sample mass leads to the conclusion that, although possible, reanalysis of these samples would be largely a waste of time. A ratio of benzo (a) and/or (e) pyrene to total particulate would be significant for comparison to other combustion sources.

Several hundred G.C. and several dozen GC/MS runs were done in the course of this investigation. Some of the data is still being organized and some analyses are still being done. Significant results and a report addendum will be presented to our Environics Division colleagues in a short time.

VI. Recommendations

A. Turbine Exhaust Samples: (1) A decision on the type of information required must be made before particulate sampling and the sampling technique can then either be biased toward; (a) real contribution of hazardous compounds to particulate pollution, or (b) precursor compound formation, by turbine engine exhaust. (2) If the goal is (a, above) then samples should be collected by air dilution cooling and sufficient residence time to duplicate the atmospheric equilibrium products. We believe that there is a definite quenching effect on PNA formation in combustion processes and that this should be avoided unless kinetic/mechanistic studies of soot formation are planned. (3) Sampling must be conducted in a way to obtain sufficient particulate loading (without vapor condensation) on filters to give sufficient concentrations for reliable analysis. We suggest 100 mg. as a reasonable lower limit. (4) Sample blanks should be handled in exactly the same manner, including heated air flow, as sample filters. (5) The method must give a reasonable degree of precision for the total particulate load on filters. (i.e., blanks should indicate no load to 0.1 mg. if samples are

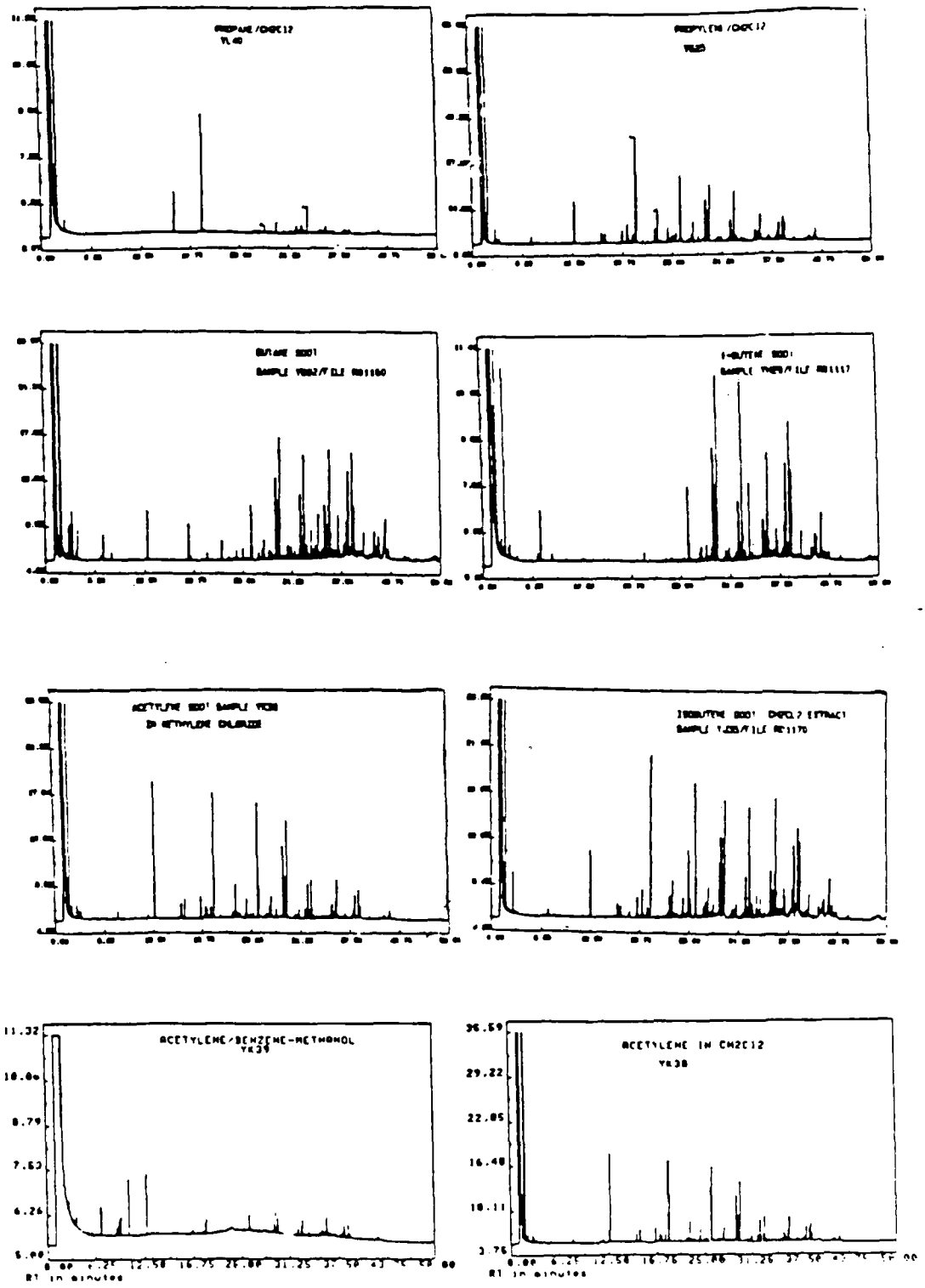
in the 100 mg range). (6) If there is to be a high volume of this type of work done, or if it should become routine, research should be done on automated supercritical fluid extraction/S.F. Chromatography/Mass Spectrometry for analysis of large numbers of samples. We believe that these rather new techniques could be automated to very high efficiency and precision if they seem economically attractive. (7) If small volumes of very light samples are done, unusual methods of concentration such as a "solid injector" or on column cyro trapping should be used.

We have been supplied (Capt. Paul Kerch) with a large number of filters (over 70) collected from J-79-GE-15A engines at various power settings, with and without ferrocene additive. The filter set contains duplicates for each collection conditions and each set has a very good sample mass correlation. The filters all have over 100 mg load and seem to have been subject to reasonable preservation. We recommend that these samples are worth funding for extraction and analysis. We have adequate facilities for extraction and concentration, appropriate standards and we can essentially dedicate a GC/MS system to analysis of extracts for as long as is required.

B. Combustor Laboratory Samples: We feel that our summer results from analysis of the particulate combustion products of pure hydrocarbons are very exciting, but that a great deal more experimentation is necessary for truly significant results. We believe that the difference in PNA contents of "burner face" and "downstream filter" samples is due to reaction quenching at the burner face. This could be tested and "quenched" products could be analyzed from a sampling probe that would give very rapid cooling at or in the sooting flame. In addition, the particulate products collected at essentially ambient conditions should be repetitively collected and analyzed. The collection should be done for precis-

ely determined combustion and collection conditions. The combustor laboratory at Tyndall A.F.B. has the capability for this sample production and collection. We will propose to extract, concentrate and analyze such samples and to attempt interpretation of the results in terms of soot formation mechanism.

Figure I
Chromatograms of Soot Extracts
from Pure Fuels



(a) No quantitative information is implied by these chromatograms.

Figure II

Chromotograms of Turbine Engine Exhaust
Particulate Extracts

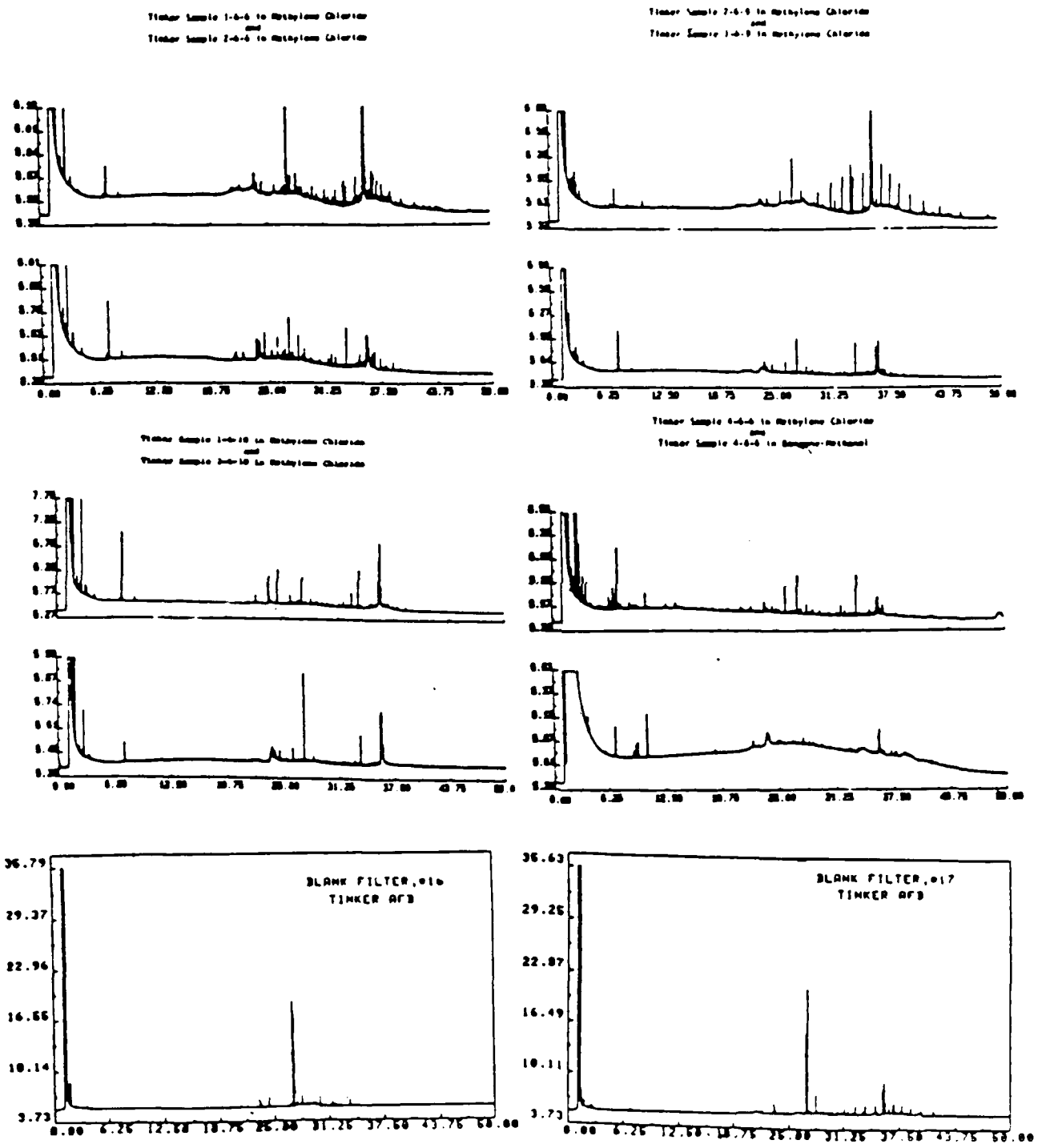


Table I

PAH's Identified by
Retention Times in Combustor
Soot Samples

Fuel	Area	PAH Identified
Propane	5516	Azulene
	1604	Benzo(ghi) Perylene
(Peaks Present- 22; Peaks Identified -2)		
Propene	33406	Napthalene
	5831	Azulene
	166825	1-4 Dimethylnapthalene
	2237	Phenanthrene
	1222	Benzo(b) Fluorene
	4370	Triphenylene
	1339	Benzo(a) anthracene
	8580	Benzo(b) fluoranthene
	21364	Benzo(e) pyrene
	2700	Benzo(a) pyrene
	29038	Benzo(ghi) perylene
	10459	Coronene
(Peaks Present- 42; Peaks Identified -12)		
Butane	25369	Napthalene
	2570	Benzo(b) fluorene
	9385	Triphenylene
	25372	Benzo(e) pyrene
	64560	Indeno (1,2,3,cd) perylene
	1704	Benzo(ghi) perylene
	8776	Coronene
(Peaks Present- 45; Peaks Identified -7)		
1-Butene	1033	1-5 Dimethylnapthalene
	7084	Fluroantnene
	2650	Triphenylene
	5578	Benzo(b) fluoranthene
	13090	Benzo(e) pyrene
	22253	Benzo(ghi) perylene
	20848	Coronene
(Peaks Present- 40; Peaks Identified -7)		
Acetylene	79830	Napthalene
	81201	1-5 Dimethylnapthalene
	15641	Fluoranthene
	18630	Benzo(e) pyrene
	1499	Benzo(a) pyrene
	5320	Triphenylene
	7695	Coronene
(Peaks Present -49; Peaks Identified -7)		
Fuel	Area	PAH Identified
Isobutylene	33598	Napthalene
	114843	1-5 Dimethylnapthalene
	2237	Fluorene
	18125	Fluoranthene
	23.07	Benz (b) fluroanthene
	8309	Chrysene
	16197	Benzo(a) anthracene
	20946	Benzo(k) fluoranthene
	3483	Benzo(e) pyrene
	79693	Benzo(ghi) perylene
	61115	Coronene
(Peaks Present-42; Peaks Identified -11)		

Table II

Relative Amounts of Major Polycyclic Aromatic Compounds Formed from Different Fuels at the Combustor Head and Found in Exhaust at Ambient Temperature(a)

PAH's	Fuels: Propane		Propene		1-Butene	
	Filter	Burner Face	Filter	Burner Face	Filter	Burner Face
Napthalene	- -	1.26×10^{7}*	- -	4.86×10^{4}*	- -	2.21×10^{6}*
Azulene	- -	8.74×10^{5}	- -	- -	- -	3.01×10^{5}
1-5 Dimethylnapthalene	4.45×10^{3}	- -	- -	3.09×10^{5}	- -	5.50×10^{6}
Anthracene	- -	- -	- -	- -	4.59×10^{3}*	- -
Fluoranthene	- -	- -	- -	- -	- -	8.59×10^{5}*
Pyrene	3.59×10^{3}*	- -	- -	- -	3.22×10^{3}*	- -
Benzo [b] Fluorene	3.04×10^{4}*	1.83×10^{5}*	1.13×10^{5}*	2.92×10^{4}	2.02×10^{4}	8.06×10^{6}
Triphenylene	6.33×10^{4}	9.60×10^{6}*	5.56×10^{5}	6.54×10^{5}	3.62×10^{4}	4.00×10^{5}*
Chrysene	1.16×10^{4}	- -	- -	- -	- -	- -
Benzo [a] Anthracene	1.42×10^{4}	2.92×10^{5}*	8.24×10^{4}	1.83×10^{4}	- -	- -
Benzo [b] Fluoranthene	- -	1.67×10^{6}*	- -	- -	6.02×10^{4}*	6.41×10^{5}
Benzo [k] Fluoranthene	2.29×10^{5}*	- -	- -	1.26×10^{5}	- -	- -
Benzo [e] Pyrene	- -	- -	1.59×10^{6}	1.25×10^{5}	1.6×10^{6}*	1.74×10^{7}*
Benzo [a] Pyrene	3.10×10^{5}*	3.45×10^{6}*	2.50×10^{5}	1.62×10^{4}*	2.68×10^{4}	9.82×10^{4}
Indeno [1,2,3,cd] Pyrene	1.51×10^{5}*	2.73×10^{6}*	- -	- -	- -	- -
Benzo [ghi] Perylene	- -	- -	1.31×10^{6}	2.21×10^{5}*	3.61×10^{3}*	3.84×10^{4}*

(a) Numbers are Simply Integrator Areas per Gram Sample.

(*) Identity confirmed by Mass Spectrometry

Table III
Analysis of Turbine Engine Exhaust Samples

Run #	Engine	Powersetting	ug/ft³ of Benzo (e) Pyrene	ft³ @ 150°C	Sample Mass Wt/Hg	Benzo Pyrenes: Benzo(a)pyrene ugPAH	Benzo(a)pyrene ppmEOM(a)	Benzo(e)pyrene ug PAH	Benzo(e)pyrene ppm EOM(a)
1-6-6	TF-41	Idle	36.03	35.15	0.12	. . .	. . .	915.0	1600.0
2-6-6	TF-41	30%	0.0761	43.20	-3.5	. . .	. . .	3.10	6.925
3-6-6	TF-41	75%	5.57	45.75	15.30	. . .	. . .	25.50	402.50
4-6-6	TF-41	Normal Rated	0.436	52.25	41.25	. . .	. . .	22.80	300.00
1-6-9	TF-30-103	30%	. . .	34.80	-0.85	3.30	214.25	. . .	. . .
2-6-9	TF-30-103	75%	1.499	58.50	10.40	3.15	?	86.50	?
3-6-9	TF-30-103	Normal Rated	0.0028	58.90	8.00	6.65	175.5	36.00	950.00
4-6-9	TF-30-103	Idle	1.132	32.70	-1.85	. . .	. . .	37.00	2435.00
1-6-10	TF-30-109	Idle	2.88	40.43	-3.95	2.15	390.00	116.75	21225
2-6-10	TF-30-109	30%	1.667	17.40	-0.60	. . .	. . .	29.00	1895
3-6-10	TF-30-109	75%	1.715	14.23	0.80	. . .	. . .	24.42	4200
4-6-10	TF-30-109	Normal Rated	. . .	51.30	0.90	0.53	27.50	. . .	. . .
5-6-10	TF-30-109	Normal Rated (roof top)	2.038	26.13	0.93	6.00	91.00	53.10	807.5
6-6-10	TF-30-109	After/burning	1.225	28.13	0.90	247.25	24725	32.0	3200

(a) Figures in ppm of extractable organic material. Data is not reliable due to very low filter loadings and difficulty equilibrating temperature/humidity conditions for weighing.

. . . not detected

References

(1) Drake, R. L. Characterization of Chemicals on Engine Exhaust Particles (final report) Dynamac Corporation, Rockville, MD. Sept. 1985.

(2) Berry, D. A., Holdren, M. W., Lyon, T. F., Riggin, R. M., and Spicer, C. W., Turbine Engine Exhaust Hydrocarbon Analysis, Tasks 1 and 2, Interim Report No. ESL-82-43, June, 1983.

(3) Spicer, C. W., Holdren, M. W., Lyon, T. F., and Riggin, R. M., Composition and Photochemical Reactivity of Turbine Engine Exhaust, Final Report No. ESL-TR-84-28, 1984.

(4) Spicer, C. W., Holdren, M. W., Lyon, T. F., and Riggin, R. M., Composition and Photochemical Reactivity of Turbine Engine Exhaust, Final Report No. ESL-TR-84-61, June 1985.

(5) Lee, M. L., and Bartle, K. D., "The Chemical Analysis of Particulate Carbon" Particulate Carbon, Formation During Combustion, Siegla, D. C. and Smith, G. W., Editors, Plenum Press, New York, N.Y., 1981.

(6) Hawthorne, S. B. and Miller, D. J., "Analytical Extractions of PAH From Environmental Solids Using Supercritical Fluids". Presentation at 192nd National Meeting of the American Chemical Society (Anaheim, CA., September 7-12, 1986) Paper #90, Division of Analytical Chemistry.

(7) Junk, G. A., and Richard, J. J., Analytical Chemistry, 1986, 58, 962-965.

(8) Lee, M. L., Vassilaros, D. L., White, C. M., and Novotny, M., Analytical Chemistry 1979, 51, 768-773.

(9) Rostad, C. E., and Pereira, W. E., Journal of High Resolution Chromatography and Chromatography Communications, 1986, 9, 328-334.

(10) Markides, K. E., Chang, H-C. Schregenberger, C. M., Tarbet, B. J., Bradshaw, J. S., and Lee, M. L., Journal of High Resolution Chromatography and Chromatography Communications, 1985, 8, 516-520.

(11) Crittenden, B. D., and Long, R., "Carcinogenesis-a Comprehensive Survey: Polynuclear Aromatic Hydrocarbons", R. Freudenthal and P. W. Jones (eds.) Raven Press, New York, (1976) p. 209.

(12) Spitzer, T., and Dannecker, W., Journal of High Resolution Chromatography and Chromatography Communications, 1982, 5, 98-99.

(13) Chesler, S. N., News Item in Chemical and Engineering News, Sept. 1, 1986, p. 25 and personal communication.

(14) Proctor, Charles and Touati, Abderrahmane, Survey of Soot Formation, I. Soot Precursor Mechanism, Contract No. F08635-83-C-0136, October, 1983.

1986 USAF - UES Summer Faculty Research Program

Sponsored by the

Air Force Office of Scientific Research

Conducted by the

Universal Energy Systems, Inc.

Final Report

Systems Effectiveness Concerning Vulnerability of Hardened Targets to a Variety of Weapons

Prepared by:	Meckinley Scott
Academic Rank:	Professor
Department and	Department of Mathematics
University:	University of Alabama
Research Location:	The Armament Laboratory,
	Eglin AFB, FL
USAF Researcher:	George C. Crews
Date:	July 30, 1986
Contract No.:	F49620-85-C-0013

Systems Effectiveness Concerning Vulnerability of Hardened Targets to a Variety of Weapons

by

Meckinley Scott

ABSTRACT

This work resulted from a study of several technical reports concerning vulnerability of hardened targets to a variety of weapons. The main objective is to refine and improve some of the methodologies in current use for evaluating target effectiveness. A target usually consists of two or more subsystems, each of which is made up of a number of components. Effectiveness of a subsystem is a random variable whose sample space is a subset of the interval [0,1]. Basic to the problem of evaluating the overall target effectiveness is the determination of the probability density functions of the random variables associated with the various subsystems of the target. In this work, appropriate methodologies are discussed to evaluate target effectiveness for a wider class of targets than those being presently considered. Computer programs presently used can easily be modified to handle a variety of more complex targets. Some areas for further research are briefly discussed.

Acknowledgments

I would like to thank the Air Force Systems Command and the Air Force office of Scientific Research for sponsorship of my research. In order to be meaningful, scientific research must be conducted in an intellectually stimulating environment. The Armament Laboratory at Eglin AFB provided this environment. I would like to thank Lieutenant Colonel John Russell and Dr. Massey Valentine for giving me the opportunity to work in the Analysis and Strategic Defense Division. Special thanks go to Mr. George Crews, Chief of the Vulnerability Assessments Branch, who provided me with all the support and directions to successfully complete this project. Mr. John Collins provided me with all the necessary technical background. His suggestions concerning areas for further research are highly appreciated. Among other persons who have also provided me with technical background include Mr. John Gagliano and Mr. David Jerome. I wish to express my sincere appreciation to Mr. Ralph McGuire, Mr. Mark Amend and Mr. Martin Schmidt for their encouragement and assistance during my entire stay here at Eglin AFB. Lastly, I would like to thank Mrs. Charlotte Montie for all the help given to me in various forms including the typing of this report.

I. Introduction

I received my Ph.D. in Statistics from the University of North Carolina, Chapel Hill, N.C. My research interests are mainly in two areas of applied probability. One of these deals with Stochastic Systems of waiting lines and the other concerns productivity and efficiency of machines operating under stochastic environments.

The research problem at the Armament Laboratory, Eglin AFB involved vulnerability assessments of hardened targets to a variety of conventional weapons. Of prime interest is the System effectiveness after weapons impact on the target. Obviously, such problems deal with evaluating the performance of complex Stochastic Systems. My research experience in applied probability and, in particular, the concepts used in analyzing various Systems concerning waiting lines and machines subject to breakdowns provided a good background to work on the problem assigned to me at Eglin Armament Laboratory.

II. Objectives of the Research Effort

The overall objective of this project is to provide appropriate methodologies for improving a computer program presently used to determine vulnerability of hardened targets to a variety of conventional weapons. The computer program entitled "Weapon Effects/Target Vulnerability (WETV) Computer Program" has a number of shortcomings. Revision and modification of the present version of WETV is needed not only for improving the accuracy of the output data but also to accommodate increasing complexities in target designs.

My individual objectives are:

1. Refine and present a more formal description of some of the methods used in WETV. The purpose is to help readers and users gain a better insight and understanding of their use.
2. Introduce more types of subsystem component configurations, and provide a general method to analyze those configurations.
3. Point out and rectify shortcomings of WETV in those cases where different components may be functionally dependent. If components are functionally dependent, a more general method of analysis is necessary.

III. General Background

As mentioned in the preceding section, the computer program WETV is used to determine the vulnerability of hardened targets to a variety of conventional weapons. Hardened targets include buried shelters, concrete bridges and command, communications, and control facilities. The WETV algorithms determine warhead penetration through soil and concrete media, detonation locations from time delay or void sensing fuzes, and component damage caused by warhead fragments, blast overpressure and conical shaped charge jets.

Use of WETV program involves generating a compatible geometric model of target components including breach plates from the so-called FASTGEN computer program. Also, weapon trajectories are simulated using shotlines generated by FASTGEN that intercept critical components. The output from WETV is the residual target performance (RTP) calculated after weapon impact on the target along one shotline.

RTP is a function of damage to each individual component, probabilities of survival for the components and the logical configurations of components required to perform necessary functions often with degraded performance.

For the purpose of developing appropriate methodologies for target effectiveness, it is sufficient to describe the structure of the target using only general terms. A target usually consists of a large number of components which interact in a complex manner to achieve a specific goal. These components are divided into a number of mutually exclusive groups called phases. Each phase must be operational (possibly with degraded effectiveness) for the target to function. Different phases operate independently of one another and they may have completely different structures. Each phase, in turn, is divided into a number of modes. A mode may require some or all of the components of a phase. Residual modal performance is strictly determined by that of its components.

IV. System Effectiveness

RTP is a function of damage to each of the individual phases. The residual performance of a phase is a random variable whose sample space is a subset of the interval [0,1]. It consists of all residual performance values of its modes and also zero. The value zero arises from the possibility when all modes of a phase are damaged.

Basic to the problems of evaluating the overall (expected) RTP is the determination of the probability density function for each of the random variables (e.g., the proportion of the number of components that survive a hit in a phase when all components are identical with no redundancy) associated with the various phases of the target. The expected residual

performance of each phase can then be determined. For a target with n phases, let X_i denote the random variable representing the residual performance of the ith phase. The RTP is a random variable (X) defined by

$$X = \prod_{i=1}^{n} X_i \qquad (1)$$

Since different phases of a target operate independently, the random variables $X_1, X_2, \ldots, X_n$ are mutually independent. It follows that the expected RTP (E) is given by

$$E = \prod_{i=1}^{n} E_i \qquad (2)$$

where $E_i = E(X_i)$, is the expected effectiveness of the ith phase.

Note that for a phase with m modes, $E_i = \sum_{j=1}^{m} x_j P(x_j)$, where x_j is the residual performance value of the jth mode of the ith phase and $P(x_j)$ is the probability of survivability of those components of this mode with value x_j for the residual performance. In the special case where all components of the jth mode are identical with no redundancy of components, x_j is the proportion of the components that survive a hit and $P(x_j)$ is the corresponding probability of survival of those components.

The configuration of the modes of a phase usually meets one of the three criteria described below. In more complex situations, it is conceivable that a combination of these criteria may be needed.

a. Hierarchical

As the name suggests, in this set-up, the first (optimal) mode will be utilized if it is available. If not, a search is made for the next best mode that survives. Note that there is a possibility that no mode survived.

Let m be the number of modes of the ith phase. Also, for $j = 1,2,\ldots m$, let e_j and p_j denote, respectively, the residual performance value and probability of survivability of the jth mode of the ith phase. Under the conditions that survivabilities of various modes of the same phase are independent events, it follows from the definition of the expected value of a random variable that

$$E_i = e_1p_1 + e_2(1-p_1)p_2 + e_3(1-p_1)(1-p_2)p_3 + \ldots$$

$$+ e_m(1-p_1)(1-p_2)\ldots(1-p_{m-1})p_m + 0(1-p_1)(1-p_2)\ldots(1-p_m).$$

This formula can be compactly written as

$$E_i = \sum_{j=1}^{m} e_j p_j \prod_{k=0}^{j-1} (1-p_k), \quad (p_0 = 0). \tag{3}$$

Note that

$$p_j \prod_{k=0}^{j-1} (1-p_k)$$

is the probability that the most efficient mode that survives is the jth one, and

$$\prod_{k=1}^{m} (1-p_k)$$

is the probability that no mode survives. Thus, in this case the probability density function of the random variable representing the residual performance of the ith phase is easily determined. The present version of WETV makes use of formula (3) but there was no mention whatsoever about possible use of the same components. In many targets of interest modal dependencies within a phase are quite prevalent, such as the use of the same components

by different modes. In these cases it would be incorrect to use formula (3).

When two or more modes utilize the same component, survivabilities of these modes may no longer be independent events. In such a case, the survivability of any mode, in general, is a composite event and conditional probabilities of various events are needed to determine the probability of survivability of the most efficient mode. To illustrate this point, let us consider a simple case where a phase has two modes, 1 and 2. Both modes require component A and, in addition, mode j, $(j = 1,2)$ requires component M_j. Mode 1 functions if and only if both components A and M_1 work and similarly for Mode 2. For convenience, we will also simply denote by A the event that component A works. The event that component A does not work will be denoted by A^c. Similar meanings are given to M_j and M_j^c, $(j = 1,2)$. Also, let U_j, $(j = 1,2)$ denote the event that the most efficient available mode is the jth one. U^c will denote the event that neither Mode 1 nor Mode 2 is available. Avoiding the use of the intersection symbol for the joint occurrence of events, we write

$$U_1 = AM_1.$$

The right-hand-side denotes the event that both components A and M_1 are working. Similarly,

$$U_2 = AM_1^cM_2,$$

where the right-hand-side denotes the event that component A works, M_1 does not and component M_2 does. The event U^c that no mode is available may be represented as the union of two disjoint events A^c and $AM_1^cM_2^c$, i.e.

$$U^c = A^c \cup AM_1^cM_2^c.$$

$AM_1^cM_2^c$ is the event where component A works but components M_1 and M_2 do not. Let $P(A) = p$, the probability that component A works, and $P(M_j/A) = g_j$ the conditional probability that component M_j works given that component A works, $(j = 1,2)$. Now,

$$P(U_1) = P(AM_1) = P(A)P(M_1/A) = pg_1$$

$$P(U_2) = P(AM_1^cM_2) = P(A)P(M_1^c/A)P(M_2/AM_1^c)$$

$$= P(A)P(M_1^c/A)P(M_2/A).$$

Since components M_1 and M_2 are assumed to operate independently, i.e.,

$$P(U_2) = p(1 - g_1)g_2.$$

Also,

$$P(U^c) = P(A^c \cup AM_1^cM_2^c) = P(A^c) + P(AM_1^cM_2^c)$$

$$= P(A^c) + P(A)P(M_1^c/A)P(M_2^c/A)$$

$$= (1 - p) + p(1 - g_1)(1 - g_2).$$

It is easy to check that

$$P(U_1) + P(U_2) + P(U^c) = 1.$$

The above example suggests that the expected phase residual phase performance should be expressed simply in its general form. More specifically, for the ith phase consisting of m modes,

$$E_i = \sum_{j=1}^{m} e_jP_j,$$

where e_j is the residual performance value of the jth mode and P_j is the probability that the jth mode is the most efficient one that survives.

Formula (3) is a special case of (4) when the modes of a phase function independently of one another. Note that for a complex System where different modes utilize common Subsystems (components), the calculations for determing the P_j's could be very laborious and time-consuming.

b. Parallel

In this set-up, all modes share in determining the residual performance value of the phase. To calculate the expected residual phase performance, all possible combinations of modes are considered. The residual performance value of each of the combinations is the sum of the residual performance values of all modes (but not greater than 1) for that combination.

For a phase with m modes, say $M_1, M_2, \ldots, M_m$, there are 2^m possible combinations for modal survivabilities. One of these is the case where none of the m modes survives. In this case, the phase effectiveness assumes a value zero with probability $P(M_1^c M_2^c \ldots M_m^c)$, where $M_1^c M_2^c \ldots M_m^c$ denotes the event where all modes are unavailable for use. An event such as $M_1 M_2 M_3^c \ldots M_m^c$ is one where only modes M_1 amd M_2 survive. $M_1 M_2 \ldots M_m$ is the event that all modes survive and so on. Using the combination symbol

$$\binom{m}{r} = \frac{m}{r!(m-r)!},$$

there are $\binom{m}{r}$ possible combinations in which exactly r modes survive out of m, $(r = 0,1,2,\ldots,m)$. For illustration purposes, consider the case where the ith phase consists of three modes with residual performance values e_1, e_2 and e_3 respectively. The expected effectiveness for this phase is

$$E = (0)P(M_1^cM_2^cM_3^c) + (e_1)P(M_1M_2^cM_3^c) + (e_2)P(M_1^cM_2M_3^c)$$

$$+ (e_3)P(M_1^cM_2^cM_3) + \text{Min}(1,e+e_2)P(M_1M_2M_3^c) +$$

$$\text{Min}(1,e_1+e_3)P(M_1M_2^cM_3) + \text{Min}(1,e_2+e_3)P(M_1^cM_2M_3)$$

$$+ \text{Min}(1,e_1+e_2+e_3)P(M_1M_2M_3), \qquad (5)$$

where Min(a,b) denotes the minimum of a and b. This formula does not assume independence of the three modes. Again here, if the modes share some common components, the probability for each case of modal survivability will have to be calculated using conditional probabilities. When modal survivabilities can be assumed to be independent and that the sum of all modal residual performance values does not exceed 1, equation (5) above simply reduces to

$$E_i = \sum_{j=1}^{m} e_j a_j, \qquad (6)$$

where e_j is the residual performance value of the jth mode and a_j is the corresponding probability of its survivability.

c. Series

For a phase in which modes are in series configuration, Subsystems that use common components are considered to belong to the same mode. Consider a phase with m modes in series configuration and let $e_1,e_2,\ldots,e_m$ be the residual performance values of the modes. The residual performance of the jth mode may itself be considered as a random variable (Y_j), $(j = 1,2,\ldots,m)$ whose sample space is $\{0,e_j\}$, a set consistsing of only two points 0 and e_j. Clearly, the random variables $Y_1,Y_2,\ldots,Y_m$ are mutually independent and

$$E(Y_j) = (e_j)r_j + (0)(1 - r_j) = e_jr_j, \ (j = 1,2,\ldots,m), \qquad (7)$$

where r_j is the probability of survivability of the jth mode. The random variable X_i which denotes the effectiveness of the ith phase is the produce of the Y_j's, i.e.,

$$X_i = \prod_{j=1}^{m} Y_j. \tag{8}$$

Thus,

$$E_i = \prod_{j=1}^{m} e_j r_j. \tag{9}$$

V. General Assumption for Configuration of Components of Subsystems

In the present version of WETV all Subsystems and components of Subsystems belonging to a mode are assumed to be in a series configuration. This assumption is quite restrictive. Equally restrictive are those targets which use parallel configuration of components. A more general configuration is the so-called k-out-of-n structure. Here the structure functions if and only if at least k of the n components function. Note that a parallel structure is a 1-out-of-n structure and a series structure is an n-out-of-n structure. An airplane which is capable of functioning if and only if two of its three engines are functioning is an example of a 2-out-of-3 System. The three engines need not necessarily be identical. To illustrate, consider a 2-out-of-3 System with components C_1, C_2 and C_3. The event that the System will function can be represented as

$$C_1C_2 \cup C_1C_3 \cup C_2C_3,$$

i.e., the functioning of at least one of the combinations C_1 and C_2 or C_1 and C_3 or C_2 and C_3. The probability of this event

$$P(C_1C_2 \cup C_1C_3 \cup C_2C_3) = P(C_1C_2) + P(C_1C_3) + P(C_2C_3) - 2P(C_1C_2C_3). \tag{10}$$

References

1. Barlow, R. E. and F. Proschand, Statistical Theory of Reliability and Life Testing, New York, Holt, Rinehart and Winston, Inc., 1975.

2. Collins, J. A., H. E. Page, and C. M. Guglielmoni, "Weapons Effects/ Target Vulnerability (WETV) Computer Program," Volume 1 - User Manual, Technical Report, Datatec Inc., Fort Walton Beach, Florida, December 1985.

3. Feller, W., An Introduction to Probability Theory and Its Applications, New York, John Wiley and Sons, 1968.

4. Gregory, L. D., "Hardened Sam Site Vulnerability Analysis (U)," Volume II - SAMV User Manual, Technical Report, Vought Corporation, Dallas, Texas, April 1981.

5. Mann, N. R., R. E. Schafer, and N. D. Singpurwalla, Methods for Statistical Analysis of Reliability and Life Data, New York, John Wiley and Sons, 1974.

6. Williams, R. L., "System Functional Modeling for Vulnerability Analysis, WETV and GO," Technical Report, Kaman Sciences Corporation, Colorado Springs, Colorado, June 1985.

7. Williams, R. L. "System Functional GO Models for Vulnerability Analysis," Technical Report, Kaman Sciences Corporation, Colorado Springs, Colorado, October 1985.

1986 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT SUMMER SUPPORT PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

Resolution of Laser Beam Intensity Spots by the Target Plate Measurement Technique

Prepared by:	Martin Andrew Shadday Jr.
Academic Rank:	Assistant Professor
Department and University:	Dept. of Mechanical Engineering, University of South Carolina
Research Location:	The Air Force Weapons Laboratory, Kirtland AFB, New Mexico
USAF Research:	First Lt. Jerry R. Couick
Date:	August 5, 1986
Contract No:	F49620-85-C-0013

Resolution of Laser Beam Intensity Spots by the Target Plate Measuremnt Technique

by

Martin A. Shadday Jr.

ABSTRACT

The intensity distribution in the beam of a high power CO_2 continuous wave laser can be measured by the target plate measurement technique, developed at the Air Force Weapons Laboratory. The laser beam is directed on to the polished surface of a thin metal plate, and its intensity distribution is measured by observing the transient response of temperature sensitive paint on the rear surface. In order to predict the intensity distribution on the front surface of the plate, one-dimensional heat conduction is assumed. The capability of this technique to resolve narrow intensity spots was investigated by modeling the heat transfer in the target plate numerically. The importance of conduction parallel to the plate surface is quantified.

Acknowledgements

I would like to thank the Air Force Systems Command and the Air Force Office of Scientific Research for sponsorship of my research. The Air Force Weapons Laboratory is a stimulating environment in which to do research. I would like to thank Dr. Patrick Yail and Major Thomas Edwards for giving me the opportunity to do some interesting and useful research. First Lieutenant Rod Couick was especially helpful. His cooperation and guidance significantly enhanced the positive aspects of this research opportunity.

I. Introduction

I received my Ph.D. in Mechanical Engineering from the University of Virginia. My thesis research was in the field of high speed rotating fluids. I am presently an Assistant Professor at the University of South Carolina, and my research has been primarily concerned with computational fluid dynamics and heat transfer.

The research problem at the Air Force Weapons Laboratory, AFWL, was the study of the spatial resolution of the high energy laser target plate measurement technique. This involved comparison of one and two-dimensional heat transfer models of the target plate. Because of my experience with numerical models, I was assigned to work on this problem.

II. Objectives of the Research Effort

Two separate projects were identified as research goals at the beginning of the Summer research period, and both tasks were satisfactorily completed. The two projects were:

1. Instrumentation and calibration of a subsonic wind tunnel.
2. Numerical modeling of the heat transfer processes during irradiation of a metal target plate by a high energy laser. The target plate measurement technique is a way of measuring the intensity distribution in the beam of a high energy laser. The numerical model of the target plate was used to determine the resolution limits of this measurement technique.

III. Subsonic Wind Tunnel

The wind tunnel in the AFWL Laser Effects Laboratory provides a subsonic airstream over targets irradiated by high power CO_2 continuous wave lasers. The wind tunnel is an open-loop blowdown type, designed to provide flow of air in the test section, at atmospheric pressure, up to Mach .9.

A pitot probe was put in the plenum chamber, upstream of the wind tunnel nozzle. This probe allows continuous measurement of the stagnation pressure, and with atmospheric pressure, calculation of the test section Mach number. A pitot/static pressure probe in the test section was used to calibrate the test section Mach number, calculated from the measured plenum stagnation pressure. The results of this calibration are shown in figure (1). The agreement between the calculated and measured Mach numbers is very good.

The boundary-layer thickness on a target in the test section was measured by traversing a pitot/static pressure probe normal to the plate. Results of a test with a free stream Mach number of .47 are shown in figure (2). The boundary-layer thickness is approximately .2 inches, and the velocity profile is fairly flat with a very steep gradient near the wall. This profile is characteristic of turbulent flow, and the Reynolds number based on the distance from the nozzle exit is approximately 10^6. This is in turbulent range.

The final test on the wind tunnel was a measure of its transient performance. The output of the plenum pitot probe was recorded by an x:y:t recorder. There is an initial pressure transient with a duration of approximately three seconds. Thereafter the test section Mach number is steady within approximately two percent.

IV. Target Plate Measurement Technique

The target plate measurement technique is a way of measuring the spatial intensity distribution in the beam of a high power continuous wave laser. The laser beam is directed on to the polished surface of a thin metal plate. The rear surface of the plate is painted with a coat of temperature sensitive paint, and the temperature response of the rear surface is recorded with a high speed movie camera.

The camera records a set of lines that are a specific isotherm at different elapsed times. With the assumption of one-dimensional heat conduction through the plate, each of the isotherms can be converted to an isointensity line on the front surface of the plate. With narrow intensity spots in the laser beam, conduction in the target plate parallel to the plate surfaces can become important. This multi-dimensional heat conduction introduces errors in the predicted laser beam intensity distribution, due to the assumption of one-dimensionality.

V. Numerical Heat Transfer Model

The objective of this research was to quantify the resolution accuracy of the target plate measurement technique of narrow intensity spots in the laser beam. This was done by modeling the heat transfer in the target plate two-dimensionally and comparing the results with those of a one-dimensional model.

The problem considered in this study is the two-dimensional transient thermal conduction in a thin metal plate, exposed to laser radiation on one side. Figure (3) is a schematic of the problem. The incident laser radiation pattern consists of a central intensity spike on a background radiation field with uniform intensity. The problem is modeled in axi-symmetric coordinates. The spatial intensity distribution is assumed to be independent of time. The radial extent of the computational domain is large in comparison with the radial dimension of the intensity spike, consequently the outer edge of the target plate is treated as an insulated boundary. Natural convection and thermal radiation losses are neglected on the front surface of the target plate. These losses are negligible in comparison with the absorbed laser energy. Natural convection and radiation losses from the rear surface of the target plate are included in the model.

The transient thermal conduction in the target plate is modeled numerically by the method of finite-differences. The model is explicit with a uniform mesh spacing. The computational grid consists of ten node points distributed through the thickness of the plate and fifty node points distributed radially. The numerical model calculates the temperature of each of the nodes at a specific time. Finite-difference equations for each of the nodes are derived from application of the conservation of energy prinicple to the incremental control volume surrounding each node:

$$\rho\, C_p\, (\text{Vol})\, \frac{\Delta T}{\Delta t} = \text{net heat rate in} \qquad (1)$$

The right side of equation (1) is a function of the node under consideration and its immediate neighbors. The explicit formulation of the finite-difference equations expresses the nodal temperature, at a new time level, in terms of the temperatures at the old time level:

$$T_{ij}^{n+1} = T_{ij}^{n} + \frac{\Delta t}{\rho\, C_p (\text{Vol})}\, (\text{net heat rate in}) \qquad (2)$$

The numerical model solves the transient problem by marching forward in time, updating the temperature distirubtion in the plate as it progresses. The model is solved on a digital computer. The result of the model is the temperature distribution in the target plate at a specified elapsed time.

VI. Results

The laser intensity distribution on the front surface of a target plate is predicted from the transient temperature distribution on the rear, assuming one-dimensional heat conduction through the plate.

One-dimensional heat transfer is necessary in order to assume that there is a one-to-one correlation between a back surface plate temperature at a specified time and an incident laser intensity on the front surface.

The resolution capability of the target plate measurement technique was determined by comparing the incident laser intensity distribution on the front surface of a target plate with the predicted intensity distribution, based on the plate rear surface temperature distribution of the axi-symmetric numerical model. The predicted intensity distribution is normalized by the difference between the peak and background intensities on the front surface of the plate..

The numerical model is dimensional, but the presentation of results is considerably simplified if teh results are non-dimensionalized. The governing equation for transient heat conduction is:

$$\frac{1}{\alpha_o}\frac{\partial T}{\partial t} = \nabla^2 T \qquad (3)$$

If this equation is non-dimensionalized by the following variables:

$$T^* = \frac{T-T_o}{q_o L/k^2} \qquad t^* = \frac{\alpha_o t}{L^2} \qquad r^* = \frac{r}{L} \qquad (4)$$

$$Z^* = \frac{z}{L}$$

The resultant differential equation is without coefficients:

$$\frac{\partial T^*}{\partial t^*} = \nabla^2 T \qquad (5)$$

The non-dimensional solution is independent of the target plate material. Constant thermal diffusivity was assumed in equation (3) in order to simplify the non-dimensionalization process. The numerical model

accounts for the temperature variation of the physical properties of the plate. The background laser intensity was varied form 7.5 kw/cm^2 to 30 kw/cm^2, the ratio of the peak spot intensity to the background intensity was varied from 1.2 to 1.8, and the model was run for both stainless steel and copper target plates. The non-dimensional numerical results were found to be independent of the laser intensity level and the plate material. The non-dimensional time accounts for the thermal diffusivity of different target plate materials.

Figures (4) through (7) show the effect of thermal conduction parallel to the plate surfaces on the resolution of a laser intensity spot by the target plate intensity measurement technique. The widths of the central intensity spots vary from one to four plate thicknesses. The spike shape for all cases is a cosine function. This shape was chosen for its computational simplicity. In each case, predicted spot profiles are shown for several different elapsed times. The dimensional times in parenthesis apply to a stainless steel plate with a thickness of .1651 cm. The ability of the target plate to resolve narrow intensity spots in a laser beam drops off drastically for spot widths less than two plate thicknesses. The resolution of the spots also drops off with time. The error in predicting the peak intensity, relative to the background, of a spot with a width of one plate thickness would be in the range of seventy or eighty percent.

Figures (8) and (9) are results for two different materials, stainless steel and copper. The central intensity spot is the same for both plates, a square pulse with a width of four plate thicknesses. The background intensity is 7.5 kw/cm^2 and the peak to background intensity ratio is 1.2. The non-dimensional elapsed times in both cases are the same: .1289, .3233, and .6446. The non-dimensional solutions for both

materials are the same. The dimensional times in seconds are also shown, and they are drastically different for the two materials. Stainless steel is a much better target plate material than copper because the same resolution of an intensity spot occurs much later, and changes in the back surface temperature of the plate takes place at a slower rate. A second factor recommending stainless steel over copper is the actual temperatures of the back surfaces of the two plates. For copper after .0076 seconds, the peak rear surface temperature has changed three degrees celsius from the initial temperature of 20 degrees, and the difference in temperature between the peak and base of the intensity spot is approximately one-half of a degree celsius. For the same non-dimensional time with stainless steel, the peak rear surface temperature is 112 degrees, and the peak to base temperature difference of the spot is approximately 15 degrees. The spot is certainly resolvable with temperature sensitive paint on the stainless steel target plate and probably not detectable at all after .0076 seconds on the copper plate. The theoretical resolution of the intensity spot is the same for both plates.

VII. Conclusions and Recommendations

There are three factors that strongly influence the ability of the target plate laser intensity measurement technique to resolve narrow intensity spots in a laser beam: plate thickness, thermal diffusivity, and time. Thin plates better resolve narrow spots and lowering the thermal diffusivity slows down the rate of temperature change on the back surface of the plate. The numerical results also demonstrate the importance of taking the data quickly. The resolution of a spot decreases with increased elapsed time. Only the effect of conduction parallel to the plate surfaces was considered in this study, and the results are independent of the manner in

which the rear surface temperature of the target plate is recorded. The results of this study are, therefore, the theoretical resolution limits of the target plate.

This study considered only the resolution of a single intensity spot in a laser beam. The resolution of multiple closely spaced spots should be investigated. This would require a three-dimensional model. The possibility of reducing the effect of multi-dimensional heat conduction in the target plate by use of composite target plates should also be studied.

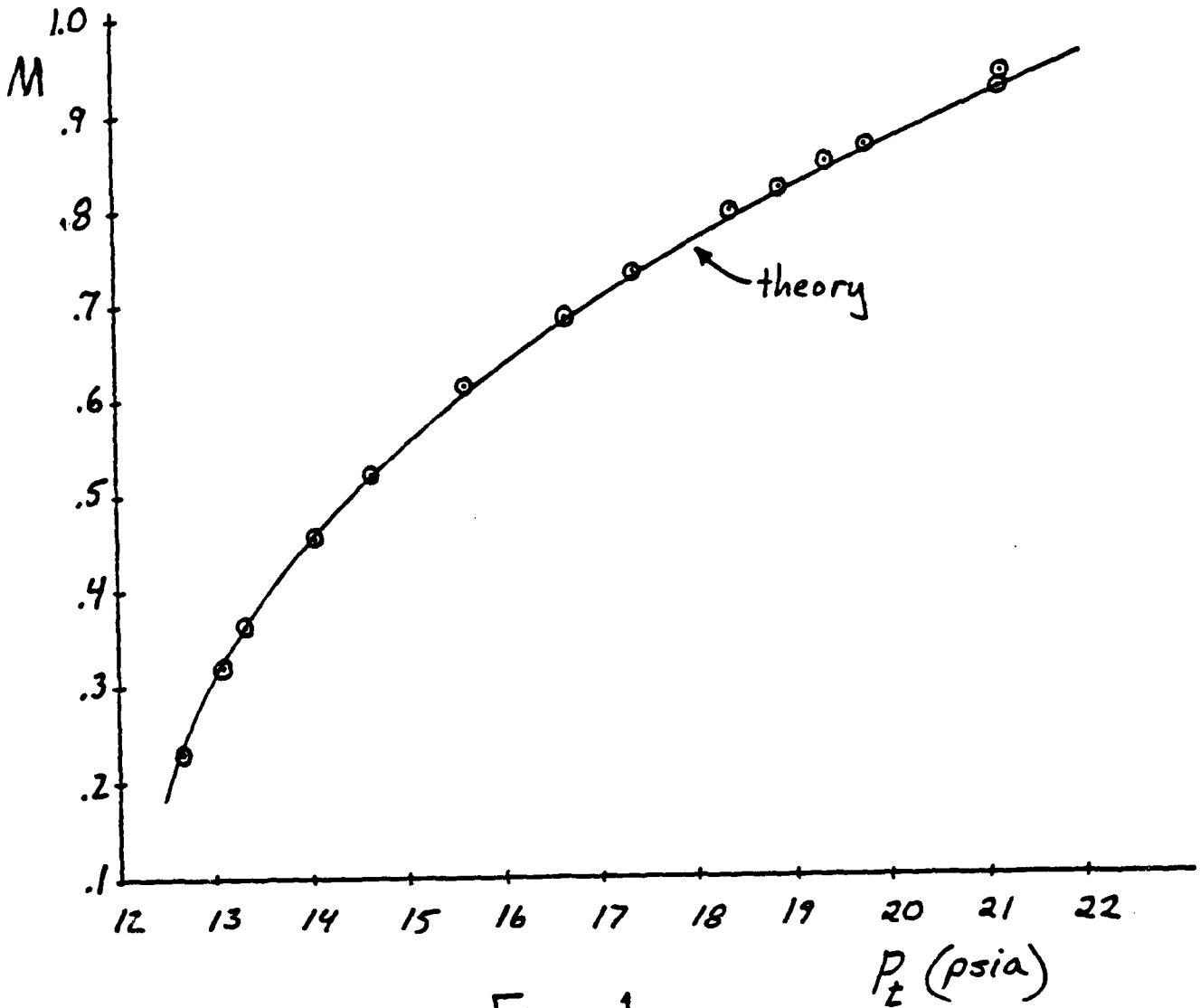


Fig. 1

Test Section Mach Number vs

Plenum Chamber Stagnation Pressure

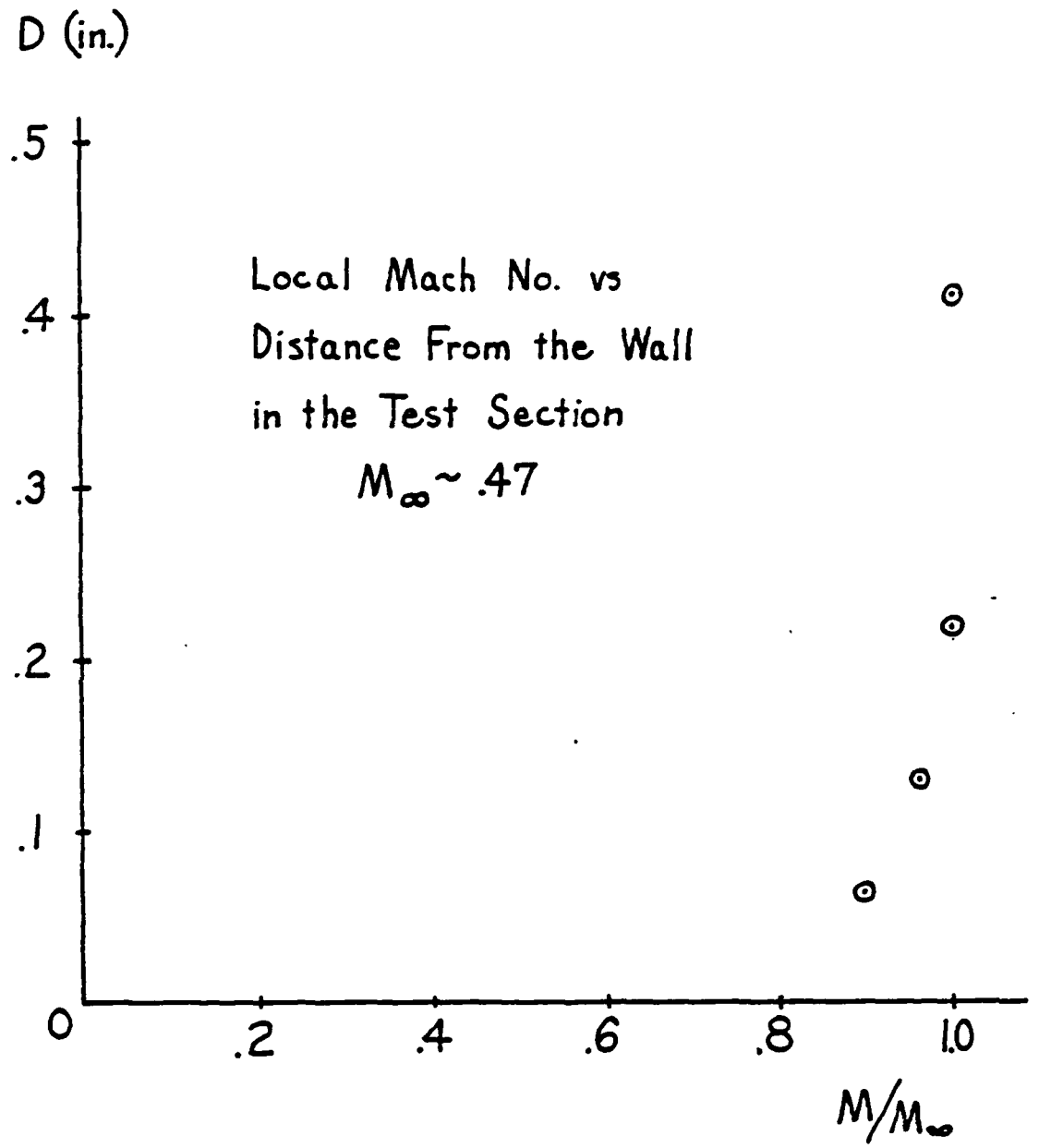


Fig. 2
Boundary-Layer Measurements
on Target Plate

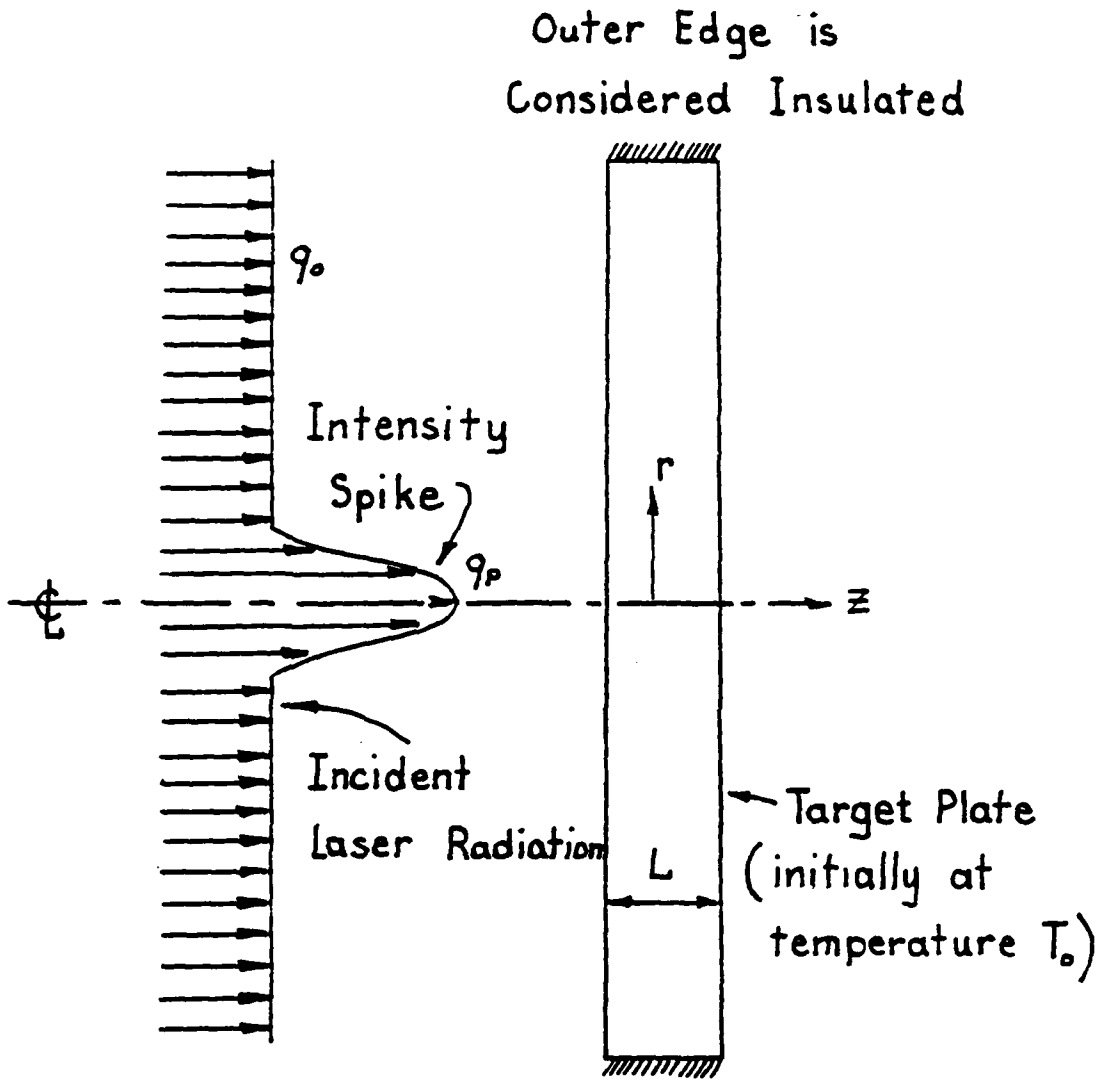


Fig. 3
Schematic of the Target Plate

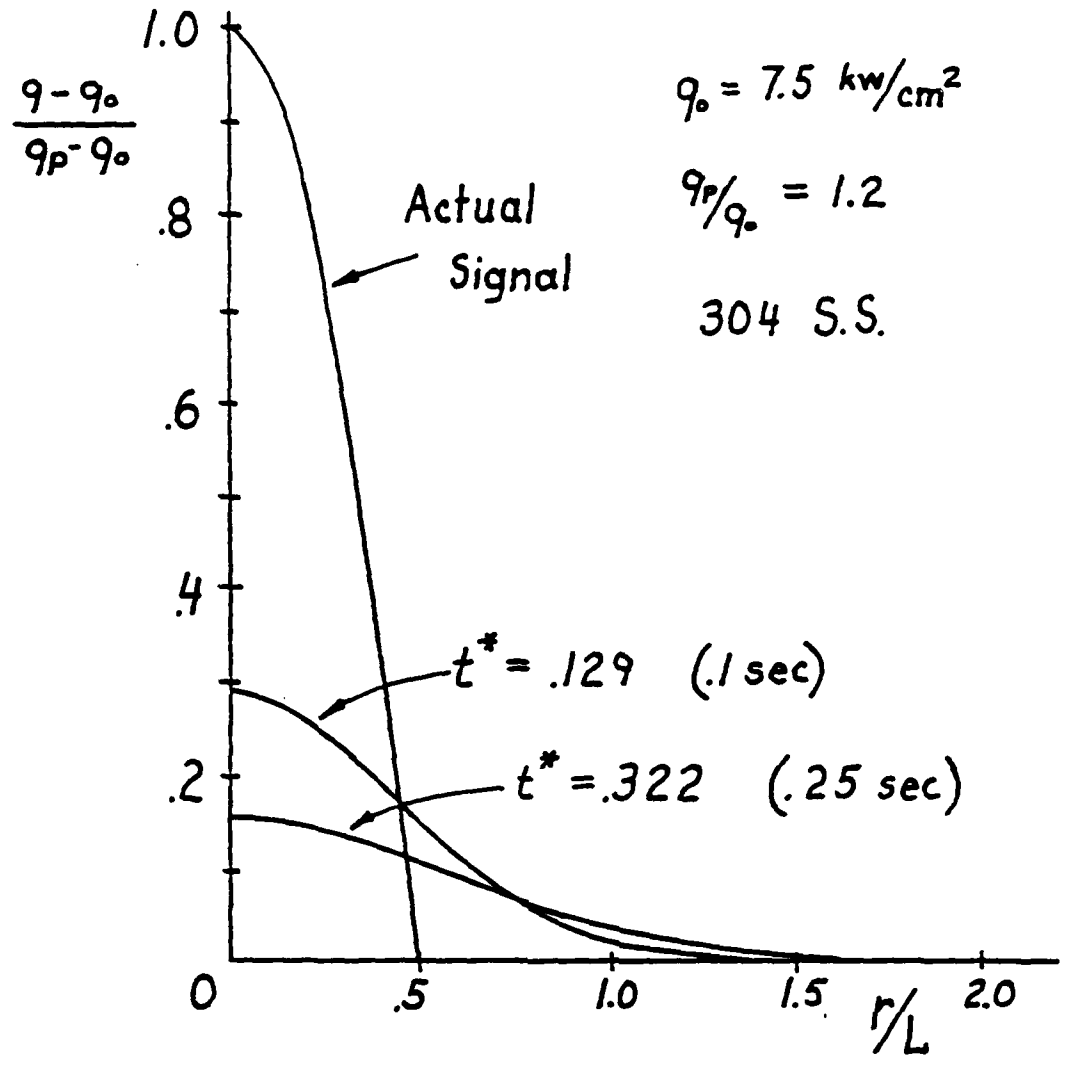


Fig. 4

Predicted Laser Intensity Spots on the Front Surface of the Target Plate. Spot Width is One Plate Thickness

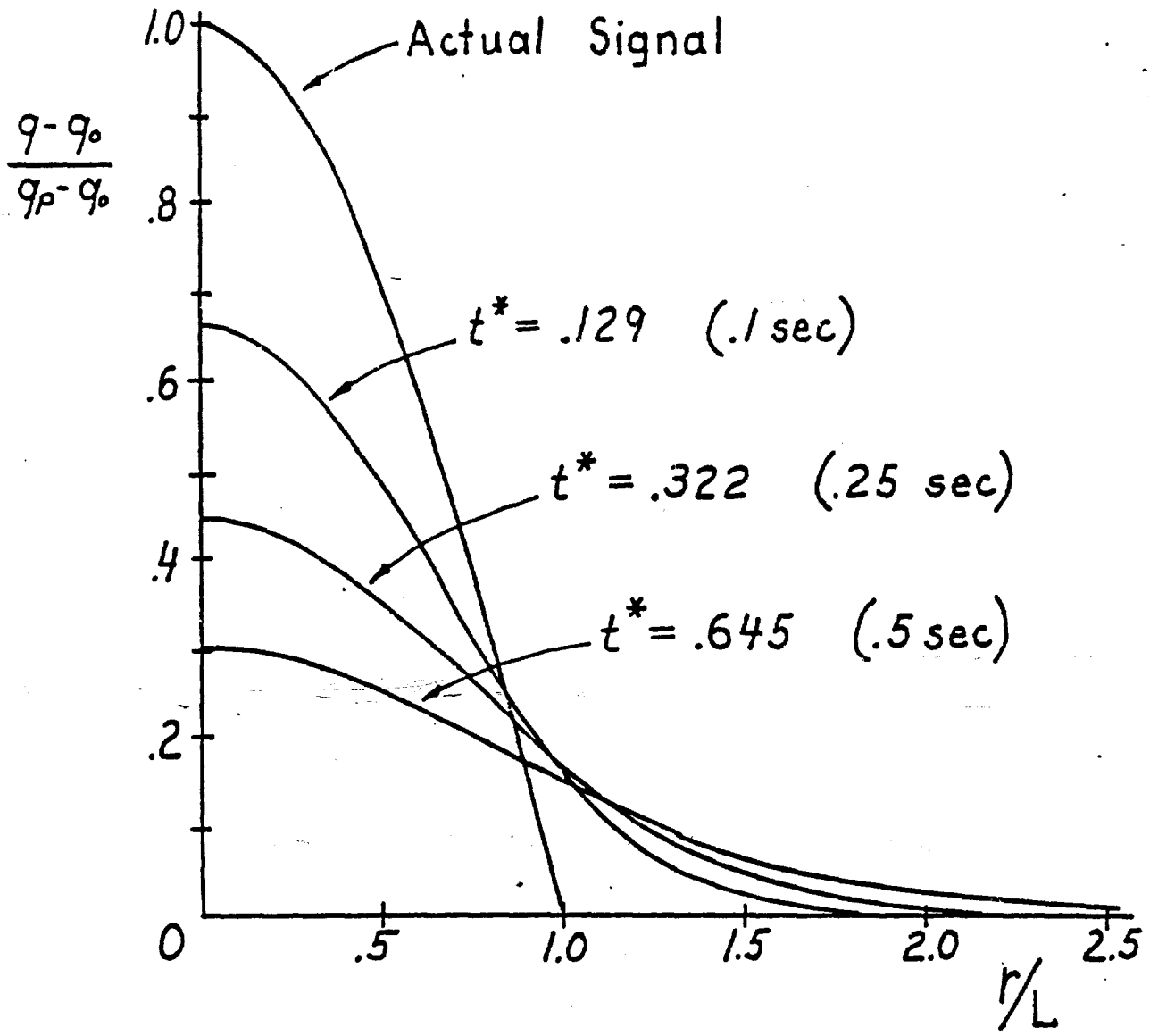


Fig. 5

Predicted Laser Intensity Spots on the Front Surface of the Target Plate

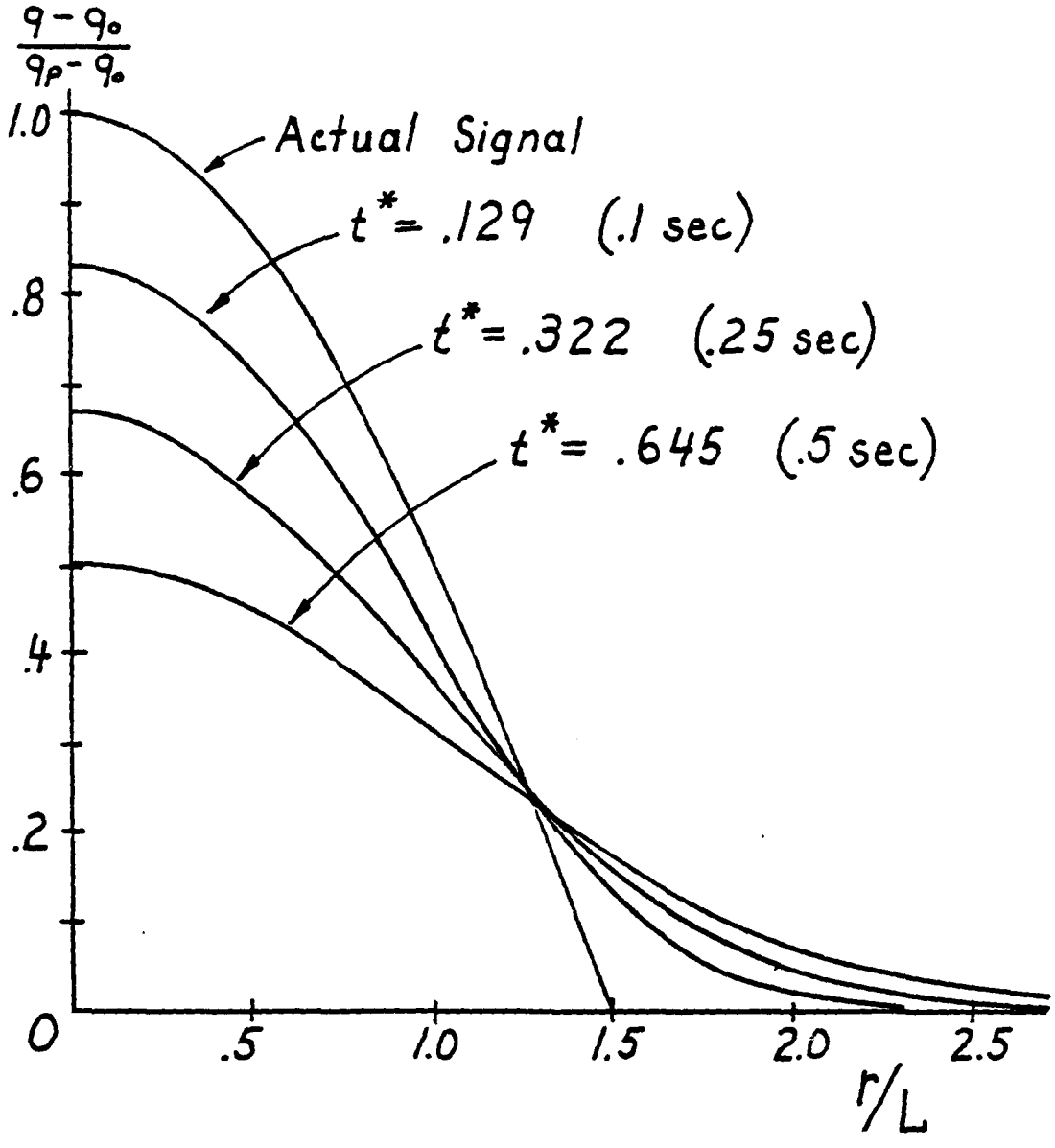


Fig. 6

Predicted Laser Intensity Spots on the Front Surface of the Target Plate

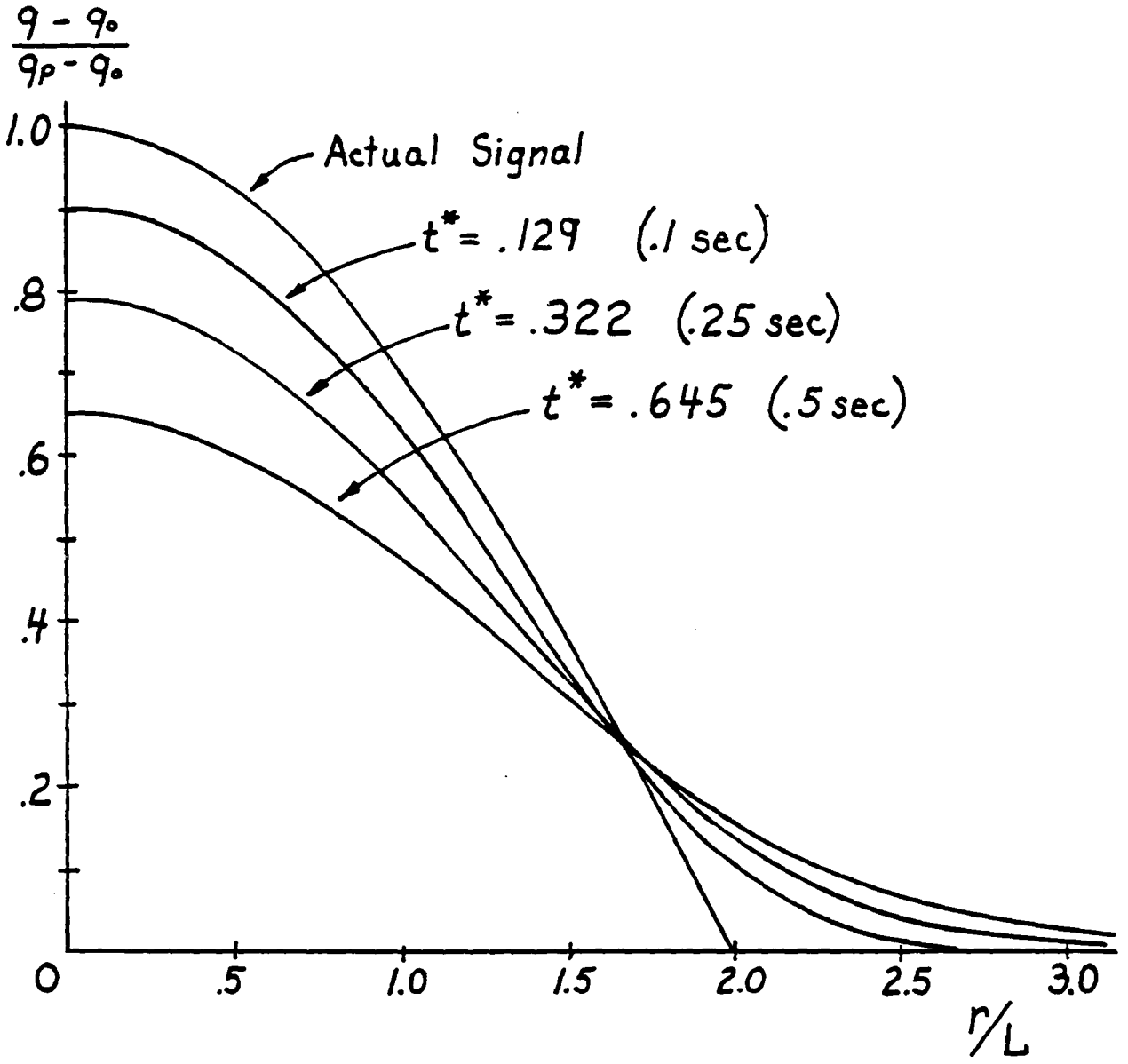


Fig. 7

Predicted Laser Intensity Spots on the Front Surface of the Target Plate

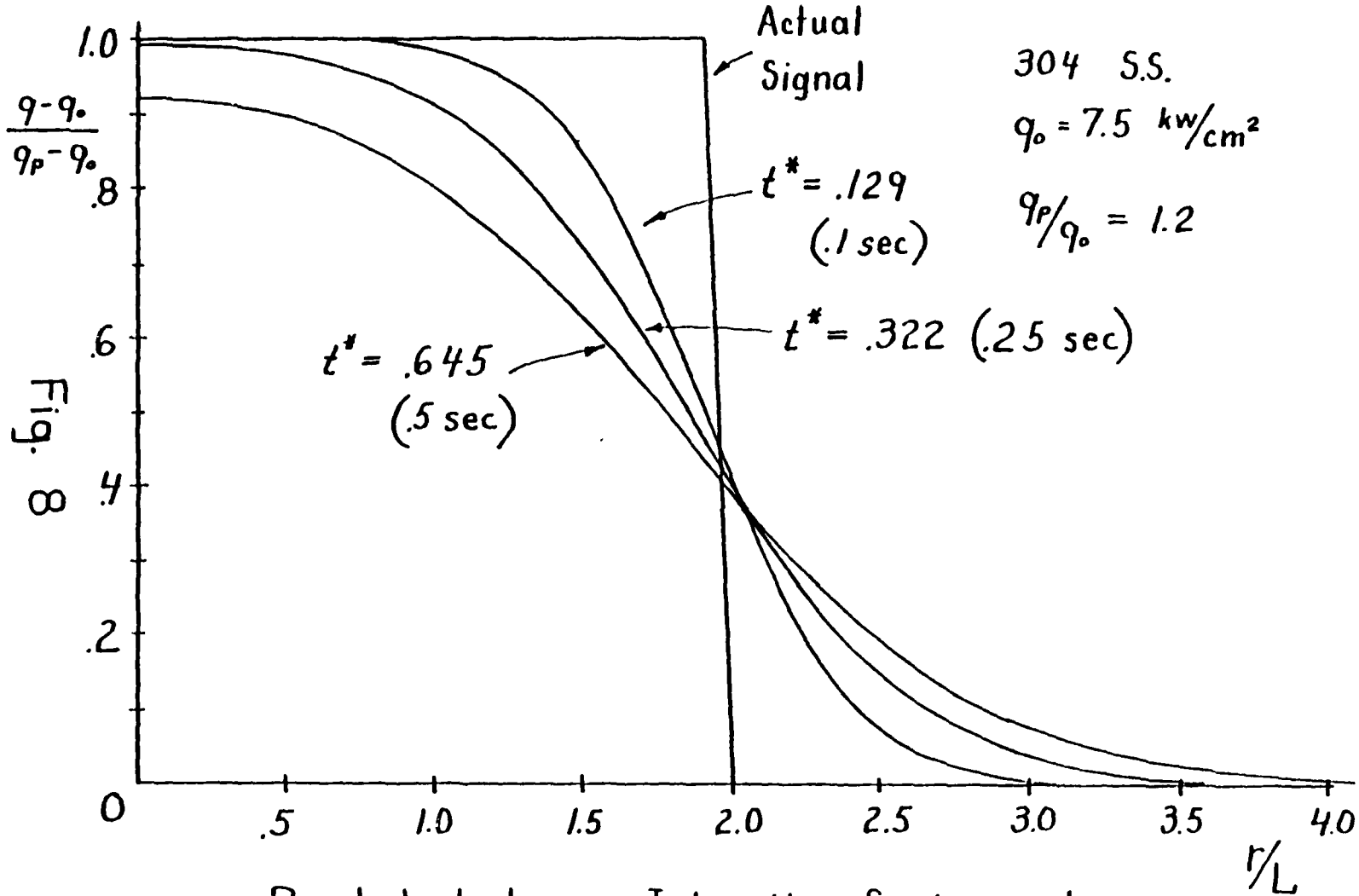


Fig. 8

Predicted Laser Intensity Spot on the Front Surface of a Stainless Steel Target Plate

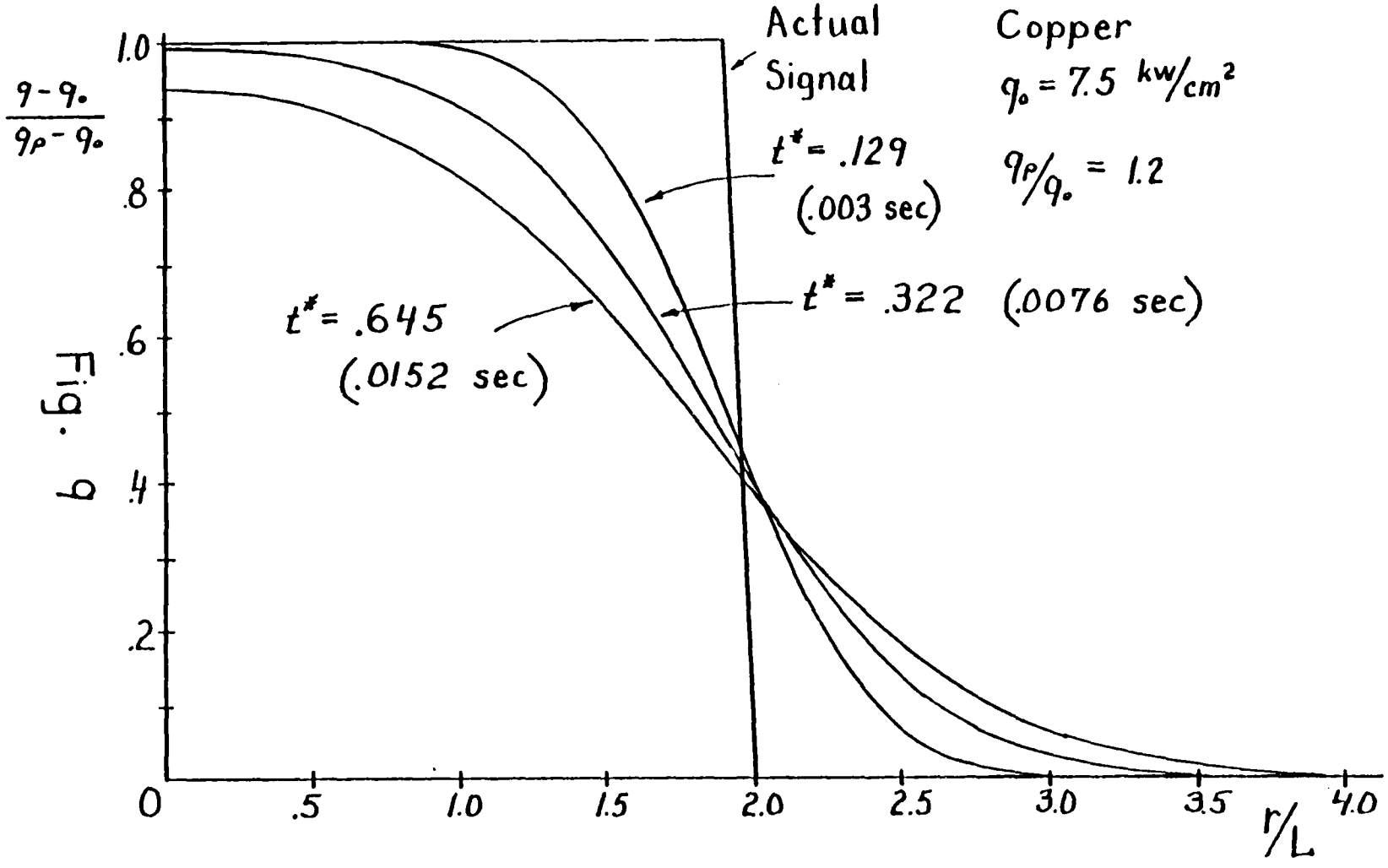


Fig. 9

Predicted Laser Intensity Spots on the Front Surface of a Copper Target Plate

1986 USAF-UES Summer Faculty Research Program/

GRADUATE STUDENT SUMMER SUPPORT PROGRAM

Sponsored by the

Air Force Office of Scientific Research

Conducted by the

Universal Energy Systems, Inc.

Final Report

Propagation of Lamb Waves in Fibrous Composite Materials

Prepared by:	Nisar Shaikh
Academic Rank:	Assistant Professor
Department and University:	Engineering Mechanics Department University of Nebraska-Lincoln
Research Location:	NDE Branch - U.S. Air Force Wright Aeronautical Laboratory
USAF Research:	Dr. T. J. Moran
Date:	July 18, 1986
Contract No:	F49620-85-C-0013

Propagation of Lamb Waves in Fibrous Composite Material

by

Nisar Shaikh

ABSTRACT

The imaging of composite materials has made tremendous progress. There is need for continuing research in wave propagation phenomena to meet the needs of ongoing technological development in explaining and characterizing various complexities.

The behavior of Leaky Lamb Waves in fibrous composite materials is studied with particular emphasis on dispersion and the corresponding energy flow directions. Some interesting results are obtained pointing to possible additional modes when propagation occurs along off-symmetry axes. The energy flow study also strongly suggests further avenues for investigation.

The report describes the experimental technique used and the necessary fixtures that were designed and fabricated. The results of various tests on Graphite-Epoxy specimen are presented.

ACKNOWLDEGEMENTS

I would like to thank the NDE Branch of the Air Force Wright Aeronautical Laboratory and the Air Force Office of Scientific Research for sponsorship of my fellowship and the staff of Universal Energy System for their able assistance and efficient handling of the contract.

My thanks are first due to Doctors Moran and Chimenti for providing me the opportunity of participation in their ongoing research program in the exciting area of Nondestructive Evaluation using Leaky Lamb Waves. My work was also facilitated by the cooperation and assistance of Fiedler, Blodgett and Buynak.

I would also like to thank all the members of UDRI, the 'on-site contractors', and in particular Bob Andrews, Mark Ruddell and Ed Klosterman. The amount of data generated would not have been possible without their help.

I. INTRODUCTION:

My background in NDE stems from my dissertation research at Stanford University in the Department of Mechanical Engineering. The main thrust of my research there was in the development of scanning with shear waves applicable to Acoustoelasticity. Subsequently my interest extended to other NDE areas, particularly those in composite materials, motivated by the interest of my host institution. University of Nebraska-Lincoln is launching an interdisciplinary research activity under the "Nebraska Center for Composite Materials".

The NDE branch at AF Wright Aeronautical Laboratory was a natural choice because of their high standing in the field, their up-to-date research facility and renowned staff. My participation exposed me to the current interest and needs of the US-AFOSR and got me involved directly in their research activity. This places me in a better position to undertake and continue the present and future research for USAF. Important benefits have been educational (I feel I have learned much about the ongoing research in this relatively short time) and the acquaintance I have made with the staff here.

OBJECTIVE OF THE RESEARCH EFFORT:

The overall objective of this particular program at the laboratory is to conduct basic research in acoustics and follow it up with the necessary technological developments for the advancement of NDE and imaging techniques to meet the nations's Air and Strategic defense needs.

Our particular objective during this summer was to further the understanding of wave propagation in fibrous composite materials which exhibit strong anisotropy. The specific questions addressed were:

1. What are the dispersion relations for the propagation of Leaky Lamb Waves launched in arbitrary directions in anisotropic materials?
2. What are the energy flow directions in the general cases of excitation of guided waves?

INTRODUCTION TO THE TOPIC AND THE APPROACH

In an isotropic material the wave propagation and the corresponding dispersion relations are well understood. However, composite materials represent two specific areas of difficulties; one is the prediction of the elastic properties and the other is the treatment of the strong anisotropy exhibited by the fibrous structure of such materials.

Chimenti and Nayfeh have modeled and computed numerically the elastic properties of fiber-resin composites and shown good experimental verification (see references at the end of this report). They have also studied experimentally the reflection and dispersion of Leaky Lamb Waves as reported in several of their publications.

The preliminary work on wave propagation in an arbitrary direction had been addressed at the lab and was in progress among other activities. Added thrust was provided to this activity by the writer, since this became his primary activity.

It was decided to conduct the research on an experimental basis to utilize the excellent facilities available here, since this was the most productive approach to attack this problem. Two different fixtures were designed for extra capabilities required for the special tests and in consideration of the efficiency needed for the large amount of data anticipated. Some literature review and explorative analytical studies were also conducted to seek theoretical guidance and corroboration for the experimental results.

EXPERIMENTAL SET UP AND TECHNIQUE:

Following is a brief description of two versatile transducer orientation devices that were constructed to carry out the special tests.

1. Opposing Transducer Scheme:

In this arrangement the specimen was held in a gimbal like fixture such that the sample can rotate in its own plane. The fixture itself is secured to a turntable enabling two rotational degrees of freedom. The transmitting and receiving transducers are placed on either side of the sample as shown in Figure 1.

The angle ϕ can be easily traversed from zero to ninety degrees allowing for the full range of corresponding phase velocities. This feature enabled measurements over a broader range. Also the angles can be changed easily making the data acquisition more efficient.

2. Azimuthal Holder:

A second fixture was designed in which the transmitting transducer is held such that it can rotate in azimuth about a fixed axis, permitting the change of angle θ, without moving the point of insonification. This device is secured to a turntable such that it can rotate with respect to the turntable, with the rotational axis of the device coinciding with that of the turntable axis. The specimen is placed and held fixed to the turntable, which rotates with respect to the receiving transducer. The device has four rotational degrees of freedom, and these angles can be changed with ease. The sketch in Figure 2 illustrates this arrangement.

EXPERIMENTS AND THE RESULTS:

Dispersion studies:

The intersection of the plane normal to the specimen and passing through the transducer makes an angle ϕ with the fiber direction. For three values of this angle, namely 0, 45 and 90 degrees, full sets of data were taken with two different measuring techniques. The data is shown superimposed on the computational results of Chimenti and Nayfeh.

Figures 3, 4 and 5 show the results with the null technique used widely here at the lab in which the interference with specular reflection is used to determine the Leaky Lamb Wave frequencies. Figures 6, 7 and 8 show similar results obtained with direct monitoring of the wave peaks received on the opposite side of the launching transducer. The results can be made more precise in this arrangement if the deconvolving of transducer response is implemented. It is interesting that in this configuration the data could be taken at the limiting value of infinite phase velocity, which corresponds to the thickness resonance.

To check each new test setup, the measurements were always made first on Aluminum samples, and Figure 9 shows the results of one such experiment on two different thicknesses of plate specimens.

Energy flow studies:

To map energy flow in various directions, direct detection of Lamb Waves is advantageous, while for dispersion studies the null technique is efficient. Figures 10a and 10b show a comparative display of received signals versus a frequency sweep. In Figure 10a, the lowest amplitudes correspond to the Leaky Lamb Wave frequencies, while in Fig. 10b the peaks appear at the same abscissa measured at x=15 mm.

Using Azimuthal Holder, the specimen is rotated with respect to the turntable to which the transmitting transducer is fixed. This allows setting of the transmitting angle ϕ_t. Now by rotating the turntable, in effect, receiving angle ϕ_r is traversed. The maximum amplitude of the peak received is considered to be the principal direction of the energy flow of the particular mode.

DISCUSSION OF THE RESULTS

The verification with earlier test results provides the confidence in the results of the new test setups and schemes. The measurements along the fiber direction (ϕ = 0) check well. The discrepancies in the measurements at ninety degrees have to be explained. The choice of material properties may be a probable cause. The assumption of transverse isotropy may be a weak one for a thin plate due to the manufacturing technique used in molding; the more general orthotropic representation may be better applicable. With modified properties the results for the normal (to fiber) direction may be improved.

The data for the launched waves at 45 degree is quite interesting. This points to the possibility that there may indeed exist additional modes, possibly due to the quasi-longitudinal and quasi-shear displacements.

The tests to profile the energy flow have been very exciting. One behavior which is quite evident, is that certain modes

persist only for a range of value of the angle ϕ. It has also been observed that the maximum energy propagation is not always in the direction of the launched waves but in different directions depending upon the particular mode. These results are considered preliminary, and further refinement of the measurement technique will be necessary to investigate these issues further.

RECOMMENDATIONS

The results have verified some of the ideas presented by Moran and Chimenti and provide direction for future work. Further investigations leading to an understanding of wave propagation in composite materials will not only aid in the present technological development of imaging but can also provide some new features that can be exploited.

Further work is necessary both in explaining the additional modes at off-angle measurements and understanding the energy flow mechanism.

REFERENCES

1. D. E. Chimenti and A. H. Nayfeh, J. Appl. Phys. 58(12), (1985).

2. D. E. Chimenti, A. H. Nayfeh and D. L. Butler, J. Appl. Phys. 53(1), (1982).

3. De. E. Chimenti, Appl. Phys. Letter, 43(1), (1983).

4. Y. Bar-Cohen and D. E. Chimenti, Rev. of Prog. in Quantitive NDE, (1985).

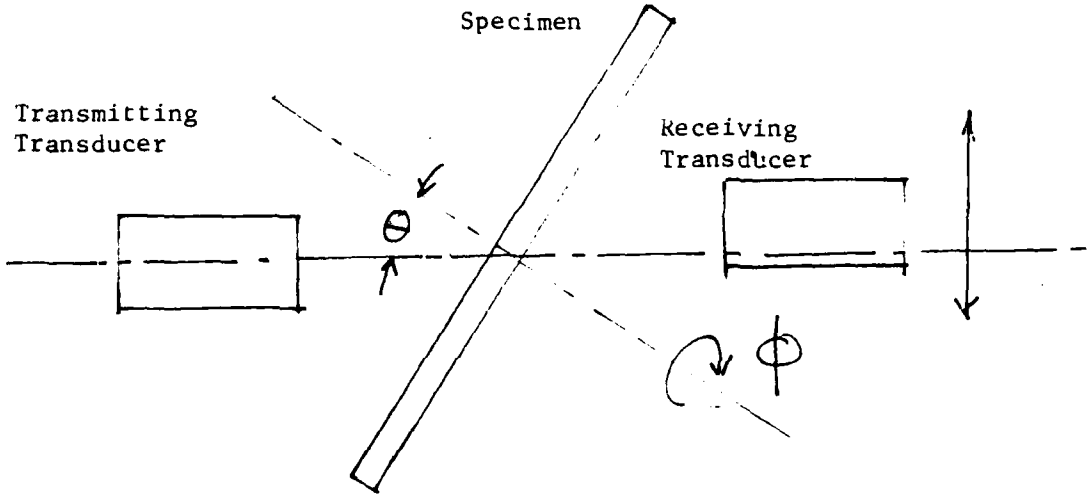


Fig. 1
OPPOSING TRANSDUCER SCHEME

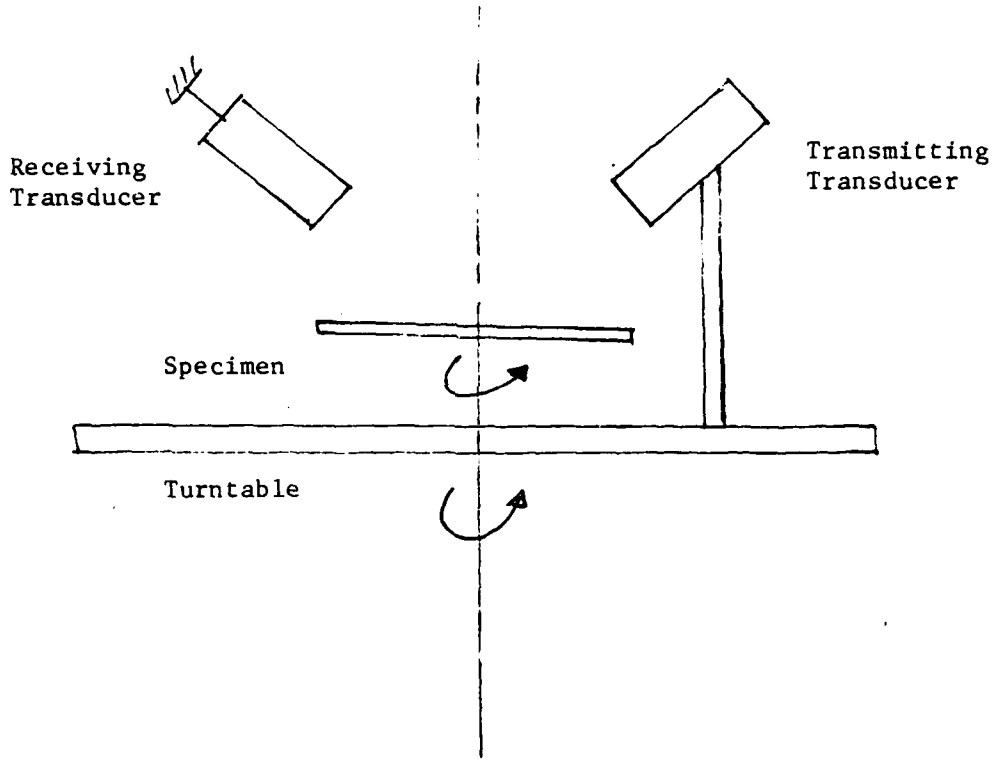
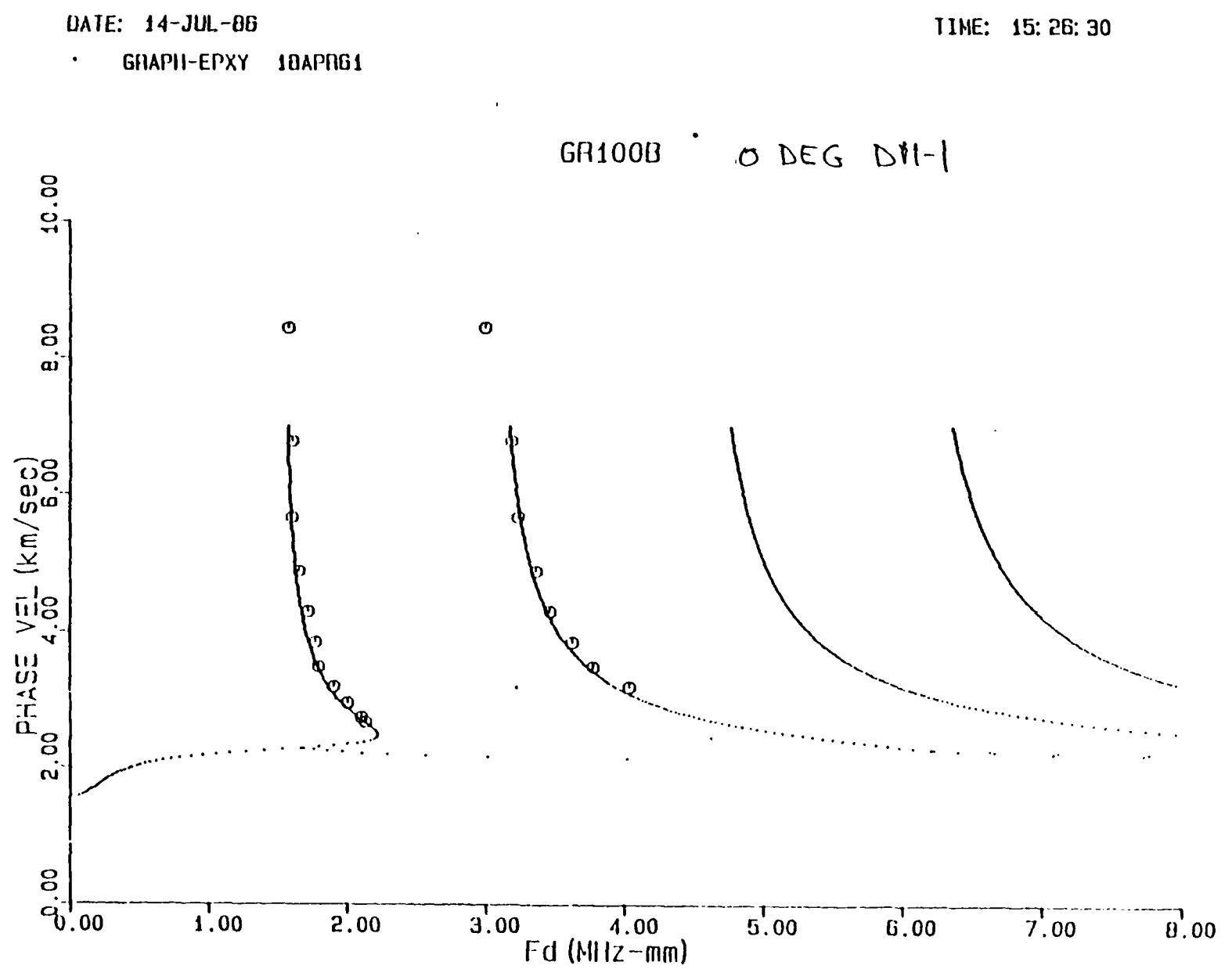


Fig. 2
AZIMUTHAL HOLDER

Fig 3


DATE: 14-JUL-86
TIME: 15:26:30
GRAPH-EPXY 10APR61
GR100B .0 DEG DM-1
PHASE VEL (km/sec)
10.00
8.00
6.00
4.00
2.00
0.00
0.00
1.00
2.00
3.00
4.00
5.00
6.00
7.00
8.00
Fd (MHz-mm)

DATE: 07-JUL-86 TIME: 14: 44: 36

GRAPH-EPXY [112, 44] 09JUN61

DATE: 10-JUL-86 TIME: 13: 35: 56

GR190B.DAT 90 DEG DM-1

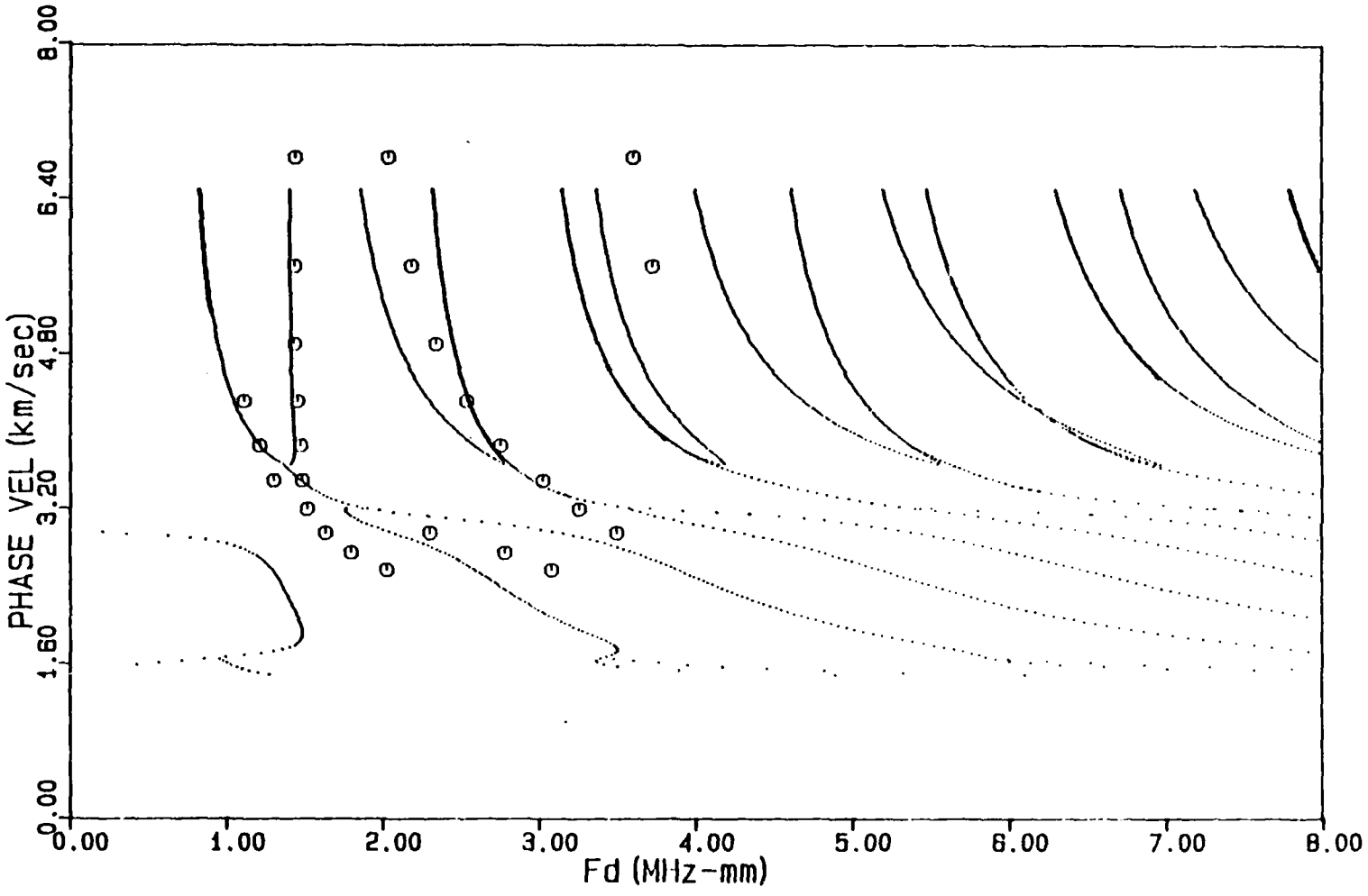


Fig 4

125-11

DATE: 07-JUL-86 TIME: 14: 23: 42

DATE: 10-JUL-86 GRX PHI=45 03JUN61 TIME: 13: 19: 22

GR145.B.

45 DEG DM-1

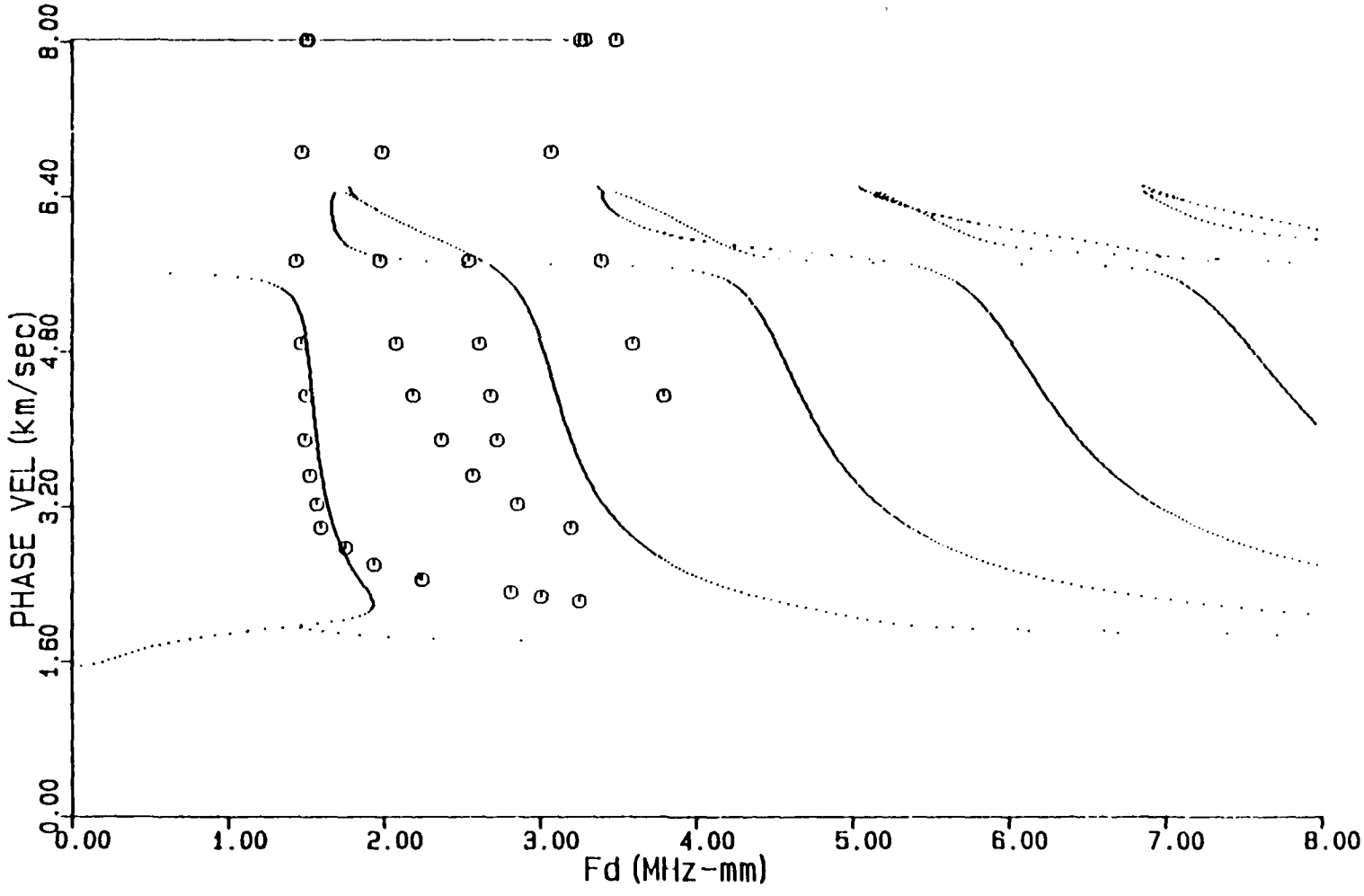


Fig 5

DATE: 14-JUL-86 TIME: 15: 26: 30

GRAPH-EPXY 18APR61

0 DEG DM-1

GR100

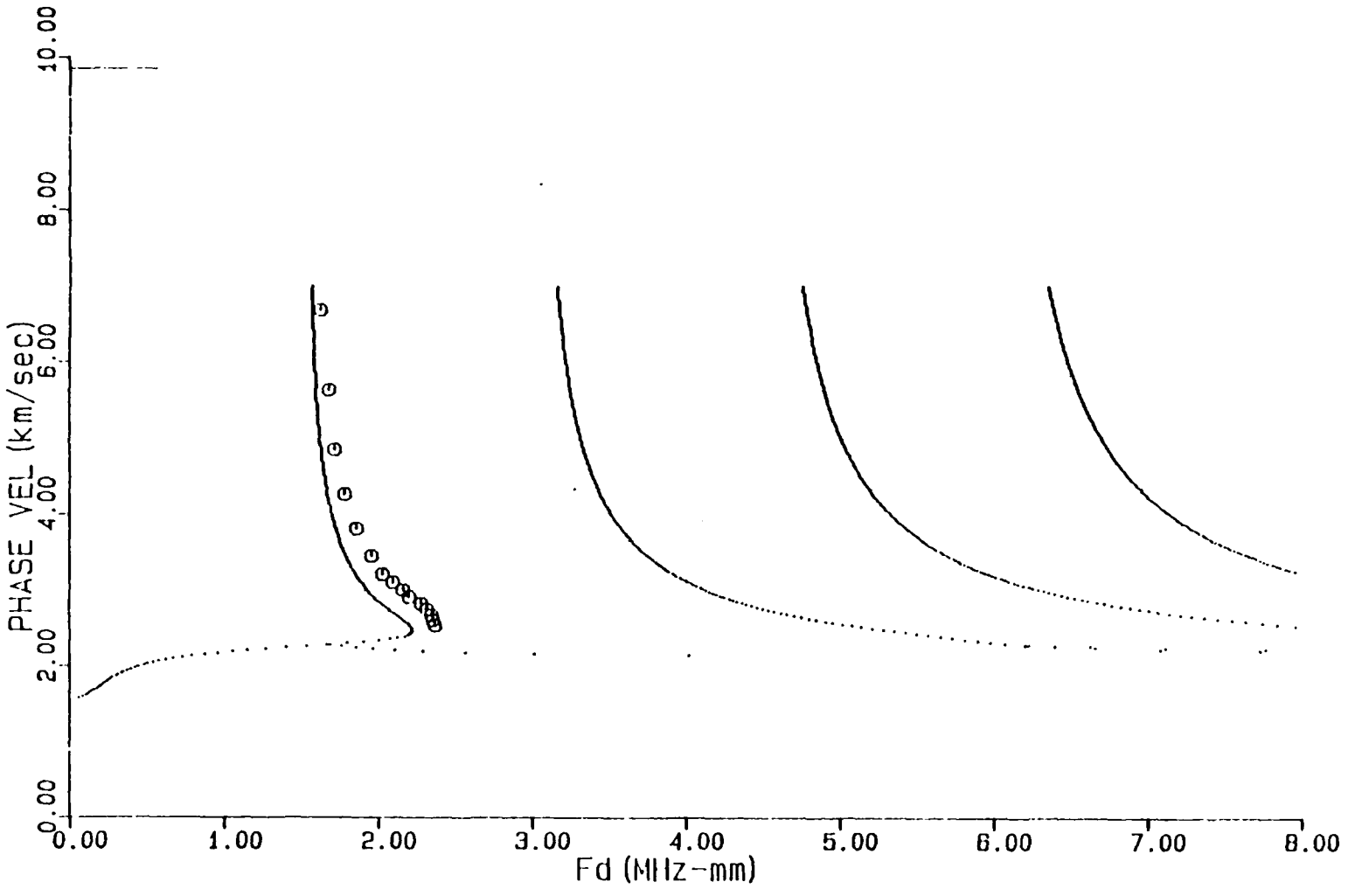


Fig 6

DATE: 14-JUL-86 TIME: 15:37:42

GRAPH-EPXY 09JUN61

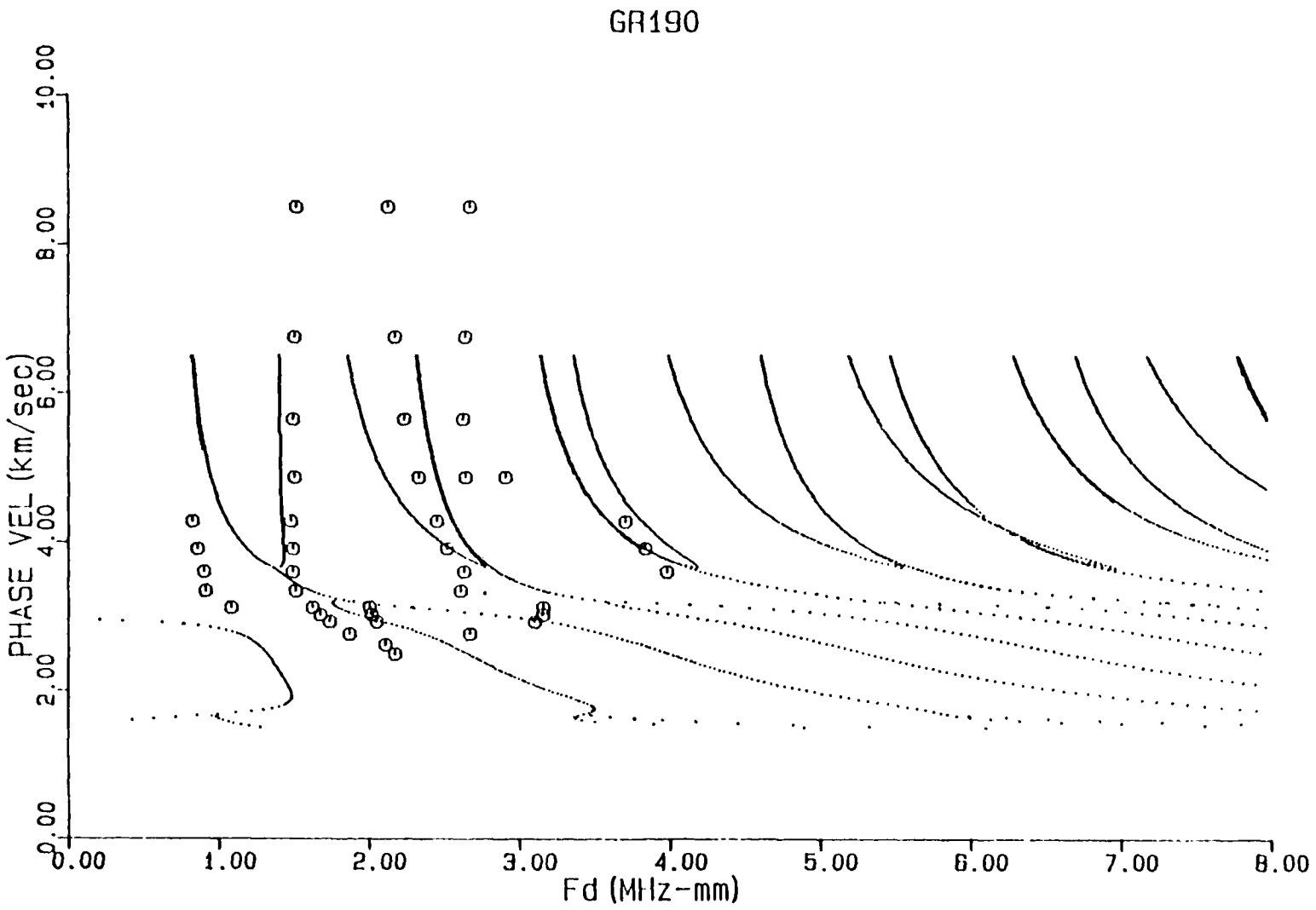


Fig. 7

DATE: 14-JUL-86 TIME: 15:56:59

GRAPH-EPXY 03JUN81

GR145

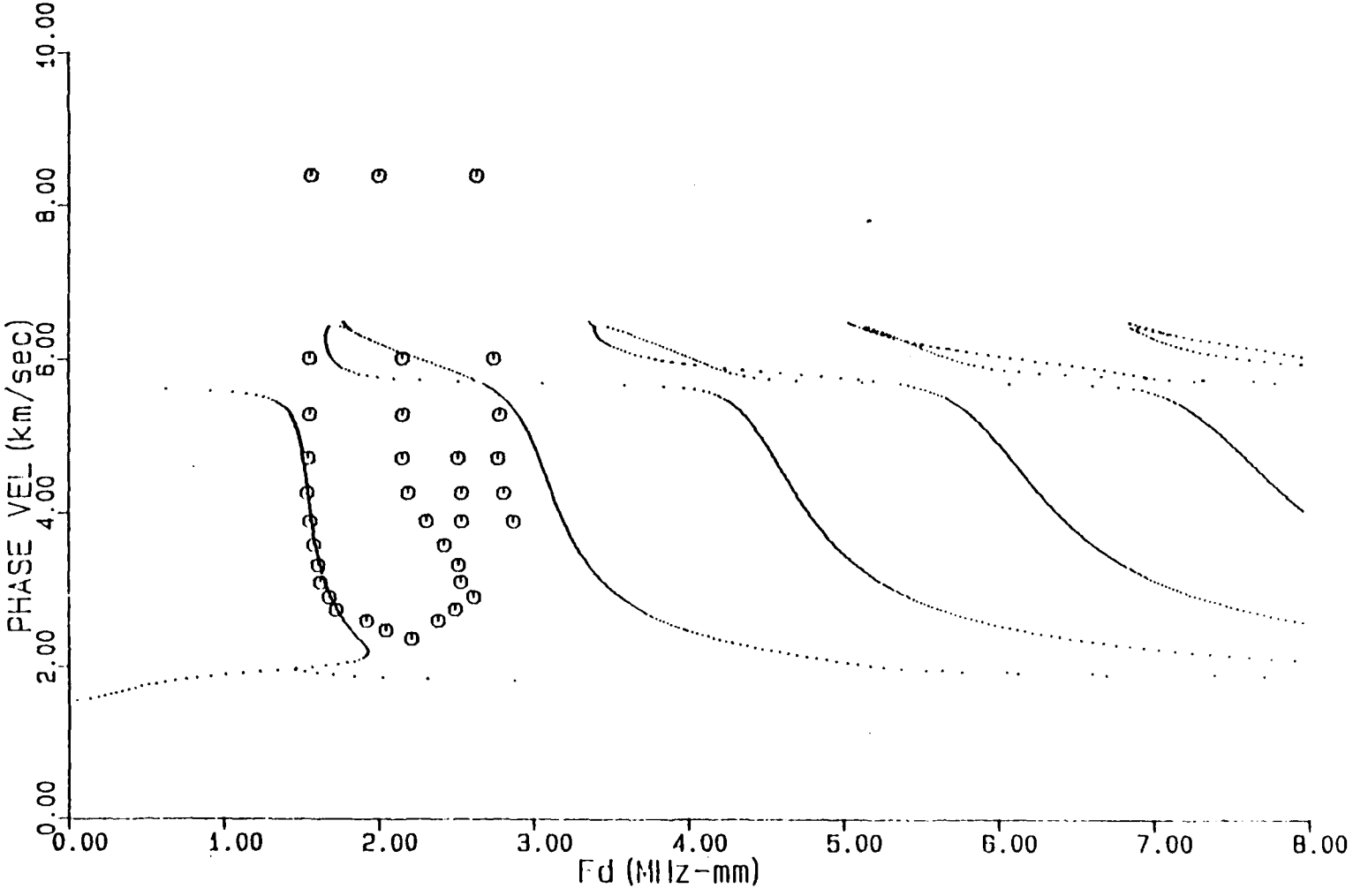


Fig. 8

$C_L = 6.30$ $C_t = 3.14$ $\rho = 2.7$

DATE: 14-FEB-86

ALUM 13FEB61

TIME: 08:34:19

\+ 1.72 mm Al

O 2.28 mm Al Plate

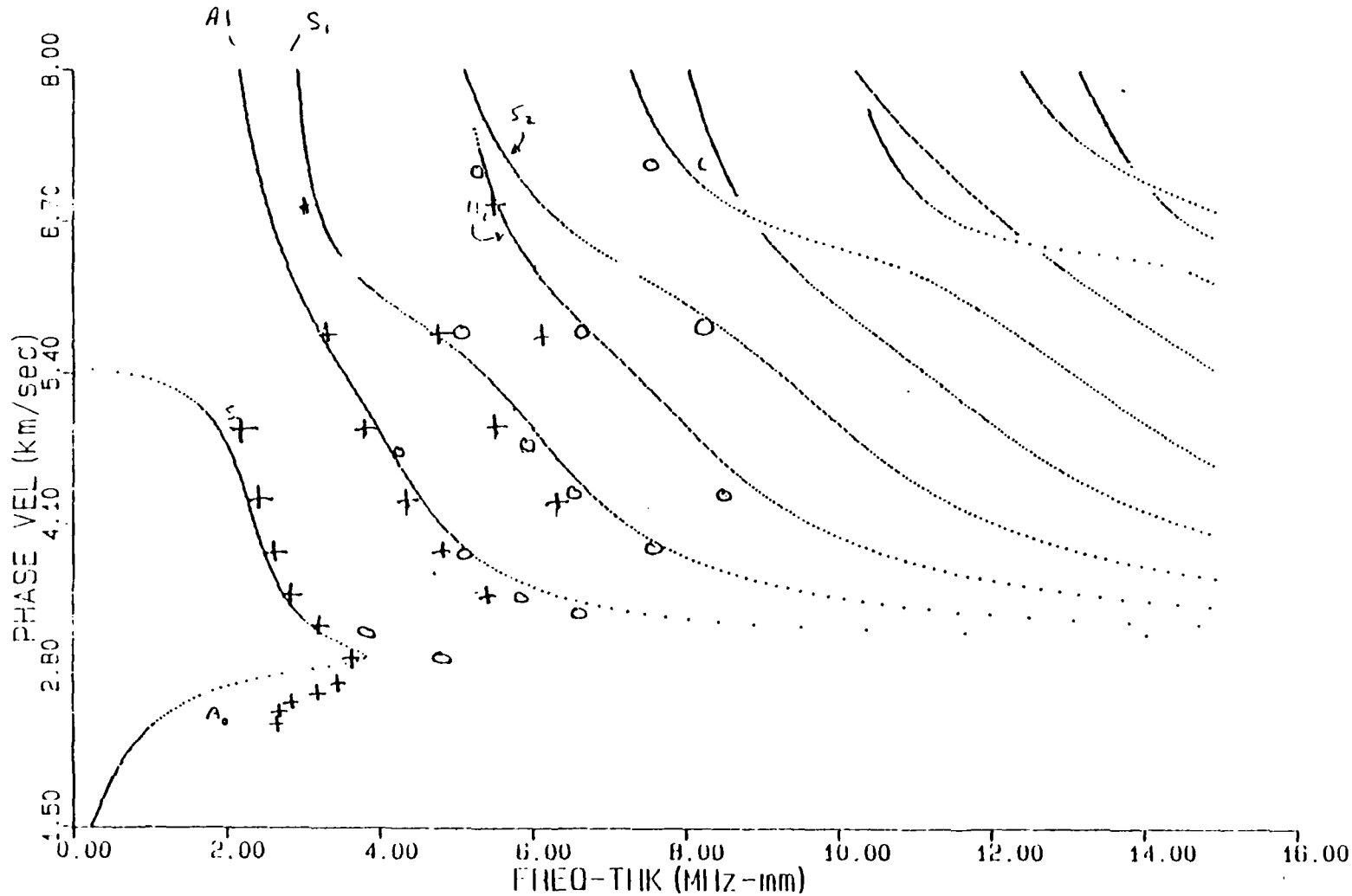


Fig. 9

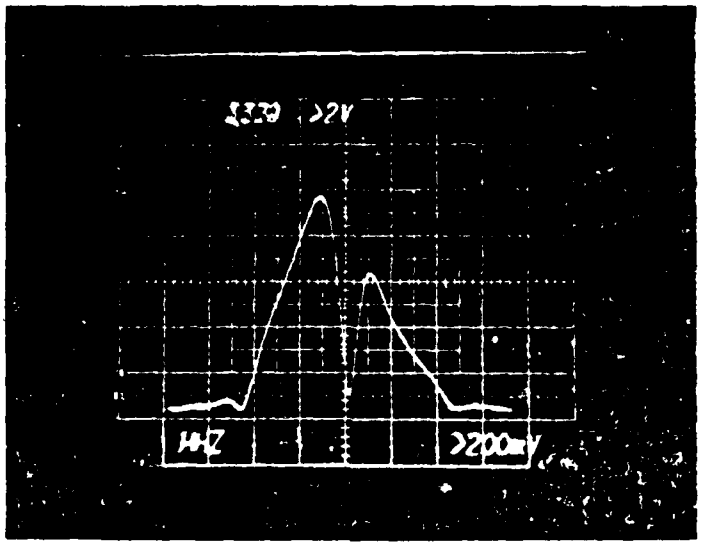


Fig 10a

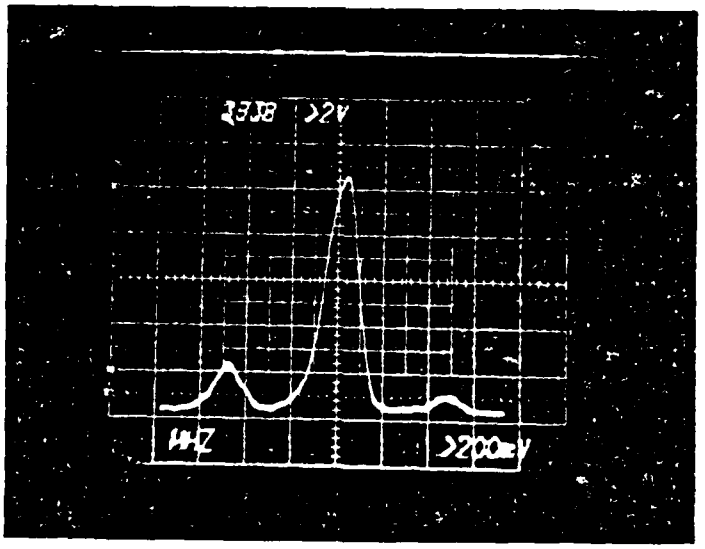


Fig 10b.
Amplitude vs. Frequency

1986 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/
GRADUATE STUDENT SUMMER SUPPORT PROGRAM

Sponsored by the
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the
Universal Energy Systems, Inc.

FINAL REPORT

Pharmacokinetics of Certain Substances of Abuse: A Review of the Literature

Prepared by:	Dolores C. Shockley
Academic Rank:	Associate Professor
Department and University:	Department of Pharmacology Meharry Medical College
Research Location:	Drug Testing Laboratory School of Aerospace Medicine Brooks AFB, TX
USAF Research:	Major (Dr.) William E. Ottinger
DATE:	September 30, 1986
Contract No:	F49620-85-C-0013

Pharmacokinetics of Certain Substances of Abuse: A Review of the Literature

by Dolores C. Shockley

ABSTRACT

This report presents a concise summary of the pharmacokinetics of certain agents that act on the central nervous system and that are abused in society. Pharmacokinetics includes the absorption, distribution, biotransformation (metabolism) and excretion of xenobiotics. A description is presented of the dynamic relationship between the substances and the body: how the body acts on the agent rather than how the agent acts on the body (pharmacodynamics). Agents included in the report are currently or soon will be determined qualitatively and quantitatively by urinalysis at the Air Force Drug Testing Laboratory. These include marihuana (THC), cocaine, barbiturates, amphetamines, opioids, lysergic acid diethylamide (LSD), and phencyclidine (PCP). Other areas addressed include some of the physicochemical properties of the agents and the influence of urinary pH on excretion.

The thrust of the report is to present relevant excerpts of what is known about the pharmacokinetics of these abused agents to assist in interpretation of urinalysis data.

ACKNOWLEDGEMENTS

I would like to thank the Air Force Systems Command and the Air Force Office of Scientific Research for sponsorship of my research. I want to thank Colonel Robert Grossner for accepting me at the Drug Testing Laboratory thereby providing me the opportunity to utilize and enhance my knowledge of drugs of abuse. I appreciate Major William Ottinger's willingness to include me in activities at the laboratory. I am especially grateful to the competent, cooperative and concerned staff at the Strughold Medical Library who helped to make my summer scientifically rewarding. This includes the entire staff and specifically Mr. Joe Franzello, Ms. Bonnie Fridley, Mr. Martinez and the Director, Mr. Fred Todd.

I. Introduction

My training and experience in pharmacology has focused on central nervous system (CNS) drugs from diversified perspectives for many years. It began with my doctoral dissertation at Purdue University where I investigated the mechanisms responsible for reversal of pain - induced modulation of physiological functions (e.g., a decrease in liver sulfhydryl concentration) by antipyretic analgesics. Since that time I have remained in this area by virtue of teaching medical and dentistry students, talking to adolescents about drug abuse, participating in media braodcasts on substance abuse, and conducting research on the effect of alpha-2 adrenoceptor agonists on the opioid withdrawal syndrome in rats (testing the biogenic amine hypothesis).

The Drug Testing Laboratory has the assignment for the Air Force of screening and confirmation of certain abused CNS drugs. Randomly collected urine specimens are tested for these agents and/or their metabolites: the kidney is the principal means of excretion of these subtances. The sojourn of a xenobiotic (foreign substance) in the body is not static but dynamic and includes the many facets of pharmacokinetics: absorption, distribution, metabolism and excretion. Proper interpretatin of urinalysis data encompasses these concepts of pharmacology/ toxicology, and even more so when adversarial proceedings or punitive actions are involved.

I was assigned to the Drug Testing Laboratory because of this mutual interest in the pharmacology/toxicology of abused drugs.

II. Objectives of the research effort

The Drug Testing Laboratory (DTL) currently does 100 percent screening of randomly collected urine specimens for metabolites of cocaine and marihuana. Imminent expansion will include the following additional CNS agents and/or

their metabolites: amphetamines, barbiturates, lysergic acid diethylamide (LSD), opioids, and phencyclidine (PCP). Specimens screened positive by radioimmunoassay (RIA) are subsequently confirmed primarily by gas chromatography/mass spectrometry (GC/MS).

The objectives of my research were:

1. to conduct a comprehensive literature search on the pharmacokinetics, physico-chemical characteristics and other parameters that might affect qualitative and/or quantitative analyses of the agents listed above;
2. to provide to DTL, in an easily accessible and interpretable format, relevant information from the literature on the pharmacokinetics, physico-chemical characteristics, effects of environmental conditions and possible substance interference related to qualitative and/or quantitative urinalysis data; and
3. to assist the staff on a regular basis, on site, with (a) clinical interpretation of urinalysis results in the context of the pharmacokinetics of the agents, and (b) methodology as influenced by <u>in vitro</u> variables, such as, compound stability, storage conditions, loss due to surface binding or interference by endogenous and/or exogenous substances.

III. a. My initial approach was the use of a computer-aided literature search service which was expanded to a regular search of Current Contents - Life sciences, Index Medicus several topics of Excerpta Medica, selected current journals that had articles on forensic science, pharmacology, toxicology, chromatography, etc.

b. From the above sources, I have a voluminous collection of important articles useful to a laboratory that tests for abused substances.

IV. a. This study was comprised of compounds representing seven different chemical types. Each category requires

extensive analysis of the available literature to formulate pharmacokinetic models. This area is imcomplete at present. I have extracted preliminary data from the literature which will be used to construct mathematical models for generalization.

b. A notebook has been prepared for DTL consisting of photocopies of key articles arranged by agent name with sub-headings, e.g., the division entitled *marihuana/marijuana (cannabinol, THC, cannabis)* consists of,

I. Chemistry
II. Pharmacokinetics
III. Methodology including extraction
IV. Stability and Binding
V. Oral Ingestion
VI. Passive Inhalation
VII. Possible substance-interference (melanin/melanin metabolites) with urinalysis

V. a. Questions that surfaced on a day-to-day basis received search priority.

b. Literature summaries that were provided during this period with interpretation of urinalysis data included (i) a comparison of the pharmacokinetics of cocaine by oral, intravenous and intranasal administration, (ii) in vitro stability and binding of THC, (iii) possible interference of melanin/melanin metabolites with THC analysis, and (iv) others.

VI. The overall objective of this section of my report is to provide a brief summary of some of the pharmacology/toxicology and physico-chemical characteristics of the drugs tested by DTL.

Agents are summarized according to CNS categories: hypnotic-sedatives (barbiturates), stimulants (cocaine, amphetamines), opioids (heroin, morphine, etc.), hallucinogens (LSD), and psychotomimetics (Marihuana, PCP).

OPIOIDS

Generic and Proprietary Name(s)	Street Name(s)
Heroin	H, Harry, Horse, Junk, Seat, Stuff, Smack, Dogie, Dope, Scag,
Morphine	Mary, M, Hard Stuff, Junk, Morph, Morpho, Miss Emma
Meperidine (Demerol, Pethidine)	Demmies
Methadone (Dolophine)	Meth, Butalgin, Dollies
Codeine	Blue Velvets
Opium	Black Pills
Oxycodone with aspirin (Percodan)	Percs
Hydromorphone (Dilaudid)	Dils
Propoxyphene (Darvon, Novrad, Darvon-N)	
Pentazocine (Talwin, Fortral)	

usually administered p.o. but i.v. i.m., SQ, or rectal; distribute into the CNS fairly rapidly; metabolized by the liver; eliminated in feces and urine.

Heroin: diacetyl morphine; known as diamorphine; name is derived from the German word "Heroisch," which means large or powerful; prepared from morphine by the acetylation of both the phenolic and the alcoholic OH groups; it is "cut" by the addition of adulterants such as milk sugar (lactose), quinine, mannitol, corn starch or other white powdery substance; "mainlining" is intravenous injection; "skin popping" is subcutaneous injection; can also be "snorted" or "sniffed" by a quick sniff into the nostrils which can cause perforation of the nasal septum.

Metabolized by the liver, kidney and tissues to monoacetyl morphine (MAM) which is hydrolyzed to morphine; urine tested for morphine; recoveries of heroin and its metabolites in the urine collected for 40.5 h after i.v. infusion of 10 mg heroin showed that excretion reached about maximum rates between 2.6 and 4.6 h - total morphine 37.4-46.4% (mean 43%), conjugated morphine 34.3-42.6% (mean 38.3% or 89% of total morphine), free morphine 3.1-5.3% (mean 4.2% or 9.75% of total morphine), 6-monoacetylmorphine 0.48-2.82% (mean 1.3%), free heroin 0.05-0.3% (mean 0.13%); some reports show that as much as 83% of total morphine is excreted as its conjugates and morphine 3-glucuromide is a major metabolite; 10 mg of pure heroin i.v. usually produces a urine positive for morphine up to 40 h; larger doses may be positive for up to 72 h; urine collected within the first 2 h after usual heroin dose may show positive for heroin and morphine but after 2-8 h will be positive for morphine only; half-life of heroin, 9 min.; half-life of MAM, 38 min.

Morphine: 6-10% excreted unchanged primarily as morphine 3-glucuromide; 90% of total excretion takes place during the first day; 7-10% appears in feces; comes from bile as conjugated morphine; enterohepatic circulation of morphine and morphine glucuronide accounts for presence of small amounts of morphine in urine for several days after the last dose.

Codeine (salts: Codeine sulfate and codeine phosphate): is methyl morphine; common ingredient of cough medicines such as Actifed-C Expectorant, Robitussin A-C, Phenergan Expectorant with Codeine, Dimetane Expectorant-DC, Terpin Hydrate and Codeine Elixir; also used in combination analgesic preparations such as Tylenol and Codeine; also used in preparations containing sympathomimetic amines such as pseudoephedrine, phenylpropanolamine (Actifed-C, Sinutab

with Codeine); orally a does of 32 mg codeine is approximately equianalgesic with 600 mg of aspirin.

Excreted as unchanged codeine, conjugated codeine, morphine (free and conjugated) and norcodeine; approximately 10% of ingested codeine is demethylated forming morphine; after 30 mg dose of injected codeine the following values were found - morphine 5-17%, norcodeine (N-demethylated codeine) 10-21% of the total, conjugated codeine 32-51% of the total, and unchanged codeine 5-10% for a 24 h period; nor-morphine also has been detected; one-half of a dose of labeled codeine was excreted within 6 h; immunoassay techniques cannot differentiate morphine from codeine - they are more sensitive to codeine than morphine.

Methadone: is dl-4,-diphenyl-6-dimethylamino-3-heptanome hydrochloride; synthesized by Germans during World War II; l-methadone has most of the analgesic activity.

N-demethylation is major metabolic pathway with cyclization to form pyrrolidines and pyrroline which are excreted in urine and bile; a small amount (less than 10%) excreted unchanged but the amount increases when the urine is acidified; primary metabolite is 2-ethylidene-1-5-dimethyl-3,3-diphenyl-pyrrolidine; the metabolite is less affected by changes in urine pH.

Meperidine: is a phenylpiperidine, ethyl-1-methyl-4-phenylpiperidine-4-carboxylate; second to morphine in use; hydrolyzed in the liver to meperidinic acid which is excreted free and conjugated; also N-demethylated to normeperidine (half-life of 15-20 h) which may be hydrolyzed to normeperidinic acid which is subsequently conjugated to glucuronide; only 5% excreted unchanged of a 175 mg dose; excretion influenced by pH: increased in highly acidic urine.

Propoxyphene: structurally similar to methadone; readily absorbed after oral administration; N-demethylated in liver to form norpropoxyphene (T1/2 is 23 h) in liver to form norprophoxyphene; 3% of 130 mg dose excreted unchanged within 24 h and 5% in 72 h; clearance of both the metabolite and parent compound is slow with repeated use and may be more rapid in smokers.

Pentazocine: belongs to benzazocine series (also known as benzomorphan series; is a weak morphine antagonist; recovery of a dose of 56.5 mg in 24 h in the urine was 61.5%, 9.5% of the dose was unchanged, 11.4% was excreted as cis-alcohol, 40% was excreted as the trans-acid metabolite; oral preparation is now marketed with naloxone.

Hydromorphone: excreted in urine as glucuromide conjugate and 6-hydroxy hydro-morphone.

Fentanyl: extensively metabolized; small quantity excreted as parent drug.

STIMULANTS

Cocaine

Generic (and Proprietary) Names	Street Names
Cocaine hydrochloride	Crack, Coke, Snow, Uptown, Toot, Blow, Flake, Nose candy, Rich man's drug, Leaf

Administration: PO,IV, smoking, intranasal (snorting); Absorption: IV administration results in a rapid onset of action, within minutes; intranasal administration results within minutes; when smoked, as a "free base," the absorption is also within 5 minutes; Distribution: Widely distributed into the lipid tissues of the body; Metabolism: Cocaine is rapidly hydrolyzed by plasma and hepatic cholinesterase; Elimination: Via the urine; Approximate

detection time in urine after use: 4-6 days; 18-27 hours (as metabolites). Tolerance.

*speedballing is the injection of a mixture of heroin and cocaine.
*free-basing is the ether extraction process used to make cocaine smokeable; explosions and fire may result from this procedure.
*crack is a highly purified cocaine HCl; so named because of the popping sound which is made when the crystals are heated; when smoked and absorbed by the pulmonary vascular bed, its effects are comparable to intravenous injection.

Generic (and Proprietary) Names	Street names
Amphetamine (Benzedrine)	Bennies, Black Beauties, Peaches, Hearts, Splash, Uppers, Christmas trees,
Dextroamphetamine (Dexedrine)	Dexies, Uppers, Oranges Pep-pills, Wakeups, Ups, Lid, Proppers
Methamphetamine (Methedrine, Desoxyn)	Crystal, Meth, Speed, Water, White Crosses

Description: weak bases; Administration: nasal, IV, PO, or smoking; Absorption: Primarily via the small intestine with a peak absorption 1-2 hours; Distribution: Water and lipid soluble with wide body distribution; Metabolism: Primarily liver; Elimination: Drug and metabolites are excreted by kidney; acidic urine increases excretion; 30% of amphetamine excreted unchanged; methamphetamine excreted mostly unchanged - small amount demethylated to amphetamine.

SEDATIVE-HYPNOTICS

Barbiturates

Generic (and Proprietary) Names	Street Names
Secobarbital (seconal)	Reds, Red Devils, F-40's
Pentobarbital (nebutal)	Yellows, Yellow Jackets

Secobarbital and Amobarbitsl (tuinal)	Blues, Blue Birds, Rainbows
Phenobarbital	Phennies, Purple Hearts, Blue Heavens
Amobarbital (amytal)	Blues, Downers
Mephobarbital, metharbital, aprobarbital, butabarbital, talbutal	

Administration: Usually oral, but can also be taken IV; Absorption: Primarily in the small intestine; Distribution: Barbiturates are bound to plasma proteins in varying degrees; Metabolism: Primarily by the liver; Elimination: Some excreted by the kidney but the only barbiturates which have significant renal excretion are phenobarbital and barbital; Approximate detection time in urine after use: varied; over 30 days for phenobarbital; Tolerance.

HALLUCINOGENS

Generic Name (proprietary name)	Street Names
LSD (Delysip)	Acid, Blotter, Sugar, Trips, Cubes, Bid D, Window panes, Blue devil, mickies

Chemically it is (+)-N,N-diethyl-lysergamide; psychological effects discovered by Hofmann in 1943; doses as low as 20 to 25 micrograms cause an effect; rapidly absorbed orally and widely distributed; found that monkeys excreted less than 1% as unchanged drug in urine in 24 h after single i.v. dose; 70% was present in bile and intestine at the end of 12 h; metabolites appear to be 2-oxy-LSD, hydroxylation of the 12-position and conjugation with glucuromic acid.

Oral dose of 1-2 ug/kg gave the following effective concentrations, 10,50 and 90%, in ng/nl in non-fatal cases:

Biospecimen	EC10	EC50	EC90 (ng/ml)
Blood n=91	0.9	3.2	5.6
Urine n=75	2.0	9.7	28.0

PSYCHOTOMIMETICS

PCP

Generic (or proprietary) Names	Street Names
Phencyclidine (sennylan)	PCP, Angel Dust (sprinkled on parsley and smoked) crystal, mist, tic and tac, rag, hog, sheets, hawaiian woodrose mist, rocket fuel, peacepill, monkey, tranquilizer, crystal joints (CJS)

Abandoned for human use in 1965 due to severe psychotomimetic actions; Administration: Smoking, injection, as well as oral and nasal routes; Absorption: Well-absorbed by all routes; Distribution: A weak base; distributed in all body tissues including fat and muscle; Metabolism: PCP is metabolized in the liver; Elimination: Excreted by the kidneys via the urine and feces; Approximate detection time in urine after use: up to 200 hours. Readily metabolized to a hydroxylated derivative and excreted in the urine as a piperidine conjugate with little pharmacological activity; acidification of urine causes a 10-fold increase in excretion at pH 6.5 and more than 100-fold increase at pH 5.5.

Marihuana (Marijuana)

Generic (and proprietary) Names	Street Names
Tetrahydrocannabinol	Locoweed, Laughing grass, Mary Jane, Weed, Grass, Joy smoke, Acapulco gold, Reefers, Tea, Pot, Love weed, Indian hay
Hashish (more potent)	Hash

The common hemp is an herbaceous annual: Cannabis sativa is the sole species and Cannabis indica and americana are some of its varieties. The resinous exudate of the tops of the female plant contains most of the active ingredients; in Middle East and N. Africa, the resin is called hashish; in India, Charas; in India, the dried leaves and flowering shoots of the female plant (less active) are called bhang, and the resinous mass from the small leaves and brackets is called Ganja; in U.S., the term marihuana refers to any part of the plant or extract that causes psychological effects. The most active ingredient of cannabis is delta-9-trans-tetrahydrocannabinol or delta-9-THC by the dibenzopyran numbering system (adopted by the National Institutes of Mental Health) NIMH and is also known as deta-1-THC according to the monoterpene numbering system (adopted by the American Chemical Society).

Administration: Smoking, PO; Absorption: Smoking is the usual route, although oral absorption is also utilized with a slow 1-6 hour peak, with only 5-20% being absorbed; Distribution: When smoked, the THC is absorbed within minutes and circulated to the brain and peripheral tissues; THC is lipid-soluble and binds to most body tissues; it is also dissolved in the lipoproteins of the blood; Metabolism: THC is metabolized by the liver; Elimination: THC is eliminated in the feces and urine; Approximate detection

time in urine after use: 120 hours (single dose); 240 hours (daily use); 336 - 720 hours (chronic daily use) -- metabolites detected in all.

A study by the NIMH showed that metabolites appear within 10 min after administration of labelled THC: 30% was excreted in urine, less than 1% was unchanged and 11-hydroxytetrahydrocannabinol was a minor metabolite; 50% of the radioactivity was recovered in feces~20% of that was 11-hydroxy-tetrahydrocannabinol.

Definition of terms used in Clinical Pharmacokinetics

Absorption - the process by which a drug gains access to the blood from its site of administration, for example oral, intravenous, or pulmonary.

Accumulation - the process by which a drug builds up in the body as a result of repeated administration of a drug before the last dose is gone. In such a situation, the plateau plasma level is not reached until three to five half-lives have passed.

Bioavailability - the percentage of an administered dose reaching the systemic circulation (obviously, i.v. is already systemic).

Distribution - the process by and to the extent which the drug travels through the body and finds its way to the site of action (receptor) or to tissues where it may be stored or metabolized.

Volume of distribution (V_d) - "apparent volume of distribution" is the dose administered divided by plasma concentration (unbound drug + drug bound to plasma proteins). Plasma concentration is expressed in ug/ml.

Time to peak - time taken for peak blood or plasma concentrations to be reached following oral ingestion of a single therapeutic dose.

Protein binding - binding to proteins in plasma and gives an indication of "free" drug which is pharmacodynamically active.

Metabolism - the biochemical processes by which the body chemically modified a drug.

Elimination - the mechanism by which the body rids itself of foreign substances.

First-order kinetics - the rate of drug accumulation, elimination, or metabolism is proportional to the amount of drug in the body.

Zero-order kinetics - a set of circumstances in which the body metabolizes a fixed amount of drug per unit time regardless of the relative concentration of the drug, e.g., alcohol.

pka - the pH at which the drug is 50% ionized

Elimination half-life -the time required for any given drug concentration in blood plasma to decrease by one half.

Blood concentration - generally accepted that the pharmacological effects of a drug are related to its concentration at the receptor site. No fixed concentration below which a compound is therapeutic and above which it is toxic. Values given are categorized broadly.

TABLE I
TOXICOLOGICAL AND PHARMACOLOGICAL DATA FOR SOME SUBSTANCES OF ABUSE

Compound	Time to Peak(hr)	Protein Binding (%)	Elimination half-life (hrs)	pka	% in urine Unchanged	Vd (l/kg)	Distribution in Blood (% in plasma)	Therapeutic (ug/ml) (Plasma)	Toxic (ug/ml) (Plasma)	Fatal Range Blood (ug/ml)	Fatal Range Liver (ug/ml)
Alfentanil	---	90	1.6± 0.2 aged	---	<1	---	7.6 ± 2.4 aged	---	---	---	---
Amphetamine	1	---	10-34	9.9	17-73	---	50	0.03-0.6	>0.5	0.5-7.0	12-14
Amobarbital	---	61	12-27	7.7	1	0.5-1.2	---	0.1-2.2	2.3-23	20-30	23-83
Butabarbital	---	---	37.5	7.9	---	0.78	---	10	30-80	---	30-148
Cocaine	0.75	---	0.75-1.25	---	---	---	variable	0.03-0.31	---	0.9-2	0.1-20
Codeine	1-2	7	3-6	8.2	5-17	2.9-3.9	80	0.03-0.1	1.3-5	1-8.8	0.6-45
Diacetylmorphine(Heroin)	---	---	0.2	7.8	0.5	---	---	---	---	0.3	>0.35
Dihydrocodeine	1	---	9	8.8	20-56	---	---	0.07-0.13	0.3-2	0.8-17	0.3-13
Fentanyl	---	80	3.7±0.4	---	8	---	13±2	1ug/ml	---	1ug/ml	---
LSD	---	---	1.7	3.3	---	---	---	0.001-0.005	>0.004	---	---
Meperidine	1-2	65-75	2.4-4	8.7	6-20	4	Age dependent	0.1	1.5	---	---
Methadone	4	60-87	18-97	8.6	5-22	5	70	0.3-0.45	---	0.4-8.8	0.5-52
Methaqualone	---	80	10-43	2.4	---	6	---	0.4-8.3	16	5-4	26-89
Morphine	0.1-0.3	26-30	1.3-3.4	8.05	3-10	3-4	---	0.08-0.1	1-5	0.5-3	0.31-5
Pentazocine	1-3	60-70	2	9	3-15	3	50	0.03-0.16	0.2	>0.3	1.2-87

Table 1 continued.

Compound	Time to Peak(hr)	Protein Binding (%)	Elimination half-life (hrs)	pka	% in urine Unchanged	Vd (l/kg)	Distribution in Blood (% in plasma)	Therapeutic (ug/ml) (Plasma)	Toxic (ug/ml) (Plasma)	Fatal Range Blood (ug/ml)	Fatal Range Liver (ug/ml)
Pentobarbital	1	60-70	20-35	8	20	0.9-1	50	1-3	2-40	10-169	20-550
Phencyclidine	2.5	65	18-24	---	---	6.2	56	0.01-0.8	---	0.1-3.0	0.9-16
Phenobarbital	6-18	32-60	48-144	7.2	27-28	---	50	10-40	40-60	69-350	123-344
Secobarbital	3	46-70	40-49	7.9	5	1.5	---	18.5.3	2-15	5-52	15-330

An exhaustive review of collected literature is prohibited by space limitation.

VII. **Recommendations**

1. The collected published data in humans on the kinetics of the agents of interest to DTL should be analyzed for possible pharmacokinetic models that might permit some generalizations to individual cases of drug abuse.

2. The collected data on pharmacokinetics, physico-chemical characteristics, etc. should be computerized for quick access, as needed, and should be maintained in a current state.

3. A study should be performed to ascertain the urinary concentration of THC metabolite after oral administration of single and multiple doses of THC (possible by using cancer patients taking THC for the nausea of chemotherapy).

The above 3 recommendations form the substance of my Mini Grant proposal.

4. That DTL request that the Strughold Aeromedical Library subscribe to the following publications:

(1) Bulletin of Narcotics (New York)

(2) Biomedical Mass Spectrometry

New Name: Biomedical and Environmental Mass Spectrometry (London)

(3) Xenobiotica

(4) Journal of toxicology: Clinical Toxicology

5. That DTL establish a small but critically selective library containing important books and periodicals in the area of substance abuse.

6. That DTL subscribe to

(1) Journal of Analytical Toxicology
(already recommended and implemented)

(2) Current Contents - Life Sciences

(3) Excerpta Media (a) Drug Abuse, (b) Toxicology, (c) Forensic Sciences.

REFERENCES

1. Ambre, J., "The Urinary Excretion of Cocaine and Metabolites in Humans: A kinetic Analysis of Published Data," J. Anal.Toxicol. 9(1985) 241-245.

2. Aqurell, S., M. Hallidin, J.E. Lindgren, A. Ohlsson, M. Widman, H. Gillespie and L. Hollister, "Pharmacokinetics and Metabolism of Delta-1 tetrahydrocannabinol and other Cannabinoids with emphasis on Man," Pharmacol. Rev. 38 (1986) 21-43.

3. Breimer, D.D., "Clinical Pharmacokinetics of Hypnotics," Clin. Pharmacokinet. 2(1977) 93-109.

4. Heusler, H., "Quantitative Analysis of Common Anesthetic Agents," J. Chromatogr. 340 (1985) 273-319.

5. Holsztynska, E.J. and E.F. Domino, "Quantitation of Phencyclidine, Its metabolites, and Derivatives by Gas Chromatography with Nitrogen - Phosphorus Detection: Application for in vivo and in vitro - Biotransformation Studies," J. Anal. Tox. 10 (1986) 107-115.

6. Lue, L.P., J.A. Scimeca, B.F. Thomas, and Billy R. Martin, "Identification and Quantification of Phencyclidine Pyrolysis Products Formed During Smoking," J. Anal. Toxicol., 10 (1986) 81-85.

7. Odar-Cedarl, I., L.O. Boreus, U. Bondesson, L. Holmberg and L. Heyner, R. "Comparison of Renal Excretion of Pethidine (Meperidine) and its metabolites in old and young patients," Eur. J. Clin. Pharmacol., 28 (1985) 171-175.

8. Shimosato, K., M. Tomita, and I. Ijir: "Rapid Determination of p-Hydroxylated Methamphetamine Metabolites by Column Liquid Chromatography - Electrochemistry," J. Chromatogr., 25 (1986) 279-286.

9. Sitar, D.S., P.C. Duke, J.A. Owen, L. Berger, and P.A. Mitenko, "Kinetic Disposition of Morphine in Young Males after Intravenous Loading and Maintenance Infusions," Can Anesth. Soc. J., 33(1986) 145-149.

10. Sunshine, Irving (Ed.) Methodology for Analytical Toxicology Volume III, Boca Raton, Florida, CRC Press, Inc., 1985.

11. Smith, R.N. and K. Robinson, "Body Fluid Levels of Lysergide (LSD)," Forensic Sci. Int. 28(1985) 229-237.

1986 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/
GRADUATE STUDENT SUMMER SUPPORT PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

UNIVERSAL ENERGY SYSTEMS, INC.

FINAL REPORT

CONTRAST SENSITIVITY AT LOW LUMINANCE LEVELS

Prepared by:	William D. Shontz
Academic Rank:	Associate Professor
Department and University:	Department of Psychology Montana State University
Research Location:	Air Force Human Resources Laboratory Operations and Training Division Williams AFB, AZ 85240-6457
USAF Research	Dr Elizabeth L. Martin
Date:	1 April, 1986
Contract No:	F49620-85-C-0013

CONTRAST SENSITIVITY AT LOW LUMINANCE LEVELS

by

William D. Shontz

ABSTRACT

This research involved the study of contrast sensitivity under four luminance levels representative of those found in flight training simulator visual display systems. Contrast sensitivities for a range of luminance levels (32.3 ft L, 7.5 ft L, 2 ft L and .5 ft L), and spatial frequencies (.5, 1, 3, 6, 11.4 and 22.8 cycles/degrees) were assessed. A sample of subjects larger than is typically found in studies of contrast sensitivity was drawn from the personnel at AFHRL, Williams AFB. Luminance level was found to produce significantly different contrast sensitivity functions (CSFs) at intermediate and higher spatial frequencies. This finding has important implications for the recommendations given to design engineers regarding trade-offs in luminance, contrast, and resolution requirements for visual displays in flight simulators. The relatively large and heterogenous subject sample permitted evaluation of the effects of several subject variables. Post hoc analyses were conducted to assess the effects of corrected vision, age, and gender. None of these variables produced significant effects on contrast sensitivity. However, the psychophysical procedure used may have permitted response biases which could obscure sensitivity differences. Recommendations on these and other issues were made.

I. INTRODUCTION:

There is a remarkable match between the areas of ongoing research at AFHRL/OTE, Williams AFB and my background in research and teaching. The extent of this match is best illustrated with a table format.

SHONTZ BACKGROUND	AFHRL RESEARCH AREA
- Eye dominance and design of helmet-mounted displays	Fiber Optic Helmet Mounted Display development program
- Value coding in Visual search	Use of color in simulator displays
- Color coding in visual search	
- Predicting operator performance in reconnaissance systems	Level of target detail requirementsfor computer image generation systems. Visual

ACKNOWLEDGMENTS

I wish to express my great appreciation for the generous support of the work described in this report by the Air Force Office of Scientific Research/Air Force Systems Command (Contract No. F49620-85-C-0013) as administered by Universal Energy Systems, Inc. This work was done while I was a Summer Research Fellow at the Air Force Human Resources Laboratory at Williams AFB, Arizona. The program has been instrumental in updating my knowledge in several areas of applied research and broadening my perspective of potential research programs I might propose to the Air Force.

I also wish to express my deep appreciation for the morale and material support provided by my colleagues, both Air Force and contractor, during my sojourn at Williams AFB. A special thanks to Dr Elizabeth Martin as my Effort Focal Point for her moral and administrative support throughout the project. I am particularly indebted to Dr Peter Crane for his help in defining the research problem initially, his continued guidance throughout the project, and his critical role as sounding board for current and future research ideas. Dr David Hubbard helped bring the project together with his data day support during the statistical analysis phase. Without him the project would have been no pun at all. Lt Kevin Dixon's unfailing enthusiasm for the project kept me on my toes during the problem resolution phase of the project, and his aid with data collection and analysis helped bring it to a timely and successful conclusion. I am also indebted to Dr George Geri for sharing his knowledge of the contrast sensitivity literature and the Optronix equipment, and to Don Fike for his equipment contribution and excellent graphics support. Last but not least I would like to express my deep appreciation to all the HRL, University of Dayton Research Institute and other contractor personnel who so generously donated their time as subjects and/or shared their knowledge of specific research areas and issues. Every person I dealt with was unfailingly kind and supportive. This made my stay a most pleasant experience.

No statement of acknowledgement of the Summer Faculty Research Program would be complete for me without a generous thank you to Rodney Darrah, Director of the program for UES. For his patient attention to my seemingly endless questions and his prompt response in resolving problems, he has my deep appreciation for helping make this both a pleasant experience and one that has been extremely useful to me professionally.

- Visual parameters in Low-Level High Speed Flight	scene requirements for simulating Low Level Flight Tasks
- Factors affecting the selection and training of forward air controllers	
- Target identification as a function of angle of regard	
- Perceptual learning theory, vision physiology, psychophysics	
- Basic sensory/perceptual processes	Brightness, contrast and resolution requirements for simulator displays

Several possible areas of research were discussed with HRL personnel prior to and during my "pre-summer" site visit in October 1985. We were seeking to identify a potential research project which met the following criteria: (1) greatest impact on ongoing research, (2) could be completed within 10 weeks, and (3) provided an opportunity to enhance my knowledge in a particular area of vision research. We decided that the research reported in this final report best met all three of these criteria.

The study reported herein was designed to provide basic data on human visual performance which would directly support a research project being initiated as a part of the training effectiveness program for simulator display designs being conducted by HRL/OTE at Williams AFB. The "parent" project will evaluate luminance, contrast and resolution parameters in a simulator environment. The study reported herein was designed to provide preliminary data on the impact of low luminance levels on contrast sensitivity. As such, it provides specific contrast threshold data on subjects who will be used again in the parent study under similar luminance conditions. The results answer a critical question regarding the impact of low luminance levels on contrast sensitivity.

II. OBJECTIVES OF THE RESEARCH EFFORT:

A preliminary and primary objective of my research effort was identified very early in my discussions with my HRL colleagues. Secondary objectives evolved as the details of research design and implementation were developed. These research objectives were as follows:

(1) To determine the effects of specific low luminance levels on contrast sensitivity functions across a wide range of spatial frequencies.

(2) To obtain contrast sensitivity data on a relatively large sample of subjects who would also represent the subject pool for the parent study.

(3) To assess the effects of several subject variables, namely gender, vision correction, and age on contrast sensitivity performance.

(4) To develop the basis for follow-on research into the relationships among a number of factors affecting contrast sensitivity and their impact on visual performance tasks in both the simulator and the real world.

Both primary and secondary objectives were met during the course of the research period.

III. BACKGROUND:

The present study represents an attempt to provide data on contrast sensitivity functions (CSFs) in support of applied research on luminance, contrast, and resolution requirements in simulator visual display systems. A good deal of theoretical work has been done on contrast sensitivity in humans, especially following the seminal work of Campbell and Robson in the mid-1960's (e.g. Campbell and Robson, 1968). Most of this work is classical psychophysical research both in terms of the paradigms used and number of subjects employed (i.e., only one or two highly trained observers - usually the authors). While such research is excellent in helping us understand the dynamics of CSFs, it is of limited use in making recommendations to design engineers regarding required characteristics of displays. The data are gathered under much too artificial an environment. As Van Nes and Bauman (1967) point out, many variables affect the quantitative nature of CSF data and in some circumstances the qualitative nature as well.

Murch and Virgin (1985) discussed the issues of resolution and addressability in the design of raster generated displays and emphasized the important role of contrast senstivity as a moderator variable in design decisions. They noted that "contrast sensitivity increases for marginally higher luminances (above the 10-15 ft L they illustrate) and decreases markedly (my emphasis) for lower luminances" (p. 103). This becomes an

important issue in the design of aircraft simulator displays. Viewing screen luminance in these systems ranges from approximately 10 ft L down to as low as .5 ft L. While the basic research data (e. g., Patel, 1966; Van Nes and Bouman, 1967) support the Murch and Virgin position in terms of qualitative effects of luminance level on contrast sensitivity, there is no reason to expect significant quantitative differences in performance at the luminance levels of interest. Stimulus presentation methods used in contrast sensitivity studies differ widely in stimulus characteristics and equipment capabilities. Viewing conditions are highly artificial. To further complicate the matter, the human visual system is not linear (Cornsweet, 1970). Therefore, These data from highly stylized basic research studies cannot be extrapolated, even to the simulator environment.

Luminance levels investigated in the present study were 32.3 ft L, 7.5 ft L, 2 ft L, and .5 ft L. To put these luminance levels in perspective, it may be noted they fall at the low end of the luminance range for photopic vision. For example, fresh snow on a clear day reflects 10,000 ft L; average earth on a clear day reflects 1,000 ft L; average earth on a cloudly day reflects 100 ft L; white paper in good reading light is 30 ft L; and white paper one foot from a standard candle less than 1 ft L (Grether and Baker, 1972). The 32 ft L level used in the present study greatly exceeds current luminance levels in simulator visual displays. On the other hand, it is about half the luminance anticipated from helmet-mounted displays for simulators now under development. As a further frame of reference for the reader, an IBM PC monitor has a maximum brightness of about 5 ft L.

The present study was undertaken to generate CSFs for specific luminance levels of interest. Viewing conditions were somewhat different from the simulator setting but were controlled to the extent necessary to obtain interpretable data. A large (by comparison) and varied subject population was used to generate the CSFs. In recruiting subjects, an attempt was made to obtain maximum overlap in subjects participating in the present study and those who will be participating in a related target recognition study in the simulator setting. Thus there should be an opportunity to further evaluate Ginsburg's (Ginsburg, Evans, Sekuler & Harp, 1982) contentions regarding the relationship between standard measures of acuity, contrast sensitivity performance and target recognition. Understanding the exact nature of this relationship is of particular importance at low levels of luminance.

In testing the effects of luminance level across spatial frequencies, we expected to obtain the ubiquitous U-shaped function found in the literature; i.e., lower contrast sensitivity at both low and high frequencies - highest sensitivity in the mid-range. The question of interest was - will there be significant differences in contrast senstivitiy among the levels of luminance under investigation? If we obtained the U-shaped function it would in a sense validate our paradigm and control procedures. If we obtained significant differences as a function of relatively modest changes in luminance level, this would have important implications for recommended trade-offs among luminance, contrast, and resolution in the design of simulator visual displays.

IV. METHOD:

Subjects: Eleven (11) male and five (5) female professionals and technicians, ages 23-60, were recruited from HRL staff and in-house contractor personnel to serve as observers in the experiment. All had 20/20 or better visual acuity either corrected or uncorrected.

Apparatus: Contrast sensitivity thresholds were obtained using the Optronix Model 200 Vision Tester which has been described elsewhere (Ginsburg and Cannon, 1983). Basically, the equipment uses a raster-scan CRT to display sinusoidal gratings at different spatial frequencies. Contrast on the video screen was defined conventionally as $(L_{max} - L_{min}/L_{max} + L_{min})$. The Optronix display was adjusted to an average screen luminance of 32.3 ft L with peak contrast of .493. The system was then internally calibrated according to the appropriate calibrating procedure for the device. Luminance levels were manipulated in a manner similar to Patel (1966) by having the observers wear goggles containing neutral density filters to produce the desired luminance level when viewing the screen. Kodak Wratten No. 96 gelatin neutral density filters were combined to provide densities of .6, 1.2 and 1.8. These together with a no filter condition provided the four levels of luminance desired. Average screen luminances desired were 32, 8, 1.7, and .5 ft L. Average luminances achieved were 32.3, 7.5, 2, and .5 ft L, respectively. Peak contrasts at these luminances ranged from .493 to .480. A check of the luminance levels after data collection was completed showed the luminance levels and peak contrasts to have held within $\pm$ 2% or less at all levels. A Spectra Pritchard Model 1980A photometer was used to calibrate the Vision Tester to

the non-standard luminance levels used in the study.

Viewing distance to the screen was 3m. Viewing distance was monitored during data collection but no head rest or bite-bar was used. Spatial frequencies used were .5, 1, 3, 6, 11.4 and 22.8 cycles/degree. These frequencies were optimum in terms of capabilities with the raster scan for precision presentation of sinusoidal grating patterns according to the operating manual for the Optronix Vision Tester. Some greater flexibility in spatial frequencies used would have been desirable, especially at the high end of the scale, but limitations in the equipment precluded this.

Design: A within-subject 4x6 factorial design was used to test for effects of luminance level and spatial frequency. Gender, vision correction and age were used as covariates. Four different sequences of luminance level were determined using a latin square design with four subjects in each order group. The order in which spatial frequencies were presented was randomized within each luminance level for each subject.

Procedure: Subjects were brought into the experimental chamber and seated in a chair so that viewing distance to the screen was 3m. At 3m the 22.5 cm x 28.5 cm screen subtended visual angles of 4.3^{o}x 5.4^{o}. Lights were turned out and S's were adapted to the level of luminance to be used in the first data session for at least 5 minutes (longer for the low luminance levels). During this time the goggles were donned by the subjects and the experimenter began giving the instructions. Subjects were given a brief explanation of the general nature of the experiment and were instructed in the sequencing of the experimental task, cues to expect, and how and when to respond. Subjects were told to establish a criterion for when the sinusoidal grating was "just-visible." They were not coached as to what the screen should look like at this "just-visible" point.

The experimental task proceeded as follows. At the beginning of each block of 10 trials a 1 sec. preview at peak contrast was given of the spatial frequency to be presented during that block of trials. The preview was followed within 1-2 seconds by the first trial of the block. Each new trial began 1-2 sec after the subject had responded until all 10 trials had been presented. A 10-15 second break occurred at the end of each block of trials while the vision tester printed out log contast thresholds for each trial and computed means, standard deviation, and sensitivity values for the 10 trials. Stimulus gratings were presented using the Method of Increasing

Contrast (Kohfeld, 1985). This particular method, which is essentially the ascending portion of the Method of Limits, was judged by Kohfeld to be as free of response bias as any of the methods of stimulus presentation available in the Vision Tester's preprogrammed modes. Onset contrast varied randomly between 0 and .001 on each trial with contrast increasing at a rate of 4 decibels per second. Thus the time till the grating became visible varied slightly from trial to trial precluding the use of a timed interval to respond.

Subjects were given three blocks of training trials at the beginning of the session using the luminance level of the first set of data trials. Three spatial frequencies were used for training which represented the range of frequencies used for the data trials, but did not duplicate any of these frequencies. During the training trials subjects were to sharpen their concept of "just-visible." They were reminded throughout the data sessions to be as consistent as possible in their judgments of "just-visible." At the end of the training trials the sequence was interrupted and subjects were asked if they had indeed established their criteria for "just-visible." All subjects reported having established their criterion by this time. Response consistency during the latter training trials confirmed this. A 3-5 minute break was given after each three blocks of trials and a 5-10 minute break given while a second sequence of six blocks of trials was programmed into the Vision Tester. Data collection for each subject required two 45 minute sessions, usually conducted on succeeding days at approximately the same time of day. The procedure including the use of training trials, was the same on both days.

A post experiment questionnaire was administered to each subject following data collection. The PEQ dealt with the following issues:

(1) Whether the goggles were a distraction in performing the tasks/

(2) The consistency with which subjects felt they had maintained their criterion throughout data collection.

(3) Descriptive statements of the appearance of the screen when the grating was "just visible."

(4) Any variation in the appearance of the screen at "just visible" as a function of spatial frequency.

The goggles were not a distraction. Other findings are summarized in the Results and Discussion Sections.

V. RESULTS

Reduction of the data for analysis included computing the 20% trimmed means (Rosenberger and Gasko, 1983) for each luminance level by spatial frequency combination for each subject. These trimmed means then became the unit of analysis for the ANOVA subsequently performed. Computing the 20% trimmed means in this case involved rank ordering the 10 log contrast threshhold values obtained from each subject for each luminance/frequency combination and eliminating the two highest and two lowest values. A simple mean was then computed of the remaining six observations. This procedure improves the stability of the means and hence the estimate sensitivity. Thus, trials in which the subject lets his or her criterion change markedly or fails to respond in a timely manner are discarded. Subjective reports from observers following data collection suggest these things did happen occasionally.

Figure 1 illustrates the results of the experimental relationship between luminance level and spatial frequency. Qualitatively, the data relate well to the CSFs found in the literature, i.e., the distinctive U-shaped functions. This suggests that the general research paradigm and control procedures were appropriate and adequate, respectively. Of the main effects and interactions of interest, only three were significant. The order in which subjects were exposed to the different luminance levels had no effect on performance. Luminance level produced a systematic (with an exception noted later) and significant effect on contrast sensitivity, F, $(3, 36) = 36.41$ $p < .001$. As was expected, the effect of Spatial Frequency was significant, F $(5, 60) = 173.71$, $p < .001$. The Luminance X Spatial Frequency interaction was also significant, F $(15, 180) = 31.16$, $p < .001$. This latter relationship is illustrated by the way the CSFs "fan out" as luminance level decreases and spatial frequency increases. Similar relationships have been found by other investigators (C.F. Van Nes and Bouman, 1966). The differential impact of luminance level is greatest at the intermediate spatial frequencies. At the lowest frequencies (.5 and 1 cycle/degree) there is essentially no effect due to luminance level. At high spatial frequency levels (22.8 cycles/degrees and beyond), the curves begin to converge. This latter tendency is less evident than at the low frequencies in Figure 1. However, pilot data taken before beginning the main experiment indicated that at frequencies beyond 22.8 c/deg, the curves

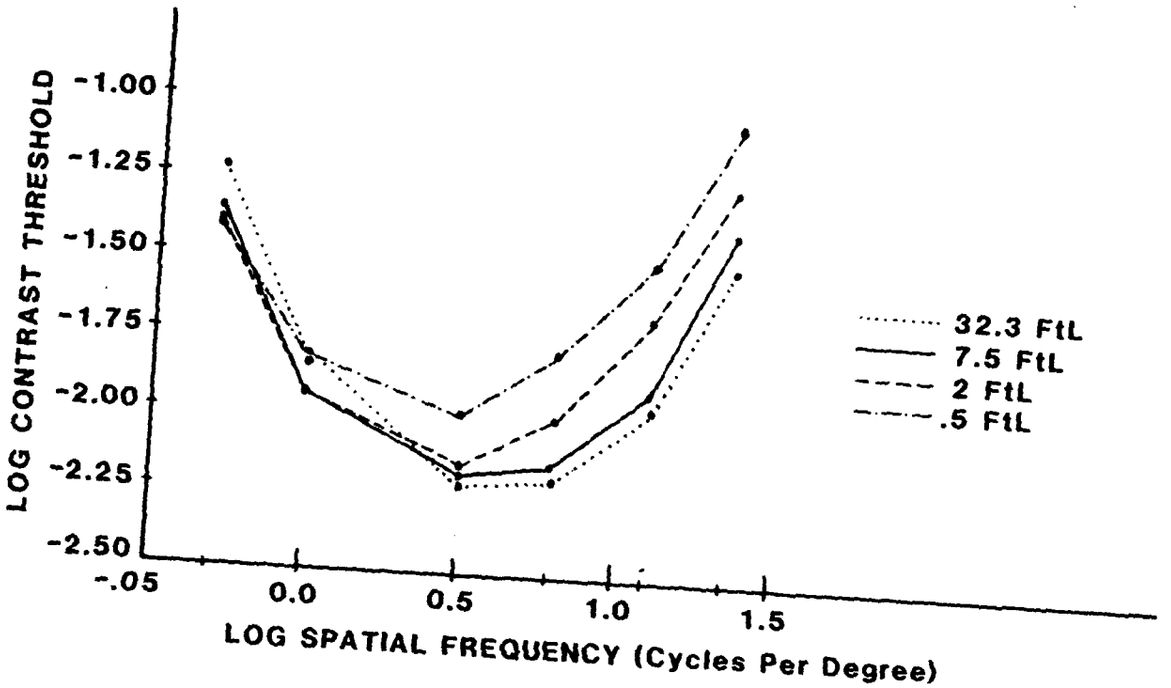


Figure 1. Contrast threshold functions illustrating the relationship between luminance level and spatial frequency

do rapidly converge. Nevertheless, a pair-wise comparison of the thresholds at 22.8 c/deg using the Bonferoni test showed contrast sensitivities at all levels of luminance to be different at the .05 level.

The single qualitative departure of the data from that are typically found in the literature was the performance at the highest luminance level and lowest spatial frequencies. This departure was quite consistent across subjects. A pair-wise comparison of means at the lowest spatial frequency (.5 c/deg) using the Bonferoni test showed the difference between thresholds at 32.3 ft L and the other three luminance levels to be significant at the .05 level. The comparison at 1 c/deg showed the difference in performance between the 32.3 ft L condition and both the 1.5 ft L and 2 ft L conditions to be significant at the .05 level as well. The difference between 32.3 ft L and .5 ft L was not significant.

Post hoc analyses were conducted to test for corrected vs. uncorrected vision and for gender effects. Neither of these variables produced significant differences in performance although the data were highly consistent.

A Pearson Product Moment correlation was calculated for age vs. contrast sensitivity. The resulting $r = -.07$ indicates no relationship between age and sensitivity for the sample. These results do not coincide with those of Owsley, Sekuler and Siemsen (1983).

Responses to the post experiment questionnaire revealed three conditions which should be given additional attention in future research. The first of these had to do with the appearance of the screen when subjects determined the grating to be "just visible" at the higher (11.4 and 22.8 c/deg) spatial frequencies. Subjects almost universally reported that at these frequencies grating did not appear to be present in a uniform manner across the screen. Rather, the grating was detected by noting its presence in patches over the screen. The location of the patches of grating was not consistent across subjects. Some subjects looked for the "patches" in the middle of the screen, while for other subjects they appeared in different areas of the screen. I have found no mention of this phenomenon in the literature.

The second condition, less widely reported, was that of a criterion shift between the first and second data collection session. Of those people reporting a shift, all reported it to be in the conservative direction. The third condition was noted by all but was of concern only to a few. Because of the nature of the program controlling stimulus presentation in the

Optronix Vision Tester system, the range of starting contrasts for each trial was always the same. This meant that at low luminance levels and high spatial frequency the subject had to wait a much longer time for the grating to appear than at either higher luminance levels or the mid-range of spatial frequencies. With more flexibility in programming of the stimulus presentation and pilot data for guidance, the approximate time to threshhold could be stabilized somewhat.

VI. DISCUSSION:

The primary objective of the experiment was to evaluate the effect of luminance level on contrast sensitivity as a function of spatial frequency within a range of luminance levels representative of those found in aircraft simulator displays. The results indicated that luminance level was indeed a limiting factor in resolution especially at higher spatial frequencies. Murch and Virgin (1985) suggested in passing that performance would decrease markedly for lower luminance but their primary focus was on issues of resolution and addressability in CRT display design. It is now clear that luminance level and contrast, as well as resolution, must be given serious consideration in display design. The limits of the trade-offs which may be possible among luminance level, contrast and resolution remained to be defined. As Patel (1966) notes, contrast sensitivity can only be enhanced to a point, i.e., when contrast reaches 100%. Beyond this, luminance must be increased to take advantage of system resolution. At presently achieved luminance levels in simulator visual display systems, there is considerably more resolution than observers can use at low contast.

The non-linearity of the human visual system in response to changes in luminance and contrast are illustrated by the data from the present study. This will complicate the task of developing predictive formulae for design engineers. For example, in the present study the greatest differential effects of luminance on contrast sensitivity were found at the intermediate spatial frequencies. It is in this area in particular that Ginsburg, et al (1982) suggest there is no relationship between acuity and contrast sensitivity. Obviously, we need to know a great deal more about how contrast sensitivity affects visual performance in both simulators and real world situations. In particular, we need to relate this information in a systematic way to the effects of luminance level.

One anomoly in the data which should be given further attention was the

reversal of performance at the highest luminance level (32.3 ft L) and lowest spatial frequencies (.5 and 1 c/deg). Performance at the .5 c/deg frequency was reliably poorer (higher threshhold) than any of the other three luminance levels at that frequency. It was also poorer than the 7.5 ft L and 2 ft L luminance conditions at 1 c/deg. This reversal of expected results was very consistent across subjects and is not reported eleswhere in the literature. It may be an artifact produced by the equipment used but no other investigators using the same system (Ginsburg, et al, 1982; Kohfeld 1985, 1986; Owsley, et al, 1983) report such an effect. Explanations regarding the visibility of low-frequency sine-wave targets (McCann, Savoy & Hall, 1978) would not account for this effect either. The stability of this effect and its implications for the design and use of simulator displays should be investigated further.

Post hoc analyses were also conducted to assess the effects of several subject variables. Researchers investigating contrast sensitivity typically use such small sample sizes that information on the effects of subject variables is for the most part missing from the literature. For example, Fiorentini and Maffei (1976) report that myopic subjects though corrected to normal vision still have much lower contrast sensitivity than observers with normal vision. In the present study, half the subjects had corrected vision. Corrections were for either myopia or presbyopia. However, no effect due to vision correction was found in the present study. It may be that the degree of myopia was quite different on the average for subjects in the present study vs. those of Fiorentini and Maffei. Further, the early experience explanation for the poor performance of myopics given by Fiorentini and Maffei would not apply to the presbyopes in the present study. While this subject variable deserves further investigation, it appears that investigators can use subjects with corrected vision without any great concern about a subject variable confound.

The systematic effects of age on contrast sensitivity found by Owsley, Sekuler and Siemsen (1983) were not obtained in the present study. It may be that reliable differences due to age are only found in subjects beyond 60 years of age. In fact, a casual evaluation of the Owsley, et al data would tend to support this conclusion. Again this is good news for investigators who have access to only a limited and somewhat older (when compared to the ubiquitous introductory psychology student) subject population.

Gender effects were also evaluated in the present study. While the resulting differences were not significant, they were so consistent that this subject variable merits further investigation. At every combination of luminance and frequency, females had lower contrast threshold on the average then males. Again, this issue is not addressed in the contrast sensitivity literature. The amazing consistency of the data suggests that it should be. When using totally within subject designs the subject variables are not confounded with manipulated variables. However, one should be cognizant of the possibility of confounding in between subject-designs.

Some of the issues identified through the post-experiment questionnaire merit consideration for further investigation. The nature and implications of the appearance of the contrast sensitivity display at threshold for higher spatial frequencies should be evaluated systematically. Again, we may have been dealing with an artifact of the particular equipment used. If not, the conditions would have serious implications for the detection of high spatial frequency/low contrast targets. It may also be useful to evaluate the effects of the relatively long wait for contrast targets to reach threshold under high frequency/low luminance conditions. While this could be considered a realistic aspect of target detection there may be systematic effects on contrast sensitivity which should at least be understood.

Most of my concern in interpreting the results of this study is focused on the effects of response bias on the sensitivity data. This is not to say that the relative relationships between luminance level and spatial frequency identified in the present study are not real and reliable. However, parametric data for display design should be as free of response bias as possible. "Real world" noise can be factored into the equations later. The Method of Increasing Contrast (MIC) for stimulus presentation used in the present study was used for several reasons. Kohfeld (1985) presented evidence that the MIC technique was the most response-bias free technique of those available within the program range of the Optronix system. Further, the technique has a common paradigmatic bond with a comparison study underway on aircraft recognition as a function of display luminance, contrast and resolution. This was an important consideration with respect to further assessing the relationship between contrast sensitivity and performance on certain simulator tasks. However, the

potential for wide variation in the subjects' criterion for "just-visible" is still possible and was present to a degree in the present study. The nature and extent of response bias cannot be measured using the MIC technique. This, of course, is the major flaw of the classical psychophysical procedures if you wish to make statements about sensitivity alone. The present study specifically and intentionally avoided the artificiality of the psychophysics laboratory in many respects; this was good. However, when one increases the sample size as was done in the present study, subjects are more naive to psychophysical procedures. This is bad in that both within and between subject variability increases with respect to the measure of sensitivity due largely to response bias. We need to look at many of the variables affecting contrast sensitivity which we reported in the literature as well as those affecting applications of contrast sensitivity data to real world problems with a research paradigm free of response bias.

Use of a signal detection theory paradigm would allow assessment of both sensitivity and response bias in a systematic way. In fact, it might be very instructive to evaluate response bias as a phenomenon in and of itself as a function of many of the variables assessed in the present study.

VII. RECOMMENDATIONS:

The recommendations following from the present study are organized under three headings: application, follow-on, and clarification. Each are discussed in turn below.

Application: This experiment was designed to produce basic data on CSFs at luminance levels commensurate with those found in large field-of-view display systems present in fighter aircraft simulators using collimated light. The CSFs obtained represent threshold performance data. Threshold performance should be related to performance on an aircraft recognition task carried out in a simulator visual environment. For example, contrast threshold performance in the present study should predict to aircraft recognition performance (in terms of recognition range i.e., target size). Data from the present study indicate that recognition range should be markedly reduced when luminance levels are reduced and contrast is held constant. Direct comparisons of performance across tasks within subjects should be made.

Follow-On: It is important to understand the performance of the human visual system at threshold contrast levels in order to understand the quantitative and qualitative nature of the non-linearities in the system. In order to make recommendations to both engineers and programmers as to how to best manipulate luminance, contrast and resolution in aircraft simulator display systems, it is necessary to take a next step in describing visual system performance. This next step involves acquiring data on the visual system at supra-threshold levels of contrast. Cornsweet (1970) has discussed the basic issues involved in the contrast matching phenomenon and some of the implications for perceptual phenomena. Clearly, we need precise data on discernible levels of contrast within the luminance range of current and anticipated aircraft simulator display systems. These data would provide the basis for generating contrast control functions within the software used for scene content and target generation. They would also be necessary in specifying hardware design parameters for the displays. The same problem of non-linearity in the response of the human visual system at threshold must be dealt with at supra-threshold levels of contrast using contrast matching techniques. The present study serves as a starting point for follow-on research in this area.

Clarification: Several issues were raised in the present study which should be addressed. Fortunately, the same experimental equipment which would be used to study contrast matching performance would be easily adaptable to the design of clarification studies. The three issues demanding most immediate attention are: (1) the reversal in expected contrast sensitivity at low spatial frequencies and high luminance; (2) the role of response bias in contrast sensitivity due to subject variables, and (3) the highly stable gender differences found. These issues should be addressed in a series of studies which could easily precede an extensive investigation of supra-
threshold contrast performance using contrast matching techniques.

VII. REFERENCES:

Campbell, F.W. and Robson, J.G. (1968). Application of Fourier analysis to the visibility of ratings. Journal of Physiology, 197, 551-566.

Cornsweet, T.N. (1970). Visual Perception, New York: Academic Press.

Fiorentini, A. and Maffei, L. (1976). Spatial contrast sensitivity of myopic subjects. Vision Research, 16, 437-438.

Ginsburg, A.P., and Cannon, M.W. (1983), Comparison of three methods for rapid determination of threshold contrast sensitivity. Investigative Opthalmology and Visual Science, 24, 798-802.

Ginsburg, A.P., Evans, D.W., Sekuler, R. and Harp, S.A. (1982). Contrast sensitivity predicts pilot's performance in aircraft simulators. American Journal of Optometry and Physiological Optics, 59, 105-109.

Grether, W.F. and Baker, C.A. Visual presentation of information. In Van Cott, H.P. and Kinkade, Robert G. Human engineering guide to equipment design. (Revised edition), Washington, D.C.: US Government Printing Office

Kohfeld, D.L. (1985a). An evaluation of visual contrast sensitivity measures. In Eberts, R.E. and Eberts, C.G. (Eds) Trends in Ergonomics/Human Factors II, Elsevier Science Publishers, B.V. (North Holland).

Kohfeld, D.L. (1985b) Role of stimulus uncertainty in visual contrast sensitivity. Final Report. Summer Faculty Research Program, Sponsored by Air Force Office of Scientific Research, HRL/OTE, Williams AFB, Arizona

McCann, J.J., Savoy, R.L. and Hall, J.A. Jr. (1978) Visibility of low-frequency sine-wave targets: Dependence on number of cycles and surround parameters. Vision Research, 18, 891-894.

Murch, G. and Virgin, L.N. (1985): Resolution and addressability: How much is enough? Society for Information Display. International Symposium Digest of Technical Papers, Volume XVI.

Owsley, C., Sekuler, R. and Siemsen, D. (1983). Contrast sensitivity throughout adulthood. Vision Research, 23, 689-699.

Patel, A.S. (1966). Spatial resolution by the human visual system. The effect of mean retinal illuminance. Journal of the Optical Society of America. 56, 689-684.

Rosenburger, J.L. and Gasko, M.N. (1983). Comparing location estimates: Trimmed means, median and trimean. In Hoaglin D.C., Moseller, F., and Tukey, J.W. (Eds) Understanding robust and exploratory data analysis, New York: John Wiley & Sons, Inc.

Van Nes, F.L. and Bouman, M.A. (1967). Spatial modulation transfer in the human eye. Journal of the Optical Society of America, 57, 401-406.

1986 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT SUMMER SUPPORT PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

Aerodynamic Parameters for a Rapidly Pitching Airfoil

Prepared by:	Dr. William D. Siuru, Jr.
Academic Rank:	Senior Research Associate
Department and University:	Space and Flight Systems Laboratory University of Colorado at Colorado Springs
Research Location:	Frank J. Seiler Research Laboratory Directorate of Lasers & Aerospace Mechanics Mechanics Division
USAF Researcher:	Major John Walker
Date:	10 August 1986
Contract:	F49620-85-C-0013

Aerodynamic Parameters for a Rapidly Pitching Airfoil

by

William D. Siuru, Jr., PhD, PE

ABSTRACT

The key aerodynamic parameters, lift, pressure drag, and moment coefficient were analyzed for a rapidly pitching airfoil. The investigation was based on extensive wind tunnel data obtained for a NACA 0015 airfoil pitched at rates between 115 and 1380 degrees/second to a maximum angle of attack of 60 degrees. Maximum as well as average values were determined as a function of the non-dimensional parameter alpha +. The angles of attack at which maximum coefficient values occured were determined also as a function of alpha +. Empirical expressions were found for the key coefficients as a function of alpha +.

ACKNOWLEDGMENTS

I would like to thank the Air Force Office of Scientfic Research (Air Force Systems Command) and especially the Frank J. Seiler Research Laboratory for sponsoring me in this project. Special thanks are given to Major John Walker for his very valuable guidance and comments during the course of the effort. I also greatly appreciate the help of my son, Brian, and daughter, Andrea, who spent many tedious hours reducing data and running computer programs.

I. Introduction

I received my PhD from Arizona State University doing my dissertation research in turbulent fluid flow. During my 24 years of active duty with the USAF, I had a four year assignment as the Commander of the Frank J. Seiler Research Laboratory. During this period I became quite interested in the laboratory's research in unsteady aerodynamics, and in particular the aerodynamic phenomena of a rapidly pitching airfoil. The significant increase in lift resulting from a rapidly pitching airfoil has potential for providing highly manueverable aerospace vehicles that are of interest to the military services. However, the positive results, as well as the negative side effects, must be quantified if they are to be of value to the designer and decision maker. Thus, I was assigned the task of correlating the commonly used aerodynamics parameters from the extensive wind tunnel testing program that FJSRL has conducted over the past few years. From this analysis, then, conclusions were drawn about the phenomena.

II. Objectives of the Research Effort

The objectives of my individual research were:

1. A comprehensive review of the data obtained in wind tunnel investigations at the Frank J. Seiler Research Laboratory and at the University of Colorado.

2. Correlation and presentation of the experimental data in the form of commonly used aerodynamic parameters such as lift (C_l), pressure drag (C_d), and moment ($C_{m1/4c}$) coefficients as well as angles of attack (α) and non-dimensional time (T) for such events as maximum lift and drag. These parameters were correlated on the basis of the experimental variables, e.g. pitching rate ($\dot{\alpha}$), freestream velocity (U_∞), and finally, the non-dimensional pitch rate (α^+).

3. Draw meaningful conclusions about the fluid flow phenomena of a rapidly pitching two-dimensional airfoil.

The end result of the investigation was (1) a better understanding of the phenomena, (2) quantitative parameters useful in applying the concept, and (3) a basis for comparison for the subsequent investigations with a three-dimensional wing.

III. Approach

The methodology for obtaining experimental data has been extensively documented in references [1] through [4] and thus will not be repeated here. The test conditions used in this investigation are summarized in Table I.

The correlation of the data consisted of plotting the data as a function of the experimental variables.

Table I

TEST CASES CONSIDERED

Pitch Rate (deg/sec) / Freestream Velocity (fps)	0	115	230	460	920	1380
10		(.1) 1151	(.2) 232	(.4) 464	(.8) 928*	(1.2) 1381*
20			(.1) 231	(.2) 462 462A	(.4) 924 924A	(.6) 1386A
30						(.4) 1384
40	(0) 00			(.1) 461 461A	(.2) 922 922A	
60				(.067) 4607A		(.2) 1382 1382A
80				(.05) 4605A	(.1) 921A	

LEGEND Number in parenthesis is the non-dimensional pitch rate defined by $\alpha^+ = \dot{\alpha} c/U$ where c is the chord dimension.

A - data taken at FJSRL, rest taken at CU

* - data of limited value.

Where more than one set of test results were available at a test condition their mean value was used. In many instances, a empirical curve fit was used to determine an analytic expression. Both a linear and logarithmic least squares fit were used and the expressions for both relations are presented where appropriate.

IV. Similarity Parameters

The parameter, $\alpha^{+} = \dot{\alpha} c/U_{\infty}$ has been shown to be a good similarity parameter for this type of flow situation. Figure 1. shows how the lift, pressure drag, and moment coefficient versus angle attack curves collapse for a given value of α^{+}, in this case $\alpha^{+} = 0.1$. For other values of α^{+} the similarity is also quite clear although there are secondary Reynolds Number and pitch rate effects.

Another similarity parameter of interest is the non-dimensional time at which certain events occur, for example the angle of attack for maximum lift coefficient. This parameter is define as $T = tc/U_{\infty}$, where t is the time of the event, in this case $t = \alpha_{Clmax} / \dot{\alpha}$. When the non-dimensional time for maximum lift coefficient is plotted versus the non-dimensional pitch rate Figure 2 results. The empirical expression $T = e^{-.265} (\alpha^{+})^{+ -.86}$ was found to fit the experimental data very well.

V. Angle of Attack for Maximum Lift and Drag

The rapidly pitching airfoil not only produces significant increases in lift and drag over the same airfoil under steady state conditions, but also delays the stall condition to a higher angle of attack. Indeed, the delay in stall is the apparent reason

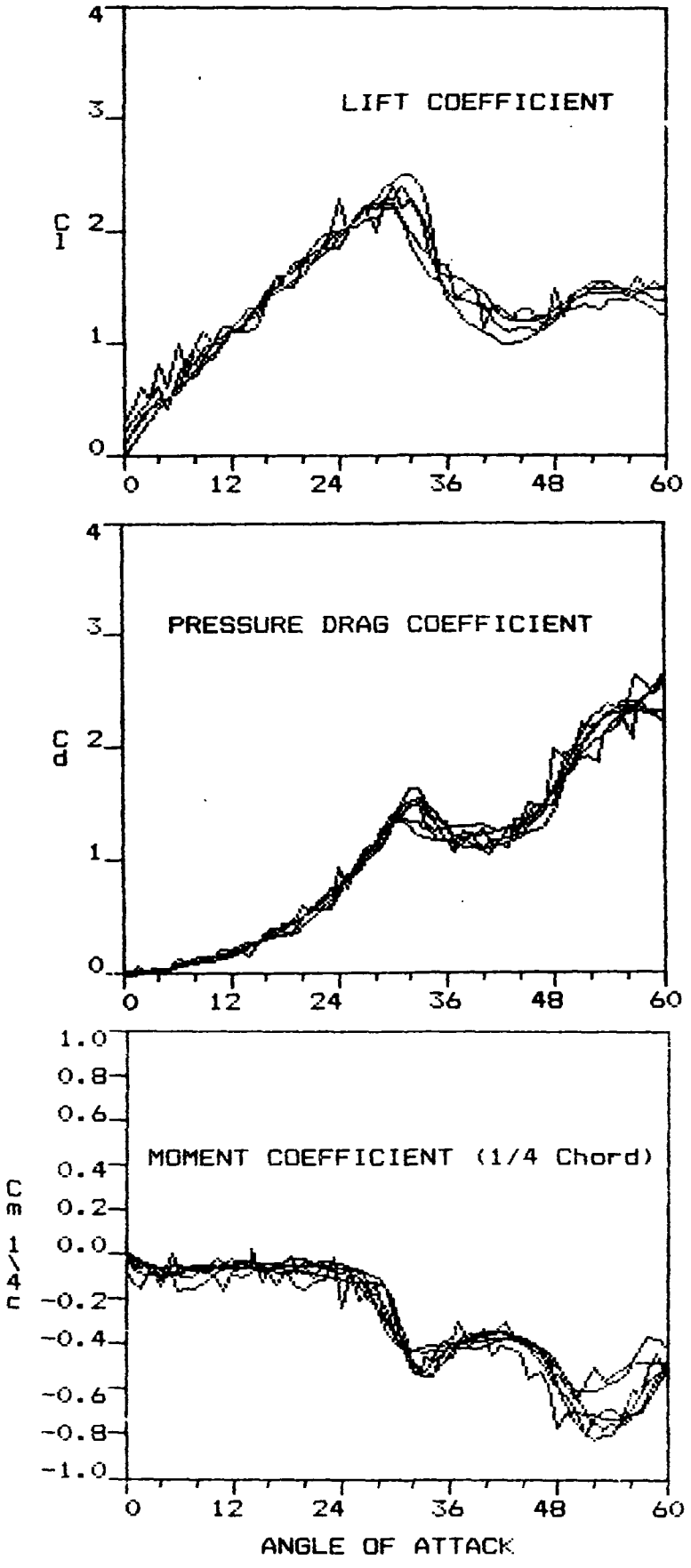


Figure 1. Lift, pressure drag, and moment coefficients as a function of angle of attack to 60 degrees for $\alpha^+ = 0.1$. (Pitch rates of 115, 230, 460, 920 deg/sec)

that the lift increases above the steady state value. Figure 3. shows the angle of attack for the maximum lift and drag coefficients. For values of α^+ of 0.1 and less, the drag coefficient grows to even a higher value near the end of the pitching motion (see Figure 1). However, for Figure 3 the drag at maximum lift was used. The same NACA 0015 airfoil has its peak lift at about 12 degrees under steady state conditions. Note also that as the non-dimensional pitch rate increases, the angle at which the peak drag occurs increases significantly above the angle for peak lift. The empirical relations

$$C_{lmax} = e^{3.90(\alpha) + .211} \quad \text{and} \quad C_{dmax} = e^{4.26(\alpha) + .333}$$

were the resulting least squares fit of the data.

VI. Maximum Lift and Pressure Drag Coefficients

Figure 4 shows the maximum lift and pressure drag coefficients. For comparison, the NACA 0015 airfoil has peak lift and drag coefficients of approximately 0.9 and 0.25 respectively under steady state conditions. Again for low values of α^+, the first peak in the drag curve was used. Both exponential and linear least squares fits were applied to the data, with the latter showing better agreement.

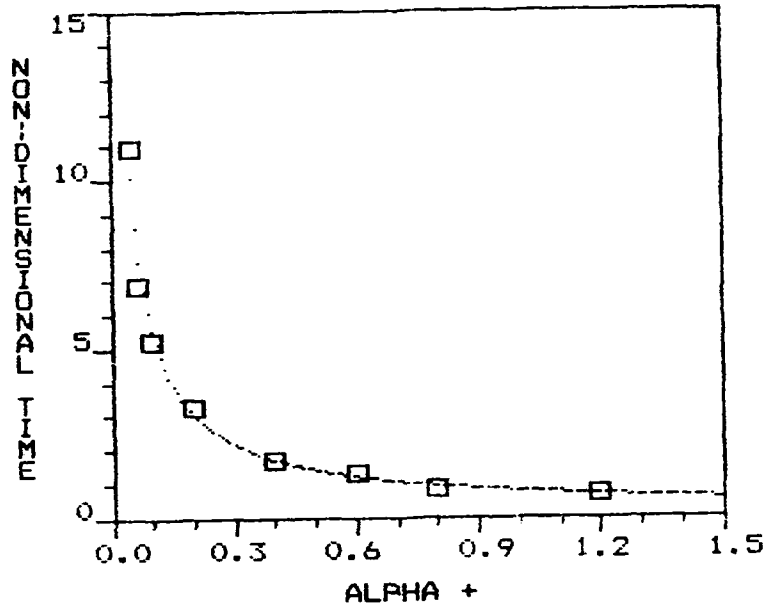


Figure 2. Non-dimensional time at which the maximum value of the lift coefficient occurs.

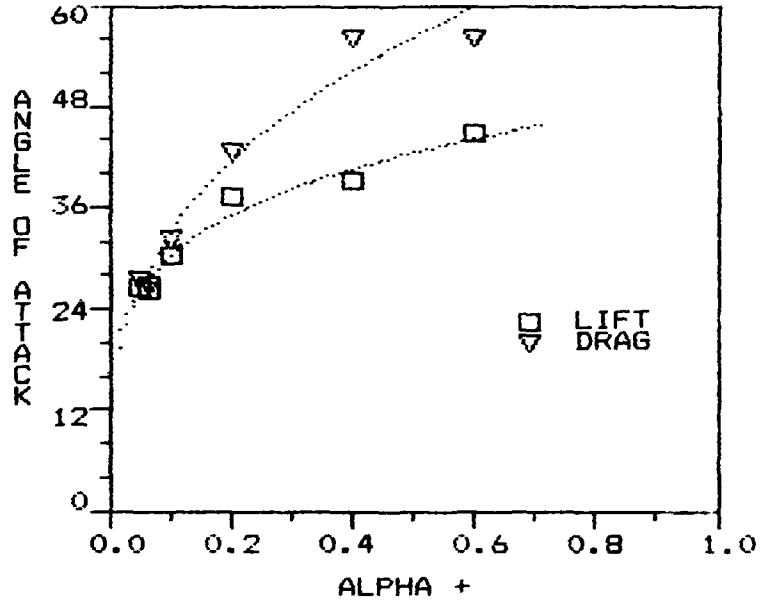


Figure 3. Angle of attack (degrees) at which the maximum lift and pressure drag occurs.

Linear:

$$C_{lmax} = 3.62\,\alpha^{+} + 1.98 \quad \text{and} \quad C_{dmax} = 6.32\,\alpha^{+} + .86$$

Exponential:

$$C_{lmax} = e^{1.49}\,(\alpha^{+})^{.263} \quad \text{and} \quad C_{dmax} = e^{1.77}\,(\alpha^{+})^{.57}$$

VII. Average Lift, Drag, and Moment Coefficients

While maximum coefficients dramatically demonstrate the potential of the concept, in actual applications the average values would be of primary importance to the designer. Thus the coefficients were integrated over the full range of pitching motion to determine the average values. Figure 5 shows the average lift coefficients as well as the maximum lift coefficient previously discussed. Two average values are shown, the average lift up to the point of maximum lift and over the full range of the pitching motion (60 degrees). It can be seen that while the average values are significantly lower than the maximum lift coefficient, they are still substantially higher than the C_l for the same airfoil under steady state conditions (0.9). Also of interest is the fact that the average lift coefficient over the full 60 degrees is not much larger than the average only up to the point of maximum lift. This would indicate that there is little advantage in rotating the airfoil beyond the point of maximum lift.

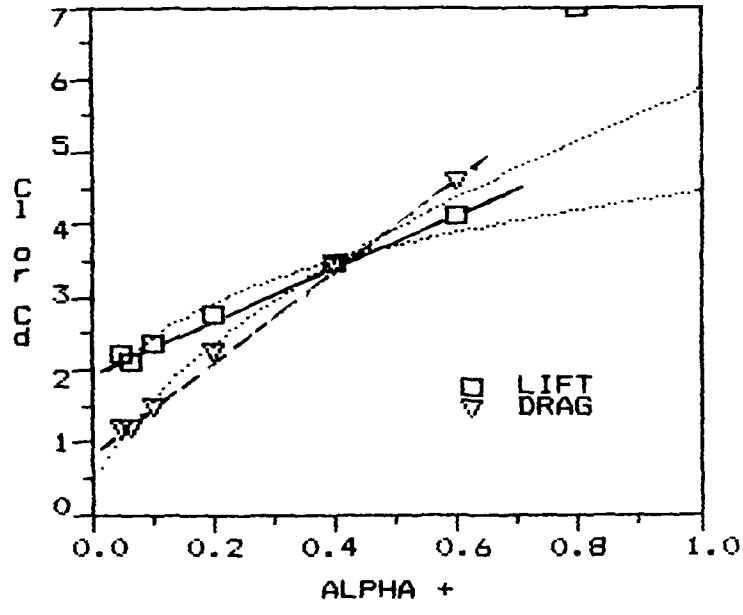


Figure 4. Maximum lift and pressure drag coefficients.

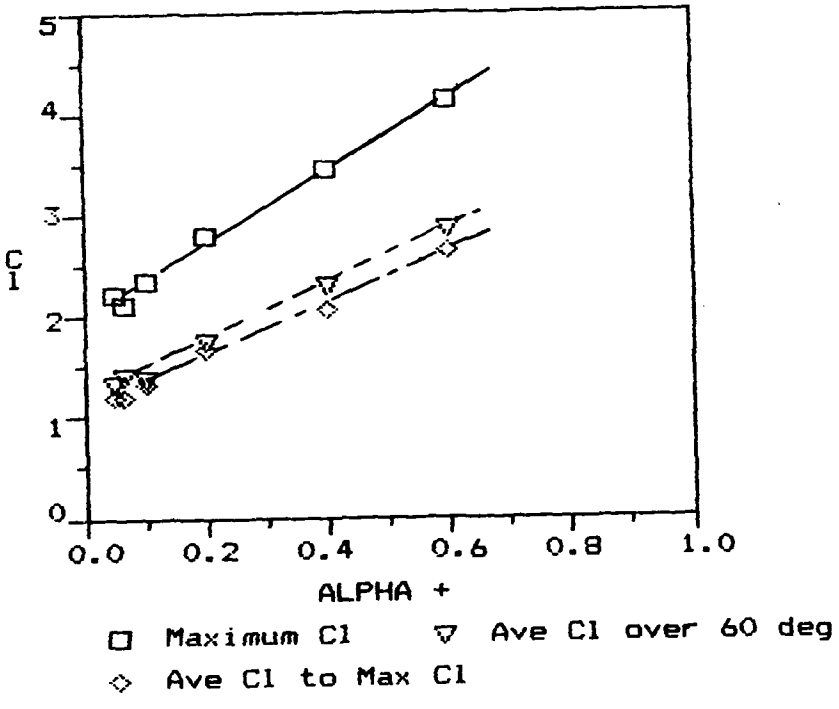


Figure 5. Average lift coefficients versus maximum lift coefficient.

The following linear expressions fit the test data most closely: for the average lift coefficient up to the point of maximum lift

$$C_{l_{ave}}{}_{max\ Cl} = 2.64\,\alpha^{+} + 1.05$$

and over the full 60 degrees of pitch

$$C_{l_{ave}}{}_{60^{o}} = 2.83\,\alpha^{+} + 1.63$$

Figure 6 shows the maximum and average pressure drag coefficients. Again the average values are much less than the maximum values and the possible advantage of only pitching to the point of maximum lift is again seen. The data also showed that the average drag up to the point of maximum lift was noticeably larger for a pitch rate of 1380 degrees/second compared to the lower pitch rates. Over the full 60 degrees of pitch; however, the difference was not noticeable. The average drag coefficient over 60 degrees can be expressed by the linear relation

$$C_{d_{ave}}{}_{60^{o}} = 1.89\,\alpha^{+} + .837$$

Figure 7 shows the average moment coefficient at the 1/4 chord point integrated up to the point of maximum lift and over the full 60 degrees of pitch.

Another aerodynamic parameter of interest is the L/D, lift to drag, ratio. Figure 8 shows this parameter for three cases (1) maximum values, (2) average values integrated to maximum lift, and (3) over 60 degrees.

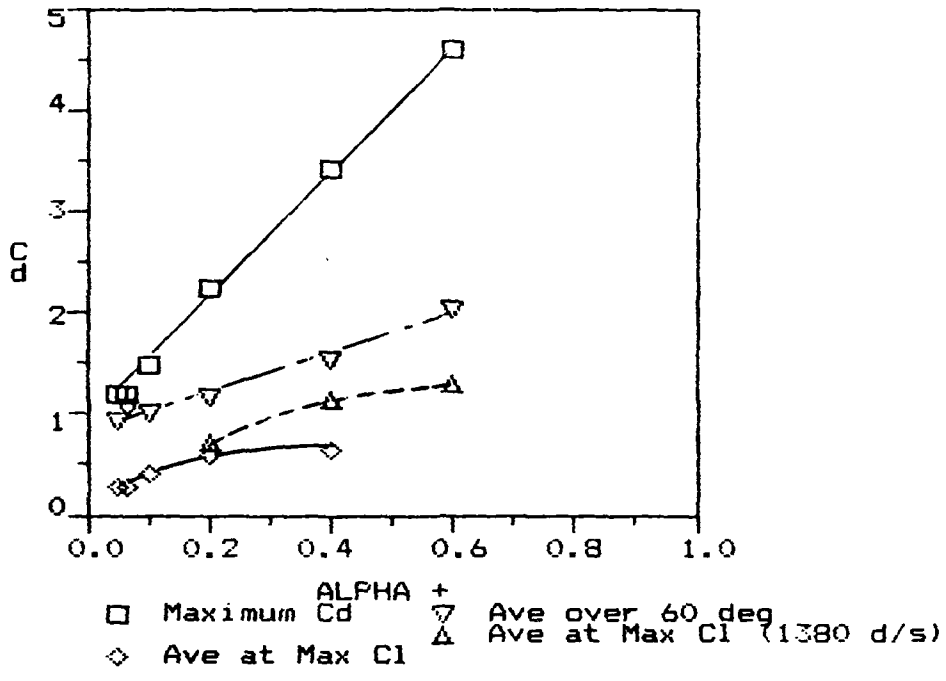


Figure 6. Maximum and average pressure drag coefficients.

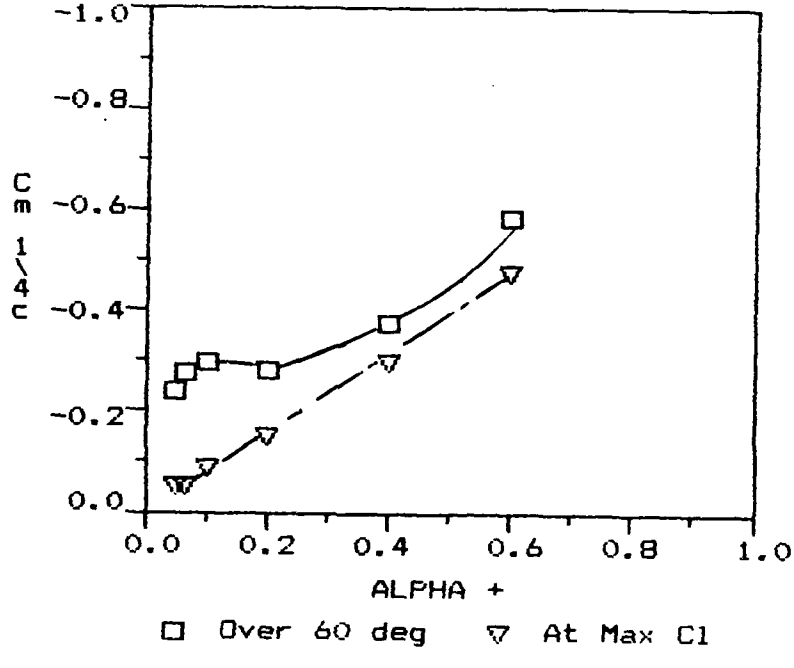


Figure 7. Average moment coefficients about the 1/4 chord.

Except at low values of α^+, the more realistic average values of L/D are greater than those determined using only the values at the maximum lift coefficient. The L/D ratios show a very clear potential advantage of only pitching the airfoil to the point of maximum lift. For the latter case, two curves are shown because the lack of similarity between the average drag coefficient at a pitch rate of 1380 deg/sec and the other lower pitch rates.

The following least squares curve fits were found to fit the data:

$$L/D\ (\text{Maximum Cl}) = e^{-.266\ (\alpha^+)\ +\ -.303}$$

$$L/D\ (\text{Ave over } 60^o) = e^{.405\ (\alpha^+)\ +\ .036}$$

$$L/D\ (\text{Ave to Max Cl})\ (\dot{\alpha} < 1380) = e^{.710\ (\alpha^+)\ +\ -.24}$$

$$L/D\ (\text{Ave to Max Cl})\ (\dot{\alpha} = 1380) = e^{.726\ (\alpha^+)\ +\ -.148}$$

VIII. Pitch Rate and Reynolds Number Effects

Figures 9 and 10 show the maximum lift and drag coefficients as a function of the pitch rate and freestream velocity. Since the experiments were conducted with the same airfoil, varying freestream velocity is synonymous with varying Reynolds Number. The figures show that both the lift and drag are less dependent on the pitch rate as the Reynolds Number increases.

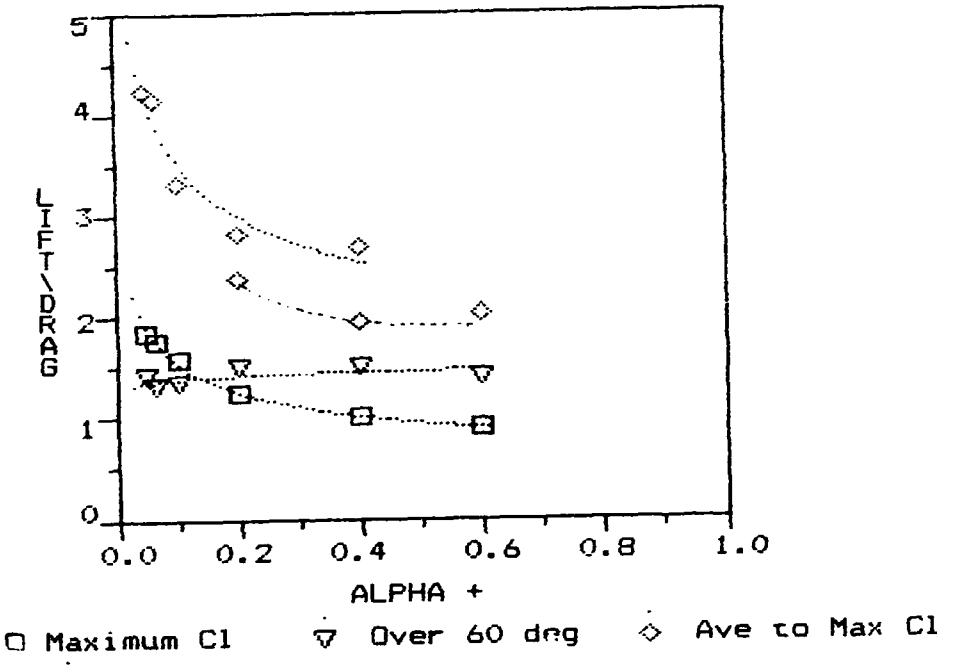


Figure 8. Lift to drag ratios. For the average to maximum Cl curves, the lower one is for the case of a pitch rate of 1380 deg/sec.

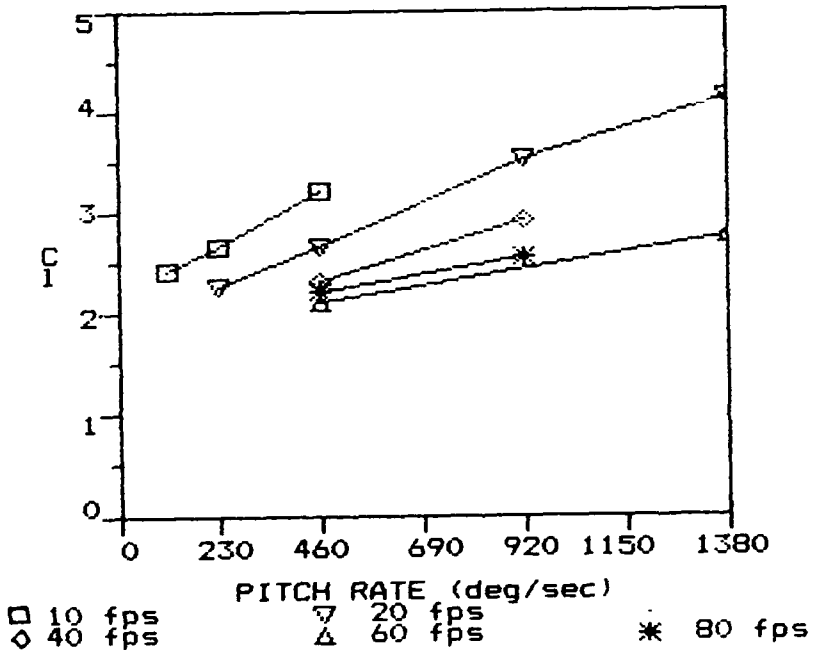


Figure 9. Lift coefficient as a function of pitch rate for various freestream velocities.

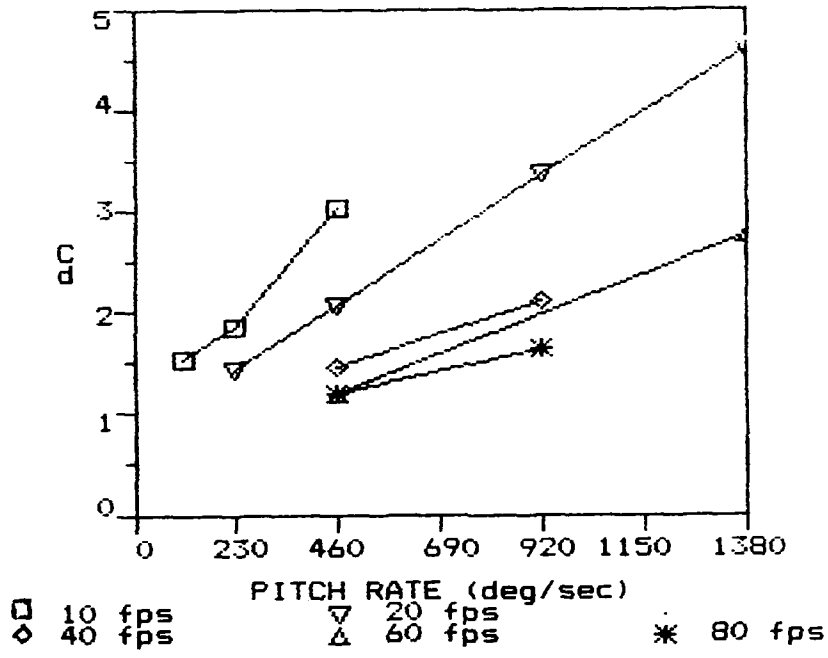


Figure 10. Pressure drag coefficient for various pitch rates and freestream velocities.

IX. Conclusions

1. The non-dimensional parameter α^+ is a valid accurate similarity parameter for the flow phenomena investigated here.

2. While the maximum lift coefficient increases dramatically by pitching the airfoil, the average lift which is of primary importance, while larger than the steady case value, is nowheres as substantial. Over the range of α^+ (0.05 to 0.6) considered here, the average lift coefficient was only about 1.25 to 3 times higher than the steady state value. By comparison, the maximum lift coefficient was 2 to 4.5 times higher.

3. It is appears possible to achieve the almost the same average lift coefficient by rotating the airfoil only up to the angle of attack for maximum lift as rotating it through the full 60 degrees.

4. The increase in lift is not without a penalty, a substantial increase in pressure drag. Indeed, the drag increases much more rapidly over steady state values compared to the lift. This can be seen by the plots of L/D. Whereas the steady state L/D of this NACA 0015 aircraft was measured to be about 3, only does the average L/D approach this value for very

small values of α^+. Again the potential advantage of pitching only to the angle of attack for the maximum lift coefficient can be seen since the high L/Ds only occur under this situation.

X. Recommendations

This research project has identified some important issues that need resolution in order to increase our understanding of the phenomena, and more importantly in pursuing its application. Thus the following experimental investigations are proposed.

1. An experimental investigation should be conducted to determine what happens when an airfoil is pitched downwards from high angles of attack in a manner to the technique used here. In a real life application the airfoil will have to be returned to its original starting position to repeat the cycle. Perhaps the gain in lift achieved in pitching the airfoil upwards might be lost in the return portion of the cycle.

2. It would be interesting to make experimental measurements at very low values of α^+. This would be useful in understanding how the apparent jump from the steady state lift value to the much higher values of lift occurs for even a relatively slowly pitching airfoil.

References

1. Helin, H.E., and J.M. Walker," Interrelated Effects of Pitch Rate and Pivot Point on Airfoil Dynamic Stall", AIAA 23rd Aerospace Sciences Meeting, January 14-17, 1985, Reno, Nevada, AIAA-85-0130.

2. Walker, J., H. Helin, and D. Chou, "Unsteady Surface Pressure Measurements on a Pitching Airfoil", AIAA Shear Flow Conference, March 12-14, 1985, Boulder, Colorado. AIAA-85-0532.

3. Walker. J.M., H.E. Helin and J.H. Strickland, "An Experimental Investigation of an Airfoil Undergoing Large Amplitude Pitching Motions", AIAA 23rd Aerospace Sciences Meeting, January 14-17, 1985, Reno, Nevada, AIAA-85-0039.

4. Walker, J .M., Forced Unsteady Vortex Flows Driven by Pitching Airfoils, PhD Dissertation, University of New Mexico, Department of Mechanical Engineering, July 1986.

1986 USAF - UES Summer Faculty Research Program

Sponsored by the

Air Force Office of Scientific Research

Conducted by the

Universal Energy Systems, Inc.

Final Report

Aircraft Sortie Effectiveness Model

Prepared by: Boghos D. Sivazlian

Academic Rank: Professor

Department and University: Department of Industrial and Systems Engineering, The University of Florida

Research Location: The Armament Laboratory/ Analysis and Strategic Defense Division, Eglin Air Force Base, Florida

USAF Research: Gerald Solomon

Date: August 18, 1986

Contract No.: F49620-85-C-0013

Aircraft Sortie Effectiveness Model

by

Boghos D. Sivazlian

ABSTRACT

A mathematical model describing the sortie of a single aircraft under enemy threats, attacking a single passive target is developed. Emphasis is placed on the determination of the probabilities associated with the various events in the sortie. These probabilities are then used to derive appropriate measures of effectiveness. The general methodology may be used to explore more complex sortie models.

Acknowledgments

I gratefully acknowledge the support of the Air Force Systems Command. The Air Force Office of Scientific Research and the Air Force Armament Laboratory in the sponsorship of my research during the Summer of 1986.

Special thanks are due to Mr. Haydon Y. Grubbs, Jr., Chief, Mr. Gerald Solomon, Technical Director and Mr. James Starling, Mathematician, all from the Weapon Effectiveness Branch, Analysis Division, Air Force Armament Laboratory (AFATL/SAE), for providing me with an excellent environment to carry on my research work, for their encouragements and for their useful comments.

Helpful discussions with Mr. Robert Hume, Mr. Everett Raspberry and Mr. Carson Sasser are also acknowledged. Finally, thanks to Ms. Bunny Hudson for typing the manuscript.

I. INTRODUCTION

In the course of the twenty years of teaching and conducting research in operations research at the University of Florida, I have particularly been interested in the development of mathematical models to quantitatively describe man-created phenomena. My previous experience at the U.S. Army MERDC, Fort Belvoir, VA, the U.S. Army White Sands Missile Range, NM, and the U.S. Air Force Armament Laboratory, Eglin AFB, FL, has provided me with ample opportunities to be associated with defense related problems and to obtain adequate expertise in applying my interest in operations research to the formulation and solution of these complex problems.

When Mr. Gerald Solomon, Technical Director of the Weapon Effectiveness Branch, Analysis and Strategic Defense Division, of the Air Force Armament Laboratory (AFATL/SAE) contacted me to work on the "Aircraft Sortie Effectiveness Model," I decided that this would provide me with an excellent opportunity to apply my knowledge and experience to an interesting and challenging problem. In particular, I found that my knowledge of stochastic processes could be utilized to model mathematically the complex problem of an aircraft sortie, and that this would constitute the first step in developing appropriate measures to evaluate the effectiveness of air-to-surface weapons.

II. OBJECTIVES OF THE RESEARCH EFFORT

An aircraft sortie is a mission or attack by an aircraft on a specified number of targets while the aircraft is operating under enemy threats in a hostile environment.

Although in a sortie the aircraft may be involved in an air-to-air combat or an air-to-surface target attack, the main consideration in the ensuing analysis is to develop models which have direct bearing to gunships used in air-to-surface missions. However, the methodology used is general enough that it can be applied to air-to-surface sorties involving any type of aircraft using any type of weapons.

Existing gunships such as the AC-130 have been successfully used in Vietnam and in Grenada. They have been effective over extended missions during daytime and nightime operations. Typical weapons carried by the gunship are the 20m/m caliber gun, the 40 m/m caliber gun and the 105 m/m howitzer, although other types of guns may be mounted on the aircraft. These gunships may be used against a variety of targets such as personnel, tanks, armored trucks, bunkers, etc..

A general question that arises is how to select the best combination of gun calibers to be mounted on a gunship in order to defeat a number of potential targets during a given sortie. The problem need not be restricted to inventoried weapons and may be extended to include the evaluation of the effectiveness of new weapons which may be under development. A possible approach to answering this question is to compare for different guns the probability of kill p_2 of a specific target. The gun that provides the largest probability of kill would logically be the one to be selected.

The primary objective of the present endeavor is the development of additional measures of effectiveness which could be used to select the best weapon from among a number of alternatives. In general, these new measures of effectiveness will be functions of a number of parameters which characterize a sortie such as frequency of encounter of enemy threats, probability of kill of the aircraft, average time elapsed between target acquisition and finally p_2, the probability of kill of a target. Since p_2 is a characteristic of the weapon/target relationship, then for two identical sorties using different weapons, the value of p_2 could be used as input to compute for each of the different weapons a particular measure of effectiveness and thus compare the two weapons. In practice, it would be more natural to evaluate the measure of effectiveness when all four parameters are simultaneously incorporated thus enabling one to compare weapons effectiveness under non-identical sorties.

Typical measures of effectiveness that may be considered are:

- probability of sortie success (mission reliability)
- probability of sortie failure
- probability that targets are destroyed
- expected number of targets destroyed
- expected cost of a sortie per expected target destroyed.
- expected gain of a sortie

Since each of the above measures must be quantitatively expressed in terms of the basic input parameters, it becomes necessary to mathematically model a sortie and to provide a probalilistic assessment of the different events associated with the sortie.

III. AIRCRAFT SORTIE MODELING

1. The Problem

We consider a sortie involving a single aircraft against a single passive target. Prior to the aircraft leaving its base, it has to be combat ready. Combat readiness may be measured by the probability that the aircraft is ready to leave the base. A probability of 1 means that the aircraft is combat ready every time it leaves the base. A probability of .50 means that the aircraft is combat ready half the time. A probability of 0 means that the aircraft is never combat ready.

When the aircraft leaves the base, it takes a time t_0 before it reaches enemy territory. We shall assume that the sortie starts at time t_0 once the aircraft reaches enemy territory and that time t_0 is the time origin for the sortie in all subsequent analysis. We ignore at this stage the possibility of an aborted mission when the aircraft is recalled to base for some reason before it reaches enemy territory.

Once the aircraft reaches enemy territory, we assume that the time T_1 to search, locate and identify the target (target acquisition) before attacking it has a negative exponential distribution with parameter μ. This $1/\mu$ is the average time it takes the aircraft to acquire the target and μdt is the probability that the target is acquired in the time interval $(t,t+dt)$. Once acquired we assume that the target is attacked by the aircraft and that the attack time is negligible. The probability of the target being killed once attacked is p_2; hence $(1 - p_2)$ is the probability that the target will not be killed once attacked.

From the moment the aircraft enters enemy territory, it is surrounded by a hostile environment which takes the form of enemy threats. We assume that the occurrence of enemy threats is a Poisson process with intensity λ. This asserts that the average number of enemy threats encounter per unit time, i.e. the frequency of threat encounter, is λ . Equivalently, λdt is the probability that the aircraft will encounter an enemy threat in the time interval (t,t+dt). Encountering more than one enemy threat over an infinitesimal time interval dt is a negligible event. Once an enemy threat is encountered, the probability that the aircraft is killed is p_1 and the probability that the aircraft is not killed is $(1 - p_1)$.

Finally, we make the assumption that the target is not part of the enemy threat (passive target) and that the occurrence of enemy threats is independent of the process involved in acquiring and attacking the target.

Note here that the total length of the sortie is measured from the time the aircraft enters enemy territory to the time it leaves enemy territory.

The object of the studied model is to derive mathematical expressions for four probability quantities in terms of the input parameters and the time t elapsed since the start of the sortie. These quantities are:

a. The probability that at time t, the target is not killed and the aircraft is not killed;

b. The probability that at time t, the target is not killed and the aircraft is killed;

c. The probability that at time t, the target is killed and the aircraft is not killed;

d. The probability that at time t, the target is killed and the aircraft is killed.

The sum of these probabilities must add up to unity. Once their expressions are established the various measures of effectiveness considered may be mathematically formulated in terms of one or more of these probabilities.

2. The Symbols

t = time parameter; the time origin is taken to be the instant the aircraft enters enemy territory

$\lambda\, dt$ = probability that in the time interval $(t,t+dt)$ the aircraft will encounter an enemy threat; λ is the frequency of occurrence of enemy threat encounter (number of threats/unit time)

μdt = probability that in the time interval $(t,t+dt)$ the target is acquired and attacked; $1/\mu$ is the average time it takes from time origin to acquire and attack the target

p_1 = probability that the aircraft is killed once it encounters an enemy threat

$\alpha = \lambda\, p_1$

p_2 = probability that the target is killed once attacked

$\beta = \mu\, p_2$

π_0 = probability that the aircraft is combat ready

3. The Model

We shall model the sortie as a two-dimensional continuous parameter Markov chain in which we compute P(i,j,t), the probability that at time t there are i (i = 0,1) remaining targets and that the aircraft is in state j (j = 0,1). j = 0 corresponds to the state "aircraft killed," and j = 1 corresponds to the state "aircraft not killed." For example, P(0,1,t) is the probability that at time t the target is killed (no remaining targets) and the aircraft is not killed. Note here that there are four probabilities to be computed:

P(1,1,t), P(1,0,t), P(0,1,t) and P(0,0,t).

The most general initial conditions for this problem are

$$\begin{aligned} &P(1,1,0) = \Pi_0 \quad ; \; P(1,0,0) = 1 - \Pi_0 \\ &P(0,1,0) = 0 \quad ; \; P(0,0,0) = 0 \end{aligned} \tag{1}$$

Note here that if at time origin the aircraft is combat ready, it is operative (equivalent to "not killed") and its state is 1. If the aircraft is not combat ready, it is not operative or out of action (equivalent to "killed") and its state is 0.

When the aircraft is 100% combat ready, the initial conditions become.

$$P(1,1,0) = 1 \quad ; \; P(1,0,0) = 0$$

$$P(0,1,0) = 0 \quad ; \; P(0,0,0) = 0$$

We relate the probability of events at time t+dt to the probability of events at time t.

a. Equation for $P(1,1,t)$

$$P(1,1,t+dt) = P(1,1,t)\,(1-\lambda dt)\,(1-\mu dt)$$
$$+ P(1,1,t)\,(1-\lambda dt)\,\mu dt(1-p_2)$$
$$+ P(1,1,t)\,\lambda dt(1-p_1)\,(1-\mu dt)$$
$$+ P(1,1,t)\,\lambda dt\,(1-p_1)\,\mu dt\,(1-p_2) + o(dt) \qquad (2)$$

To compute the probability that at time t+dt there is one remaining target and the aircraft is not killed, we note that this is equal to the probability that at time t there is one remaining target and the aircraft is not killed, times the sum of the probabilities of each of the following mutually exclusive and exhaustive events taking place in the time interval (t,t+dt):

i - no enemy threat, target not attacked

ii - no enemy threat, target attacked, target not killed

iii - enemy threat, aircraft not killed, target not attacked

iv - enemy threat, aircraft not killed, target attacked, target not killed

Equation (2) may be written

$$P(1,1,t+dt) = P(1,1,t) - (\lambda+\mu)\,P(1,1,t)dt$$
$$+\mu(1-p_2)\,P(1,1,t)dt + \lambda(1-p_1)\,P(1,1,t)dt + o(dt)$$

or $$\frac{P(1,1,t+dt) - P(1,1,t)}{dt} = -(\lambda p_1 + \mu p_2)\,P(1,1,t) + \frac{o(dt)}{dt}$$

Taking limits as $dt \to 0$ we obtain

$$\frac{dP(1,1,t)}{dt} = -(\lambda p_1 + \mu p_2)\,P(1,1,t) \qquad (3)$$

subject to the initial condition specified by (1), namely

$$P(1,1,0) = \Pi_0 \qquad (4)$$

The solution of (3) subject to (4) is

$$P(1,1,t) = \Pi_0\, e^{-(\lambda p_1 + \mu p_2)t} \qquad (5)$$

For convenience, we let

$$\alpha = \lambda p_1 \quad \text{and} \quad \beta = \mu p_2 \tag{6}$$

So that (5) becomes

$$P(1,1,t) = \pi_o e^{-(\alpha+\beta)t} \tag{7}$$

This represents the probability that t time units after the start of the sortie the target is not killed and the aircraft in not killed and it is given as a function of the five parameters π_o, λ, p_1, μ and p_2.

In a similar manner, one obtains expressions for P(0,1,t), P(1,0,t) and P(0,0,t).

b. Equation for P(1,0,t)

$$P(1,0,t+dt) = P(1,0,t) + P(1,1,t)\lambda\, dtp_1$$

$$\frac{P(1,0,t)}{dt} = \alpha P(1,1,t) \tag{8}$$

subject to $P(1,0,0) = 1-\pi_o$

Using (7) in (8) yields

$$\frac{dP(1,0,t)}{dt} = \alpha \pi_o e^{-(\alpha+\beta)t}$$

or $$P(1,0,t) = \pi_o \frac{\alpha}{\alpha+\beta}\left[1 - e^{-(\alpha+\beta)t}\right] + 1-\pi_o \tag{9}$$

This represents the probability that at time t following the start of the sortie the target is not killed but the aircraft is killed.

c. Equation for P(0,1,t)

$$P(0,1,t+dt) = P(0,1,t)(1-\lambda dt) + P(0,1,t)\lambda dt(1-p_1)$$
$$+ P(1,1,t)(1-\lambda dt)\mu\, dtp_2$$
$$+ P(1,1,t)\lambda dt(1-p_1)\mu\, dtp_2$$

or $$\frac{dP(0,1,t)}{dt} = -\alpha P(0,1,t) + \beta P(1,1,t) \tag{10}$$

subject to $P(0,1,0) = 0$

Using (7) in (10) yields

$$\frac{dP(0,1,t)}{dt} + \alpha P(0,1,t) = \beta \Pi_0 e^{-(\alpha+\beta)t}$$

whose solution is

$$P(0,1,t) = \Pi_0 e^{-\alpha t}(1-e^{-\beta t}) \tag{11}$$

This gives the probability that at time t the target is killed and the aircraft is not killed

d. Equation for P(0,0,t)

$$\begin{aligned} P(0,0,t+dt) &= P(0,0,t) + P(0,1,t)\,\lambda\, dtp_1 \\ &\quad + P(1,1,t)\,\mu\, dtp_2\,\lambda\, dtp_1 \\ \frac{dP(0,0,t)}{dt} &= \alpha P(0,1,t) \end{aligned} \tag{12}$$

subject to $P(0,0,t) = 0$

Using (11) in (12) with the initial condition yields

$$\begin{aligned} P(0,0,t) &= \int_0^t \alpha \Pi_0 e^{-\alpha u}\left(1-e^{-\beta u}\right) du \\ &= \Pi_0\left(1-e^{-\alpha t}\right) - \Pi_0 \frac{\alpha}{\alpha+\beta}\left[1-e^{-(\alpha+\beta)t}\right] \end{aligned} \tag{13}$$

This gives an expression for the probability that at time t, the target is killed and the aircraft is killed.

e. Summary

$$\begin{aligned} P(1,1,t) &= \Pi_0 e^{-(\alpha+\beta)t} \\ P(1,0,t) &= \Pi_0 \frac{\alpha}{\alpha+\beta}\left[1-e^{-(\alpha+\beta)t}\right] + 1-\Pi_0 \\ P(0,1,t) &= \Pi_0 e^{-\alpha t}\left(1-e^{-\beta t}\right) \\ P(0,0,t) &= \Pi_0\left(1-e^{-\alpha t}\right) - \Pi_0 \frac{\alpha}{\alpha+\beta}\left[1-e^{-(\alpha+\beta)t}\right] \end{aligned} \tag{14}$$

Note that the sum of the four probability expressions adds up to unity.

In the event the aircraft is 100% combat ready, $\pi_0 = 1$ and we obtain

$$
\begin{aligned}
P(1,1,t) &= e^{-(\alpha+\beta)t} \\
P(1,0,t) &= \frac{\alpha}{\alpha+\beta}\left[1-e^{-(\alpha+\beta)t}\right] \\
P(0,1,t) &= e^{-\alpha t}\left(1-e^{-\beta t}\right) \\
P(0,0,t) &= 1-e^{-\alpha t}-\frac{\alpha}{\alpha+\beta}\left[1-e^{-(\alpha+\beta)t}\right]
\end{aligned}
\tag{15}
$$

IV. MEASURES OF EFFECTIVENESS

We consider eight measures of effectiveness. These are:

1. Probability of sortie success (mission reliability)
2. Probability of sortie failure
3. Probability that the target is killed
4. Expected number of targets killed
5. Expected fraction of targets killed
6. Expected number of attacks on the target
7. Probability that the aircraft has left enemy territory at time t following the start of the sortie after killing the target.
8. Expected duration of the sortie

All eight measures of effectiveness are cost free and are expressed in terms of the time elapsed t since the start of the sortie, i.e. since the instant of time the aircraft penetrates enemy territory. For convenience we consider only the case when $\pi_0 = 1$ (100% combat readiness). The formulas can readily be adjusted for other values of π_0.

1. Probability of Sortie Success (Mission Reliability)

This may be defined in two different ways

a. It is the probability that the aircraft is not killed

$$P(1,1,t) + P(0,1,t) = e^{-\alpha t} \tag{16}$$

b. It is the probability that the target is killed and the aircraft is not killed

$$P(0,1,t) = e^{-\alpha t}(1-e^{-\beta t}) \tag{17}$$

2. Probability of Sortie Failure

This may be defined in three different ways

a. It is the probability that the aircraft is killed

$$P(1,0,t) + P(0,0,t) = 1-e^{-\alpha t} \tag{18}$$

b. It is the probability that the target is not killed and the aircraft is killed

$$P(1,0,t) = \frac{\alpha}{\alpha+\beta}\left[1- e^{-(\alpha+\beta)t}\right] \tag{19}$$

c. It is the probability that the target is not killed

$$P(1,1,t) + P(1,0,t) = \frac{\alpha}{\alpha+\beta} + \frac{\alpha}{\alpha+\beta} e^{-(\alpha+\beta)t} \tag{20}$$

3. Probability that the Target is Killed

$$P(0,1,t) + P(0,0,t) = e^{-\alpha t}(1-e^{-\beta t})$$

$$+ (1-e^{-\alpha t}) - \frac{\alpha}{\alpha+\beta}\left[1 - e^{-(\alpha+\beta)t}\right]$$

$$= \frac{\beta}{\alpha+\beta}\left[1 - e^{-(\alpha+\beta)t}\right] \tag{21}$$

4. Expected Number of Targets Killed

Let N(t) be the number of target killed at time t

E[·] be the expectation operator

Then $P\{N(t) = 0\} = P(1,1,t) + P(1,0,t)$

$P\{N(t) = 1\} = P(0,1,t) + P(0,0,t)$

The expected number of target killed is:

$$E[N(t)] = \sum_{m=0}^{m=1} m\, P\{N(t) = m\}$$

$$= P(0,1,t) + P(0,0,t)$$

$$= \frac{\beta}{\alpha+\beta}\left[1 - e^{-(\alpha+\beta)t}\right] \qquad (22)$$

5. Expected Fraction of Targets Killed

Since there is only one target involved, the expected fraction of target killed is:

$$\frac{E[N(t)]}{1} = \frac{\beta}{\alpha+\beta}\left[1 - e^{-(\alpha+\beta)t}\right] \qquad (23)$$

6. Expected Number of Attacks on the Target

We make the assumption that no more than one attack may occur on the target. In practice, this corresponds to the aircraft making a single pass attack on the target. The expected number of attacks in the interval (0,t) is therefore the probability that the target is attacked by time t.

The probability that the target is acquired (hence attacked) between x and x+dx is equal to the product of the following:

a. the probability that the aircraft is not killed on or before x, i.e. $e^{-\alpha x}$;

b. the probability that the target is acquired between x and x+dx, which is $e^{-\mu x}dx$

Hence, the probability that the target is acquired (hence attacked) on or before t is

$$\int_0^t e^{-\alpha x} e^{-\mu x} \mu\, dx = \frac{\mu}{\alpha+\mu}\left[1 - e^{-(\alpha+\mu)t}\right] \qquad (24)$$

7. Probability that the Aircraft has left Enemy Territory by time t Following the Start of the Sortie after Killing the Target

This may be shown to be

$$\Pi(t) = e^{-\kappa t}\left[1 - \frac{\beta\nu}{\beta-\nu}\left(\frac{1}{\nu}e^{-\nu t} - \frac{1}{\beta}e^{-\beta t}\right)\right] \qquad (25)$$

where it is assumed that the time elapsed from target attack to departure from enemy territory has a negative exponential distribution with parameter γ .

8. Expected Duration of the Sortie

This is equal to

$$\frac{1}{\kappa}\left(1 - e^{-\kappa t}\right) \qquad (26)$$

VI. RECOMMENDATIONS

The results of the present research may be readily implemented to determine a variety of measures of effectiveness for a particular weapon used in air-to-surface attack and to select weapons from among a set of existing or proposed alternatives. Considerations have been given to both the attrition factor of the attacking aircraft and the kill of the passive target. As a minimum, the following input parameters should be specified:

a. the frequency of occurrence of enemy threat encounter (number of threats per unit time).

b. the probability that the aircraft is killed once it encounters enemy threats.

c. the average time necessary to acquire and attack a target from start of sortie.

d. the probability that the target is killed once attacked.

e. the sortie time.

Depending on the selected measure of effectiveness, additional input parameters may be required. This would be particularly the case in a cost effectiveness measure.

The sortie model that was mathematically formulated and solved assumed a single aircraft attacking a single passive target. These assumptions may be relaxed generating several problem variants. Future research in sortie modeling may involve the following:

1. Single aircraft attacking multiple targets
2. Multiple aircraft each attacking a single target
3. Multiple aircraft each attacking several targets
4. Incorporating active rather than passive targets
5. Target prioritization
6. Allocation of aircraft and weapons to targets
7. Aircraft attack involving multiple passes
8. Incorporation of sortie distance as an independent variable in addition to sortie time.

1986 USAF-UES Summer Faculty Research Program/

Graduate Student Summer Support Program

Sponsored by the

Air Force Office of Scientific Research

Conducted by the

Universal Energy Systems, Inc.

Final Report

Combustion Under Supercritical State and Influence of Radiation on Droplet Combustion

Prepared by:	Siavash H. Sohrab
Academic Rank:	Assistant Professor
Department and University:	Mechanical and Nuclear Engineering Northwestern University
Research Location:	Air Force Rocket Propulsion Laboratory Edwards AFB, CA
AFRPL:	Mr. Michael Powell
Date:	September 25, 1986
Contract No.:	F49620-85-C-0013

Combustion Under Supercritical State and Influence of Radiation on Droplet Combustion

by

S. H. Sohrab

ABSTRACT

The influence of radiant heat transfer on combustion of single hydrocarbon liquid droplet in stagnant oxidizing atmosphere is analytically investigated. The method of matched asymptotic technique is applied based on high temperature sensitivities of chemical reaction and radiation processes. Also, the influence of radiation as it enters a partially opaque liquid droplet is analytically studied. The results show how the critical Damkohler number at extinction is modified because of the radiation from the flame zone. The existing theoretical and experimental investigations in the literature on supercritical droplet combustion are reviewed. Finally, an outline of particular research areas in need of further experimental investigations is presented.

I. INTRODUCTION

Analytical and experimental studies of droplet or spray combustion have primarily considered low pressure, subcritical environments. However, combustion of sprays of liquid fuels in rocket motors occur at high supercritical conditions. Because of the substantial modification of physico-chemical properties of reactants as the critical point is exceeded, the extension of the results from subcritical studies to the actual engine performance is quite complex if not altogether impossible. Hence, the study of liquid droplet combustion under supercritical state is an important problem in need of further consideration.

In prior investigations of liquid fuel spray combustion, the role of radiant heat transfer has often been neglected. However, because of the large temperatures, 5000 K, encountered in liquid rocket engines, it is expected that radiant transport of energy cannot be properly neglected. This is especially true for engines using hydrocarbon propellants, since both droplet absorption of radiation as well as the radiation from soot particles will be significant. The research program presented herein is complementary to the spray combustion study initiated earlier [1] and is motivated by the above mentioned considerations.

II. Objectives of the Research Effort

In general, the research program reported herein aims at achieving an improved understanding of liquid hydrocarbon spray combustion in rocket motors. The particular objectives of the analytical and experimental phases of the research efforts are outlined below.

1 - To analytically investigate the influence of radiation heat transfer on the combustion of droplets of hydrocarbon fuels such as kerosene, heptane

or liquified methane and propane. The effects of radiation from the flame zone as well as that which enters the droplet are considered.

2 - To review the literature on supercritical droplet combustion and identify particular problem areas which are in need of further experimental investigation.

III. Influence of Radiation on Droplet Combustion

The quasi-steady combustion of a spherical fuel droplet within stagnant oxidizing atmosphere is considered. A detailed description of the model is shown in Fig. 1. The conservation equations for energy and reactant concentrations are:

$$\nabla \cdot (\rho' v' C_p' T') - \lambda' \nabla^2 T' = -q^o \omega - F'(T') \tag{1}$$

$$\nabla \cdot (\rho' v' Y_i') - \rho' D' \nabla^2 Y_i' = \nu_i W_i \omega \qquad i = F, O \tag{2}$$

where primes denote dimensional quantities. Here, ρ', v', C_p', T', λ', D' and Y_i' refer to the density, radial velocity, specific heat, temperature, thermal conductivity, binary diffusion coefficient and mass fraction of specie i. Also, q^o is the heat release per mole of fuel consumed. The reaction rate ω is assumed to follow Arrhenius kinetics,

$$\omega = \frac{B' \rho'^2}{W_F W_O} Y_O' Y_F' e^{-E/RT'} \tag{3}$$

where B', W_i, E and R are the frequency factor, molecular weight of specie i, activation energy and the gas constant. The function $F'(T')$ is the rate of radiant heat loss per unit volume to be defined later.

The mass conservation equation under quasi-steady burning then requires a constant burning rate

$$\dot{m} = 4\pi r'^2 \rho' v' = \text{Constant} \tag{4}$$

Defining the nondimensional quantities:

$$Y_F = Y_F'/Y_{Fu}' \quad , \quad Y_O = Y_O'/\nu Y_{Fu}' \quad , \quad \nu = \nu_O W_O/\nu_F W_F$$

$$T = T' C_p'/Q\, Y_{Fu}' \quad , \quad Q = q^o/\nu_F W_F \quad , \quad r = r'/r_s'$$

with ν_i denoting the stoichimetric coefficient of specie i, and r_s' is the droplet radius, we obtain from (1)-(4):

$$\frac{dT}{d\xi} + \frac{d^2T}{d\xi^2} = \frac{r^4}{M^2}\hat{W} + \frac{r^4}{M^2}\hat{F}(T) \tag{5}$$

$$\frac{dY_i}{d\xi} + \frac{d^2Y_i}{d\xi^2} = -\frac{r^4}{M^2}\hat{W} \qquad i = F,\ O \tag{6}$$

where $\xi \equiv M/r$, $M \equiv \rho' v' r'^2 C_p'/r_s' \lambda'$ and

$$\hat{W} = \frac{r_s'^2 \nu \nu_F Y_{Fu}' B' C_p' \rho'^2}{\lambda' W_O'} Y_O Y_F\, e^{-T_a/T} \tag{7}$$

$$\hat{F}(T) = \frac{4\sigma' C_p' (T\, Q'\, Y_{Fu}'/C_p')^4}{Q'\, Y_{Fu}'\, \ell_p' \lambda'} \tag{8}$$

In Eq. (8), σ' and ℓ_p' respectively refer to the Stefan-Boltzmann constant and

Planck mean absorption length [2]. Equations (5)-(6) are subject to the boundary conditions:

$$\xi = \xi_s \qquad Y_{O_s} = 0 \quad , \quad dY_F/d\xi = 1 - Y_{F_s} \quad , \quad \frac{dT}{d\xi} = -L$$

$$\xi = 0 \qquad Y_O - \alpha = T - T_\infty = Y_F = 0 \tag{9}$$

where $\alpha \equiv Y'_{O\infty}/\nu Y'_{Fu}$, and we have assumed that oxygen and fuel are respectively absent at the droplet surface and the far field. Here, L is the nondimensional heat of vaporization at the droplet boiling temperature.

The solution of Eqs. (5)-(9) are obtained by dividing the flow field into five regions as shown in Fig. 1. There are two convective-diffusive zones, separated by two thin radiative-diffusive zones which are in turn separated by a much thinner reactive-diffusive zone. The analysis then follows that considered elsewhere [3] for counterflow diffusion flames. The general description of various possible burning regimes of diffusion flame structures were considered by Liñán [4] and Liñán and Crespo [5]. Here, the most relevant burning regime, namely equilibrium regime will be considered. Hence, in the Burke-Schumann limit, the flame sheet assumption is valid such that $Y_F Y_O = 0$ everywhere.

From Eqs. (6)-(8) we then obtain:

$$Y_F - Y_O = 1 - (1 + \alpha)e^{-\xi} \tag{10}$$

with the flame position $Y_F = Y_O = 0$ given by

$$e^{-\xi_f} = 1/(1 + \alpha) \tag{11}$$

and droplet surface is $e^{-\xi_f} = (1 - Y_{F_s})/(1 + \alpha)$.

a) Convective-Diffusive Zones

In regions I and V, Fig. 1, the temperature is too low such that both chemistry and radiant heat transfer are negligible, $\hat{W} \sim \hat{F} \sim 0$. The solution of Eqs. (5) and (9) are:

$$T^- = T_f - \left(\frac{L}{1-Y_{F_s}}\right) + \frac{L(1+\alpha)}{1-Y_{F_s}} e^{-\xi} \qquad \xi > \xi_f$$

$$T^+ = T_\infty + (T_f - T_\infty)\left(\frac{1+\alpha}{\alpha}\right)(1 - r^{-\xi}) \qquad \xi < \xi_f \tag{12}$$

Here, (-) and (+) refer to the fuel and oxidizer side of the flame. The flame temperature T_f is not known and will be subsequently determined by relation to radiant heat loss.

b) Radiative-Diffusive Zones

In the thin radiative zones, $O(1/\gamma_f)$, the radiation is significant but reaction is frozen [3]. Here, γ is a measure of the temperature sensitivity of $\hat{F}(T)$ defined as $\gamma \equiv \ell nF/\ell nT$ which itself has weak temperature dependence, $g \equiv \ell n\gamma/\ell nT = O(1)$. In the limit $\gamma_F \to \infty$, it can be shown that [3]

$$\hat{F}(T) = \hat{F}_f e^{\phi} \tag{13}$$

where $\phi \equiv (T - T_f)\gamma_f/T_f$. Introducing the stretched coordinate $x = (\xi - \xi_f)\gamma_f/T_f$, Eq. (5) is expressed in the first order in $1/\gamma_f$ as

$$\frac{d^2\phi}{dx^2} = \frac{M^2 T_f \hat{F}_f}{\xi_f^4 \gamma_f} e^{\phi} \tag{14}$$

The first integral of Eq. (14) subject to matching with outer solutions in Eq. (12), give

$$\left.\frac{d\phi}{dx}\right|_{0^-} = \left[\Delta^- + \left(\frac{L}{1-Y_{F_s}}\right)^2\right]^{1/2}$$
$$\left.\frac{d\phi}{dx}\right|_{0^+} = \left[\Delta^+ + \left(\frac{T_f-T_\infty}{\alpha}\right)^2\right]^{1/2} \qquad (15)$$

where

$$\Delta^{\mp} = 2M^2 T_f \hat{F}_f/(\xi_f^4 Y_f^{\mp}) \qquad (16)$$

are the fuel side (-) and oxidizer side (+) radiant loss parameters. Finally, in the flame sheet limit, the integral of Eqs. (5)-(6), and the results in Eq. (16) give:

$$\left[\Delta^+ + \left(\frac{T_f-T_\infty}{\alpha}\right)^2\right]^{1/2} + \left[\Delta^- + \left(\frac{L}{1-Y_{F_s}}\right)^2\right]^{1/2} = 1 \qquad (17)$$

which relates T_f to the parameters Δ^- and Δ^+. For adiabatic situation $\Delta^- = \Delta^+ = 0$, Eq. (17) results in

$$T_{f_a} = \frac{\alpha T_s + \alpha Y_{F_s} + T_\infty Y_{F_s}}{Y_{F_s} + \alpha}$$

which is the expected adiabatic flame temperatures [6].

c) Reactive-Diffusive Zone

In the realistic limit $T_f^2/T_a \ll T_f/\gamma_f \ll 1$ [3], the reaction zone will be

adiabatic since it is too thin to raidate appreciable energy. The analysis closely follows that of counterflow diffusion flames [3]. Defining the Zeldovich number $\beta \equiv \epsilon^{-1} \equiv T_a/T_f^2 \gg 1$, coordinate $\eta = (\xi - \xi_f)\beta$, and the expansions

$$T = T_f + \epsilon\theta + \ldots.$$
$$Y_F = \epsilon Y_1 + \ldots.$$
$$Y_O = \epsilon Y_2 + \ldots.$$

Eqs. (5) and (6) result in:

$$\frac{d^2\theta}{d\eta^2} = -\frac{d^2Y_j}{d\eta^2} = \Lambda\, Y_1 Y_2 e^{\theta} \qquad j = 1, 2 \tag{18}$$

where

$$\Lambda = \frac{M^2 \nu\nu_F r_s'^2 Y_{Fu}' C_p' B' \rho'^2}{\xi_f^4 \beta^3 \lambda' W_O'} e^{-T_a/T_f} .$$

The analysis of Eq. (18) and matching with outer solutions given in Eq. (15) follows that treated previously [3]. There exists a minimum critical value of Damkohler number Λ_E, below which the flame extinguishes [3]

$$\Lambda_E = e(m/2)[1 - 2m + 1.04m^2 + 0.44m^3] \tag{19}$$

where $m \equiv 1 - G^-$ and $G^- \equiv d\phi/dx\Big|_{0^-}$ which is given in Eq. (15).

The influence of upstream (-) and downstream (+) radiation on the flame temperature T_f was shown in Eq. (17). The modification of the extinction Damkohler number Λ_E occurs through both T_f as well as G^- defined above. In

Fig. 2 the extinction condition Λ_E is schematically shown in T_f versus Λ diagrams with $\bar{\Delta}$ as a parameter. It is found that radiant loss results in larger Λ_E such that the flame extinguishes at larger droplet radius r_s'.

IV. Radiation Entering the Liquid Droplet

Some preliminary study is made on the effects of radiant flux as it enters a droplet. A spherically symmetric and uniform radiant flux Q_R which impinges on the droplet is assumed. The energy conservation equation then gives

$$\frac{M}{r^2}\frac{dT}{dr} + \frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dT}{dr}\right) = Q_R\mu\, e^{-\mu(r-1)} \tag{20}$$

where μ is related to the total absorption coefficient of the liquid fuel. Here, absorption of radiation is assumed to follow Beer's law [2]. If the convective motion of the liquid within the droplet is neglected, $M = 0$, the solution of Eq. (20) satisfying the boundary condition $\frac{dT}{dr} = 0$ at $r = 0$, is

$$T = -Q_R\, e^{-\mu}\left[\frac{e^{\mu r}}{\mu} + \frac{2(1-e^{\mu r})}{r\mu^2}\right] . \tag{21}$$

The above temperature profiles result in droplet center temperature, $T(0) = \frac{Q_R}{\mu} e^{-\mu}$ which exceeds the initial temperature of the liquid fuel. Equation (21) expresses the exponential rise of the droplet temperature as a function of radius, due to the absorption of incoming radiation. Since infrared absorptivity of liquid hydrocarbons such as kerosene are appreciable, see Fig. 3, it is expected that radiation absorption by droplets cannot be properly neglected. The analysis of this problem is presently being further investigated.

V. Supercritical Droplet Combustion

The first theory of supercritical droplet combustion was introduced by Spalding [7]. This work considered a model where the droplet in crossing the critical state ($P = P_c$, $T = T_c$) suddenly becomes a puff of vapor. Thus, combustion of a point source associated with the puff of fuel was studied. The asymptotic analysis showed that the droplet burning time t_b, decreased with pressure as $\bar{p}^{-1/3}$ attaining a minimum at $p = P_c$. This model was later extended by Rosner [8] who considered a distributed source model.

In spite of the urgent need, relatively few experimental studies have been reported on supercritical droplet combustion. The pioneering experimental and theoretical investigations by Faeth et al. [9] and Lazara and Faeth [10] greatly contributed to the understanding of equilibrium droplet temperature as well as droplet burning time when suspended droplets undergo a free fall. The resulting data for octane and decane droplet combustion were found to be in good qualitative agreement with more recent theoretical study by Tarifa et al. [11]. Recently, Kadota and Hiroyasu [12] studied combustion of various liquid fuels under pressures slightly above P_c. However, their results showed that t_b continues to decrease for $P > P_c$ contrary to the experimental and theoretical findings mentioned above.

Most experimental investigations have considered combustion of stationary, suspended droplets in absence of external convective flow. Obviously, droplets within liquid rocket motors experience large convective effects. Such effects are aggravated under supercritical state since the surface tensions are greatly reduced. Also, because of the strong emission from soot particles, the physical geometry of the droplets during the combustion time could not be clearly visualized [13]. In fact, the sooting

characteristics of liquid hydrocarbon droplets at high pressures have not been addressed. In view of the above considerations, those aspects of the problem which require further experimental investigation are outlined in the following.

1. Visual observation of the physical characteristics of droplets as the critical point is traversed by photography and cinematography using laser sheet lighting. The critical point is traversed in the forward and backward, $(P > P_c, T < T_c) \pm (P > P_c, T > T_c)$ directions.
2. Determination of droplet temperature history for droplets suspended on thermocouple junction.
3. Evaluation of the influences of prescribed convective flow on droplet evaporation and combustion at high pressures.
4. Examination of two droplet collisions under supercritical state in absence of combustion.
5. Study of combustion characteristics of multi-component fuels under supercritical condition. Both miscible and immiscible components need be studied. Also, fuels with widely different critical pressure and temperature require consideration in order to illuminate effects on atomization when critical state of one component but not the other is traversed.
6. Combustion of small droplets, $< 200\ \mu$, require further investigations.
7. Combustion of counterflow diffusion flames at supercritical pressure should be investigated. This study will clarify the flame structure and extinction characteristics at high pressures encountered in rocket motors.

VI. Recommendations for Further Research

The problem of influence of radiant heat transfer treated herein is

considered as an initial step in understanding of the complex phenomena. Some radiative properties of liquid hydrocarbons were found in the literature and such search needs to be continued. The impact of the radiation which enters the droplet as well as the influence of soot on radiation exchange require urgent considerations. Also, spectral analysis should be considered since both gaseous emission and liquid absorption have strong spectral dependence.

Future considerations of supercritical droplet combustion were outlined in Section V. A proposal is being prepared under AFOSR Mini-Grant program to address some of the research topics discussed herein. More detailed description of the experimental program will be presented in this proposal.

ACKNOWLEDGEMENTS

This research was sponsored by USAF-UES Summer Faculty Research Program under Contract No. F49620-85-C-0013. The work was performed at Air Force Rocket Propulsion Laboratory under technical monitoring of Mr. Michael Powell to whom the author expresses his deep appreciation. The hospitality of the personnel at AFRPL is kindly appreciated.

REFERENCES

1. Sohrab, S. H., "Studies on Combustion of Liquid Fuel Sprays in Stagnation Flows," USAF-UES Final Report, AFRPL, September 1985.

2. Penner, S. S. and Olfe, D. B., Radiation and Re-entry, McGraw-Hill, 1972, pp. 708-730.

3. Sohrab, S. H., Liñán, A. and Williams, F. A., "Asymptotic Theory of Diffusion-Flame Extinction with Radiant Loss from the Flame Zone," Combust. Sci. Tech. 27, 143 (1982).

4. Liñán, A., "The Asymptotic Structure of Counterflow Diffusion Flames for Large Activation Energies," Acta Astonautica 1, 1007 (1974).

5. Liñán, A. and Crespo, A., "An Asymptotic Analysis of Unsteady Diffusion Flames for Large Activation Energies," Combust. Sci. Tech. 14, 95 (1976).

6. Williams, F. A., Combustion Theory, The Benjamin/Cummings, Palo-Alto, 1985.

7. Spalding, D. B., "Theory of Particle Combustion at High Pressures," ARS J. 29

8. Rosner, D. E., "On Liquid Droplet Combustion at High Pressures," AIAA J. 5, 168 (1967).

9. Faeth, G. M., Dominics, D. P., Tulpinsky, J. F. and Olson, D. R., "Supercritical Bipropellant Droplet Combustion," 12th Symposium (Int.) on Combustion, The Combustion Institute, 1969, p. 9.

10. Lazara, R. S. and Faeth, G. M., "Bipropellant Droplet Combustion in the Vicinity of the Critical Point," 13th Symposium (Int.) on Combustion, The Combustion Institute, 1971, p. 801.

11. Tarifa, C. S., Crespo, A. and Fraga, E., "Theoretical Model for the Combustion of Droplet in Supercritical Conditions and Gas Pockets," Acta Astronautica 17, 635 (1972).

12. Kadota, T. and Hiroyasu, H., "Combustion of a Fuel Droplet in Supercritical Gaseous Environments," 18th Symposium (Int.) on Combustion, The Combustion Institute, 1981, p. 275.

13. Natarajan, R. and Brzustowski, T. A., "Some New Observations on the Combustion of Hydrocarbon Droplets at Elevated Pressures," Combust. Sci. Tech. 2, 259 (1970).

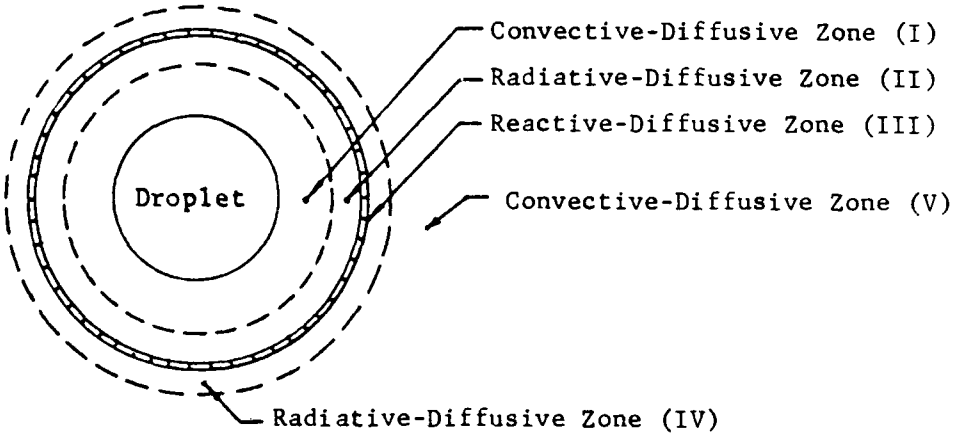


Fig. 1. Schematic of various zones for droplet burning model.

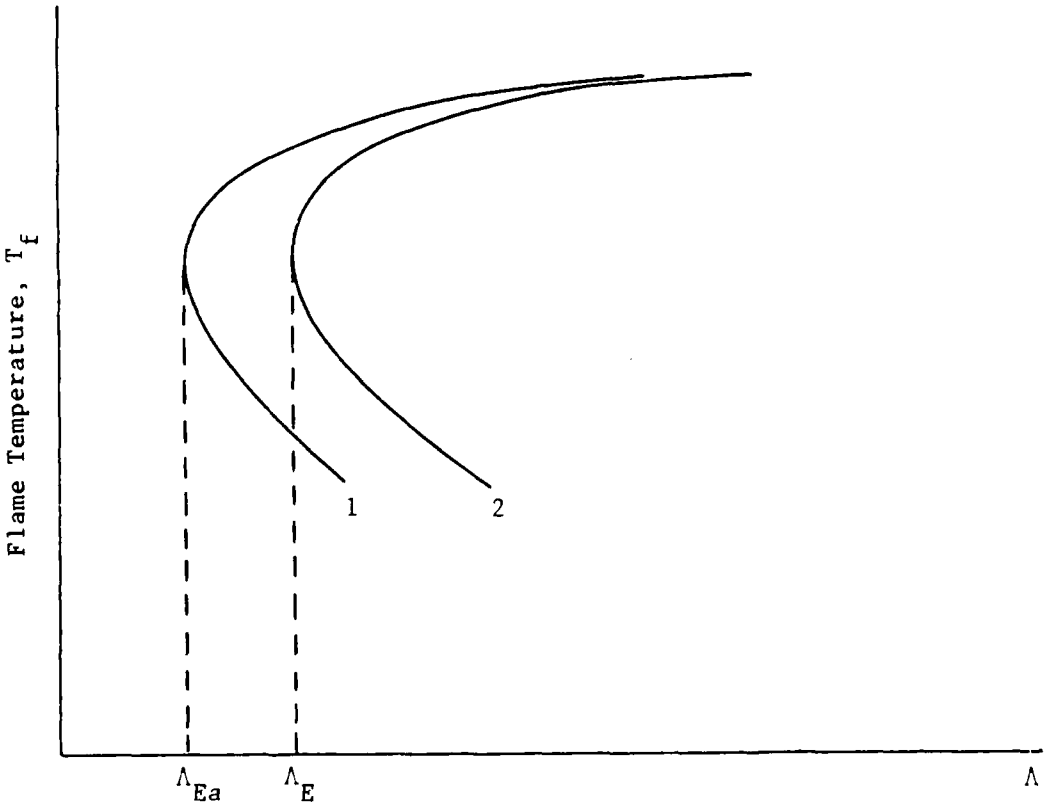


Fig. 2. Schematic of Damkohler number versus flame temperature (1) adiabatic, $\Delta^- = 0$ and (2) non-adiabatic $\Delta^- \neq 0$.

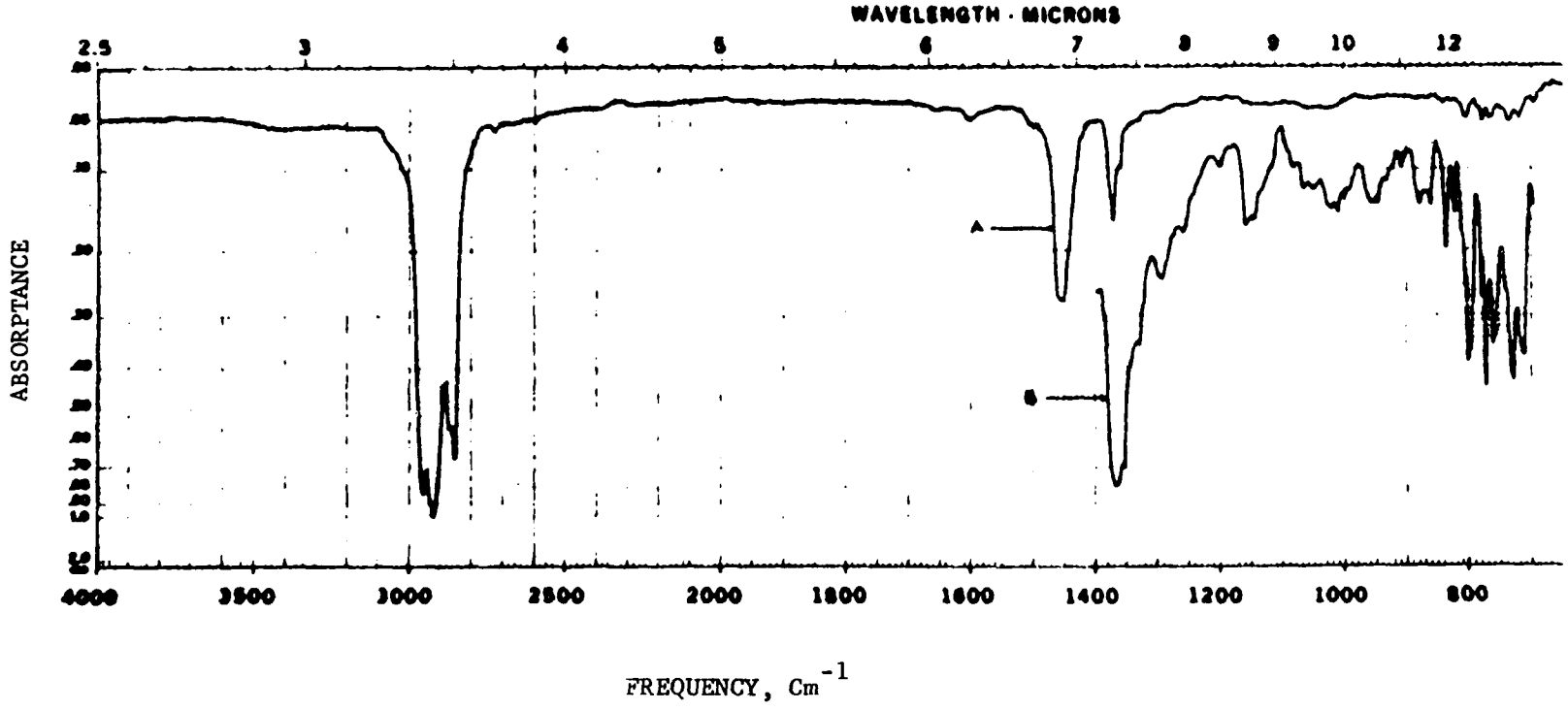


Fig. 3. Infrared spectra of kerosene. (A) liquid film between salts. (B) 0.10 mm cell.

1986 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT SUMMER SUPPORT PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

X-ray Topographic Chracterization of Si and GaAs

Prepared by:	Stuart R. Stock
Academic Rank:	Assistant Professor
Department and	School of Materials Engineering
University:	Georgia Institute of Technology
Research Location:	Materials Laboratory (AFWAL/MLPO) Wright-Patterson AFB, Ohio 45433
USAF Researcher:	Dr. W. C. Mitchel

Date:	September 24, 1986
Contract No:	F49620-85-C-0013

X-ray Topographic Characterization of Si and GaAs

by

S. R. Stock

Abstract

Synchrotron White Beam Topography (SWBT) was used to study the dislocation structure in a number of GaAs and Si specimens. The topographs revealed considerably higher dislocation densities in conventionally prepared single crystal and polycrystal GaAs than in single crystal GaAs grown in a low thermal gradient. Wafers of GaAs alloyed with In were also examined. Lower dislocation densities were observed, and the dislocations were widely spaced in slip bands. Two types of concentric striations were observed (one of which has not apparently been observed previously): the normal diffuse striations and narrow, sharply-defined striations. An epilayer of Ga-doped CVD Si on an (001) Si substrate and an ion-implanted Si wafer were also studied.

This study provides the foundation for use of x-ray diffraction topography as a problem solving tool in the Materials Laboratory. If electrical measurements, for example, reveal anomalous properties, x-ray topography can reveal the responsible crystallographic structure; and this data will allow rational redefinition of process parameters. Areas where topography will make an immediate impact are suggested.

I. Introduction

Dr. Stock received his Ph.D. from the Department of Metallurgy and Mining Engineering at the University of Illinois at Urbana-Champaign, and he used x-ray diffraction topography, an x-ray imaging analog of transmission electron microscopy (TEM), to study the early stages of plastic deformation in notched niobium crystals. Unlike TEM, x-ray topography is extremely sensitive to very small strains and it ideal for studying crystals with very low densities of dislocations. As a result, x-ray topography is frequently used to study semiconductor materials. In the two years that Dr. Stock has been at Georgia Tech, he has initiated research programs in x-ray topographic characterization of epitaxial layers of II-VI semiconductors on III-V substrates.

The production of high quality GaAs and ion-implanted Si for detectors is a major activity in the Laser Physics and Optics Branch (MLPO) of the Materials Laboratory. Emphasis is on electronic properties and characterization of point defects in these materials. The presence of only a few dislocations can drastically degrade the performance of devices, however, and this produced MLPO's interest in x-ray topography.

II. Objectives of the Research Effort

The original goals of this research were to use double crystal topography to characterize plastic deformation of GaAs and to examine solid solution hardening of GaAs. Between the author's preliminary visit in early April and the beginning of the summer research period in mid June, the research goals were modified because studies similar to those planned were published in the interim[1].

The modified goals were to show uses of synchrotron x-ray topography in research in the Materials Laboratory. Specifically, the variation in dislocation arrangement and density across wafers of GaAs produced by different growth methods were to be studied. Ion implanted Si and CVD Si on Si specimens were also to be examined.

III. Synchrotron White Beam Topography

X-ray diffraction topography uses a parallel x-ray beam and a photographic emulsion to image the spatial variation of a specimen's diffracting power. The attributes of synchrotron x-radiation which make it so attractive for topography include: the inherently high collimation, the extremely high intensity compared to conventional x-ray sources and the continuous spectrum of x-ray wavelengths available. In synchrotron white beam

topography (SWBT), the parallel beam of polychromatic radiation typically has dimensions on the order of 1 cm^2. Each set of atomic planes selects the appropriate wavelengths for diffraction (i.e. the wavelength which satisfies Bragg's law), and a Laue pattern of diffraction spots results. Figure 1 illustrates SWBT for the transmission setting which was used for these experiments. Each of the diffraction spots has dimensions of about 1 cm^2 because that is the area of the specimen irradiated, and each is an individual topograph.

Dislocations, stacking faults and other defects displace atoms from their ideal lattice positions. Local variation in the diffraction conditions result in contrast within spots. For example, a dislocation causes a phase shift in the diffracted radiation which is proportional to the value of $\underline{h}.\underline{R}$, where $\underline{h}$ is the diffraction vector and $\underline{R}$ is the displacement field associated with the dislocation. No contrast is present for a screw dislocation if $\underline{h}.\underline{b} = 0$, and for an edge dislocation if $\underline{h}.\underline{b} = 0$ and $\underline{h}.\underline{b} \times \underline{u} = 0$, where $\underline{b}$ is the Burgers vector and $\underline{u}$ is the dislocation line vector: the defect is invisible. Use of the invisibility criterion is the principle means of identifying the character of defects, and SWBT is an extremely effective means of characterizing defects because multiple $\underline{h}$ are observed with a single exposure.

Under conditions of low x-ray absorption, dislocations diffract higher intensities than the surrounding material. Images of dislocations on topographs are darker than background. Under conditions of heavy absorption, anomalous transmission [2] occurs, and dislocations prevent diffracted intensity from reaching the film. In this case the dislocation images are lighter than background. Examples of each kind of contrast are seen in Fig. 2.

Each topograph is a one-to-one mapping of the specimen, with no enlargement. Geometry and the nature of the strain fields encountered limit spatial resolution to about 1 μm. As a result, one must use extremely fine-grained emulsions and enlarge each topograph with optical microscopy.

All of the topographs in this report were recorded at the Stanford Synchrotron Research Laboratory (SSRL Beam Line II-4) during collaborative experiments with Dr. Zophia Rek of SSRL.

IV. Observation of GaAs Crystals

Five GaAs crystals were examined, each characteristic of different growth conditions or processes. All were LEC grown and the wafers were 3" in diameter. Two crystals

alloyed with In were examined, one in which the alloying suppressed dislocation formation and one which was heavily dislocated. One specimen was taken from an unalloyed polycrystalline boule. Another wafer was a single crystal grown under normal thermal conditions, and the final crystal was grown in a low thermal gradient.

We sought a slice of each wafer from edge to center so as to examine the radial variation of dislocation configuration and density. Subsequent to cleaving each slice from the wafer, it was mechanically polished until its thicknesses was between 100 and 300 μm. Thinning was required to optimize contrast, that is to obtain a thickness equal to the inverse of the linear absorption coefficient. For wavelengths between 1.2 and 0.8 A this thickness is on the order of 65 μm. Some of the slices fractured during mechanical polishing, however, and it was impossible to reconstruct the relative positions of the fragments. Each crystal was polished chemically to remove surface damage, and with the exception of occasional scratches, the polish was successful.

A. Polycrystalline GaAs

Grains in the polycrystalline material varied considerably in the quality. Images of some grains were distorted by asterism while neighboring grains produced sharply defined topographs with very low dislocation densities. Figure 2a shows a topograph recorded under conditions of high absorption: the dislocation density must be very low if individual dislocation images can be resolved. Figure 2b is of a larger grain, diffracting under low absorption conditions, and a much higher dislocation density is apparent. Dislocations were more evenly spaced and the density was lower in the smaller grains. Walls of tangled dislocations were very prominent in the larger grains, and distances between walls were as large as 1 mm.

One is led to speculate that grain boundaries provide significant stress relief, if the grains are not too large. This suggests that polycrystalline boules might be preferable to single crystal boules of GaAs for applications where large areas of a single orientation are not required, i.e. for chips with relatively small active areas.

B. Conventional GaAs

A section from the edge to the center of the wafer was available from this material. The typical well defined dislocation wall structure was observed throughout this slice. Wall separations varied between about 0.25 and 0.5

mm with the largest spacings (and lowest dislocation densities) occurring midway between the edge and the center of the wafer (Fig. 3a). Individual dislocations could be observed in the interior of the larger sub-grains (Fig. 3b) and the dislocation structure is very similar to that which has been described in the literature[3].

It is interesting in that contrast changes from that characteristic of low absorption to that typical of high absorption in some of the topographs (Fig. 3c):. Apparently the crystal is slightly warped, and wavelengths near the absorption edges of Ga or As (1.196 or 1.045 A, respectively) are selected at different positions across the specimen. Contrast may be present in topographs due to fluctuations in absorption (termed topographic EXAFS fringes). These fringes have been observed in topographs of niobium [4], and their use was suggested as a means of simultaneously probing chemical and crystallographic defect distributions. Densitometric studies of these samples may indicate, for example, whether fluctuations in Ga concentration exist near dislocations or dislocation walls.

C. GaAs Grown Under Low Thermal Gradients

Topographs of GaAs grown using a low thermal gradient reveal that this material is considerably superior to conventional GaAs. Figure 4a is a composite of four topographs which nearly covers the 1.5" radial slice. Several scratches have not been completely removed and are labeled numerically for reference. The subgrain dimensions are considerably greater (1 mm) than in the conventional material, and the variation in the dislocation structure from edge to center of the wafer was less pronounced. The only significant change in structure occurs near the edge and center of the wafer; Fig. 4a (right side) and b show topographs from these regions taken under conditions of high and low absorption, respectively. The open structure began about 5.6 mm from the edge of the specimen, and the cell size increased from less than 0.7 mm to greater than 1.2 mm. Occasional anomalously long, straight dislocation walls were observed (Fig. 4a and b); their origin is obscure.

Fewer dislocations comprised each cell wall than in the conventionally grown material, and individual dislocations could be observed in some of the walls. Numerous dislocations emanate from the walls, and individual segments apparently comprise a sizable fraction of the total dislocation density. Generally the dislocations are straight and are as long as 0.5 mm (Fig. 4c) although curved segments are visible in other reflections from the same volume of material (Fig. 4d). In several reflections complex contrast, characteristic of intermediate absorption conditions, was observed.

Verification of the wavelength diffracted by each spot must wait for the indexing of SWBT pattern. Identification of Burgers vectors for individual segments must also await indexing. The large number of topographs recorded has required the entire summer to enlarge, and the remaining analysis will require the author's attention over the next six months.

D. In Alloyed GaAs

Two types of In alloyed GaAs were examined. One crystal was a 1.5" slice of a wafer, characterized by the vendor as highly dislocated. The second crystal was much smaller and was supposed to be nearly free of dislocations. Our topographs reveal both assessments to be correct. The dislocated material contains two orthogonal sets of diffuse bands which are probably dislocation tangles lying along favorable crystallographic directions (Fig. 5a). A few individual dislocation images can be resolved in Fig. 5b, but it is clear that the addition of In was unsuccessful in preventing dislocation formation.

The second GaAs crystal is an example of the successful suppression of dislocations through In alloying. A number of scratches are present on the surface and led to very broad, dark images. Some dislocation slip bands were observed, but large regions of the crystal contained no dislocations (Fig. 6). Also of interest are the two sets of circular striations. The broad bands of contrast with diffuse borders are associated with the strains accompanying chemical segregation [5]. A second set of narrow, well-defined and concentric bands were also observed (Fig. 7). So far as the author is aware, these narrow features have not been reported elsewhere and can not be explained at present.

A set of topographs recorded with different diffraction vectors are shown in Fig. 7. Identification of the Burgers vectors of these dislocations and of their slip planes must await indexing of the pattern. Two slip systems are apparently active, and the dislocations in each system are parallel, indicating a specific crystallographic orientation is favored. Some of the topographs appear to be recorded with a range of wavelengths spanning the absorption edge(s) of Ga or As, and the striation images may be affected by EXAFS.

V. Observation of Si Specimens

The standard Si wafer thicknesses are ideal for x-ray topography, and no further specimen preparation was necessary. The exit surface for the x-ray beams was always chosen to be the side having the feature(s) of most

interest (e.g. the CVD layer); this ensures that contrast from the feature(s) is not obscured by that from the rest of the specimen. Two types of Si specimens were examined: 6 um thick Si:Ga CVD layers on Si and Si + Ga implanted Si.

A. Si and Si + Ga Implanted Si

These specimens were first implanted with Si, and subsequently contacts for resistivity measurements were formed by Ga implantation (lower corners of Fig. 8). No distinct features were observed in topographs of the Si implanted regions. Stippled contrast was present, however, and it may be due incomplete annealing of the implantation damage. This contrast is typical of damage which cannot be resolved by x-ray topography. The Ga implanted pads show similar contrast in their interior, and the In soldered connections are white areas in the corners. Considerable strain at the edge of the Ga implanted areas is revealed by the black and white contrast.

B. CVD Si:Ga on (001) Si

The CVD layer was 6 um thick and unexpectedly contained a significant number of dislocations. The dislocation density was non-uniform, Fig. 9. The two topographs shown were recorded with diffraction vectors from two perpendicular zones, and it is interesting to note that the images of dislocation tangles are quite different. In Fig. 9a the tangles appear to run vertically while in Fig. 9b they run horizontally; in both cases the images are predominantly parallel to the diffraction vector.

VI. Recommendations

The experiments described above were routine studies which would be a normal part of an electronic materials growth program. One would expect few surprises, and this was indeed the case for the unalloyed GaAs. Material grown with low thermal gradients was superior, in terms of dislocation density, to conventionally grown GaAs. The quality of some of the grains of the polycrystalline GaAs, however, was surprising: the dislocation densities could be lower than in single-crystal GaAs grown by conventional techniques. Other unexpected results were the observation of a new type of striation in In alloyed GaAs and of dislocations in a Ga doped Si CVD layer on (001) Si.

The unexpected results are a strong argument for establishing x-ray topographic capabilities. The ability to correlate anomalous properties of specific specimens and crystallographic features is very important. One must understand the microstructure if processing is to eliminate

the anomaly. X-ray topography is nondestructive and a specimen can be examined at different steps during processing. This will have a major impact in debugging complex processes.

Development of in-house topographic expertise at the Materials Laboratory appears to be infeasible at present, however, due to staff and equipment constraints. If x-ray topography continues to be available, its use and impact will continue to increase, and this capability will eventually be required. In the interim the author recommends that topographic characterization of semiconductor materials be continued by outside investigators, with emphasis in areas identified during discussions with MLPO staff and contractors*. It is his opinion that the best vehicle for this continued work is research under a continuation grant which is an integral part of the summer faculty program.

Acknowledgments

Support of the Air Force Systems Command and of the Air Force Office of Scientific Research is gratefully acknowledged. The author is also grateful to the Materials Laboratory for making the synchrotron research at Stanford possible. The staff of MLPO and MLPJ and associated contractors were exceedingly helpful; Drs. Bill Mitchel and Pat Hemenger made the completion of this project possible. Dr. Wade Adams and his group generously made their equipment available. To these individuals, named and unnamed, he offers his thanks.

The work reported herein was partially done at SSRL which is supported by the Department of Energy, Office of Basic Energy Sciences; and the National Institutes of Health, Biotechnology Resource Program, Division of Research Resources.

---

*These topics are identification of anomalous electrical regions in ion implanted Si, elucidation of the striation structures in In alloyed GaAs, characterization of dislocation density as a function of Ga doping in CVD Si on Si (if specimens are available) and development of a technique for characterizing strain distribution and magnitude around ion implanted regions in Si wafers.

References

1. Von Neida, A. R. and Jordan, A. S., "Reducing Dislocations in GAAs and InP," Journal of Metals, June 1985, pp. 35-40.

2. Tanner, B. K., X-ray Diffraction Topography, Oxford, England, Pergamon Press, Ltd., 1976.

3. Halliwell, M. A. G., Childs, J. B. and O'Hara, S., "Reflection X-ray Topography of Compound Semiconductors," in Gallium Arsenide and Related Compounds 1972 pp. 98-105.

4. Bowen, D. K., Stock, S. R., Davies, S. T. Pantos, E., Birnbaum, H. K. and Chen, Haydn, "Topographic EXAFS," Nature 309 (1984) pp. 336-8.

5. Kubena, J. and Hlavka, J., "Space Correlation of Microdefects with Recombination of Excess Carriers in CZ-Si," Physica Status Solidi A 89 (1985) pp. K23-5.

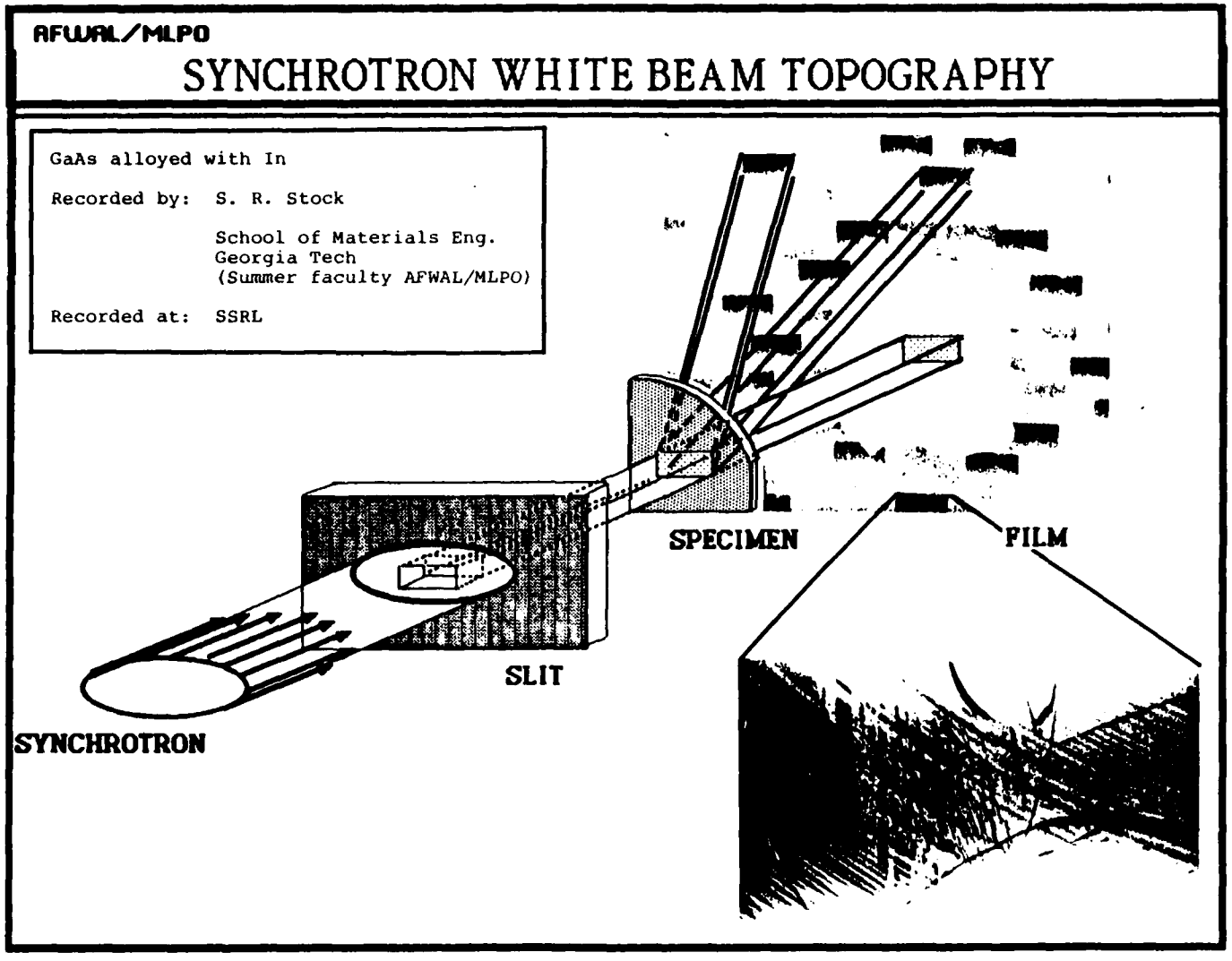

AFWAL/MLPO
SYNCHROTRON WHITE BEAM TOPOGRAPHY
GaAs alloyed with In
Recorded by: S. R. Stock
School of Materials Eng.
Georgia Tech
(Summer faculty AFWAL/MLPO)
Recorded at: SSRL
SPECIMEN
FILM
SLIT
SYNCHROTRON

Figure 1.

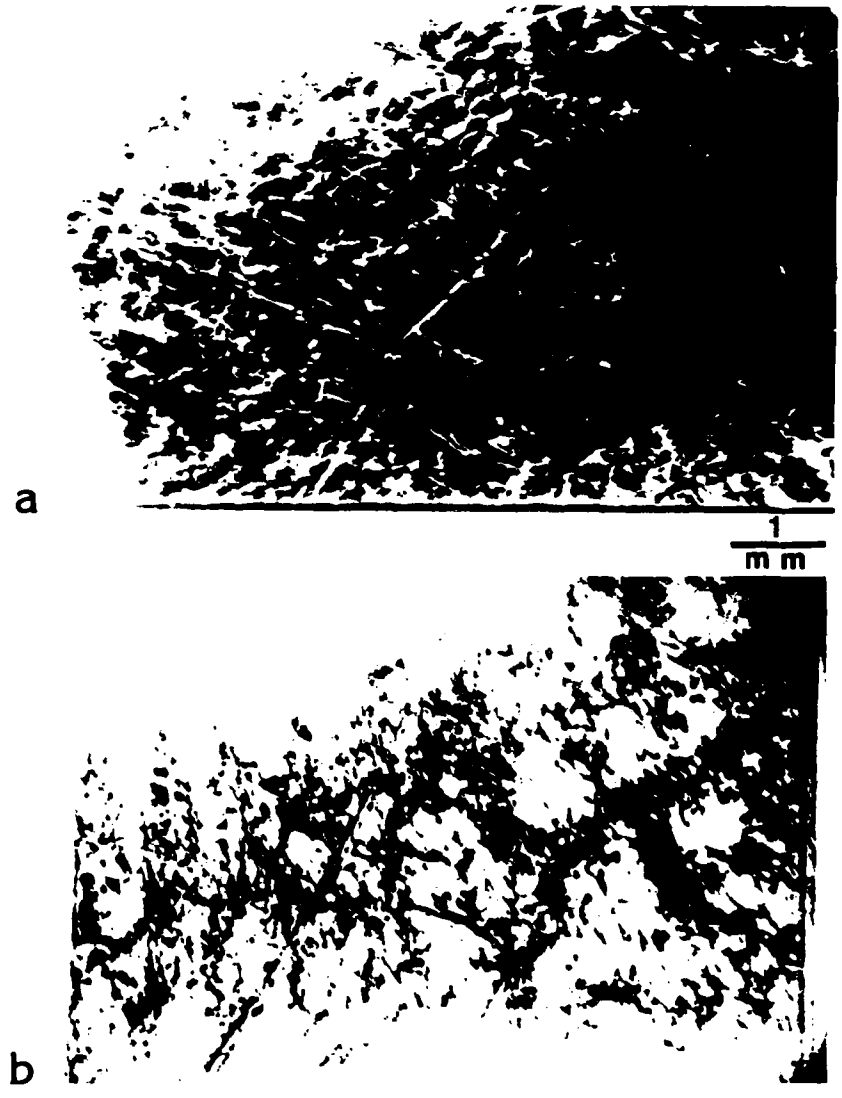


Figure 2. Synchrotron White Beam Topographs of grains from polycrystalline GaAs.

a. High absorption conditions and
b. Low absorption conditions.

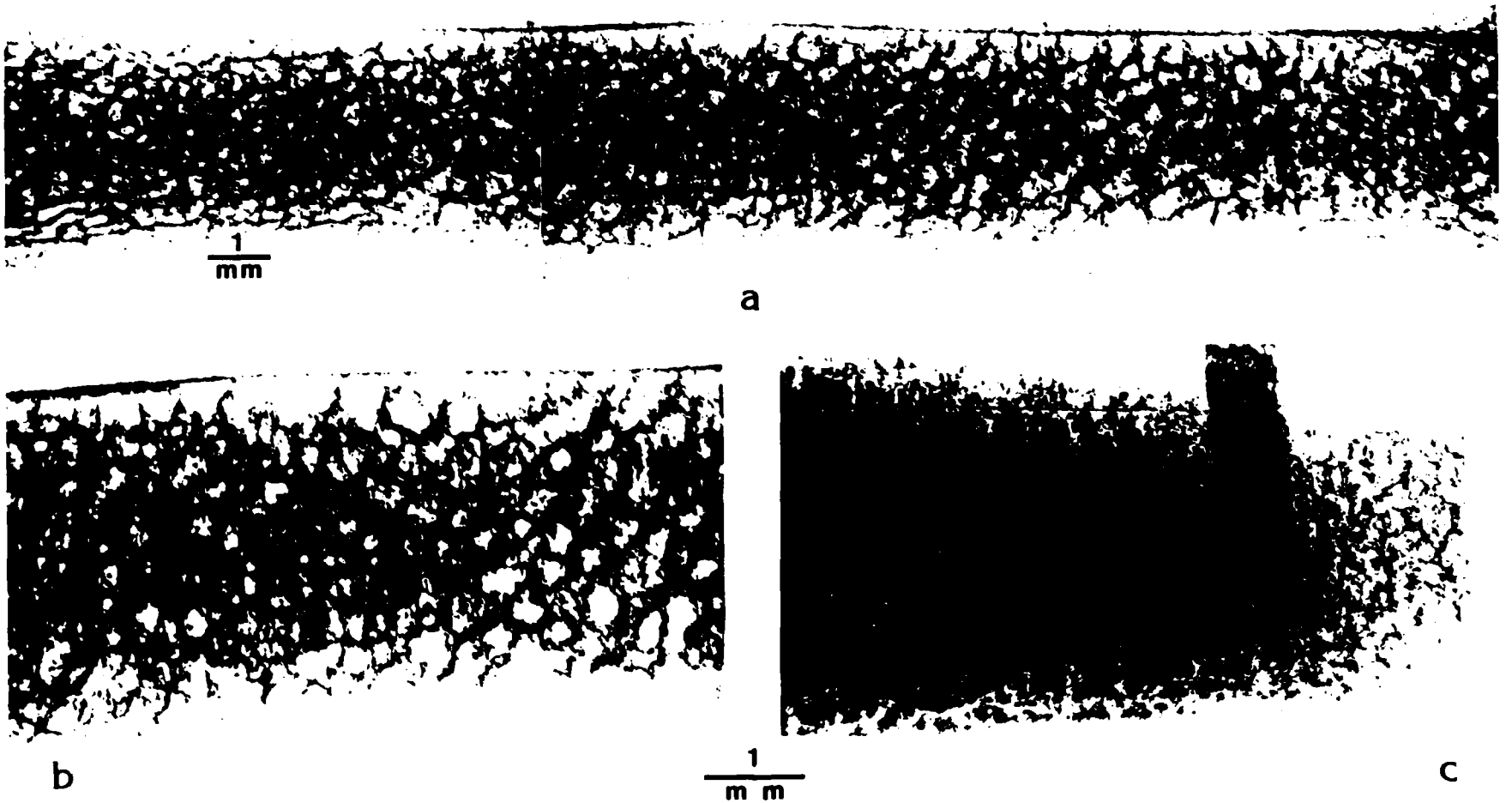


Figure 3. Dislocation Structure in Conventional GaAs.

a. Composite topograph with center at right,
b. Dislocations in cell interiors and
c. Contrast variation across crystal.

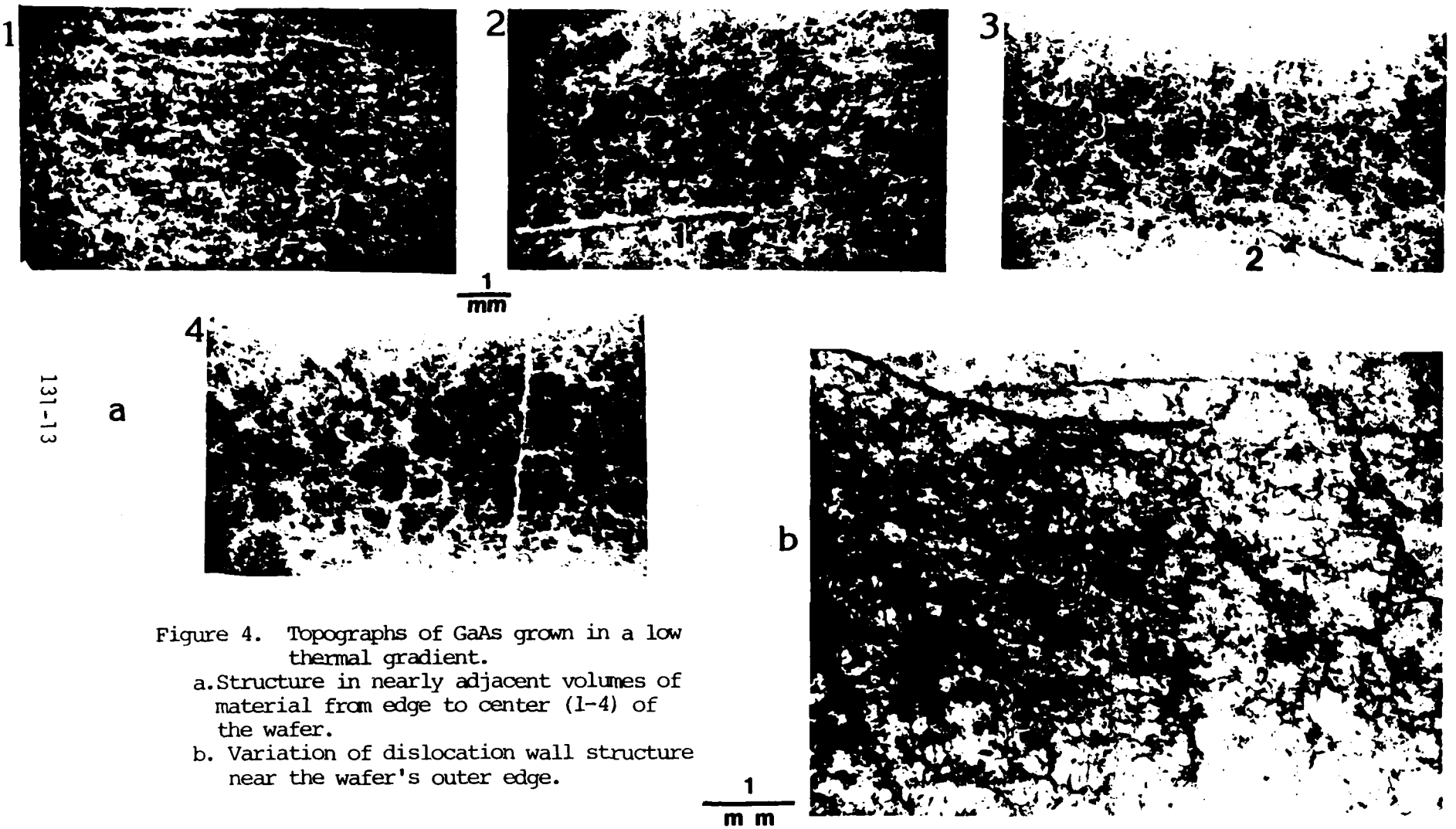


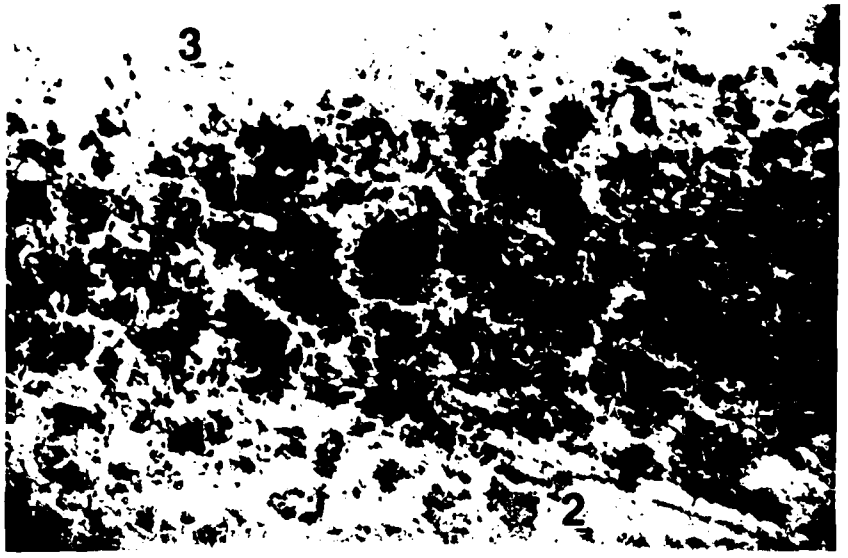
Figure 4. Topographs of GaAs grown in a low thermal gradient.
a. Structure in nearly adjacent volumes of material from edge to center (1-4) of the wafer.
b. Variation of dislocation wall structure near the wafer's outer edge.

c



1 mm

d



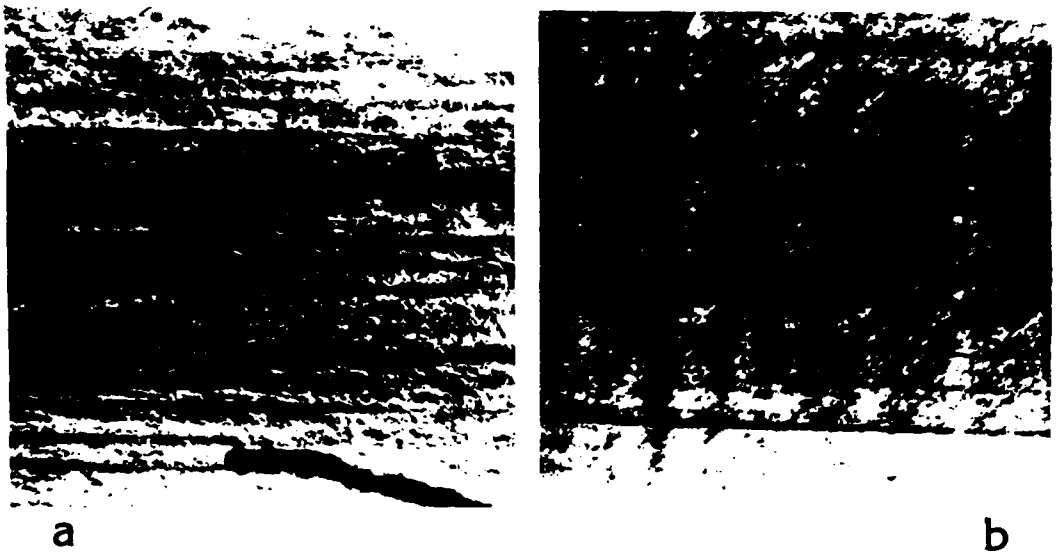


Figure 5. Topographs of dislocated In alloyed GaAs.

1
mm



Figure 6. Topograph of a Second In Alloyed GaAs Crystal showing the limited volume of dislocated material.

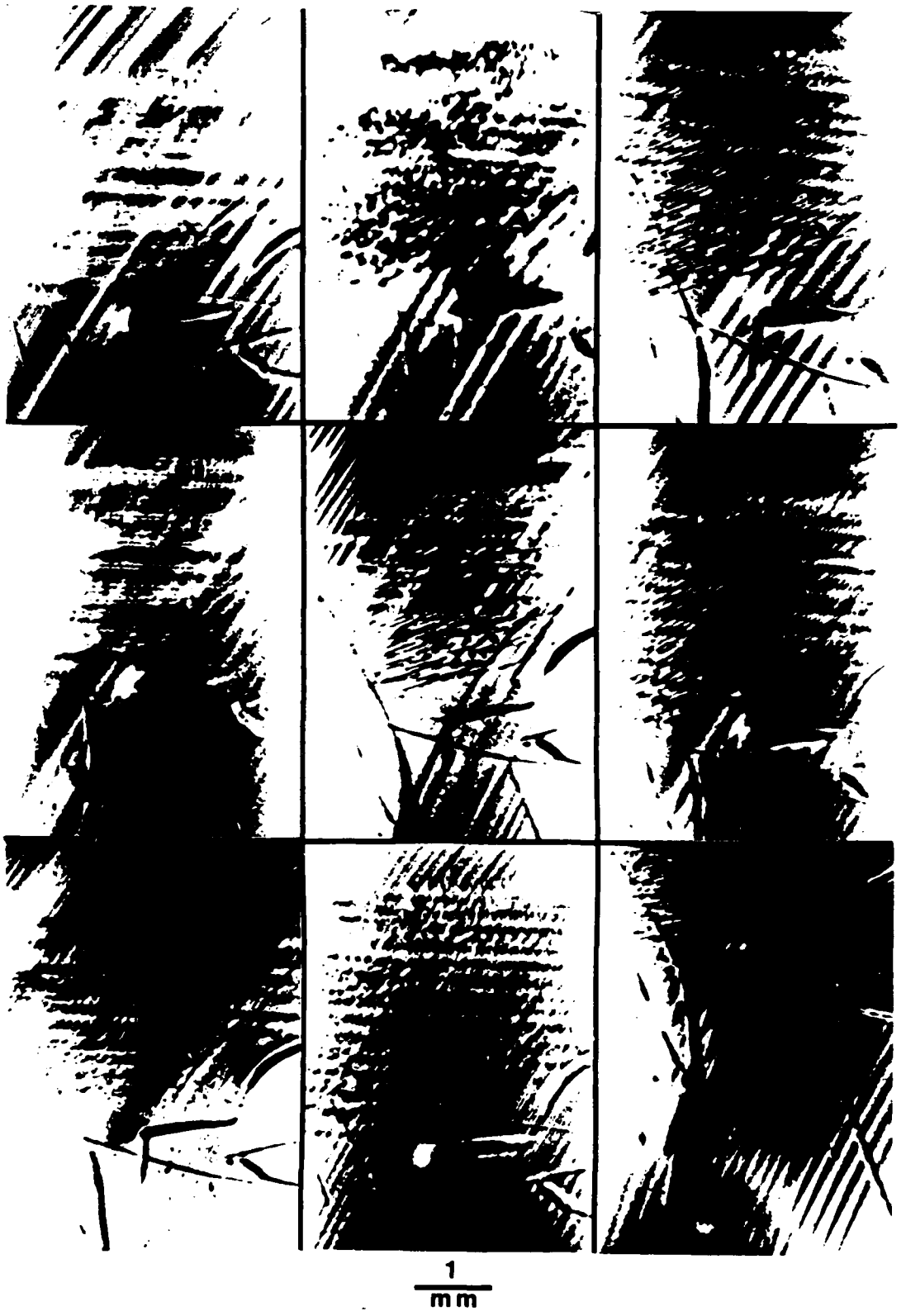
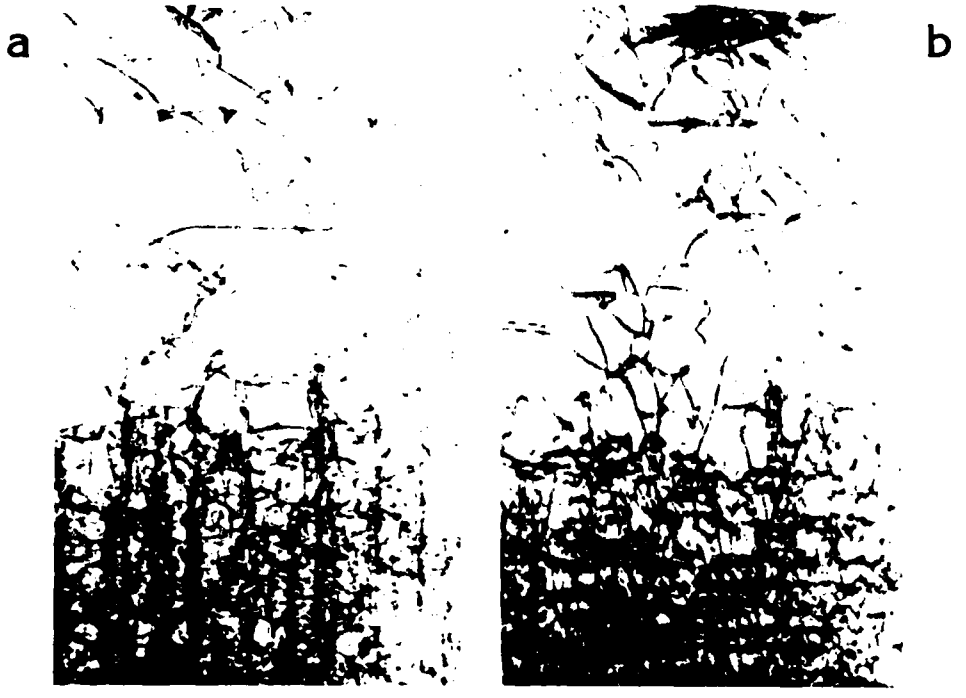


Figure 7. Topographs of In Alloyed GaAs Showing Dislocations and Striations.



1
mm

Figure 8. Topograph of Ion Implanted Si. The square regions in the lower corners have been also implanted with Ga.



1
mm

Figure 9. Topographs of CVD Si:Ga on (001)Si. These topographs are of the same volume of material and were recorded with diffraction vectors from two perpendicular zones.

1986 USAF-UES SUMMER FACULTY RESEARCH PROGRAM

GRADUATE STUDENT SUMMER SUPPORT PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

UNIVERSAL ENERGY SYSTEMS, INC.

FINAL REPORT

ASSESSMENT OF MAXIMUM ENTROPY METHOD SOFTWARE BY OPERATION ON INTERFEROGRAMS OF MODEL SPECTRA

Prepared by:	James E. Sturm
Academic Rank:	Professor
Department and	Department of Chemistry
University:	Lehigh University
Research Location:	Air Force Geophysics Laboratory
	Hanscom Air Force Base, MA
	Infrared Technology Division
	Infrared Dynamics Branch
USAF Research:	Dr. William A. M. Blumberg
Date:	September 29, 1986
Contract No.:	F49620-85-C-0013

ASSESSMENT OF MAXIMUM ENTROPY METHOD SOFTWARE BY OPERATION ON INTERFEROGRAMS OF MODEL SPECTRA

by

James E. Sturm

ABSTRACT

Software available for reduction of interferometric data by the maximum entropy method (MEM) was applied to computer-fabricated data sets representing simple delta-function line patterns. In each case a data set of noise signals with a Gaussian intensity distribution was added. The combined data set resembled in form those obtained from the AFGL LABCEDE facility.

The main MEM parameter studied was the "order of the prediction error filter." Generally, more closely-spaced lines required higher orders to achieve resolution. Small shifts in peak position were also generated. For a given peak separation the order needed increased as the number of lines in the spectrum was increased. Recovery of all lines of the model was achieved at less than half the line separation resolvable by the fast Fourier transform (FFT) algorithm. In representation of five-line spectra, the MEM generated spurious line splitting at higher orders. Although line intensities were not reproduced in all MEM spectra, integrals over the line widths appeared to match those of the models as long as no line splitting occurred.

I. INTRODUCTION -

This report covers work of the author's second successive appointment to the AFOSR-SFRP program. Both were to the Air Force Geophysical Laboratory, in the Infrared Technology Division, and specifically in the Infrared Dynamics Branch, abbreviated AFGL/LSI.

The activities of the AFGL/LSI overlap considerably with the author's background and interests. His doctoral study in the Radiation Laboratory of the University of Notre Dame was on electrical discharge and radiation chemical decomposition of ethane. Postdoctoral research at the University of Wisconsin dealt with tritium beta radiolysis of ethyl iodide. Subsequently, in parallel with his own research efforts at Lehigh University, the arthor participated at various times in studies at national and governmental laboratories. These efforts were all related to characterization of elementary processes in both photochemistry and radiation chemistry by means of both experiments and computer modeling. Teaching activities regularly included reduction and quantitative assessment of experimental data.

Choice of project for the 1986 participation was uncomplicated; it simply was agreed that more owrk was desirable on the topic embraced in the 1985 project: assessment of the maximum entropy method (MEM) of spectral analysis of interferometric data related to interferometric data of the type obtained in the AFGL LABCEDE facility. No pre-summer visit was deemed necessary. A private contractor familiar with the topic, Dr. Boh K. Yap of Yap Analytics,

Inc., Lexington, MA, was assigned to advise the author occasionally on steps to be taken or on interpretation of results obtained.

II. OBJECTIVES OF THE RESEARCH EFFORT-

Below is given a condensation of the background from which the objectives of this project are to be understood. It represents an effort to assimilate and apply an extensive, often abstract, and sometimes labyrinthine literature.

A. Spectral analysis instruments - Interferometers have established themselves as the main instruments for obtaining infrared spectroscopic data because of their advantages of efficiency, sensitivity, and, in some designs, resolution over earlier dispersive instruments (1). A major factor in the adoption of interferometers has been the increasing availability of computers for data reduction. Virtually every commercial instrument employs the so-called fast Fourier transform (FFT) to extract the power spectral density distribution (PSD) from the interferometric data. The FFT algorithm is based on a 1965 paper (2) in which the earlier work of Good (3) was shown to be adaptable to computers.

Implementation of this algorithmic breakthrough has had its price, however. A Fourier transform is defined as an integral over an infinite range of one variable (4):

$$F(\sigma) = \lim_{x' \to \infty} \int_{-x'}^{x'} f(x) \exp(-i2\pi\sigma x)\, dx \qquad [1]$$

Data collected by an instrument are necessarily finite in extent.

Integration of [1] over a finite range destroys the uniqueness of the correspondence between f(x) and F(σ). Sometimes the range of integration, still finite, is extended by assigning values beyond the range of measured values (5). When zeros are assigned, side lobes are found in the resulting F(σ) even for single-line spectra (1). Often a compromise "apodization" procedure is adopted whereby the experimental values are decreased gradually towards zero at the ends of the data sets. It is difficult to defend a priori these extensions beyond the range of measurement and also alterations of valid data.

B. Maximum Entropy Method (MEM) - While there are arguments (6) to defend the apodization algorithms, there is room for consideration of alternative approaches hopefully having a firm conceptual foundation. The literature on MEM has its origins in communication theory and seismology rather than in molecular spectroscopy. Soon after the introduction (7) of concepts on information theory, radio engineers began to pursue the related statistical definition of entropy in their search for methods to extract signals form noisy backgrounds. Frequent reference is made to the work of Burg (8) who developed an algorithm for spectral analysis of time series signals. This so-called maximum entropy method (MEM) is only beginning to be applied to problems in mass and molecular spectroscopy (9-12).

Regardless of the data reduction method, one begins with a set of measured intensities $I(x_i)$, i = 1, . . . , n. The power spectral density is defined in the MEM as that function F(σ) which has the property that the integral J (the entropy) given by

$$J = \int_{-\infty}^{\infty} \ln F(\sigma)\, d\sigma \qquad [2]$$

is a maximum subject to the following constraints which relate $F(\sigma)$ to the measured intensities:

$$I(x_i) = \int_{-\infty}^{\infty} F(\sigma) \exp(i2\pi\sigma x_i)\, d\sigma \quad , i = 1, \ldots, n \qquad [3]$$

The limits of [3] need not be infinite if $F(\sigma)$ is zero outside of a spectral range of significance.

Description of the MEM computations warrants amplification even if only a brief summary (13). A set of 2M ($M < n$) autocorrelation functions R(m) can be defined in terms of the given data:

$$R(m) = (1/m) \sum_{\text{over all products}} I_k \cdot I_{k+m} \quad , m = -M, (-M+1), .., -1, 0, 1, .., (M-1), M \qquad [4]$$

Next, a set of so-called 'lag' coefficients a_m, $m = 1, 2, \ldots, M$, is defined recursively along with a prediction-error-power term P_M of order M in terms of the corresponding values of these quantities for the previous order M-1. At each order the coefficients a_m and P_M are made to satisfy the condition that

$$P_M = (1/M) \left(\sum f_k\, f_k + \sum b_k\, b_k \right) \qquad [5]$$

is minimized where

$$f_k = I_k + \sum_{m=1}^{M} a_m \cdot I_{k-m} \quad \text{and} \quad b_k = \sum_{m=0}^{M-1} a_{M-m} \cdot I_{k-m} + I_{k-M}. \qquad [6]$$

An important implication of this condition is that P_M and the coefficients a_m are solutions of the matrix equation [7]. Once PM and the am are obtained, the final step in the MEM formulation is the generation of the intensity or power spectrum for this order M as

given by equation [8].

$$\begin{bmatrix} R(0) & R(-1) & \cdots & R(1-M) & R(M) \\ R(1) & R(0) & \cdots & R(2-M) & R(1-M) \\ & & \vdots & & \\ R(M) & R(M-1) & \cdots & R(1) & R(0) \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ \vdots \\ a_M \end{bmatrix} = \begin{bmatrix} P_M \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad [7]$$

$$I(\sigma) = \frac{P_M \Delta x}{\left| 1 + \sum_{m=1}^{M} a_m \exp(-i2\pi\sigma m \Delta x) \right|^2} \qquad [8]$$

The main achievements of the MEM shown in the few papers on molecular spectroscopy are striking improvements in resolution when compared to that achieved by the FFT reduction of the same data and sometimes with somewhat smaller data sets [9-12].

C. Statement of Research Objective - With all of its promise, superresolution capabilities of the MEM must nevertheless have their own limitations which at this stage are not clearly identified. The present study was suggested to pursue such limitations.

OBJECTIVE: to study application of the MEM to interferograms (I/Fs) representing infrared emission spectra having simple structure allowing characterization of the limitations of the MEM.

III. APPROACHES TOWARDS MEETING OBJECTIVE -

A. Facilities - As in 1985, this study made use of computers. The following hardware and software were available at AFGL/LSI:

- a network of at least three Apollo computers, disk storage, and a

hard copy printer-plotter

- compiled versions and source code (in FORTRAN) principally of two programs:

-'labcede_sturm' for I/F plots and FFT reduction

-'mem_tst_v1' intended for MEM reduction and computation and plotting of spectra. This program was modified slightly in 1986 to 'mem_tst_jes' which allowed an 8-fold higher plotting and integrating density.

-code, written by the author in 1985, for generating fabricated or model I/Fs representing:

-closely-spaced triangular line shapes. This code was not used in 1986 since a different line shape was chosen as described later.

-noise which was understood at the time to be band-limited, Gaussian-distributed, white noise. More is reported later about characterization of a set of noise data.

B. Approaches taken - Where appropriate, effort was made to build on the experience and results of the 1985 study. The study once again focused on MEM reduction of I/Fs fabricated by computer to represent spectra with well-defined and simple structures. The I/Fs always contained 2048 data points, intensities, each with mirror displacement of 0.6328×10^{-4} cm. The center or zero-path datum was the 1025th point. Spectra represented were either two-lined or five-lined groups of delta functions of which the first line was at 3456 cm^{-1} and whose line separation was a user-chosen parameter. Figure 1 illustrates the

two types of spectra.

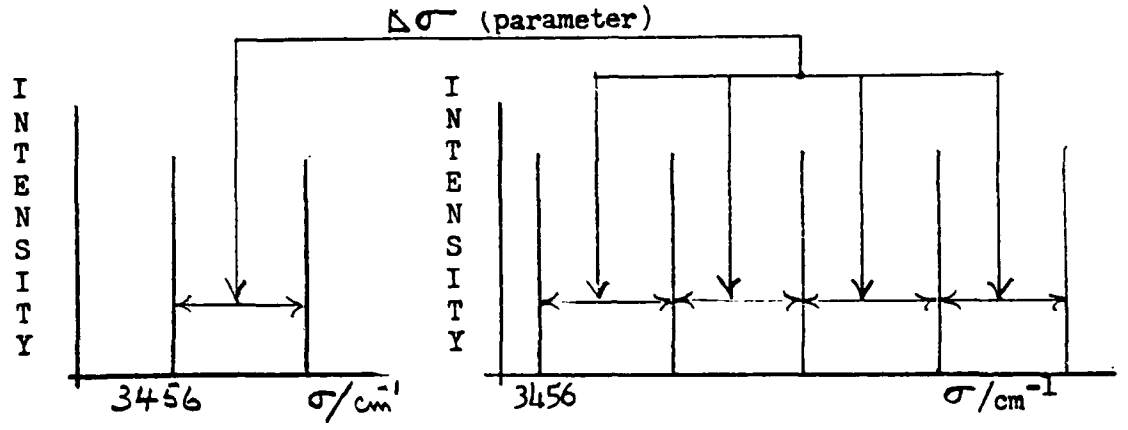


Figure 1. Model spectra represented by fabricated interferograms.

In the FFT reduction of any given I/F, zero-filling on both sides of the 2048 points was used to make a total of 8192 points. No apodization scheme was applied. The resulting FFT spectra, even though showing side lobes because of the finite extent of the I/Fs, represented the highest resolution attainable by the FFT algorithm. The given two or five lines were resolved as long as the assigned value of dsig exceeded about 15 cm^{-1}.

The main adjustable parameter in the MEM algorithm, given a fixed-length I/F, is the order M, the maximum lag, of the autocorrelation (14). The MEM software was made to operate on a given I/F in the following two manners, among others:

a) calculation of a so-called lag file containing the M coefficients in the column vector of equation [5] and also the final prediction error P_M of equation [4] but obtained through [5].

b) generation of the intensity (power) spectral plot from a log file computed in (a).

c) computation of integrals under chosen peaks of a spectrum, an operation appropriate only when the spectral lines were resolved.

The MEM is ironically least effective on completely noiseless idealized I/Fs. Operation on a pure two-line I/F with an ample separation (dsig = 40 cm-1) led to spurious, very sharp peaks especially with order M > 400. When noise signals giving S/N = 10 were added as described below, the spectrum was somewhat more reliably reproduced. Therefore, all of the fabricated I/Fs reduced by either FFT or MEM contained added noise.

IV. RESULTS -

A. Random Noise Fabrication - While the study of 1985 included a program to generate noise signals to be added to an I/F for an idealilzed spectrum, there was no means used to assess the characteristics of the noise data set. Thermal and shot noise signals are known (15) to have a Gaussian distribution of values. Consequently, a program called 'nzsort' was written to do the following operations on a data set:

- calculate the mean value and standard deviation of the data set
- sort the data set into 100 bins linearly divided between maximum and minimum values of the set
- calculate for these same bins the Gaussian populations normalized to the bin with maximum population.

With this program it was possible to make a comparison of the actual

distribution in a noise data set with that expected of a Gaussian distribution with the same standard deviation.

Operation of this program on the noise file used in the 1985 study showed that distribution, while narrowly peaked near zero (suggesting a small standard deviation), to contain enough large signals to make the computed standard deviation at least ten times the half width of the distribution. Consequently a new noise generator program 'fabnoise_86' was written to carry out the following sequence of operations:

1. to provide an array - a pool - of up to 20000 addresses divided into 100 bins equally spaced in signal magnitude

2. to fill this array with positive signal values x distributed according to the Gaussian distribution function

$$f(x) = [\text{const.}/(2\sqrt{2\pi\sigma_x^2})] \exp[-x^2/2\sigma_x^2]$$

where the term 'const.' was chosen to limit the integral $f(x)\,dx$ to no more than 20000 entries summed over all bins. The standard deviation σ_x was chosen by the user and x ranged over $5\sigma_x$.

3. to draw values from this pool at random

4. to choose, also at random, + or - sign for each value drawn

5. to store the values drawn, with their signs, in an array of 2048 addresses to constitute a noise file.

Noise files so generated were tested by 'nzsort' and were found to have distributions consistent with the expected Gaussian distribution for the same σ_x.

A program 'nzyspec' written in 1985 was then used to generate an

I/F file representing a chosen idealized interferogram with noise added. It involved merging, summing, a spectral I/F with a noise file point-by-point with a factor 'facnz' included to allow choice of the S/N ratio. It was these merged files which were reduced by the FFT or MEM operations.

B. Performance of the MEM - Presented here are several figures in each of which the FFT spectrum is compared with those obtained by the MEM on the same I/F. Integrals under the line profiles are given when the MEM successfully resolved the peaks. Comparison of integrals should be scrutinized only approximately since: (a) the Laurentian line shape of MEM spectra () admits of some overlap even when peaks are visibly resolved, and (b) from one spectrum to another the step size used in the integration routine may have been changed.

1. Two line Spectra Three I/Fs were generated to have line separations of 20, 10, and 5 cm^{-1}. Figure 2 shows that when dsig = 20 cm-1 the FFT can resolve the two lines. Resolution of these lines by the MEM depends on the parameter M and is successful with M > 100 with no splitting shown at M = 800, the highest order run in this case. When dsig is decreased to 10 cm^{-1}, the FFT can no longer distinguish two lines. The MEM, however, can resolve them reliably if the order M > 200 but begins to show splitting at orders about 800 , as shown in Figure 3. The MEM is capable of resolving the lines even when dsig = 5 cm^{-1}, but line splitting begins to compete so that there is only a narrow range of orders around M = 400 capable of reliable spectral regeneration. See Figure 4.

2. Five-line Spectra - A pair of I/Fs was generated to have five lines each separated by 20 cm^{-1} in one case or by 10 cm^{-1} in the other case. In both of these I/Fs the noise file used was the same as that added to the two-line I/Fs above. The S/N ratio was ~100, i.e., 5/2 times that involved in the two-line cases. Results are shown in Figures 5 and 6. Again, the FFT spectrum was capable of resolving the 20 cm^{-1} separation but not the 10 cm-1 separation. In both of these cases, the MEM results showed only a narrow range of orders capable of resolving the lines before the onset of spurious splitting. The peak intensities of the outermost lines of the group are at least five times those of the inner three lines, but the latter are broader. All of the integrals under the lines of the group are about equal.

A third five-line I/F was generated with a S/N ratio of ~10. As shown in Figure 7, the FFT spectrum is hardly different from that at S/N ~100. The MEM capability of resolving the lines is improved only in that the onset of spurious splitting is less pronounced here than at the lower noise level.

V. SUMMARY AND RECOMMENDATIONS -

In summary, this effort generated examples of I/Fs with characteristics in the boundary region between accurate and unsuccessful MEM regeneration of the corresponding spectra. The algorithm used was capable of resolution greater than FFT resolution by a factor of 2 to 3. Spurious lines appeared as the order M of the prediction error filter was increased. In five-line groupings, the central three lines were less intense but broader than the two end

lines.

It was remarked earlier that the superresolution limitations are not yet clearly identified in the literature on MEM especially when applied to I/Fs. The scope of MEM is so manifold that much room is left for further characterization before the method can be used without relying on other data reduction methods for sorting information content from artifacts.

Three relevant reports (14-16) issued by the AFGL warrant attention in the present context. They all include studies which made use of fabricated I/Fs to test the performance of the MEM, but the MEM algorithms used differed from that used in this study. Instead of using the subroutine MEMPR (13), these studies used algorithms based on the so-called Yule-Walker equations. The latter have matrix elements which are already autocorrelation functions. The results showed considerable success of the MEM, at least the versions used in these studies. Short interferograms of less than 500 points were found to reproduce more than 40 lines separated on the average by less than 2 cm^{-1}. Similar operation with the Yule-Walker algorithm on the interferograms used in the present study may help to explain the discrepancy in performance of the two versions of the MEM.

ACKNOWLEDGEMENTS

The author wishes to express his gratitude to the Air Force Systems Command, Air Force Office of Scientific Research for sponsorship of the SFRP Program and to the Air Force Geophysics Laboratory, Infrared

Technology Division, Infrared Dynamics Branch, for cordial accommodation. He is grateful for the support of his application by department Chairperson, Professor G. Doyle Daves of Lehigh University. Special thanks are due Dr. William A. M. Blumberg for suggesting the project and to Dr. Boh K. Yap of Yap Analytics, Inc. for guidance during critical stages of the effort. Dr. Dorothy Flanagan and Sgt. Debra Douglass were very helpful in searching and linking files in the Apollo computer system as was Ann Lozier of Realtime Engineers, Inc. in adapting software to the needs of the project. Stimulating discussions were held with Drs. Alfred Rahbee and Paul Fugere of AFGL. Help by Dr. Agnes Bain in lisiting possible living quarters and area activities is appreciated.

REFERENCES

1. Griffiths, P. R., ed., Transform Techniques in Chemistry, New York, Plenum Press, 1978.

2. Cooley, J. W., and J. W. Tukey, "An Algorithm for the Machine Calculation of Complex Fourier Series," Math. Computation, 19, 297 (1965).

3. Good, I. J., "The Interaction Algorithm and Practical Fourier Series," J. Roy. Statist. Soc., Ser. B., 20, 361 (1958).

4. Bracewell, R., The Fourier Transform and its Applications, New York, McGraw-Hill, 1965; also: 2nd ed., 1978.

5. Haykin, S., ed., Nonlinear Methods in Spectral Analysis, Springer-Verlag, 1979.

6. Connes, J., "Computing Problems in Fourier Spectroscopy," in G. A. Vanasse, A. T. Stair, Jr., and D. J. Baker, eds., Aspen International Conference on Fourier Spectroscopy, 1979.

7. Shannon, C. E., and W. Weaver, The Mathematical Theory of Communication, Urbana, The Univ. of Illinois Press, 1949.

8. Burg, J. P., "Maximum Entropy Spectral Analysis," paper presented at the 37th Annual International Meeting, Soc of Exploratory Geophysics, Oklahoma City, Oct. 31, 1967; also: Burg, J. P., Ph.D. Dissertation, Dept. of Geophysics, Stanford Univ., 1975.

9. Kawata, S., K. Minami and S. Minami, "Superresolution of Fourier Transform Spectroscopy Data by the Maximum Entropy Method," Applied Optics, 24(2), 162 (1985).

10. Rahbee, A., "Application of Maximum Entropy Spectral Analysis to

Fourier Transform Mass Spectrometry," Chem. Phys. Lett., 117(1), 352 (1985).

11. Ni, F., and H. A. Scheraga, "Resolution Enhancement in Spectroscopy by Maximum Entropy Fourier Self-Deconvolution, with Applications to Raman Spectra of Peptides and Proteins," J. Raman Spect., 16(5), 337 (1986).

12. Viti, V., P. Barone, L. Guidoni and E. Massaro, "Maximum Entropy Spectral Analysis of ^{31}P NMR Signals from Human Cells," J. Magn. Res., 67, 91 (1986).

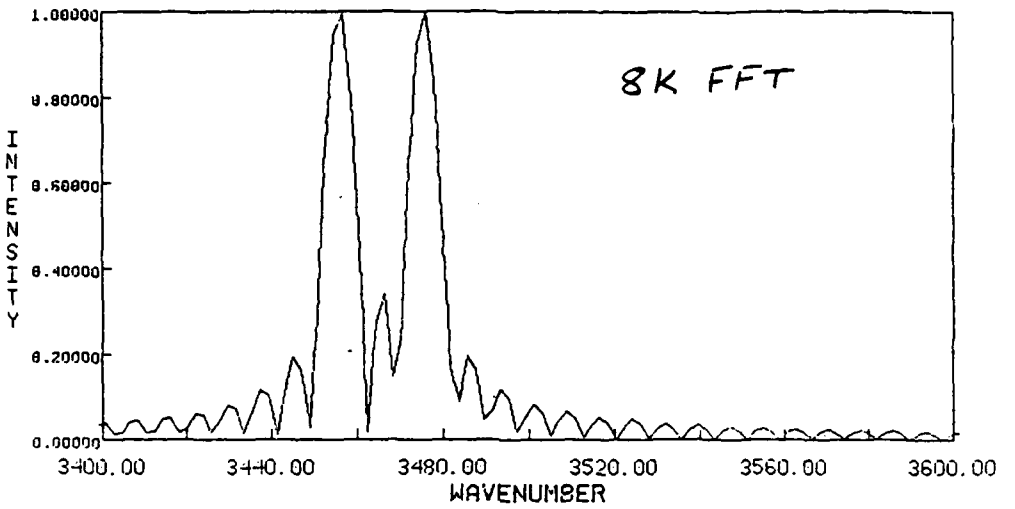
13. Ulrych, T. J., and T. N. Bishop, "Maximum Entropy Spectral Analysis and Autoregressive Decomposition," in D. G. Childers, ed., Modern Spectrum Analysis, New York, IEEE Press, 1978.

14. Lee, Y. W., Statistical Theory of Communication, Wiley, 1960.

15. Dewan, E. M., "A Review of Maximum Entropy Spectral Analysis and Applications to Fourier Spectroscopy," AFGL-TR-85-0091 Environmental Res. Paper No. 911, 3 April, 1985.

16. Richards, A. R., S. H. Delay and W. F. Grieder, "AR Methods for Spectral Estimation from Interferograms," Utah State Univ. Contract USU 85-038 Final Report, 15 May, 1986.

17. Zachor, A. S., and D. S. Smith, "A Study of the Maximum Entropy Method of Power Spectrum Estimation as Applied to Interferometer Data," Utah State Univ. Subcontract 85-077 Final Report, Feb. 20, 1986.



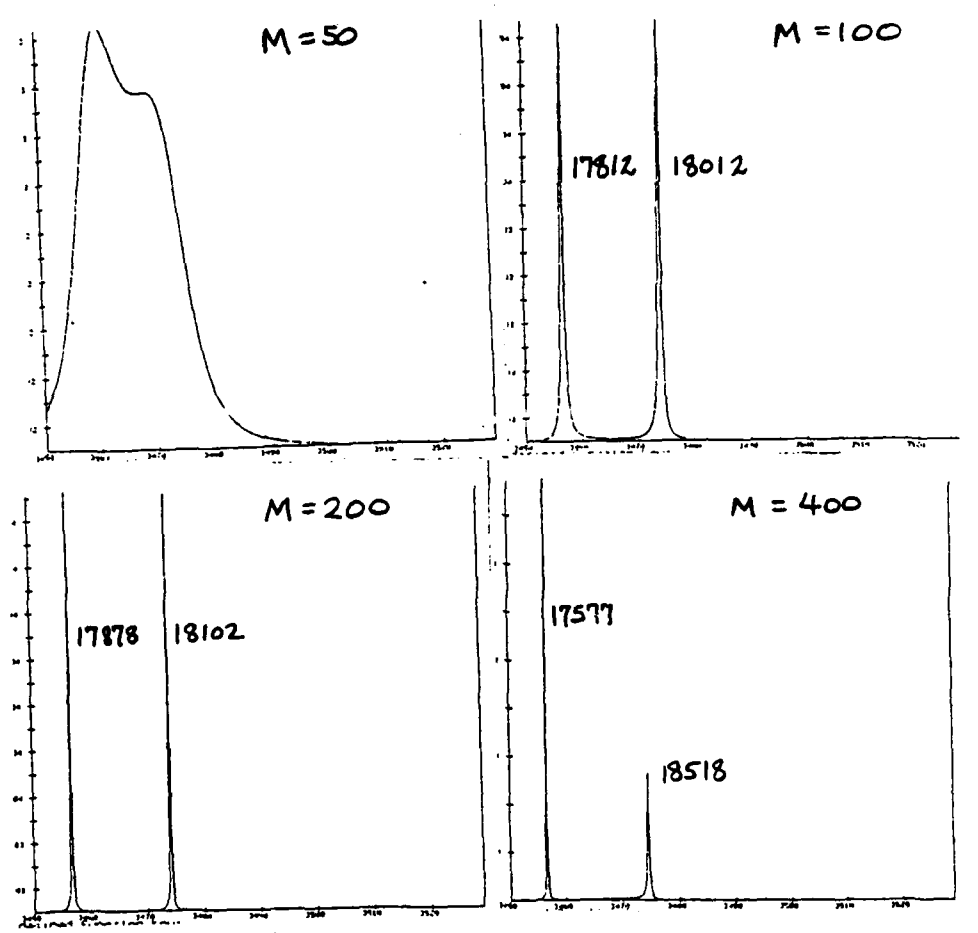
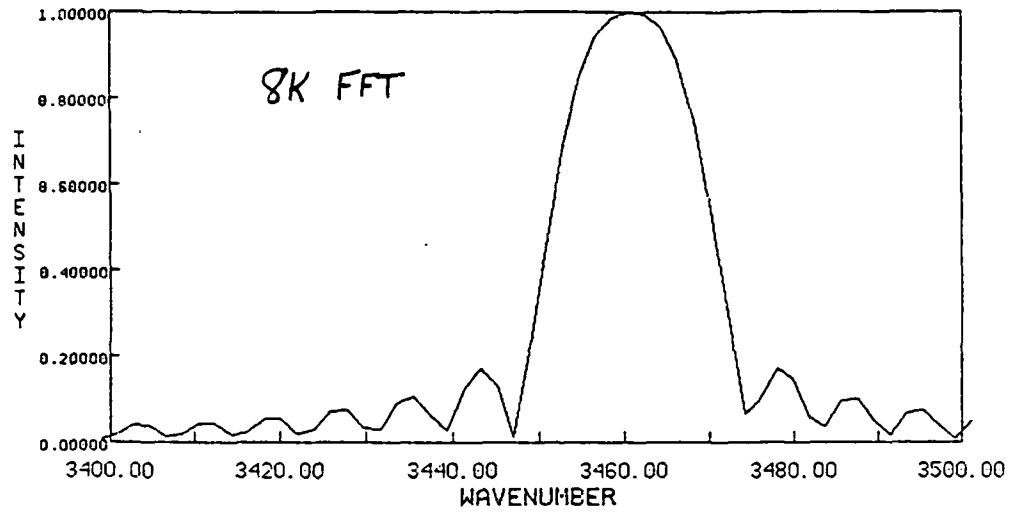
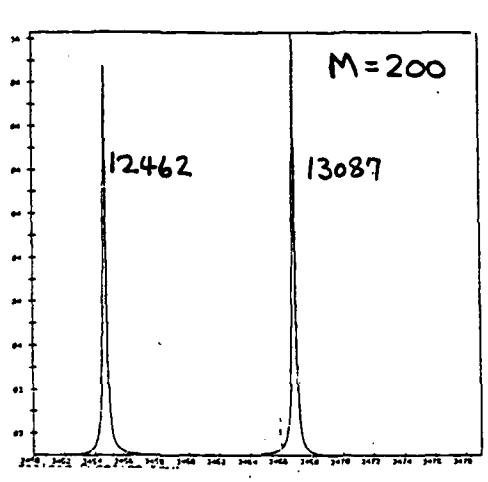
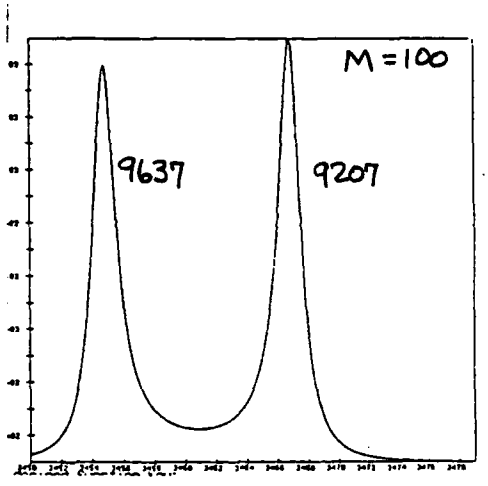


Figure 2. FFT and MEM results on a 2-line I/F with $\Delta\sigma$ = 20 cm^{-1}.







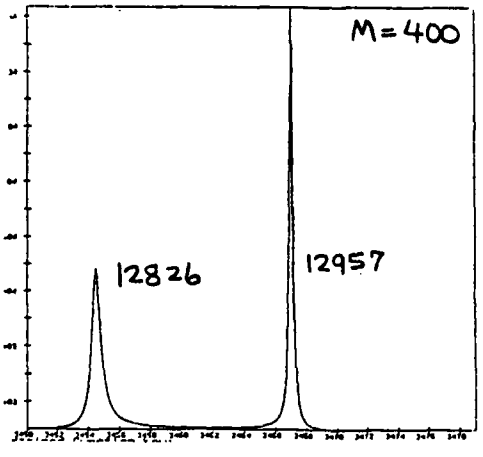


Figure 3. FFT and MEM results on a 2-line I/F with $\Delta\sigma$ = 10 cm^{-1}.

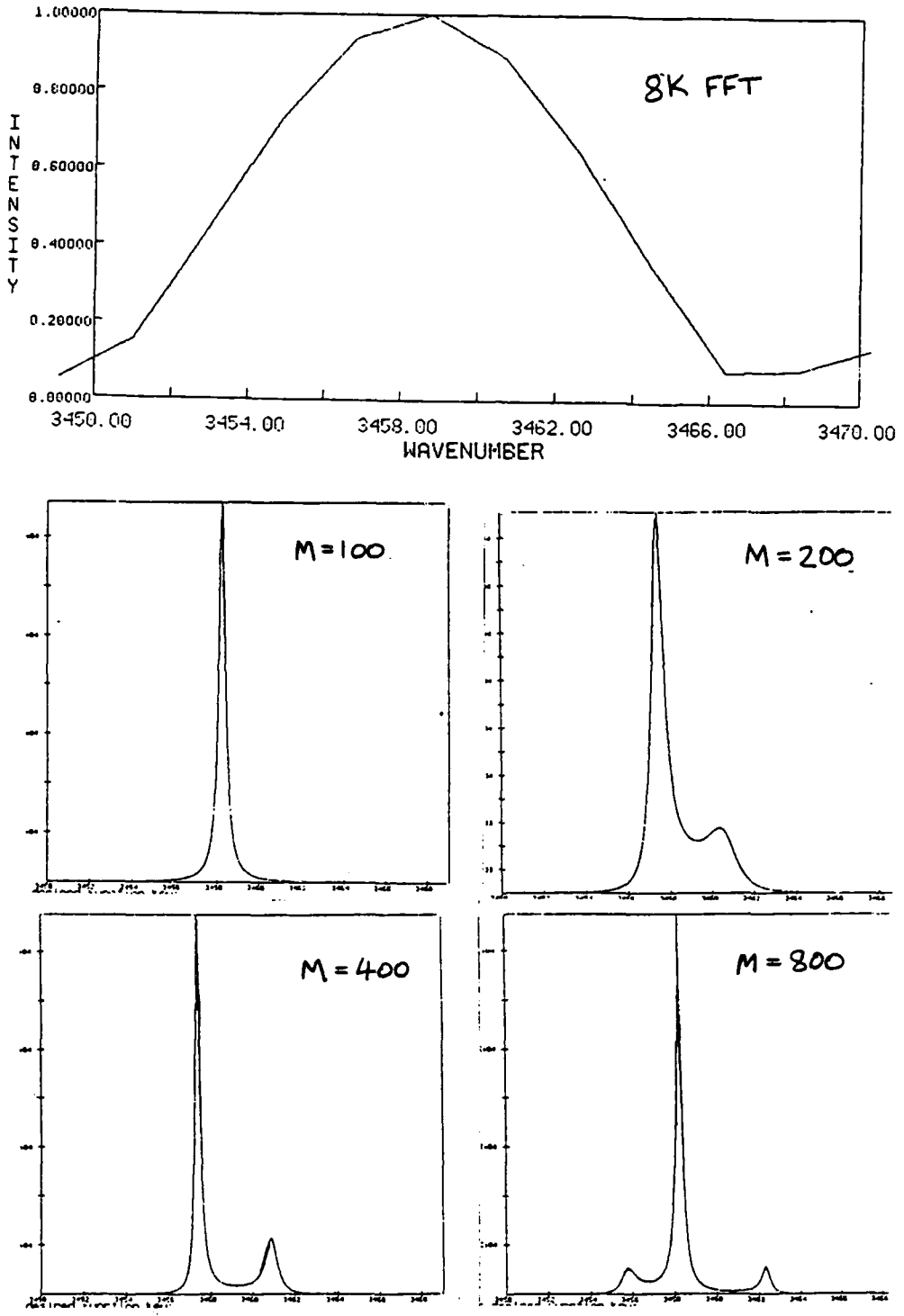


Figure 4. FFT and MEM results on a 2-line I/F with $\Delta\sigma = 5\ cm^{-1}$.

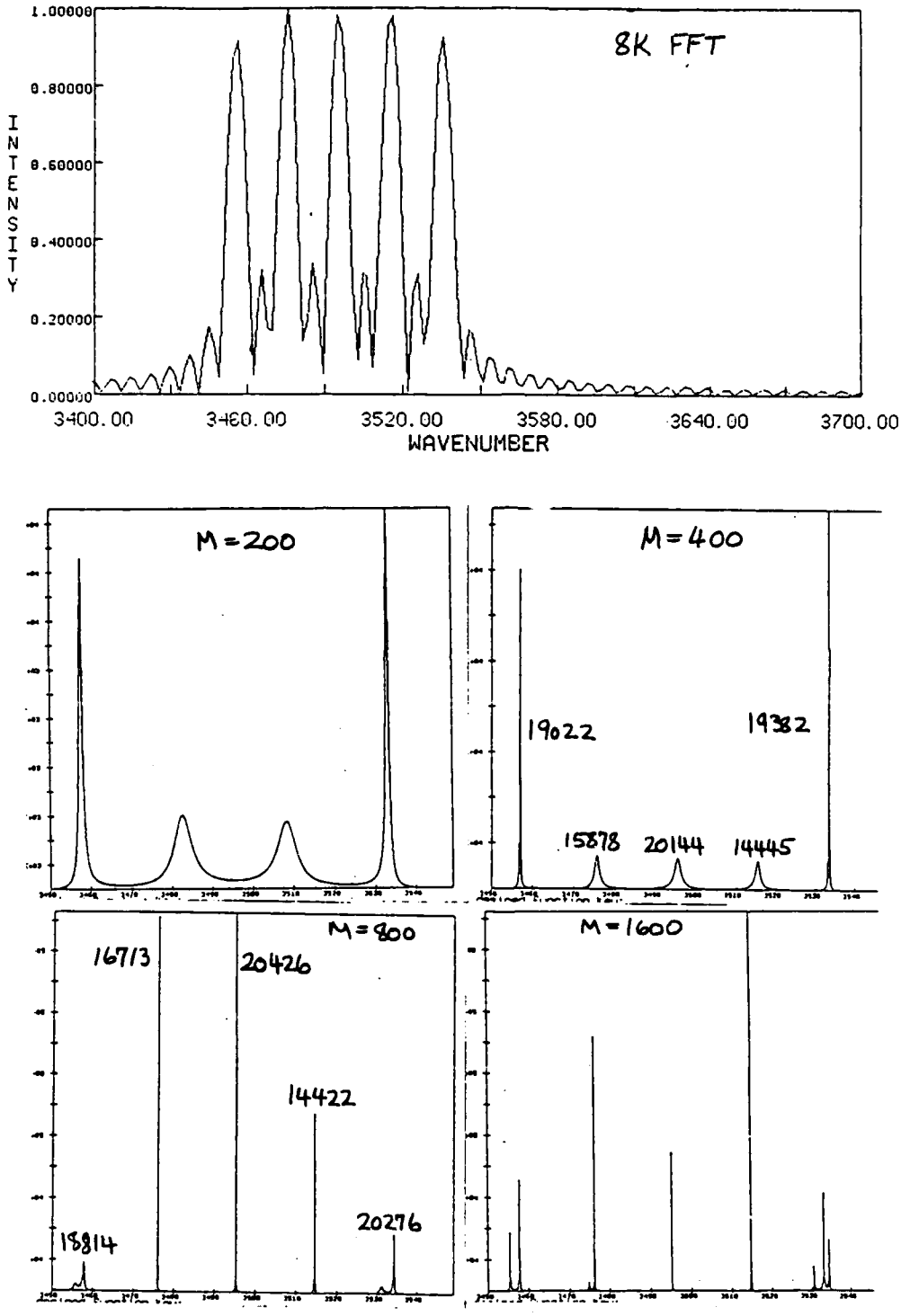
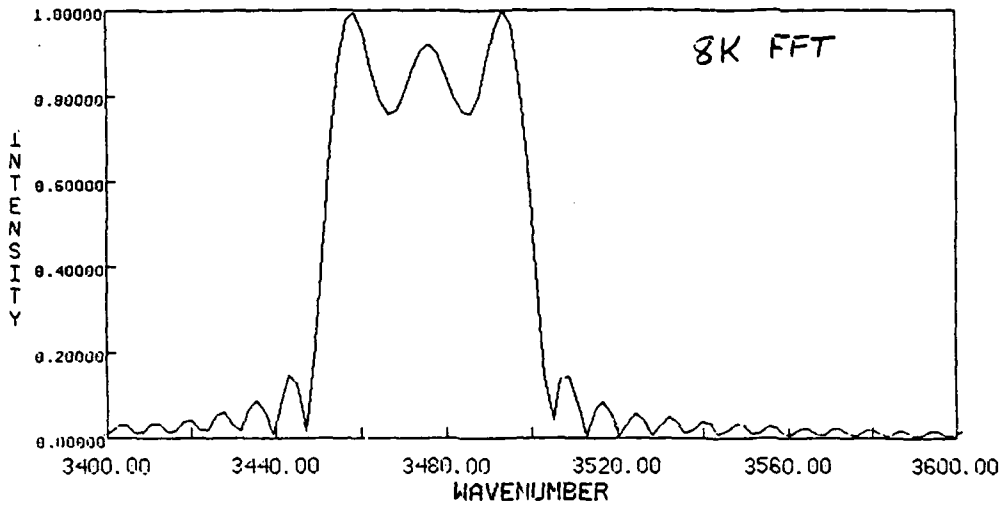


Figure 5. FFT and MEM results on a 5-line I/F with $\Delta\sigma$ = 20 cm^{-1}.



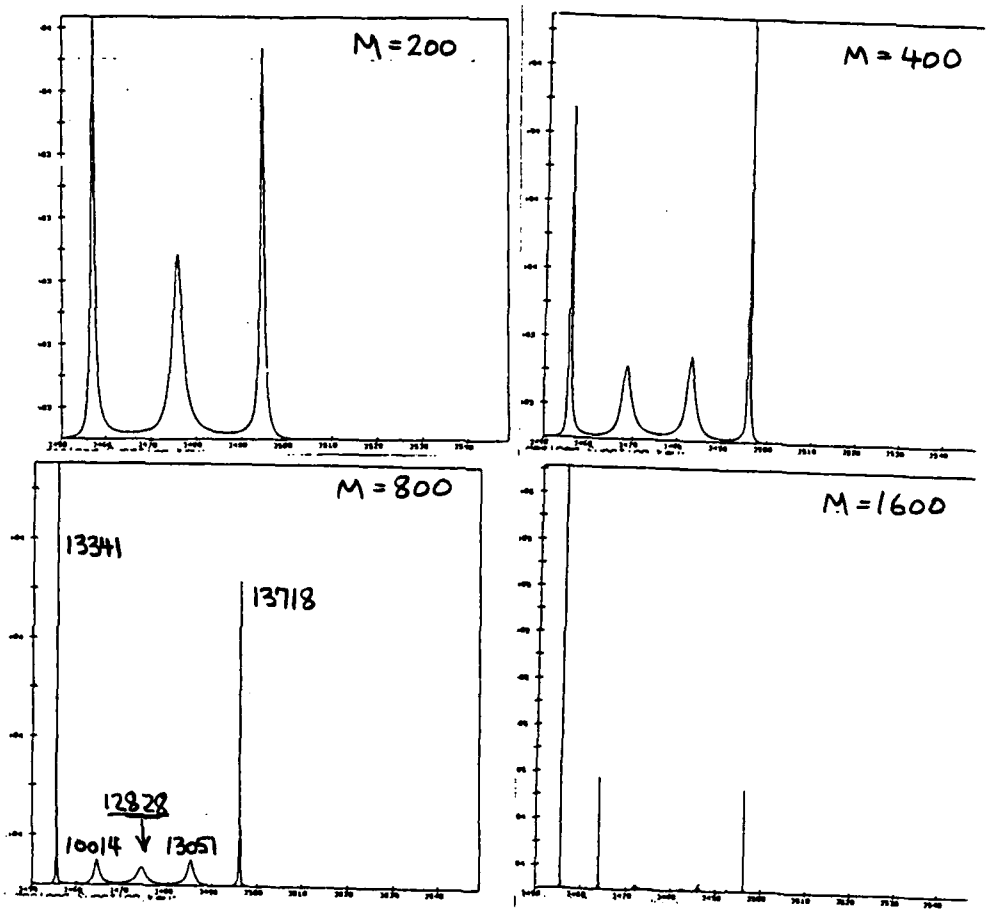
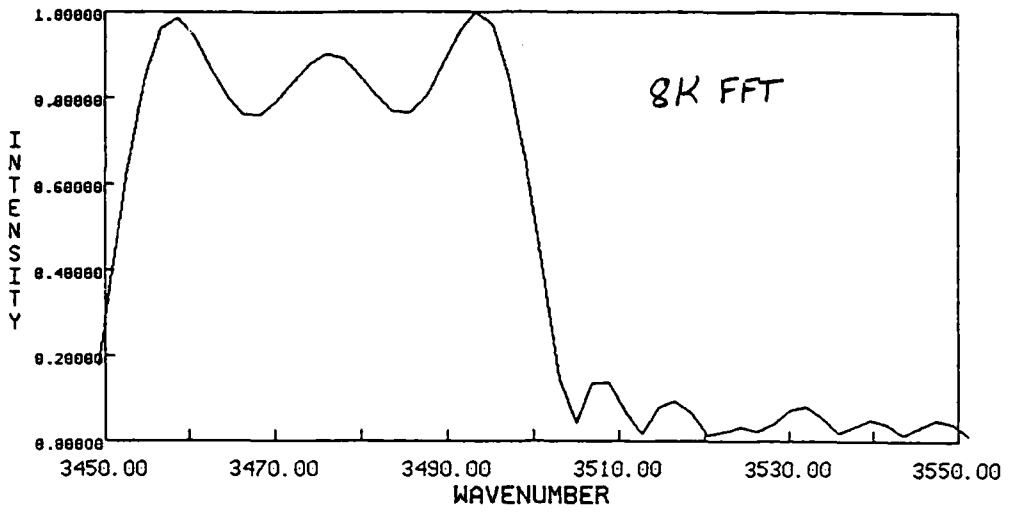


Figure 6. FFT and MEM results on a 5-line I/F with $\Delta\sigma$ = 10 cm^{-1}.



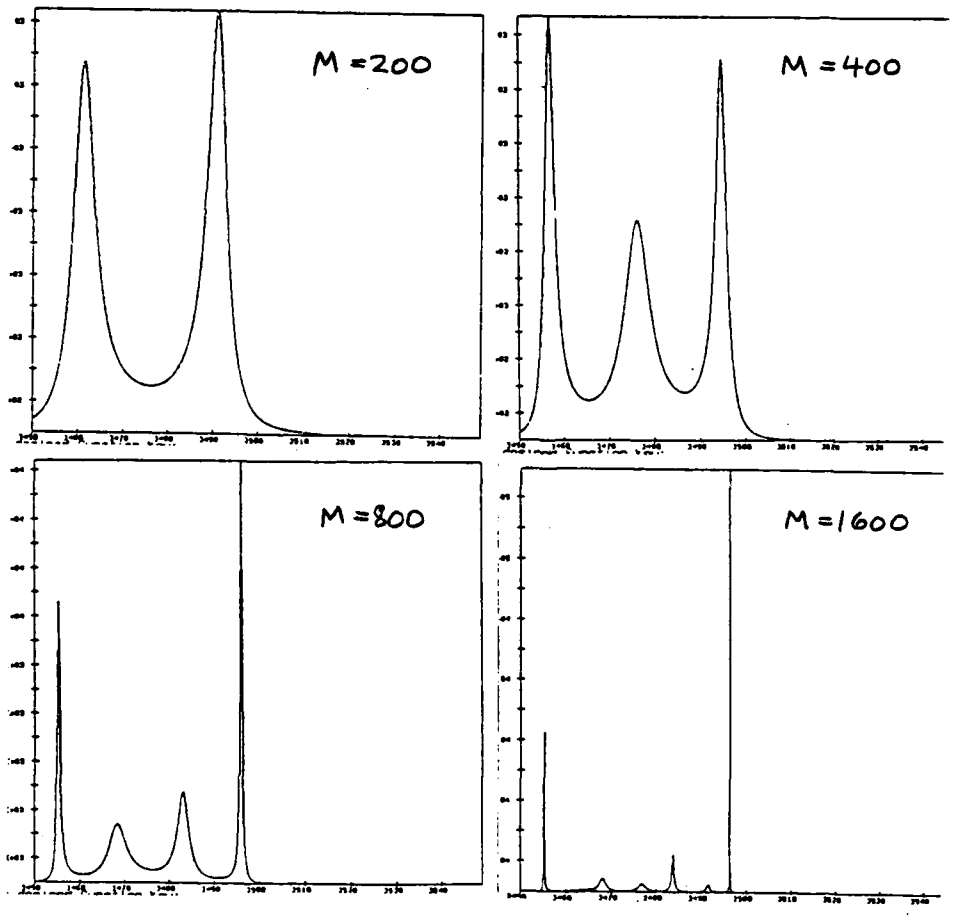


Figure 7. Results similar to Fig. 6 but with S/N ratio ~ 10.

1986 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT SUMMER SUPPORT PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

AI and Large-Scale Systems Approaches to Enhanced Situation Awareness in Missile Warning Systems

Prepared by:	Edgar C. Tacker
Academic Rank:	Professor
Department and University:	Electrical Engineering Department The University of Tulsa Tulsa, Oklahoma
Research Location:	Technology Development Branch Human Engineering Division Harry G. Armstrong Aerospace Medical Research Laboratory Wright-Patterson Air Force Base Dayton, Ohio
USAF Research:	Michael D. Mc Neese
Date:	8 August 1986
Contract No.:	F49620-85-C-0013

AI and Large-Scale Systems Approaches to Enhanced Situation Awareness in Missile Warning Systems

by

Edgar C. Tacker

ABSTRACT

A "top-down" view of a missile warning C^3 system was described and then particularized to apply to the SIMCOPE missile warning problem. The cognitive tasks required of the missile warning officer (MWO) were analyzed and modeled via an "information processing matrix". A "descriptor language" for precisely formulating hypotheses relative to the MWO decision making process was outlined and a particular type of decision-aiding expert system was recommended for enhancing the MWO' situation awareness. Several natural extensions of this research were described, and an appropriate sequencing of these proposed research topics was given.

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to the US Air Force Office of Scientific Research for sponsoring this research and to the Technology Development Branch at the AAMRL laboratory for providing the research topic and a working atmosphere that was friendly, supportive, and one in which free inquiry and creativity can flourish. It was a thoroughly enjoyable experience for me.

I would also like to thank the staff at Universal Energy Systems for their friendly and professional handling of the business aspects of this program.

Finally, I would like to thank my wife Florence for taking care of the many family matters back in Tulsa so that I could concentrate on this research.

I INTRODUCTION

In recent years my research interest has evolved toward the artificial intelligence/cognitive psychology area and applications to command & control problems. I had previously performed research in computer control systems, stochastic optimal control and estimation, interconnected electric power systems, and generic large-scale decision systems, In 1982-83 I performed research (1,2) in applying large-scale systems and queueing theory to command & control problems associated with nuclear attack-class submarines, and initiated the personal goal of defining a generic paradigm for intelligent decision-making in complex command & control systems. Since that time I have been integrating AI ideas into the framework of that paradigm. The research in decision-aiding in C^3 systems at the Technology Development Branch of the Human Engineering Division at the Harry G. Armstrong Aerospace Medical Research Laboratory (AAMRL) seemed to offer me an opportunity to continue my research program within the context of problems of significant practical and national import.

II OBJECTIVES OF THE RESEARCH

In 1984 the Technology Development Branch at AAMRL initiated a plan to apply AI techniques to a missile warning problem simulated via their SIMCOPE facility. As the first phase of this research, they decided to develop an expert system(ES) that had explanation capability. The resulting expert system, SENTINEL, is described in (3). SENTINEL was a forward-chaining system using a "community of blackboards" approach, and provided a two-level "why-how" explanation capability. This ES was designed to provide the missile warning officer(MWO) assistance toward meeting the continuing need to maintain situation awareness.

The objective of my research this summer was to build upon what was learned in the first phase and to develop approaches that could lead to an expert system or systems that would further the effort toward enhancing the operator's situation awareness.

III THE GENERIC MISSILE WARNING PROBLEM

Let us take a "top-down" view of a generic Missile Warning System (MWS). The "ceiling" is establishedwithin the mind of the Ultimate Decision Maker(UDM) residing within the Command Defense Center(CDC). This UDM seeks at all times to establish an accurate view of the "situation" from a global perspective, for this is the only consistent precursor of intelligent informed decision-making.

Aiding the UDM are a set of Local Decision Makers(LDM's). These LDM's seek to establish an accurate view of the "situation" from their own particular local perspectives. The UDM serves as the coordinator[1] in two distinct ways: (1) He[2] specifies a "Readiness State" and a set of descriptors to the LDM's and requests that the LDM's determine "local values" for these descriptors and then to forward these values (in the form of "reports") to him in a timely manner; and (2) He fuses these individual sets of descriptor values into a "world-view" of the current "situation" of the overall system.

In carrying out Task 1, each LDM will have access to two types of data: (1) System Measurement Data (SMD); and (2) Context Data (CD). The LDM uses these data plus his "personal world knowledge" to obtain a localized view of the "situation". The requested descriptor values are then obtained as "projections" from this localized view of the situation. That is, one can think of each slot value in the report form as being the "projection" of this local view of the situation along a particular "coordinate axis" selected by the UDM(who designed the report forms).

Before moving on, it should be mentioned that each descriptor is intended to describe a particular aspect of the situation, thus the

---

1 In some C^3 decision problems there will also be communication "down" (i.e., from the UDM to the LDM's) and "across" (i.e., between the LDM's) the system (e.g., see (1,2,4)).

2 Throughout this report the reader should feel free to substitute 'she' for 'he', 'her' for 'him', etc.

choice of the terminology "descriptor". Moreover, a descriptor is to be thought of as a type of information processing function, one that plays a key role in defining the aforementioned generic paradigm for intelligent decision-making in complex C^3 systems.

IV THE SIMCOPE MISSILE WARNING PROBLEM

One of the most important of the LDM's is the Missile Warning Officer. The role of the MWO and the environment within which he works has been simulated within SIMCOPE[3]. There are two principal players (countries), US and THEM, in the SIMCOPE "world". There are two Advanced Detection System (ADS) satellites (North and South) positioned over THEM, and seven Barrier Surveillance System (BSS) radar sites (#1-7) located on the border of US adjacent to THEM. There are two types of missiles that THEM are capable of launching onto US. Type 1 is a high accuracy non-maneuverable single warhead missile having an IR intensity signature in the (normalized) range of 4 ± 2, while Type 2 is a low accuracy multiple warhead missile of limited maneuverability having an IR intensity signature in the (normalized) range of 6 ± 1. There are eight "known" lauch sites in THEM; two are known to launch Type 1 missiles, two are known to launch Type 2 missiles, while the types of launchersin the other four launch sites are unknown. There are also six "suspected" launch sites in THEM. Several targets in US are distinguished as "high value targets"---the C^3 processing sites and the capitol of US. The nine surveillance sensors exist for the purpose of gathering system measurement data and then communicating associated messages directly to the MWO. If, indeed, a missile were to be launched from THEM and subsequently were to penetrate the air space of US, and if all sensors along the resulting trajectory and all required communication links were operating as they should, then the MWO should receive the following sequence of three messages:[4]

---

3 COPE: C^3 Operator Performance Evaluation.
SIMCOPE: A dedicated computer-driven facility at AAMRL that simulates a C^3 command post similar to that of an MWO. The SIMCOPE manual states that "scenarios used in SIMCOPE are fictional but believed to be credible representations of real missions"

4 The descriptions of the seven data fields of all message types differ, sometimes greatly, from those in the SIMCOPE manual. A similar comment applies with respect to the descriptions for other SIMCOPE messages and event reports to be considered later in this report.

1st	ADS-1 Message	"Pass 1" Data
		(1) Detection Source (North or South Satellite)
		(2) Time of Detection
		(3) IR Intensity (0-9, integer values)
		(4) Launch Latitude
		(5) Launch Longitude
		(6) Predicted Time of "Pass 2" Data
		(7) Sensor Site Confidence Measure (0-9, integer values)
2nd	ADS-2 Message	"Pass 2" Data
		(1) Detection Source (North or South Satellite)
		(2) Time of Detection
		(3) Missile Type
		(4) Heading
		(5) Predicted BSS site to Detect
		(6) Predicted Time of BSS Data
		(7) Sensor Site Confidence Measure (0-9, integer values)
3rd	BSS Message	BSS Data
		(1) Detection Source (1-7)
		(2) Time of Detection
		(3) Number of Objects
		(4) Heading (of "center of pack")
		(5) Trajectory Stability (Yes/No)
		(6) Predicted Time of Missile Impact
		(7) Sensor Site Confidence Measure (0-9, integer values)

This "actual launch" scenario wherein this sequence of three messages is received is only one of a myriad of possible scenarios that could transpire. There could be fires or explosions in THEM that might be "detected" by one of the ADS satellites and then lead to a "false alarm" at the Pass 1 stage. There can also be sensor malfunctions that can cause false alarms, ambiguous or erroneous data, loss of data, etc. The MWO can misinterpret data because of various human factors limitations. Further, there are some "built-in" ambiguities, e.g., an IR intensity level of either 5 or 6 can correspond to either a Type 1 missile or a Type 2 missile. There can also be correlated errors, e.g., heading errors propagating into extrapolated errors in predicting impact points, etc.

It is easy to see that there are many opportunities for the MWO to experience erroneous, incomplete, ambiguous, etc. data, and, due to human factors limitations, there will also be many opportunities for "operator-generated" errors (omission and commission), especially when the MWO is subjected to heavy workloads.

To assist the MWO, two types of context data, over and above the UDM-supplied readiness state, are supplied. In particular, the status of all sensors and the communication network is continuously monitored, and System Status (SYS) messages are sent to the MWO when appropriate. The other contextual source of informations concerns Intelligence (INT) messages. Such messages provide a component of "world knowledge" to the MWO. INT messages are sent to the MWO whenever pertinent intelligence information becomes available. These two message types are of the following form:

System Status (SYS) Message	SYS Data
	(1) Type of Outage (Routine/ Emergency)
	(2) System Affected (Satellite, Radar, C^3 Network, etc.)
	(3) Subsystem Affected (Specific Sensor, C^3 Site, etc.)
	(4) Reason for Outage (Test, Failure, Sabotage, etc.).
	(5) Time of Outage
	(6) Estimated Time of Return to Service
	(7) Current Capability (Full, Partial, Zero)

Intelligence (INT) Message	INT Data
	(1) Initiator of Action (Country)
	(2) Type of Action
	(3) Recipient of Action (Country)
	(4) Expected or Known Time of Action
	(5) Result of Action
	(6) Reporting Source (OPEN or INTEL)
	(7) Time of Previous Correlated Message (if any).

If one pictures himself seated at the SIMCOPE command post, he will observe two large color displays. Normally, one of the screens will feature a "detail map" (a plan view of the two countries US and THEM with launch sites, targets, satellite, and radar coverages, missile "footprints", etc. clearly marked), and the other screen displays messages, message sequences, flashing selection keys, report forms with slots to be filled in, etc..

The principal tasks that the MWO has to perform in this environment is to fill out three types of reports---ADS-1, ADS-2, and BSS---as accurately and intelligently as possible, and then to forward them to the UDM as soon as possible. The MWO uses the data in the corresponding ADS-1, ADS-2, and BSS messages, together with the available contextual data (Readiness State, SYS and INT messages), plus his "world knowledge" and intuition in carrying out these tasks.

V ANALYSIS OF THE SIMCOPE MISSILE WARNING PROBLEM

It is easy to see that the MWO can become confused, or simply overloaded, etc. and can experience, at times, a sharp increase in error rate[5] in the process of filling out these reports. So, even before considering the detailed procedures in filling out the reports, we can see that the MWO can certainly profit from some "intelligent" assistance. This observation will guide us throughout all that follows.

Upon looking at the specific requirements given in the SIMCOPE manual it can be seen that there are instances in which the MWO is requested to fill in report entries for which:

(i) No human judgement is actually needed, or

(ii) Very little human judgement is needed, or

(iii) The human judgement required depends upon a reasonably complicatedweighting of evidence that can only be "verified" (or "rechecked") by performing a series of key selections, screen scannings, etc..

All of these factors would seem[6] to decrease the effectiveness of the MWO's

---

5 In the controlled experiment simulation environment of SIMCOPE, one can know the correct response, and can thus determine true error rates.

6 It would be interesting to devise some psychological experiments to test these hypotheses. It would seem that there would be two types of factors to be considered: (1) MWO perception of (i)-(iii), and (2) resulting effect on MWO's performance.

decision-making and the subsequent quality of his reports. It seems clear that few, if any, operators would choose to have unnecessary cognitive loads placed upon them, and insofar as they perceived (i)-(iii) (especially during high load periods), they mightexperience feelings and attitudes (though perhaps imperceptible at the conscious level) that would be counterproductive to task performance.

The first order of business would seem to be that of addressing the problems associated with (i)-(iii). This could be done by modifying SIMCOPE, by adding an expert system (ES) or by a combination of these two. The approach considered herein involves utilizing an ES to address (i)-(iii). An actual implementation of this could, of course, approach it differently.

Let us now concentrate on the task of seeking to improve MWO performance. As already alluded to, a great deal is expected of the MWO. He is expected to acknowledge all five types of messages as they arrive, and to "track" the ADS and BSS messages within the context[7] of previous INT, SYS, ADS, and BSS messages. This "tracking" involves piecing together "events" (an abbreviation for "missile events"), wherein the MWO is always "integrating" data and seeking to recognize ADS-1, ADS-2, BSS message sequences that "match" (i.e., that correspond to the 1st, 2nd, and 3rd "samplings", respectively, of an actual missile trajectory passing from THEM to US). This constant "integration" requires a high level of attention to be maintained by the MWO, and any proposed decision-aid must take this into account.

To better understand the duties of the MWO, one may think of the MWO as, effectively, an intelligent information processor whose inputs are Event[8], SYS , and INT messages, and whose outputs are Event Reports. As discussed earlier, the purpose of these reports is to provide a "local" component toward the sequential updating of the situation awareness of the UDM.

The formats of the three types of event reports are given next[9]. Note

---

7 Henceforth, inclusion of "readiness state" as a component of context will become tacit.

8 By "Event" is meant any element in the set ('ADS-1', 'ADS-2', 'BSS').

9 The SIMCOPE manual provides additional details.

that under "comments" the entries in the column to the extreme right correspond either to automatically filled in values, or to the previously discussed items (i)-(iii), or to new items (iv), (v). The next section will provide a unified method of describing all of these items.

	ADS-1 Report		Comments	
	ADS PASS 1 EVENT:	_____	Add 1 to event count.	(i)
1.	DETECTED:	ZZ:ZZ:ZZ	Sensor data, filled in auto.	--
2.	ADS:	ZZZZZ	Sensor data, filled in auto.	--
3.	REGION:	_____	Launch Site (Region of THEM)?	(i)
4.	SITE:	_____	(KNOWN, SUSPECTED, UNKNOWN)?	(i)
5.	LAUNCHER TYPE:	_____	(TYPE 1,2, INDISTINCT, UNKNOWN)?	(i)
6.	MWO CONFIDENCE:	_____	0-9 (Call it ADS1CONFIDENCE)	(iv)

	ADS-2 Report		Comments	
	ADS PASS 2 EVENT:	_____	Corresp. to which ADS-1 report?	(iii)
1.	DETECTED:	ZZ:ZZ:ZZ	Sensor data, filled in auto.	--
2.	ADS:	ZZZZZ	Sensor data, filled in auto.	--
3.	EVENT TYPE:	_____	(HOSTILE,TEST,UNRESOLVED,UNKNOWN)?	(iv)
4.	BSS DETECT:	_____	(1-7:BSS Site #, 0: None)?	(ii)
5.	THREAT:	_____	(YES,NO)?	(iv)
6.	MWO CONFIDENCE:	_____	0-9 (Call it ADS2CONFIDENCE)	(iv)

	BSS Report		Comments	
	BSS EVENT:	_____	Corresp. to which ADS-2 report?	(iii)
1.	DETECTED:	ZZ:ZZ:ZZ	Sensor data, filled in auto.	--
2.	BSS:	Z	Sensor data, filled in auto.	--
3.	THREAT:	WWW	Filled in-from matching ADS-2 Rpt.	(v)
4.	LAUNCHER TYPE:	WWWW	Filled in-from matching ADS-1 Rpt.	(v)
5.	OBJECTS TRACKED:	Z	Sensor data, filled in auto.	--
6.	POSSIBLE TARGET:	_____	Most likely target?	(ii)
7.	EST. IMPACT TIME:	VV:VV:VV	Filled in auto.	--
8.	MWO CONFIDENCE:	_____	0-9 (Call it BSSCONFIDENCE).	(iv).

VI A DESCRIPTOR LANGUAGE FOR MODELING THE MISSILE WARNING DECISION PROCESS

Let us now follow up on the "information processing view" of the MWO reporting function. Figure 1 facilitates that view. The input "messages" constitute sets of rows and the output "reports" constitute sets of columns in this "information processing matrix (function)". We shall denote by M_{MWO} the function represented by the matrix in Figure 1, and by (M_{MWO}) the matrix itself.

M_{MWO} is intended to serve as an abstract[10] representation of the MWO reporting task. The matrix representation (M_{MWO}) serves as a "cognitive interface" between the abstract mathematical function and a concrete realization for that function.

Much of the structure of M_{MWO} is implied rather than explicit in Figure 1. In particular, the operational aspect of this IP matrix demands the incorporation of the temporal aspect of the information processing.

One way to achieve this is to define three types of "component" functions:

I^k: An instantaneous[11] function, whose domain is a set of event elements (rows in (M_{MWO}) corresponding to the most recent ADS or BSS message inputs); and whose range forms a component of descriptor D^k,

S^k: A state function, whose domain is a set of history elements (rows in (M_{MWO}) corresponding to present and past ADS, BSS, INT, or SYS message inputs); and whose range forms a component of descriptor D^k,

A^k: An algorithmic component that performs data[12] processing tasks such as table-look-ups, comparisons, metric calculations, function minimizations, etc..

10 That is, as an abstract mathematical function.

11 That is, instantaneous modulo the amount of time it takes the MWO to process the information in this "row-to-column" transformation.

12 As opposed to information processing.

SYSTEM MEASUREMENT DATA	DESCRIPTORS: ADS1-EVENTNUMBER -REGION -LAUNCHERSITETYPE -LAUNCHERTYPE -CONFIDENCE	ADS2-EVENTNUMBER -EVENTHYPOTH -BSSDETECTMWO -THREAT -CONFIDENCE	BSS-EVENTNUMBER -THREAT -LAUNCHERTYPE -POSSIBLETARGETS -CONFIDENCE
ADS1-DETECTIONSOURCE -DETECTIONTIME -IRINTENSITY -LAUNCHLATITUDE -LAUNCHLONGITUDE -ETAPASS2 -SENSORSITECONFID			
ADS2-DETECTIONSOURCE -DETECTIONTIME -LAUNCHERTYPE -HEADING -BSSDETECT -ETABSS -SENSORSITECONFID	X	X	
BSS-DETECTIONSOURCE -DETECTIONTIME -NUMBEROFOBJECTS -HEADING -TRAJECTORYSTAB -ETAOFIMPACT -SENSORSITECONFID	X	X	X
INT-ACTIONINITIATOR -ACTIONTYPE -ACTIONRECIPIENT -ACTIONTIME -ACTIONRESULT -INFOSOURCE -CORRELMSGTIME			
SYS-OUTAGECRITICALITY -SYSTEMAFFECTED -SUBSYSTEMAFFECTED -OUTAGEREASON -OUTAGETIME -ETARESTORESERVICE -STATUSSENSORCAPAB			

(FROM MESSAGES)

Row 1 ... 35

Column 1 ... 15

(FOR REPORTS)

FIGURE 1 The MWO "Information Processing Matrix" Model

Using this notation, we may write

$$D^k = \sum_a I_a^k \oplus \sum_b S_b^k \oplus A^k . \qquad (1)$$

In words, equation (1) says that the kth descriptor, D^k, is composed of #a instantaneous components, #b state components, and an algorithmic component, where '#' denotes 'number of elements of'. Note that equation (1) features two types of decomposition. First, there is the instantaneous/state type, wherein the I_a^k's and the A^ks are instantaneous components and the S_b^k's are the state components. Then there is the heuristic/algorithmic type decomposition, wherein the I_a^k's and the S_b^k's are the heuristic components and the A^ks are the algorithmic components.

We now wish to translate the expression in equation (1) to "matrix notation". In this regard, let R_i denote the ith row of the matrix (M_{MWO}) of Figure 1, and let C_j denote the jth column of (M_{MWO}). Equation (1) in matrix notation is

$$D^k = C_k = \sum_{R_a} I_a^k \oplus \sum_{R_b} S_b^k \oplus A^k, \qquad (1')$$

where R_a is a set of rows of (M_{MWO}), as is R_b.

We have now developed the genesis of a "descriptor language". Via the use of this language we shall be able to precisely formulate hypotheses relative to the structure of the MWO[13] decision making process. The key to the effectiveness of this language is its compactness and ease of specification. These beneficial attributes in turn were the result of our "*reduction*":

'decision-making' → 'information processing' → 'information processing matrix' .

Let us now become somewhat more specific. In particular let us give the form of each descriptor, expressed in this new language. First, we shall

---

13 The generic form of this language should extend the domain of applicability to an extremely wide class of decision making models.

for convenience replace $\oplus$ with + from now on.

$$C_1 = I_2^1 + A^1$$

$$C_2 = I_4^2 + I_5^2 + A^2$$

$$C_3 = I_4^3 + I_5^3 + A^3$$

$$C_4 = I_3^4 + A^4$$

$$C_5 = \sum_{k=1}^{7} I_k^5 + \sum_{k=22}^{28} S_k^5 + \sum_{k=29}^{35} S_k^5 + A^5$$

$$C_6 = S_2^6 + I_9^6 + A^6$$

$$C_7 = \sum_{k=1}^{7} S_k^7 + \sum_{k=8}^{14} I_k^7 + \sum_{k=22}^{28} S_k^7 + \sum_{k=29}^{35} S_k^7 + A^7 \quad \text{(GO)}^{14}$$

$$C_8 = I_{10}^8 + I_{11}^8 + I_{12}^8 + A^8$$

$$C_9 = \sum_{k=1}^{7} S_k^9 + \sum_{k=8}^{14} S_k^9 + \sum_{k=15}^{21} I_k^9 + \sum_{k=22}^{28} S_k^9 \quad \text{(GO)}$$
$$+ \sum_{k=29}^{35} S_k^9 + A^9$$

$$C_{10} = \sum_{k=1}^{7} S_k^{10} + \sum_{k=8}^{14} I_k^{10} + \sum_{k=22}^{28} S_k^{10} + \sum_{k=29}^{35} S_k^{10} + A^{10}$$

$$C_{11} = S_2^{11} + S_9^{11} + I_{16}^{11} + A^{11}$$

$$C_{12} = \sum_{k=1}^{7} S_k^{12} + \sum_{k=8}^{14} S_k^{12} + \sum_{k=15}^{21} I_k^{12} + \sum_{k=22}^{28} S_k^{12} \quad \text{(GO)}$$
$$+ \sum_{k=29}^{35} S_k^{12} + A^{12}$$

$$C_{13} = S_9^{13} + S_{10}^{13} + S_{11}^{13} + I_{16}^{13} + I_{18}^{19}$$
$$+ I_{19}^{13} + A^{13}$$

---

14 The goal-oriented descriptors are labeled (GO).

$$C_{14} = \sum_{k=15}^{20} I_k^{14} + A^{14}$$

$$C_{15} = \sum_{k=1}^{7} S_k^{15} + \sum_{k=8}^{14} S_k^{15} + \sum_{k=15}^{21} I_k^{15} + \sum_{k=22}^{28} S_k^{15} + \sum_{k=28}^{35} S_k^{15} + A^{15}$$

This particular description should be considered as tentative. Other components may need to be added and some of the terms included can be simplified. Also, one may wish to utilize C_5 as a "partial sufficient statistic" for C_{10}; and, likewise, (C_5,C_{10}) or C_{10} alone might be used to serve as partial sufficient statistics for C_{15}. That is, one may use $f_{10}(C_5)$ for the first term in the expression for C_{10}, and to either use $f_{15}(C_5,C_{10})$ or $g_{15}(C_{10})$ in place of the first two terms in the expression for C_{15}.

VII THE GOAL-ORIENTED DESCRIPTORS

Let us now consider in more detail the goal-oriented (GO) descriptors. First, C_7 (or D^7), ADS2-EVENTHYPOTH, has to do with classifying the "predicted result"[15]. That is, is this event to be considered "HOSTILE" or not? SIMCOPE partitions the complement of HOSTILE into three mutually exclusive states: TEST, UNRESOLVED, and UNKNOWN.
Descriptor D^9 (or C_9), ADS2-THREAT, is decomposable into "THREAT" and its complement at the ADS2 (i.e., at the "Pass 2") level; while Descriptor D^{12} (or C_{12}), BSS-THREAT, does the same thing, only at the BSS level.

The most important decision that the MWO must make is whether or not an event should be classified as a THREAT[16]. As the MWO performs his duties of filling out event reports he is continuously "tracking" the accumulating "evidence" of the various detections, always seeking to determine whether or not a THREAT exists. Figure 2 depicts the missile warning decision process from this "goal-oriented" point of view.

---

15 Usually, this coincides with the "intent" of THEM, hence the use of the SIMCOPE terminology "HOSTILE".

16 Note that, semantically, the SIMCOPE terminology 'THREAT' should be read as 'attack on US likely' or something to that effect.

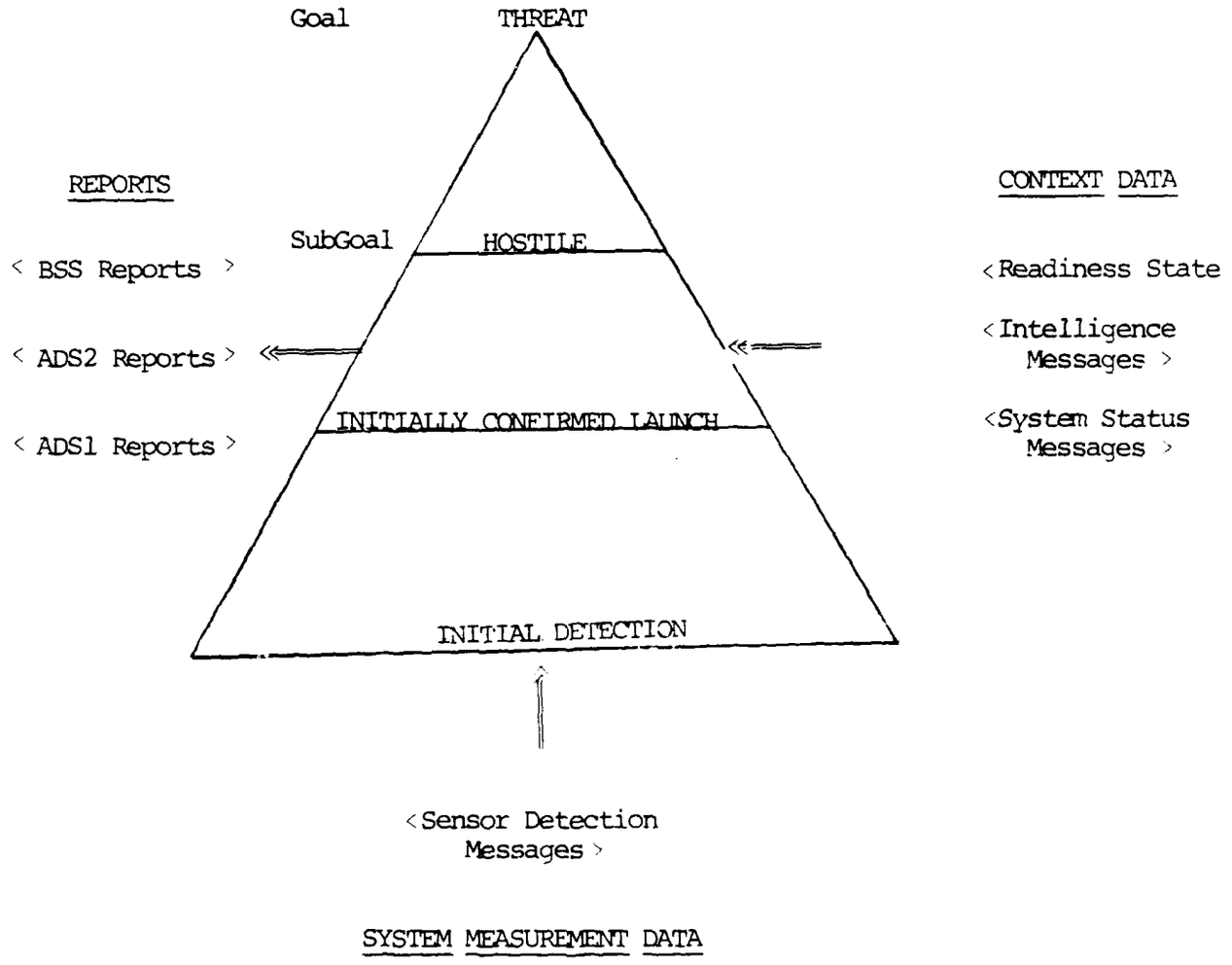


FIGURE 2 Goal-Oriented View of the Missile Warning Decision Process

It seems clear that the MWO is making a particular application of "opportunistic problem solving(OPS)" in that he is interleaving: (1) a data-driven forward-chaining operating mode when filling out the non-GO report slots, with (2) a goal-directed backward-chaining operating mode when filling the GO report slots. It also seems clear that the ES can do no less than this if it is to be useful in enhancing the MWO's situation awareness. In fact, it would seem that anything less than this would disorient the operator rather than aid him, thereby possibly leading the operator to completely reject the ES as a real-time decision aid.

VIII RECOMMENDATIONS

The research performed this summer has a number of natural extensions. Five of these will be given below[17]. However, they should not be thought of as being so easily separated as their distinct enumeration might suggest. In particular, in carrying out the research in any one of these five topics one would undoubtedly advance the research in the other four.

Follow-On Research Topics:

(R1) Generalize the concepts of 'descriptor' and 'information processing matrix' to their natural conclusion---a complete[18] formal descriptor language.

(R2) Utilize the language of (R1) together with knowledge engineering and/or cognitive experimentation to determine a particular implementation of equation (1) for the SMWP[19].

(R3) Develop the structure of an expert system of the opportunistic problem solving type that would be appropriate for use with single-level decision systems.

---

17 Among the natural extensions not explicitly included herein are: mutual (ES/Operator) situation awareness, expert system teams at both the local and global levels, etc., ---all the way to battle management problems.

18 Modulo Gödel of course.

19 SIMCOPE Missile Warning Problem.

(R4) Develop a generic top-down modular expert system structure that explicitly provides for the on-going modifications that necessarily accompany active research in intelligent decision-aiding.

(R5) Design an expert system specifically for the SMWP by particularizing parameters, functions, etc. within the generic structure of (R4).

I am very interested in this research and wish to pursue all five of these topics (and, later, the other natural extensions of this research). The most appropriate sequence for this pursuit would seem to be (R4) followed by (R3) and (R5). This would then provide the additional increment in informed intuition necessary to drive the abstraction to the form required by (R1). In turn, this would delineate the point at which the abstract model would require experimental data to fill in the knowledge gaps that will inevitably become apparent when applying this abstract model to a specific practical problem such as in (R2).

REFERENCES

1. Tacker, E. C., "Basic Research on the Analysis of Submarine Command and Control Systems: A Large-Scale Systems Approach", Naval Underwater Systems Center Contract, Final Report, 1983.

2. Tacker, E. C., M. T. Silvia, and F. C. Spicola, "Nonclassical Queueing: A New Language for the Modeling and Analysis of Command and Control Systems", Proc. IEEE Technology for an Efficient Tomorrow Conf., Houston, Texas, March, 1983.

3. Tobat, D. L., S. K. Rogers, and S. E. Cross, "SENTINEL: An Expert System Decision Aid for a Command, Control and Communication Operator", SPIE:ORLANDO, April/86, Vol. 635 Applications of Artificial Intelligence, Vol. 3, pp. 30-34.

4. Tacker, E. C., C. W. Sanders, and T. D. Linton, "Some Results in Decentralized Filtering and Control", International Federation of Automatic Control Symposium on Large-Scale Systems, Udine, Italy, 1976.

1986 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT SUMMER SUPPORT PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

Feasibility Investigation of Single-step Nitrations of Organometallics by Nitronium Triflate

Prepared by: Dr. Nicholas E. Takach

Academic Rank: Assistant Professor of Chemistry

Department and University: Chemistry Department, The University of Tulsa

Research Location: Air Force Rocket Propulsion Laboratory, Liquid Rocket Division, Chemistry and Materials Branch, Basic Chemical Research Section, Edwards AFB, CA

USAF Researcher: Dr. Robert D. Chapman

Date: September 30, 1986

Contract No: F49620-85-C-0013

FEASIBILITY INVESTIGATION OF SINGLE-STEP NITRATIONS OF ORGANOMETALLICS BY NITRONIUM TRIFLATE

by

Dr. Nicholas E. Takach

ABSTRACT

The attempted single-step nitration of n-butyllithium by nitronium tetrafluoroborate and by nitronium triflate is described. The use of ultrasonic energy to activate these heterogeneous systems is discussed. The desired product, 1-nitrobutane, could not be isolated in any of the reactions involving either nitronium salt, despite widespread variations in the reaction parameters. Instead, evidence exists that 1-nitrobutane reacted with additional n-butyllithium to form the nitronate salt, $CH_3CH_2CH_2CH{=}NO_2^-Li^+$. The results suggest that formation of the nitronate salt is faster than that of 1-nitrobutane. The attempted nitration of t-butyllithium by nitronium tetrafluoroborate also failed to yield the desired product. With t-butyllithium, nitronate salts cannot form but other side products are possible. The procedure used for preparing nitronium triflate was not reproducible due to the exothermicity of the reaction and the difficulty in trapping the intermediate product, N_2O_5. Suggestions are made for improvement of the synthesis.

Acknowledgements

The author wishes to express his gratitude to the Air Force Systems Command and the Air Force Office of Scientific Research for providing the opportunity to participate in the Summer Faculty Research Program. Also deserving thanks, for being the vehicle of the program, is Universal Energy Systems, particularly R. Darrah (Program Manager) and S. Espy (Program Administrator).

The author would like to thank the personnel at the research site. Two individuals provided invaluable assistance: Dr. R. Chapman, who initiated the project and supplied background information and preliminary results; 1st Lt. J. Andreshak, who helped me find my way around the lab and was always willing to give technical assistance.

The author also appreciated the comaraderie and help of the other scientists and support staff, including J. Nakamura (Branch Chief) and W. Roe (Effort Focal Point).

Special thanks are due to P. Jacobs (University of Tulsa) for finding time to type the manuscript on short notice.

Finally, I would like to thank my family for their willingness to relocate, especially my wife, Lilla, who put considerable effort into managing all aspects of the relocation.

I. INTRODUCTION

The author received his Ph.D. in Inorganic Chemistry, from the University of Nevada-Reno. His dissertation research involved the synthesis and spectroscopic, kinetic, and mechanistic studies of transition metal organometallic compounds. Upon completion of his graduate studies, the author accepted a postdoctoral fellowship whose specific objective was to participate in a research exchange program by spending one year in residence at the University of Chemical Industries in Veszprem, Hungary. This NSF-sponsored project allowed the author to gain experience in asymmetric synthesis using highly air-sensitive organometallic rhodium compounds as homogeneous catalysts.

The project on which the author was chosen to work at the AFRPL of Edwards AFB aimed at developing a new nonacidic nitration technique. The substrate organometallics and the nitrating agents used in this study are all very sensitive to moisture and require special handling under inert atmosphere.

II. OBJECTIVES OF THE RESEARCH EFFORT

The primary goal of the project was to develop a new nonacidic nitration technique for organometallics using nitronium triflate reagent, $NO_2^+OSO_2CF_3^-$ ("NO_2OTf"). The concept is simple: the metal atom of model organometallics is displaced by NO_2 to produce nitro-organics.

It was first proposed by Scott Shackelford (Major, USAF) when he set up the Basic Chemical Research Section at the Rocket Propulsion Lab of Edwards AFB, as part of a broader program to study the applicability of triflate chemistry toward developing useful chemical transformations. Addition of a nitro substituent to unsaturated compounds using NO_2OTf and nitronium salts in general has been reported[1]; however, reactions in which a metal is displaced by NO_2 in a single step, as described above, have not yet been accomplished.

The preliminary objective was to evaluate the general feasibility of such a transformation with model compounds representing three categories of organomecallics: nitroaliphatics, nitroalkenes, and nitroaromatics. Initial experiments in each category were to use nitronium tetrafluoroborate (NTFB) as the nitrating agent because it is commercially available while NO_2OTf must be synthesized. Although studies of comparative nitrating ability indicate that NO_2OTf is superior to NTFB and other intronium salts,[1,2] it is generally accepted that the attacking species in any case is NO_2^+, the nitronium ion.[2]

During the course of the project, it became necessary to prepare more NO_2OTF. Attempts by the author to reproduce a synthesis reported by Dr. Robert D. Chapman (RDC), Research Chemist at Edwards AFB, were unsuccessful. When subsequent attempts by RDC were also unsuccessful, it was concluded that a necessary additional goal of the project was to re-evaluate and modify the synthesis until it was reproducible.

III. NITRATION STUDIES

A. Preliminary Experiments/Background

The first attempt to nitrate an organometallic compound using NO_2OTf was reported by RDC in February, 1985.[3] The specific system and reaction conditions are summarized below:

$$NO_2OTf + CH_3CH_2CH_2CH_2Li \xrightarrow[\text{ultrasound}]{\text{hexane, } -80°C} CH_3CH_2CH_2CH_2NO_2 \qquad (1)$$

The low temperature was presumably a precaution against an unknown rate of reaction. The choices of solvent were limited. Nitronium salts not only have very low solubility in organic solvents, they also react with many of them.

For instance, NO_2OTf is stable but insoluble in THF at temperatures just above the latter's melting point, then decomposes when the THF is warmed to room temperature.[4] Because of their inertness aliphatic hydrocarbons have frequently been used as the solvent for nitrations by nitronium salts. Hexane was convenient for this case because $CH_3CH_2CH_2CH_2Li$ ("n-BuLi") is sold as a hexane solution.

Because of the heterogeneity of the system, the reaction vessel was immersed in an ultrasonic cleaning bath. The promotion of chemical processes is an application of ultrasonic energy that is still in its infancy and one whose mechanism is not completely known. It seems to involve a process called cavitation. This phenomenon occurs when the alternate expansion and compression waves of the sound in a solvent cause minute vapor-filled bubbles to form and then collapse. The collapsing bubbles generate powerful shock waves which give rise to temperatures on the order of several thousand °C and pressures up to 10^4 atm, but only for a few nanoseconds. Thus, even though the solution may be at room temperature and pressure, the net effect, on a microscopic scale, is like running a reaction in a high-temperature autoclave.[5] Pioneers in the field of "sonochemistry" believe that heterogeneous reactions will greatly benefit from sonication.[5,6]

The product of RDC's first reaction was not the desired one. Rather, its ^{13}C nmr spectrum was consistent with that of a nitronate salt. This was not surprising inasmuch as nitronate salts readily form when nitroalkanes are in the presence of strong bases.[7] In this particular case the reaction is

$$CH_3CH_2CH_2CH_2NO_2 + CH_3CH_2CH_2CH_2Li \longrightarrow CH_3CH_2CH_2CH{=}NO_2^-LI^+ + CH_3CH_2CH_2CH_3 \quad (2)$$

Nitrobutane can be regenerated by treating the above nitronate salt with dilute acid[7] but, because other nitronate systems may be acid-sensitive, it was desirable to devise a technique which avoided the need for acidification. For this reason, RDC had a special vessel prepared (by Jurgen Linke of Jet Propulsion Laboratory in Pasadena, Ca.). The vessel simulates a flow system in which the nitrobutane product is removed as n-BuLi solution is slowly added (schematic drawing and experimental details in next section). Using the new vessel ("u/s vessel"), the reaction represented by Equation (1) was repeated under identical conditions, except that NTFB was the nitrating agent instead of NO_2OTf. With the failure to isolate a reaction product of any type, it was concluded that the reaction was not fast enough at low temperature. Repeating the reaction again, this time at room temperature, surprisingly yielded a product that was consistent with a nitroalkene, according to its color, odor and ^1H and ^{13}C nmr spectra. This highly unexpected outcome might have resulted from dinitration (effected by NO_2^+), with subsequent dehydronitrosation ($-HNO_2$) induced by the strongly basic butyl anion. Indeed, there is precedent for the interaction of nitronium salts with the heteropolar C=N double bond in nitroalkane salts to form dinitrocompounds.[8] Because the use of ultrasound to carry out nitrations heterogeneously in the new u/s vessel appeared to have potential, a more powerful ultrasonic processor was ordered (Heat Systems-Ultrasonics W-225). Shortly after the processor was delivered, the author of this report arrived at the research site to join the project.

B. Results and Discussion

The author elected to continue attempts to nitrate n-butyllithium. To conserve the dwindling supply of NO_2OTf, NTFB was used as the nitrating agent. The reaction was carried out in the custom-made u/s vessel which looks approximately as shown in Figure 1 on the following page.

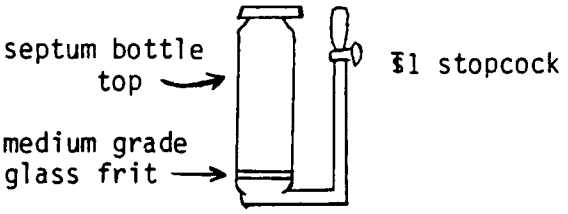


Figure 1: Ultrasound ("u/s") Reaction Vessel

The vessel was designed to fit inside the Heat Systems--Ultrasonics Sonicator cup horn. The components of the sonication setup are shown in Figure 2 and a blow-up of the cup horn is shown in Figure 3. (both on following page).

Since it is unlikely that the reader is familiar with the use of high-power sonication equipment, a detailed description of a typical nitration experiment, from tuning of the ultrasonic generator to work-up of the reaction product, will be given. The description will also explain the use of the u/s vessel and provide a context for discussion of modifications used in subsequent experiments.

NTFB (1.014g. 7.6mmol) was weighed into the u/s vessel in a glove box. The vessel was then filled with n-hexane and capped with a teflon-lined rubber septum encased in aluminum, using a special capping tool. Aldrich 1.58M n-butyllithium in hexanes (4.7mL, 7.5mmol) was filtered through an FEP Millipore filter in a Swinny Adapter into a 30-mL serum bottle, diluted with 15-mL n-hexane and capped. The u/s vessel and serum bottle were left in the glove box until the sonicator was ready to be used--i.e., it was tuned.

The cup horn was installed on the convertor and then tuned according to the procedures in the Cup Horn Instructions (supplemental to the applications information contained in the Sonicator Series Application Notes AN-6). After the tuning procedure, the u/s vessel was centered about 3-4mm above the radiating surface of the horn. It is vibration of this surface,

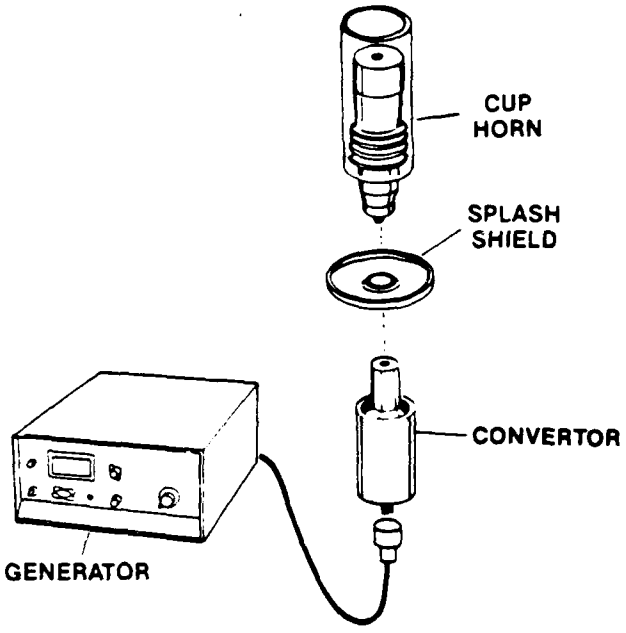


Figure 2. Equipment Used for Ultrasonic Reactions

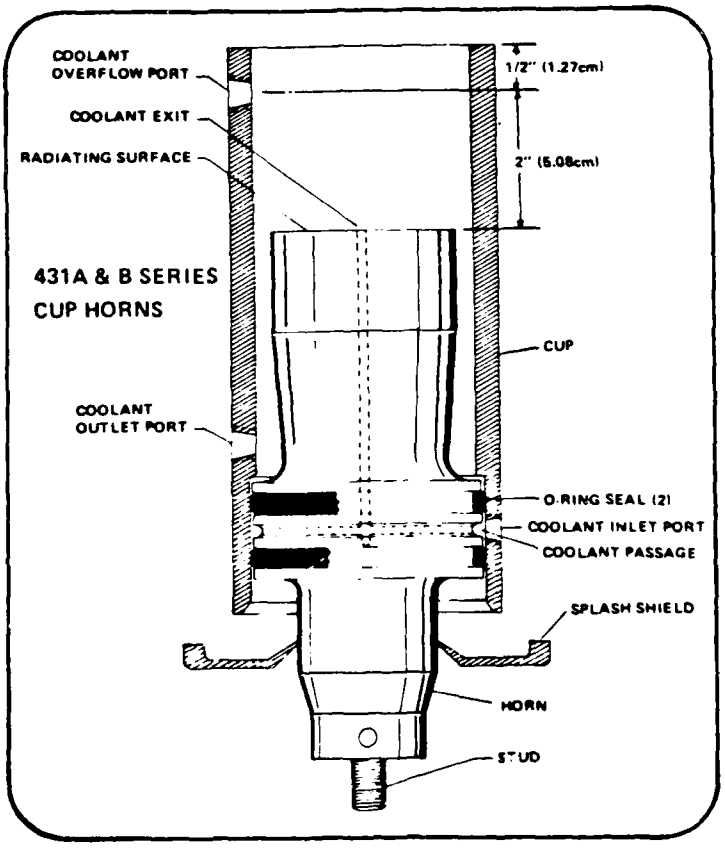


Figure 3. Expanded View of Sonicator Cup Horn

driven by a lead zirconate titanate transducer and producing sound waves in the 20-kHz frequency region, that causes cavitation of the n-hexane inside the u/s vessel. The side arm of the vessel was fitted with teflon tubing that lead to a 100-mL, two-neck, round-bottom receiving flask. With the generator output set to the manufacturer's recommended level (depends on the type of coolant used in the cup--in this experiement it was ice water) dropwise injection of the n-BuLi solution by a syringe needle was started. As the solution was injected, liquid pressure inside the vessel forced liquid through the frit, up the side arm, through the open stopcock, and eventually into the receiving flask. Injection of the n-BuLi solution took about 20 min. Then pure n-hexane was injected to force the remaining reaction mixture through the frit. Before this was accomplished, the frit became clogged and no more liquid could pass through. Rotoevaporation of the yellow liquid in the receiving flask resulted in a light brown powder that was insoluble in D_2O as well as several organic solvents. After some of the solid finally dissolved in a mixture of C_6D_{12}, $CDCl_3$ and D_6-DMSO, an ^1H nmr spectrum was taken. The only peaks were broad resonances in the upfield region, suggesting the presence of an amorphous polymeric substance, congruent with the product's insolubility.

The attempted nitration of n-BuLi by NTFB was repeated several times. Modification of the reaction parameters fell into five major categories: 1. temperature; 2. flow-conditions (i.e., increasing the rate of flow through the u/s vessel by either exerting a negative pressure on the receiving vessel or a positive pressure of N_2 gas on the u/s vessel); 3. nature of the solvent; 4. sonication time (in some experiments, the n-BuLi solution was divided into aliquots that were subjected to sonication for varying amounts of time); 5. sonication power. The results of all the above modifications were the same: failure to obtain the desired product. Also, at some stage of every experiment, clogging of the frit halted passage of liquid from the u/s vessel to the receiving vessel.

At this point it was thought that an impurity in the NTFB might be related to the plugging of the frit and, more importantly, to the inability to obtain the desired product. Contributing to this suspicion was the fact that even the purest grade of NTFB commercially available (97%) contains inherent impurities. For this reason it was decided to commence the experiments in which NO_2OTf was the nitrating agent. Although this compound, prepared by RDC, was not subjected to chemical analysis, it was assumed to be purer than NTFB, largely on the basis of close agreement between its melting point and the published value.[9]

For comparison purposes, the organolithium chosen for initial studies was once again n-BuLi. Several experiments were performed employing essentially the same conditions and subsequent modifications that were described in reactions using NTFB. The results were also familiar: there was no evidence that the desired product had formed in any of the experiments and clogging of the frit occurred each time during the stage when pure solvent was injected to force remaining reaction mixture through the vessel.

The overall conclusion from the attempted nitrations of n-BuLi is that inability to isolate the desired product, 1-nitrobutane, appears to be a problem of relative rates. That is, reaction of 1-nitrobutane with additional n-BuLi to form the nitronate salt occurs faster than the reaction by which 1-nitrobutane is initially formed. To support this conclusion, the solid that remained above the frit of the u/s vessel was analyzed after one of the attempted nitrations using NO_2OTf. The ^1H-nmr spectrum of this solid did in fact agree with an earlier spectrum obtained by RDC of the lithium nitronate salt derived from 1-nitrobutane. It is also reasonable to conclude that as the amount of nitronate salt increased, it progressively clogged the pores of the frit until eventually liquid was unable to pass through.

At this point, approximately half the summer research period was over and the supply of NO_2OTf was exhausted. After the many futile attempts to nitrate n-BuLi, it was decided to try another substrate. Of particular interest was t-butyllithium ("t-BuLi"). Because this compound has no alpha hydrogens (i.e., alpha to the lithium), nitronate salt formation is not possible. Thus, the feasibility of the direct transformation of an alkyllithium to a nitroalkane could be tested. Although the u/s vessel offered no advantage, it was used for convenience (i.e., the frit allowed a built-in mode for filtering the reaction mixture). The reaction conditions were similar to those of earlier experiments, with one main exception: because there was no known need for a flow system, the u/s vessel was not filled to the top with pure solvent. A minor variation was the nature of the solvent. Pentane was used because t-BuLi is commercially available as a pentane solution. An approximately equimolar mixture of t-BuLi and NTFB (NO_2OTF was temporarily unavailable) was sonicated at 0°C for approximately 30 minutes. Then the reaction mixture was forced through the frit and into the receiving flask by applying a positive pressure of N_2 gas on the u/s vessel. Work-up of the reaction product yielded the residue of t-BuLi; in other words, no reaction occurred.

The reaction was repeated with one major modification: the "forced-diffusion" flow system (i.e., forcing of solution through the frit of the u/s vessel due to liquid pressure) was again used. Even though nitronate salt formation cannot occur, another secondary reaction is conceivable:

$$(CH_3)_3C\text{-}NO_2 + (CH_3)_3C\text{-}Li \xrightarrow{-HNO_2} (CH_3)_2C{=}CH_2 + LiNO_2 + (CH_3)_3C\text{-}H \quad (3)$$

Thus, it may once again be desirable to remove product ($(CH_3)_3C\text{-}NO_2$) from the vicinity of t-BuLi to prevent further reaction. Surprisingly, the frit of the u/s vessel again became blocked during the injection of pure solvent (to force the reaction solution through). However, the blocking was only partial and, with extra effort, liquid could be forced through the frit. It is presumed that in this case the clogging was due largely, if not wholely, to a salt by-product, such as $LiNO_2$. Unfortunately, workup of the reaction product provided no evidence that anything other than unreacted t-BuLi made it into the receiving flask.

IV. SYNTHESIS OF NO_2OTf

A. Background/Preliminary Results

As mentioned in the preceding section the small available supply of NO_2OTf was used up at the midpoint of the author's summer research period. Even though a preparation of NO_2OTf was reported as early as 1973[2], its synthesis is nontrivial. Following a wide variety of reported methods and modifications thereof, RDC attempted to prepare the compound 13 different times and on each occasion found the results unsatisfactory for one reason or another.[10] Nonetheless, the fourteenth attempt produced essentially a quantitative yield of pure material and, as such, no difficulties were anticipated.

The experimental setup of the successful method was designed by RDC to allow two reported procedures[11,12] to be carried out in sequence. With RDC's method, N_2O_5 was first generated by the dehydration of 100% nitric acid using P_2O_5 under an ozone atmosphere. The N_2O_5 vapor thus generated was carried by the ozone stream into another flask (connected by teflon tubing) where it was trapped as it bubbled through 1,2-dichloroethane ("$EtCl_2$") that was cooled to -30°. The flask containing the solution of N_2O_5 in $EtCl_2$ was then transferred to the glove box where the dropwise addition of triflic acid ("HOTf") produced NO_2OTf and HNO_3. The overall

reaction is summarized by Equation (4).

$$P_2O_5 + HNO_3 \xrightarrow{O_3} N_2O_5 + HOTf \xrightarrow{EtCl_2} NO_2OTF + HNO_3 \qquad (4)$$

B. Results/Discussion

The author found the generation of N_2O_5 to be a very difficult process to control due to its highly exothermic nature and, consequently, its unpredictability. To "tame" the reaction, the procedure calls for first freezing the HNO_3 with liquid nitrogen. Then the frozen mass is covered with P_2O_5 powder and allowed to thaw at a natural rate, all the while passing an atmosphere consisting of 6% (w/w) ozone (in oxygen) through the flask. As the HNO_3 melts and reacts with P_2O_5, a tremendous amount of heat is liberated. If the reaction becomes too rapid, copious amounts of NO_2 and some N_2O_4 are formed and quickly swept away into the receiving flask by the O_3/O_2 flow before the ozone has a chance to oxidize the lower oxides of nitrogen to N_2O_5. The presence of lower oxides in the receiving flask dooms the reaction because they lead to the eventual formation of nitrosonium triflate, NO^+OTf^-, which cannot be separated from NO_2OTf.[9]
In the author's first three attempts to repeat RDC's synthesis, the reaction between P_2O_5 and HNO_3 went out of control and, it was subsequently determined, produced the undesirable mixture of nitrosonium triflate and nitronium triflate. The unpredictability of this step is illustrated by the fact that in the third attempt, the P_2O_5 sat on top of the frozen HNO_3 for one hour before the latter started to melt, but in the second attempt the reaction commenced within two minutes.

In the author's fourth attempt to prepare NO_2OTf, the other major difficulty associated with this method was discovered. Namely, despite the large stoichiometric excess of HNO_3 relative to NOTf, there is no way of predicting just how much N_2O_5 actually gets trapped in the receiving flask. There are two main reasons for this. First, N_2O_5 has limited solubility in $EtCl_2$; on occasion it was possible to actually see N_2O_5 vapor exiting the bubbler that acted as a vent for the receiving flask. Second, it is not possible to know how much of the HNO_3 is actually converted to N_2O_5 or when the conversion is complete. If <u>less</u> than a stoichiometric amount of N_2O_5 is available for reaction with HOTf, the excess HOTf can react with the HNO_3 byproduct of Equation (4) to form an inseparable mixture of nitronium triflate and hydronium triflate, according to Equation (5)[2]

$$2HOTf + HNO_3 \longrightarrow NO_2^+OTf^- + HNO_3OTf^- \qquad (5)$$

After the author's fourth unsuccessful attempt, RDC tried to repeat his earlier successful preparation of NO_2OTf. The reaction between HNO_3 and P_2O_5 seemed to be going smoothly until RDC stirred the mixture. Suddenly, the reaction became very vigorous and the mixture started to boil, simultaneously producing orange "clouds" of NO_2 that were quickly swept into the receiving flask. At that point the reaction was terminated.

The author and RDC worked together on the next attempt and a method for controlling the dehydration of HNO_3 seemed to be found. The teflon blades of a mechanical stirring shaft were immersed into the HNO_3 before it was frozen. As the HNO_3 began to melt and react with P_2O_5, the shaft was carefully pulled out of the HNO_3. The resulting cavity in the mostly frozen HNO_3 was filled with P_2O_5 when the sides of flask were tapped. This technique accomplished two purposes. First, it provided a means of slowly mixing the reactants. Second, the reacting surface of HNO_3 was smaller and thus easier to control. If orange vapors indicative of NO_2 formation were observed, the flow rate of the ozonator was reduced to give the ozone time to oxidize the NO_2 to N_2O_5 before the former was swept out of the flask. To provide a richer oxidizing atmosphere, the ozonator settings were also changed to produce the highest concentration of ozone possible.

Despite these improvements, difficulties were still encountered in the second stage of this attempt and in the subsequent two attempts, which were once again performed by the author alone. However, the experience gained in the last few attempts provided the framework for recommendations found in the next section.

V. RECOMMENDATIONS

A. Nitration studies

The results obtained by the author indicate that at least for n-BuLi and t-BuLi, it might not be possible to obtain the corresponding nitro compounds in a single step. Reports in the literature suggest that this may also be true for model organometallics representing the other categories of interest listed in the objectives. For instance, the nitration of vinyllithium, the simplest alkenyllithium, may be complicated by the known fascile polymerization of nitroethylene.[12] Cyclohexenyllithium should be less susceptible to polymerization because of steric factors, but is also known to form nitronate salts[13]. If phenyllithium can be converted to nitrobenzene, the strongly electron-withdrawing nitro substituent of the latter might result in the formation of an adduct (called a "Meisenheimer complex") between it and another phenyl anion.[14]

Even if undesirable reactions such as those mentioned might occur, the objectives listed in Part III of this report should be completed. And if, for instance, nitronate salts do form in a system, it would still be of interest to develop methods for their fascile conversion to the corresponding nitro-organic. It would also be interesting from a basic research standpoint to see if competing reactions could be avoided by modifications in the chemical structures of the substrates. For example, could 2,2,4,4-tetramethylcyclohexenyllithium (in which the nitronate salt should not form) be converted directly to the corresponding nitro compound by NO_2OTf?

Finally more work is needed to investigate the effects of varying the ultrasonic parameters. Does ultrasound really promote the types of nitrations that are the focus of this project, does it hinder them, or is it even necessary at all?

If the synthesis of more NO_2OTf had been successful, the author would have attempted the nitration of n-BuLi by the former <u>without</u> the use of ultrasound. According to Coon, et al., the reaction of NO_2OTf with aromatic hydrocarbons is one of the few examples in which a reagent that is insoluble in the reaction medium is used to effect nitration.[2]

B. Synthesis of NO_2OTf

With the problems associated with N_2O_5 generation reasonably well-resolved, the author feels the key to making the synthesis of NO_2OTf reproducible lies in ensuring that enough N_2O_5 is available in the $EtCl_2$ to react with all the added HOTf. Using an even larger excess of HNO_3 is a possibility, but one that is wasteful. It would be preferable to improve the efficiency of the method used to trap the N_2O_5. A method worth trying might be to have two or more trapping flasks containing $EtCl_2$ connected in series. After one is confident that N_2O_5 generation is complete, the contents of the trapping flasks could be combined and reduced in volume before adding HOTf.

REFERENCES

1. Guk, Yu. V., M. A. Ilyushin, E. L. Golod, and B. V. Gidaspov, "Nitronium Salts in Organic Chemistry," Russ. Chem. Rev., 52, 1983, 284.

2. Coon, C. L., W. G. Blucher, and M. E. Hill, "Aromatic Nitration with Nitric Acid and Trifluoromethane Sulfonic Acid," J. Org. Chem., 38, 1973, 4243.

3. Chapman, R.D., AFRPL/LKLR, Edwards AFB, CA, "New Synthetic Techniques for Advanced Propellant Ingredients -- 2303M1RD," R & D Record Book, Vol. IX, 1984.

4. Chapman, R. D., R & D Record Book, Vo. XII, 1986.

5. Graff, G. M., "Putting Chemical Reactions on a Sound Footing," Chemical Engineering, March, 1985, 22.

6. Han, B. H., and P. Boudjouk, "Organic Sonochemistry. Ultrasonic Acceleration of the Reduction of Simple and Deactivated Aryl Halides Using Lithium Aluminum Hydride," Tetrahedron Letters, 23, 1982, 1643.

7. Goldwhite, H., Rodd's Chemistry of Carbon Compounds, 2nd Ed., Edited by S. Coffey, Chapter 6, Amsterdam, Elsevier Publishing Co., 1964, 93.

8. Olsen, R. E., D. W. Fish, and E. E. Hamel, "Modern Chemistry of Rocket Fuels," Adv. Chem. Ser., 54, 1965, 48.

9. Chapman, R.D., AFRPL, Edwards AFB, CA. private communication.

10. Gruenhut, N. S., M. Goldfrank, M. L. Cushing, and G. V. Caesar, "Preparation of Nitrogen(V) Oxide, Inorg. Syn., 3, New York, McGraw Hill, 78.

11. Effenberger, F., and J. Geke," Aromatic Electrophillic Substitution. II Preparation of Nitronium Trifluoromethane Sulfonate. Comparing Nitrification with Nitronium Salts," Synthesis, 1, 1975, 40.

12. Jagur-Gradzinski, J., "Nitroethylene Polymers, "Encyclopedia of Polymer Science Technology, Vol. 9, New York, John Wiley and Sons, Inc., 1968, 315.

13. Nielsen, A. T., "The Isomeric Dinitrocyclohexenes. III. Chemistry of the Nitronate Ions," J. Org. Chem., 27, 1962, 2001.

14. Carey, F. A., and R. J. Sundberg, Advanced Organic Chemistry, Part B: Reactions and Synthesis, Chapter 7, New York, Plenum Press, 1977, 280.

1986 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT SUMMER SUPPORT PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

UNIVERSAL ENERGY SYSTEMS, INC.

FINAL REPORT

DISSIPATION OF PLASMA CLOUD GENERATED BY THIRD STAGE ROCKET SEPARATION

Prepared by:	Arjun Tan
Academic Rank:	Associate Professor
Department and	Department of Physics
University:	Alabama A & M University
Research Location:	Arnold Engineering Development Center
	Arnold Air Force Station, TN 37389
USAF Researchers:	Dr. W. K. McGregor and Dr. C. C. Limbaugh
Date:	July 31, 1986.
Contract No.:	F49620-85-C-0013

DISSIPATION OF PLASMA CLOUD GENERATED BY THIRD STAGE ROCKET SEPARATION

by

Arjun Tan

ABSTRACT

The plasma cloud generated by third stage rocket separation by means of explosive bolt pyrotechnic causes temporary disruption in the radar signals tracking the space vehicle and the burnt-up third stage. This study describes a theoretical investigation of the dissipation of the plasma cloud thus generated. It is estimated that around 10^{24} positive and negative ion pairs are generated in a typical third stage rocket separation using the common explosive of gunpowder. The dissipation of plasma due to diffusion and recombination separately are studied. It is shown that when loss processes are neglected, the plasma disperses in a Gaussian mode. The evolution of the cloud in the first 10 seconds after explosion is determined for both isotropic and anisotropic media and by using Sutton's formula. A notable feature of the anisotropic diffusion is the elongation of the cloud in the direction of the larger diffusion coefficient. The effects of recombination processes are studied for the quadratic, linear and altitude-dependent loss rates. For quadratic loss, the ion density decreases dramatically in the first second, but then declines more gradually to levels depending solely on the loss coefficient. The destruction of ions is exponential and generally more severe in the linear case, but the loss rate decreases significantly if the altitude-dependence is taken into consideration. It is suggested that the dissipation of the plasma due to the combined effects of diffusion and chemical loss be investigated as a follow-up study.

1. INTRODUCTION

In the launch of intercontinental ballistic missiles, the burnt-up rockets are usually separated from the rest of the vehicle by means of explosive bolt pyrotechnics. This explosion generates a plasma cloud consisting of positive and negative ions. For the explosive gunpowder (KNO_3), the positive ions are most likely to be K^+, produced by thermal ionization of potassium, which has a relatively low ionization potential. In the lower atmosphere, the negative ions are mostly molecular ions (denoted by X^-), since the electrons stripped from potassium would quickly attach themselves to neutral molecules. The high-density ion cloud thus produced can cause temporary disruption of radar signals tracking the vehicle and the burnt-up rocket. The problem becomes most acute in the third stage rocket separation because of the following reasons: (1) the target is smaller; (2) the distance is greater; and (3) the recombination rate is slower at higher altitudes.

2. OBJECTIVES

In this research project, we study the dispersion and dissipation of the plasma cloud generated by the third stage rocket separation by means of explosive bolts. In particular, we investigate the following:
(1) How does the plasma cloud dissipate seconds after the explosion;
(2) How does the peak ion density vary with time; and
(3) What role does the recombination process play in the decay of the cloud following the explosion.

3. BACKGROUND MATERIAL RELEVANT TO THIS STUDY

3.1. Trajectory Parameters and Coordinate system

The trajectory parameters of a space vehicle can be determined from the universal trajectory curve (see, for example, Haviland and House, 1965). For a typical ICBM of 5000 nautical mile range, the velocity V and flight path angle θ at burnout are respectively 7 km/s and 25° approximately, the former being 88.5% of the minimum circular orbit velocity (cf. Haviland

and House, 1965). In 10 seconds following the burnout, the re-entry vehicle travels about 70 km, which translates to a 19.58 km ascent and 63.44 km horizontal traverse. For such distances, the flight path and the arc on the earth's surface can be closely approximated by straight lines (cf. Haviland and House, 1965). The flight path trajectory up to 10 seconds after the third stage separation is shown in Fig.1. The separation is assumed to take place at 60 km altitude.

The separation usually causes some additional boost to the re-entry vehicle. However, it is a reasonable assumption that the cloud produced by the explosive fragments travel with the speed of the vehicle before the separation, viz., 7 km/s. For this study, we choose our coordinate system to move with the explosive cloud with its x-coordinate in the direction of motion and the y-coordinate in the vertical plane (Fig.1).

3.2. The Atmospheric Environment

The region of the atmosphere from 50 km to about 85 km, where the temperature decreases approximately linearly with height, is called the 'mesosphere'. The coldest temperature in the atmosphere is reached at the 'mesopause', where it is nearly constant from 85 km to 90 km. The mesosphere and mesopause are parts of the 'homosphere' (0 km - 100 km), where the atmosphere is well-mixed by turbulence and the mixing ratios of the atmospheric constituents are fixed. The number density and pressure at various altitudes are readily available from the U.S. Standard Atmosphere (1976).

3.3. The Background Ionosphere

Overlapping the mesosphere and the mesopause is the D region ionosphere (50 km - 90 km). Notwithstanding the Sporadic C, this is the lowest region of the ionosphere and is believed to be produced by solar and galactic cosmic rays, X rays and meteor showers. It is mostly composed of negative molecular ions and positive (including metallic) ions. The ion density generally increases with height, attaining a value of about 10^3 cm^{-3} at 90 km during daytime. Owing to the fast recombination rates

of molecular ions, the D region is severely depleted at night. The chemistry of the D region is extremely complex and still not fully understood (cf. Banks and Kockarts, 1973).

3.4. Estimation of the number of ions produced

An estimated 1 to 2 lbs of explosives are required to explode the 16 bolts normally used in the third stage rocket separation. Assuming potassium nitrate (molecular weight 101 amu) to be the major explosive, the number of molecules in 1 lb of KNO_3 is $N_A \times 453.6/101 = 2.69 \times 10^{24}$, N_A being Avogadro's number. Assuming a 100% dissociation of KNO_3 and a 100% ionization of K, we estimate the number of K^+ ions produced using 1 lb of explosive to be 2.69×10^{24}. In the atmospheric heights under consideration, the electrons stripped from the potassium atoms would quickly attach themselves to neutral molecules to produce an equal number of negative ions. In this study, the number of positive and negative ions produced by the explosion is assumed to be 10^{24}.

3.5. Eddy Diffusion Coefficient

Although molecular and thermal diffusion coefficients can be estimated from the kinetic theory, the eddy diffusion coefficient is a much more speculative parameter. For example, it is still not clear what values should be used in the lower atmosphere of the earth although values between 10^6 and 10^7 cm^2/s are possible (Banks and Kockarts, 1973). It is suggested that in the lower atmospheres of Jupiter, Venus and Mars, large values of the eddy diffusion coefficient of the order of 10^8 cm^2/s are possible (Hunten, 1976). The eddy diffusion coefficient represents, to a certain extent, the degree of turbulence in the atmosphere and is responsible for the uniform mixing ratios of the constituents in the homosphere. It must be assigned high values for our study, where besides the shock wave generated by the motion of the vehicle, additional blast wave is created by the explosive charges. In this report, we have used the value of 2×10^8 cm^2/s for the isotropic case and values of 2×10^8 and 2×10^9 cm^2/s for the anisotropic case.

3.6. Loss Processes

It was stated before that the ion chemistry in the D region is extremely complicated. For this study, the loss of ions is most likely to proceed via the following reactions:

$$K^+ + X^- \rightarrow K + X \,, \qquad (3.1)$$

and

$$K^+ + M \rightarrow K + M^+ \,, \qquad (3.2)$$

followed by

$$M^+ + X^- \rightarrow M + X \,. \qquad (3.3)$$

where M stands for the neutral molecule. The charge neutrality condition implies that the rate of loss of ions due to Eq. (3.1) is proportional to the square of the number density of ions (quadratic loss). If reaction (3.3) is much faster than reaction (3.2), then the loss rate is linear. The reaction rate of (3.1) is customarily taken as 1×10^{-8} cm^3/s (cf. Kvifte, 1973), although the rate of similar reactions usually decrease with temperature (cf. Jensen and Jones, 1978).

4. DISPERSION OF IONS DUE TO DIFFUSION

Once produced by explosion, the positive and negative ions would rapidly disperse outwards. In this section, the ions are treated as indestructible particles, i.e., their loss processes are disregarded. This could be a fair approximation if the recombination rates are much slower compared with the time of observation as could be expected in the upper ionospheric heights. However, it will be clear in the following section that this is actually a poor approximation for mesospheric heights. Nevertheless, this approximation is useful in the simplification of the problem and for studying the general trends of the dispersion.

The second assumption is the instantaneous point source approximation. Assuming the explosion to take place in a matter of microseconds, the distance traversed in that time is of the order of a meter, which is comparable to the diameter of the vehicle. Considering the large distance of observation (over 60 km), the instantaneous point source approximation is deemed a fair one.

4.1. Isotropic Diffusion

In the first instance, we assume the diffusion coefficient to be homogeneous and isotropic having the value $K_x = K_y = K_z = K = 2 \times 10^8$ cm^2/s. In that case, the diffusion equation in a frame of reference moving with the source has the Fickian form

$$\frac{DN}{Dt} = K \left(\frac{\partial^2 N}{\partial x^2} + \frac{\partial^2 N}{\partial y^2} + \frac{\partial^2 N}{\partial z^2} \right) . \qquad (4.1)$$

Here N is the number density of the diffusing material (ions) and the l.h.s. represents the rate of change of N is the moving coordinate system. The solution of Eq.(4.1) has the 3-dimensional Gaussian form (cf. Pasquill, 1974):

$$N(r,t) = \frac{Q}{(2\pi)^{3/2} \sigma^3} e^{-\frac{r^2}{2\sigma^2}} , \qquad (4.2)$$

where r is the radial distance, Q the total number of particles (10^{24}) and $\sigma = \sqrt{2Kt}$ is the standard deviation. According to Eq.(4.2), the number density falls off as $t^{-3/2}$. The standard deviation, i.e., the radius of a sphere containing 68.3% of the ionization varies as the square root of the time after explosion. Figure 2 shows the outlines of the spheres having diameters 2σ (i.e., containing 95.4% of the ions), every second, up to 10 seconds faollowing the explosion. The radial distribution of the ions after 1, 5 and 10 seconds of explosion are shown in Fig.3. Contour plots of the ion densities from Fig.3 are drawn in Fig.4.

4.2. Anisotropic Diffusion

In the second instance, we assume the diffusion coefficient to be anisotropic. Since more turbulence is generated along the path of the vehicle, we choose the values $K_x = 2 \times 10^9$ cm^2/s and $K_y = K_z = 2 \times 10^8$ cm^2/s. Equation (4.1) then takes the form

$$\frac{DN}{Dt} = K_x \frac{\partial^2 N}{\partial x^2} + K_y \frac{\partial^2 N}{\partial y^2} + K_z \frac{\partial^2 N}{\partial z^2} . \qquad (4.3)$$

The solution of Eq.(4.3) is found to be (cf. Giere, 1977):

$$N(x,y,z,t) = \frac{Q}{(2\pi)^{3/2}\sigma_x\sigma_y\sigma_z} e^{-\frac{x^2}{2\sigma_x^2}-\frac{y^2}{2\sigma_y^2}-\frac{z^2}{2\sigma_z^2}}, \qquad (4.4)$$

where $\sigma_x = \sqrt{2K_x t}$ etc. are the respective standard deviations along x, y and z directions. Figure 5 depicts the outlines of the cloud containing 95.4% of the ionization after 1 to 10 seconds of explosion. The prime feature of the anisotropic diffusion is the elongation of the cloud along the x-direction by a factor of $\sigma_x/\sigma_y = \sqrt{K_x/K_y}$.

4.3. Diffusion Based Upon Sutton's Formula

According to Sutton (1953), a reasonably accurate description of diffusion in the atmosphere for distances up to hundreds of kilometers is

$$\sigma^2 = k(\bar{u}T)^{7/4}, \qquad (4.5)$$

where $\bar{u}$ is the average velocity in the x-direction, T the time taken by the particles to travel the standard deviation and k is a constant. If we equate T to t, the dispersion of ions in our problem is faster than that obtained earlier. Figure 6 shows the new outlines of the plasma cloud in accordance to Sutton's formula.

5. EFFECTS OF CHEMICAL LOSS

In this section, we study the decline of the peak ion density at the center of the cloud due to various chemical loss processes, without diffusion. The initial number density is taken as $N_o = 7.9367 \times 10^9$ cm^{-3}, which is the number density after 1s in §4.1.

5.1. Quadratic Loss Effect

Neglecting the diffusion term but retaining the quadratic loss term, we have

$$\frac{DN}{Dt} = -\alpha N^2, \qquad (5.1)$$

α being the quadratic loss rate. Integrating Eq.(5.1) from the initial condition $N = N_o$ at $t = 0$, we get

$$N = \frac{N_o}{1 + N_o t} \quad , \tag{5.2}$$

If $N_o \alpha t >> 1$, then $Nt = 1/\alpha$ = constant, the equation of a rectangular hyperbola. This illustrates an interesting case, where the number density falls rapidly in the first second regardless of N_o, but then declines much more gradually to levels depending solely upon α. Figure 7 shows this variation for three values of α. Also shown in the figure is the variation of the peak ion density due to diffusion alone, for comparison.

5.2. Linear Loss Effect

For the case of linear loss rate, we have

$$\frac{DN}{Dt} = - \beta N \quad , \tag{5.3}$$

where β, the linear loss rate is also proportional to M, the number density of the neutrals. Integrating from the same initial conditions, we get

$$N = N_o e^{-\beta t} \quad . \tag{5.4}$$

The value of β is not readily available in the literature. In this study, we have chosen some arbitrary values for β. Figure 8 shows that the decline of the peak ion density can be dramatic for large values of β. Note that the declines are actually exponential although they appear as straight lines in the logarithmic plot.

5.3. Altitude-Dependent Loss Rate

Since the neutral density in the atmosphere falls off approximately exponentially with height, we replace β by $\beta \exp(-h/H)$, where h is the vertical height and H the atmospheric scale height. For the moving plasma cloud, we can replace h by t from the relation $h/t = V \sin\theta$, whence we get $\beta\exp(-\gamma t)$, with $\gamma = V \sin\theta/H$. Taking the average scale height in the mesosphere as H = 6.5267 km (from U. S. Standard Atmosphere, 1976), we

get $\gamma = .3\ s^{-1}$. The equation of continuity of ions is thus

$$\frac{DN}{Dt} = -\beta e^{-\gamma t} N \quad . \tag{5.5}$$

Once again, by integrating from the same initial conditions, we find

$$N = N_o\, e^{\frac{\beta}{\gamma}\left(e^{-\gamma t} - 1\right)} \quad . \tag{5.6}$$

Figure 9 shows the decline of the peak ion density for the altutude-dependent linear loss rate. Compared with Fig.8,we find that the rate of decrease slows down considerably as the cloud moves upwards from the denser atmosphere below.

6. RECOMMENDATIONS

The main recommendations for follow-up study are the following:
(1) Investigate the dissipation of the plasma cloud due to the combined effects of diffusion and chemical loss, which involves solving the differential equation

$$\frac{DN}{Dt} = K \nabla^2 N - \alpha N^2 - \beta e^{-\gamma t} N \quad ; \tag{6.1}$$

(2) Properly identify the loss processes and estimate the loss rates;
(3) Estimate a reasonable value of the diffusion coefficient;
(4) Repeat this study for other types of explosives used such as Triaminobenzene (TATB), Hexanitrostilbene (HNS) and Hexanitrodiphenyl (DIPAM) (cf. Urbanski, 1984); and in such case,
(5) Incorporate ion chemistry, if possible.

ACKNOWLEDGMENTS

It is a pleasure on the part of the author to tender his sincere thanks to the following institutions and individuals:

(1) The United States Air Force Systems Command and The Office of Scientific Research for sponsoring the Summer Faculty Research Program;

(2) Universal Energy Systems, Inc. for successfully conducting this program;

(3) Arnold Engineering Development Center, Arnold Air Force Station, where this study was carried out;

(4) Mr. Marshall K. Kingery, the focal point at Arnold Engineering Development Center, for his effort in coordinating the program;

(5) Dr. (Mac) W. K. McGregor of Sverdrup Technologies, Inc. for introducing the research topic; and

(6) Dr. (Chad) C. C. Limbaugh of Sverdrup Technologies, Inc. for helpful collaboration.

REFERENCES

(1) P. M. Banks and G. Kockarts, Aeronomy, Academic Press, New York, 1973.

(2) A. C. Giere, NASA Tech. Memo. 78128, 1977.

(3) R. P. Haviland and C. M. House, Handbook of Satellites and Space Vehicles, Van Nostrand Co., Inc., Princeton, 1965.

(4) D. M. Hunten in Jupiter, The University of Arizona Press, 1976.

(5) D. E. Jensen and G. A. Jones, Combust. Flame, 32, 1, 1978.

(6) G. Kvifte in Physics and Chemistry of Upper Atmospheres, Reidel, New York, 1973.

(7) F. Pasquill, Atmospheric Diffusion, Ellis Horwood Ltd., New York, 1974.

(8) O. G. Sutton, Micrometerorology, McGraw-Hill, New York, 1953.

(9) U. S. Standard Atmosphere, U. S. Govt. Ptg. Office, Washington, 1976.

(10) T. Urbanski, Chemistry and Technology of Explosives, Vol.4, Pergamon Press, Oxford, 1984.

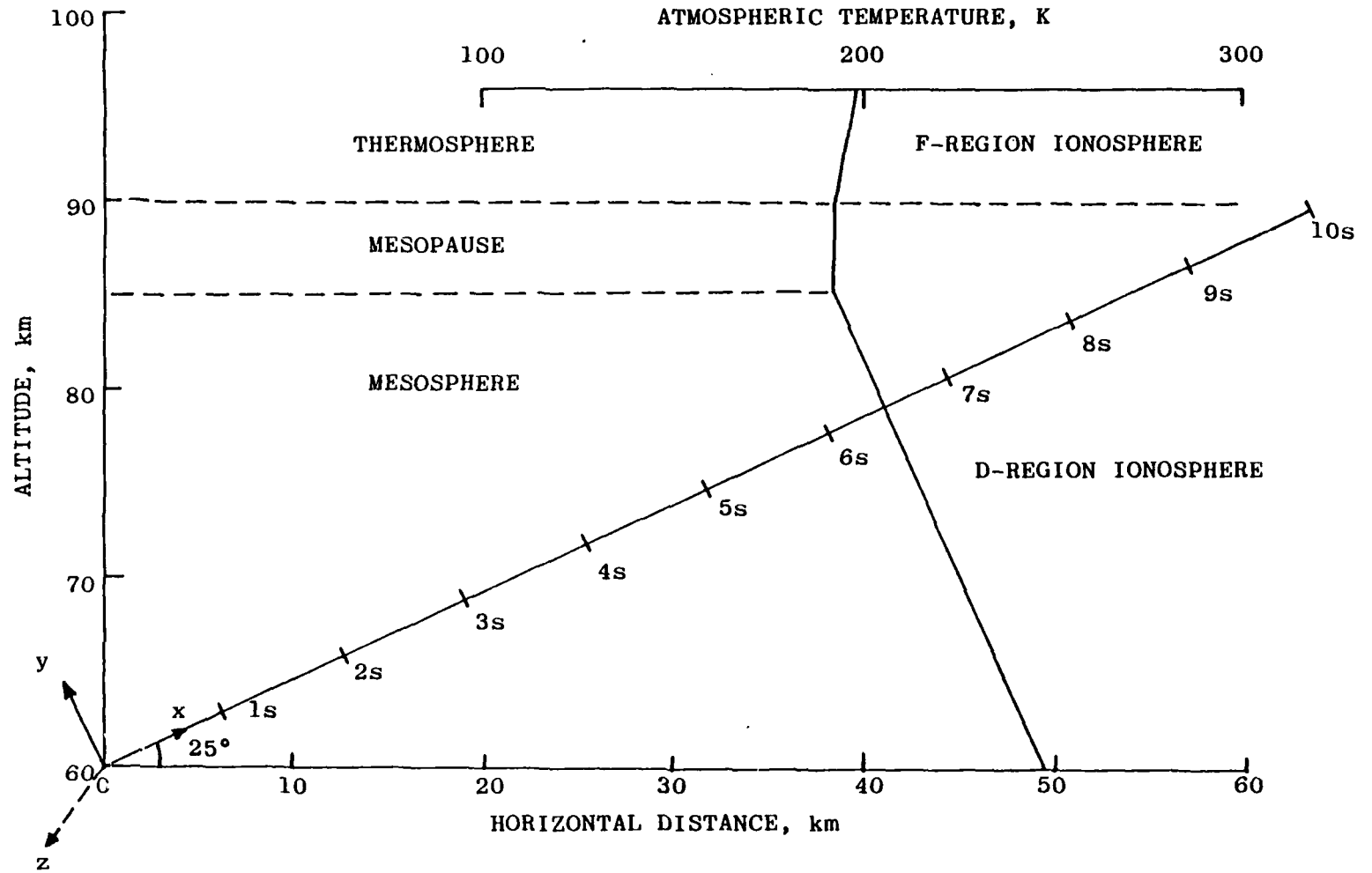


Fig.1. ICBM trajectory following third stage separation. Also shown are coordinate system chosen, atmospheric temperature profile and neutral atmospheric and ionospheric regions.

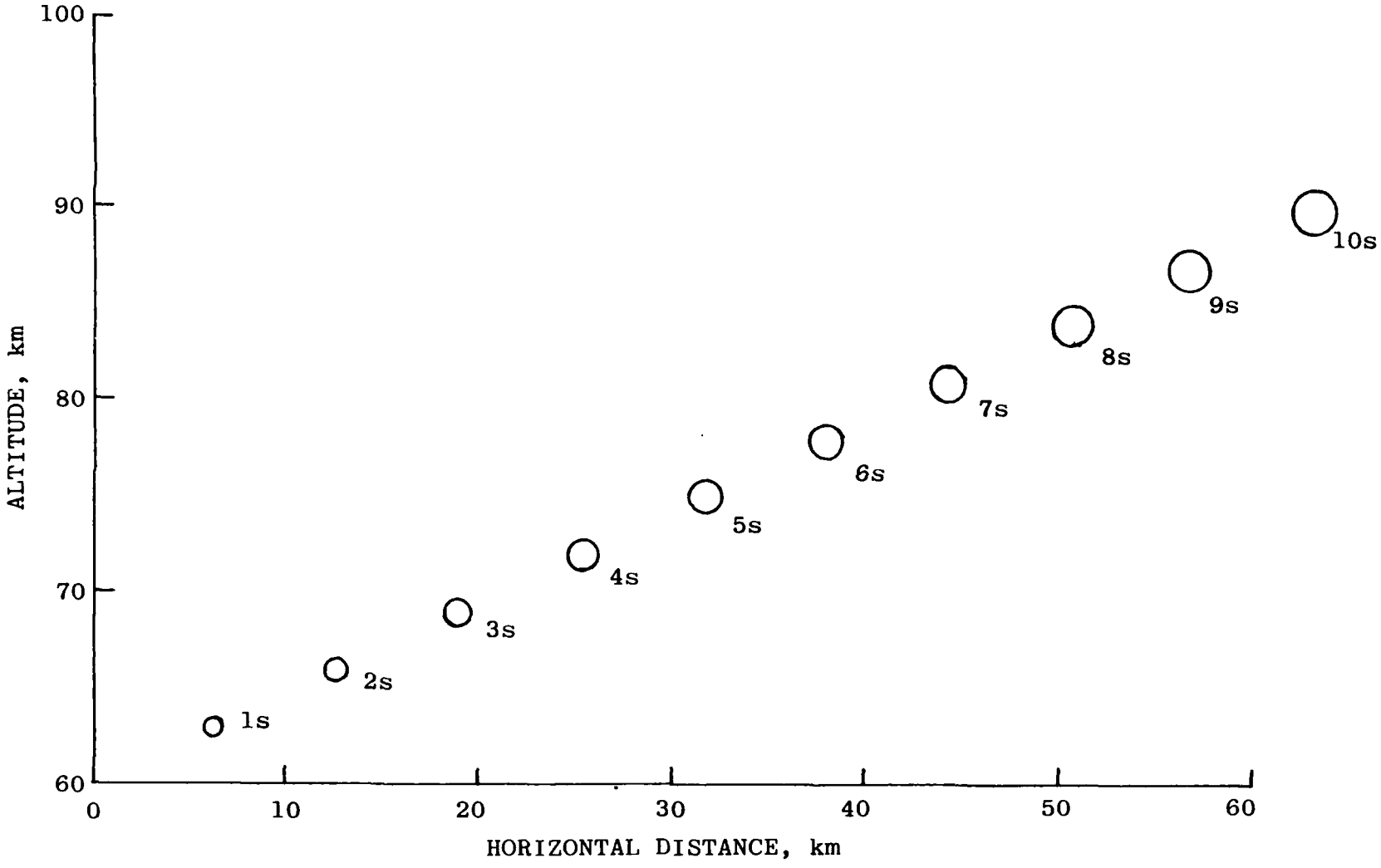


Fig.2. Outlines of plasma cloud containing 95.4% of ionization in an isotropic diffusive medium with eddy diffusion coefficient of 2×10^8 cm^2/s. The destruction of ions is neglected.

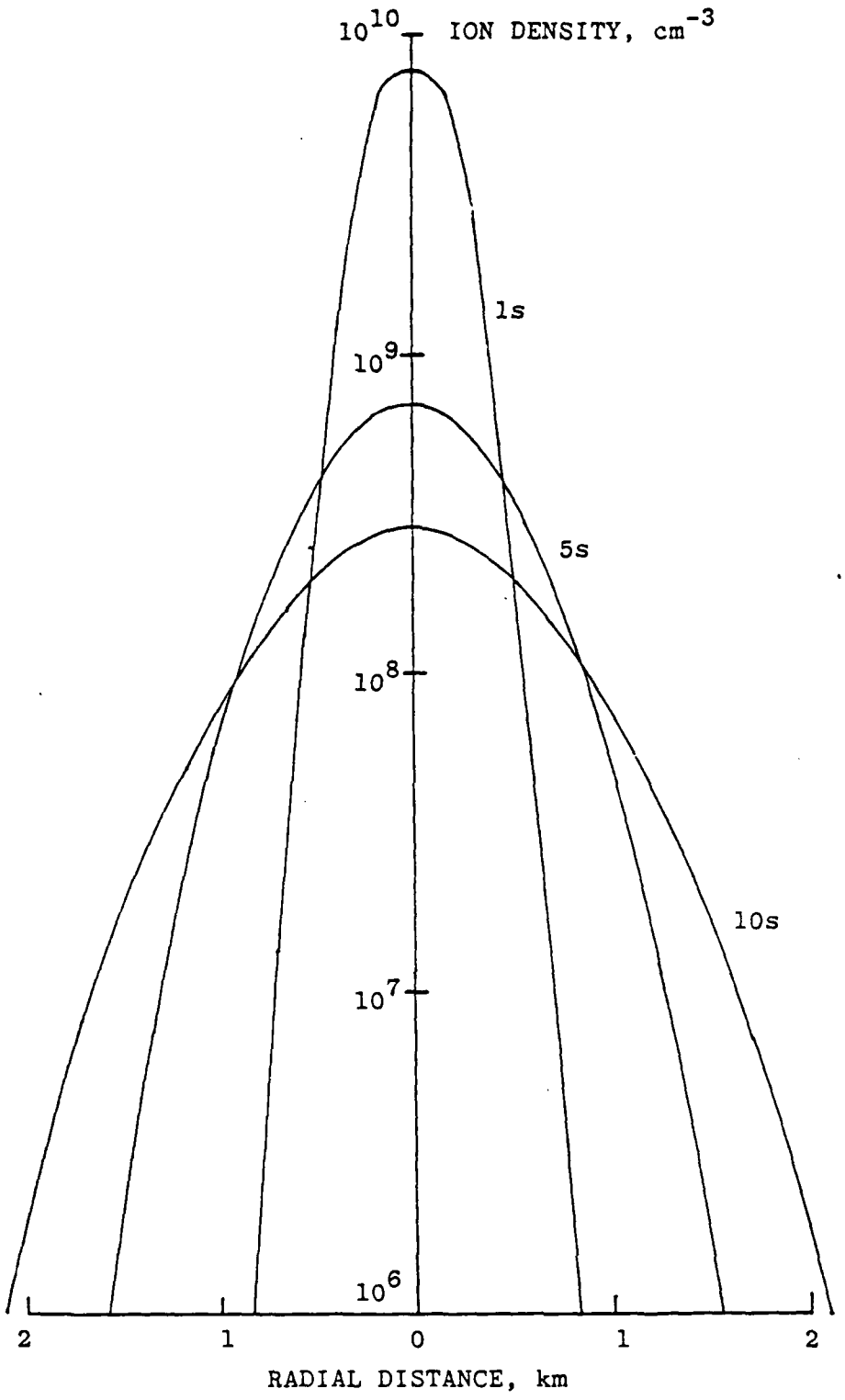


Fig.3. Radial distribution of ions after 1, 5 and 10 secs. of explosion in logarithmic scale.

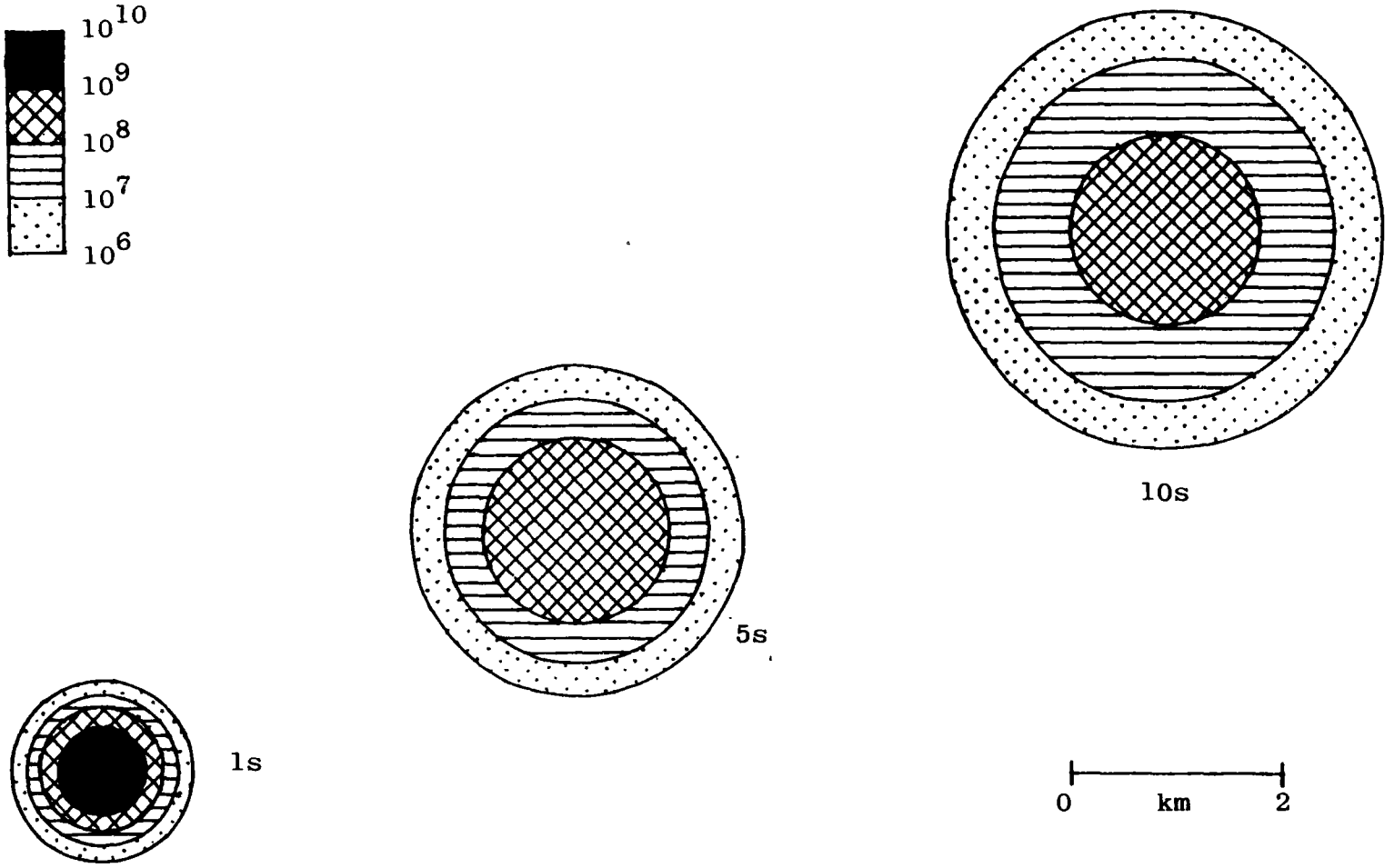


Fig.4. Contour plots of ion density after 1,5 and 10 secs. of explosion.

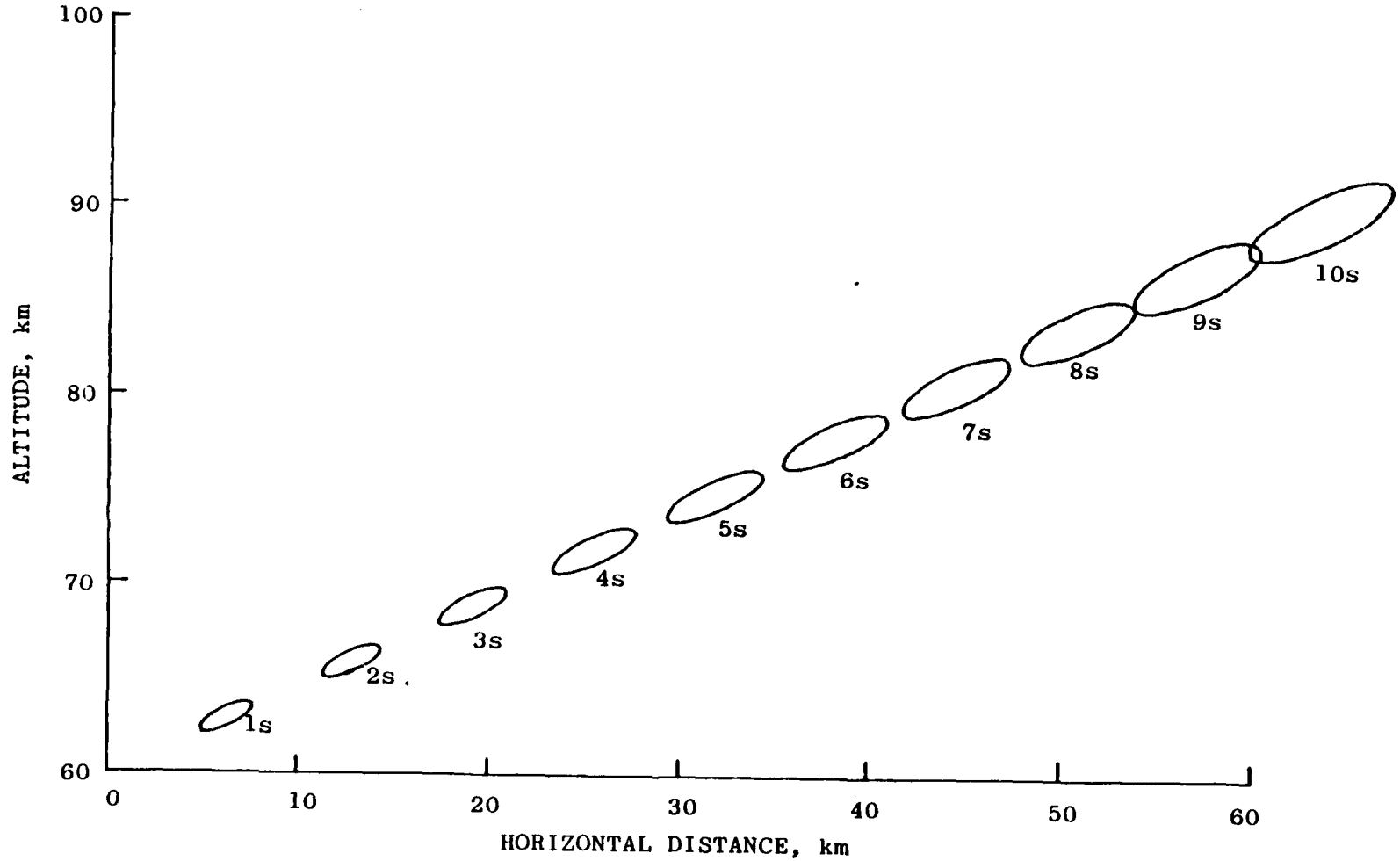


Fig. 5. Outlines of plasma cloud containing 95.4% of ionization in an anisotropic medium with larger diffusion coefficient along the direction of motion.

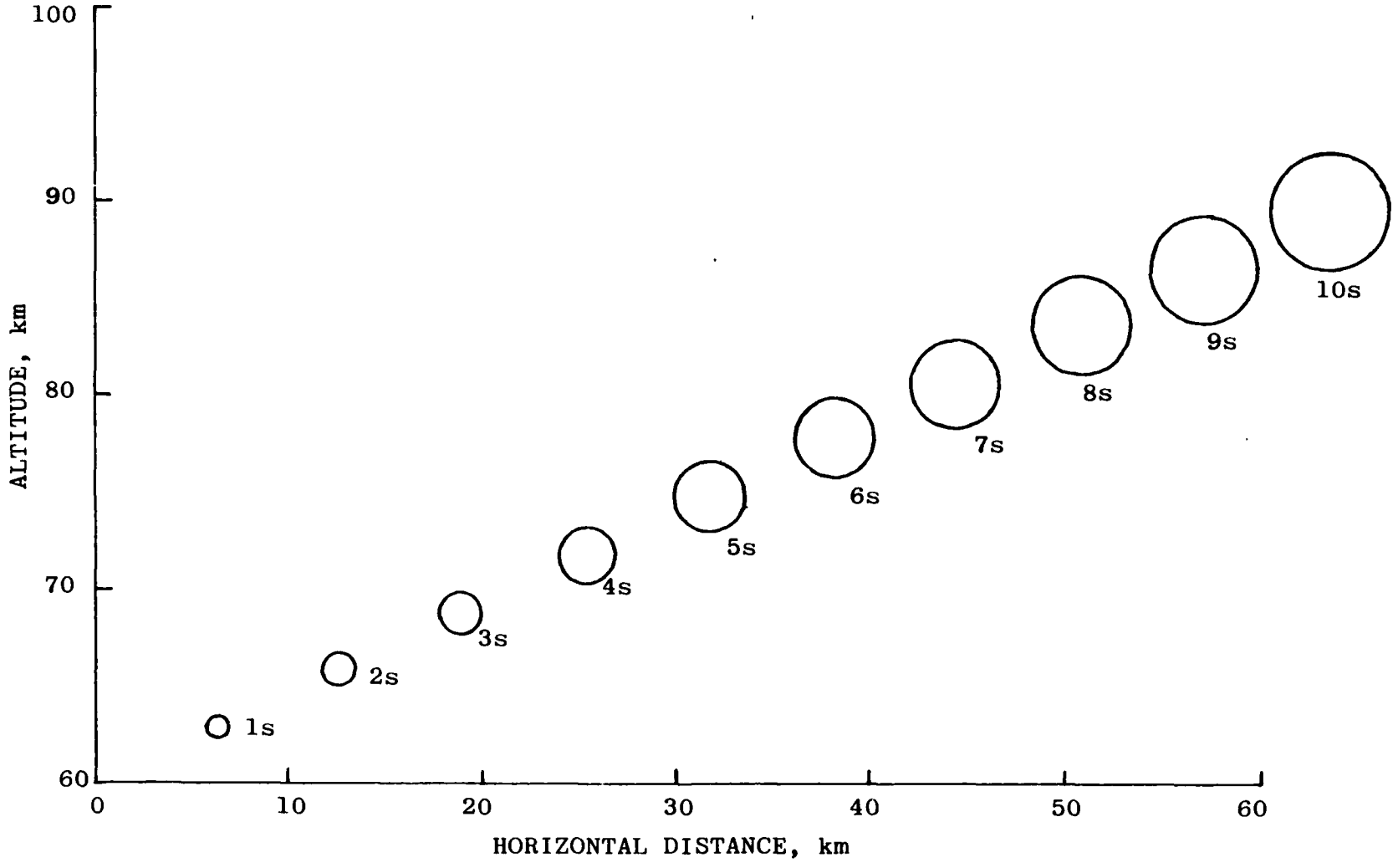


Fig.6. Outlines of plasma cloud containing 95.4% of ionization in isotropic medium according to Sutton's formula.

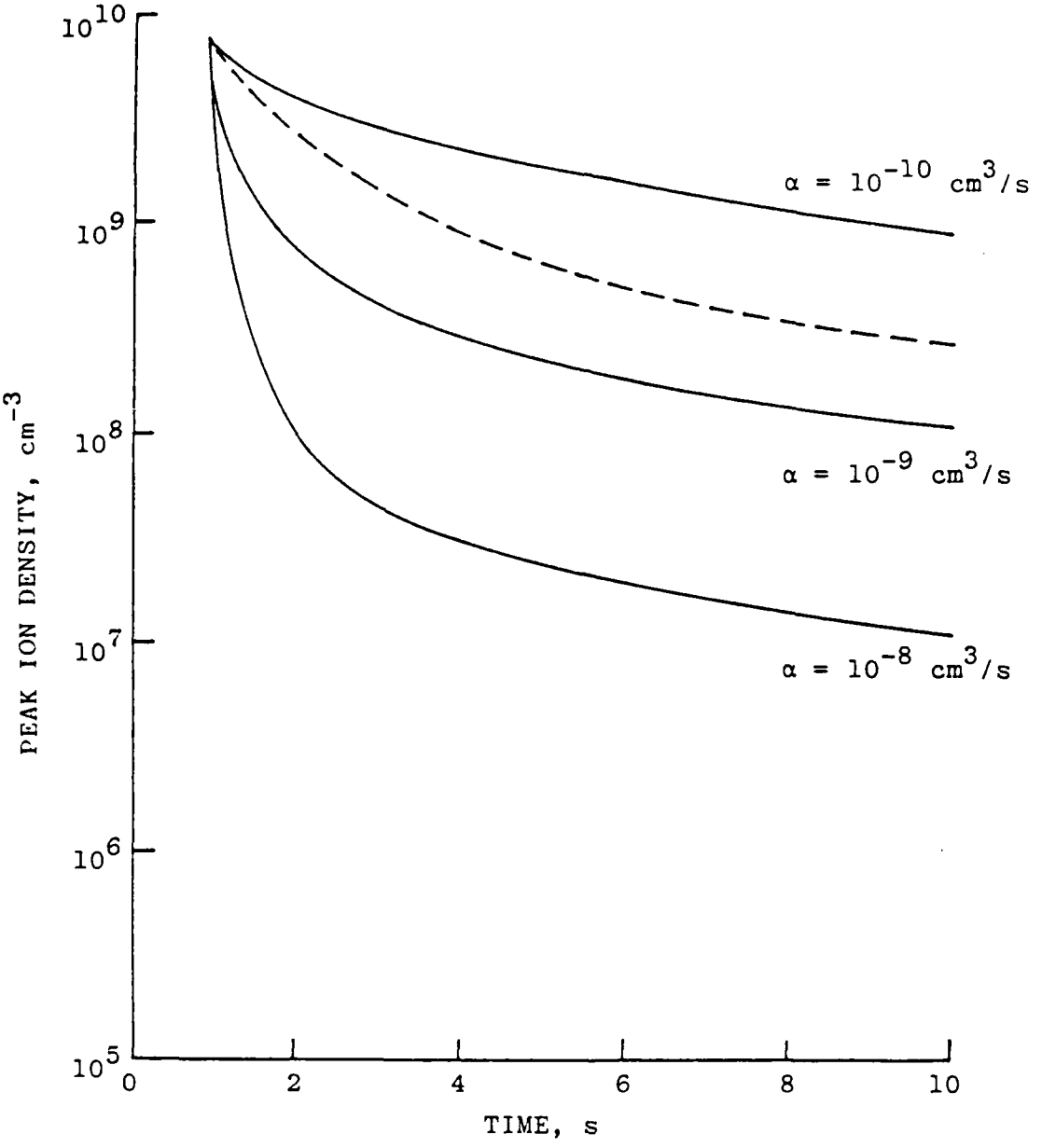


Fig.7. Variation of peak ion density (at the center of the cloud) with time due to diffusion (— — —) and quadratic loss process (———).

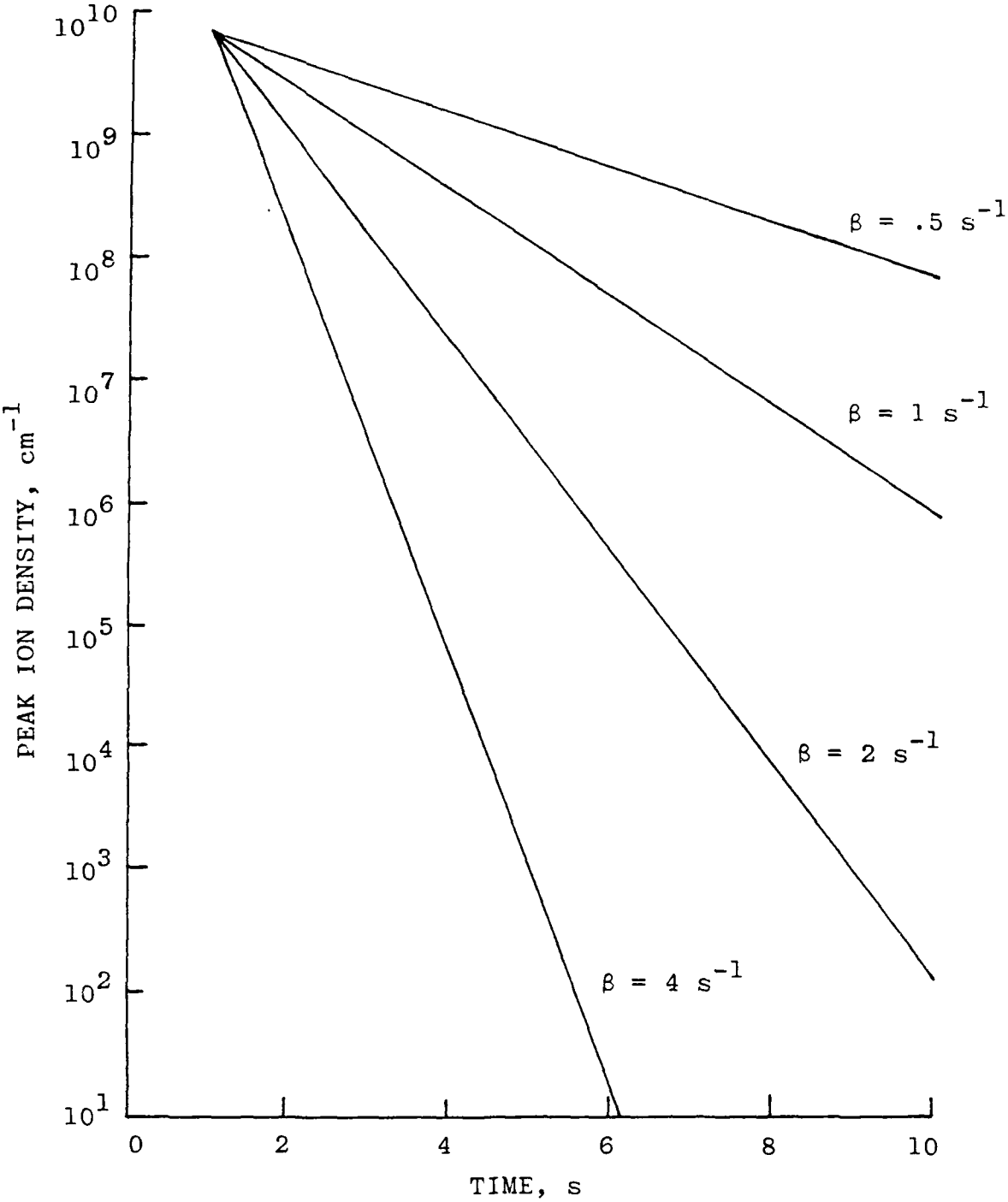


Fig.8. Variation of peak ion density (at the center of the cloud) with time due to linear loss rate.

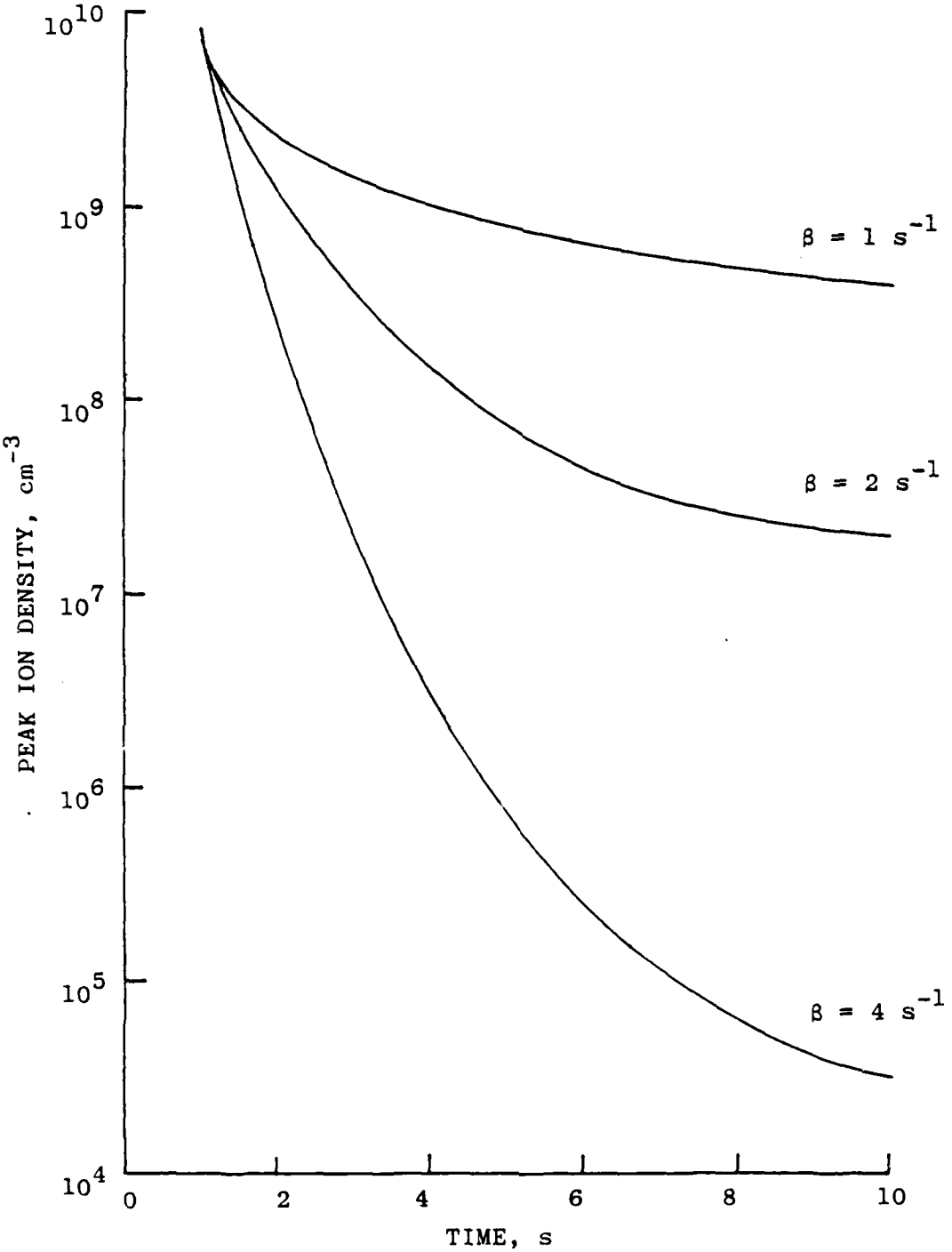


Fig.9. Variation of peak ion density (at the center of the cloud) with time due to altitude-dependent loss rate.

1986 USAF-UES SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

SURFACE ROUGHNESS EFFECTS ON HEAT TRANSFER AND SKIN FRICTION

Prepared by:	Robert P. Taylor
Academic Rank:	Assistant Professor
Department and University:	Mechanical and Nuclear Engineering Department Mississippi State University
Research Location:	Air Force Wright Aeronautical Laboratories Aero-Propulsion Laboratory Turbine Engine Division, Components Branch
USAF Researcher:	Dr. Richard B. Rivir

Date:	September 2, 1986
Contract No.:	F49620-85-C-0013

SURFACE ROUGHNESS EFFECTS ON HEAT TRANSFER AND SKIN FRICTION

by

Robert P. Taylor

ABSTRACT

Results are presented from a theoretical study of the effects of surface roughness on adiabatic wall temperature. This study was undertaken to determine the appropriate conditions required to use a turbulent circular Couette flow as an experimental test bed to study the effects of surface roughness on adiabatic wall temperature. It was concluded that the circular Couette flow is not an appropriate test bed for this purpose.

In addition, results are presented from profilometer measurements of surface roughness of inservice turbine blades. Data reduction techniques are developed to compute the roughness height distribution functions, auto correlation functions and power spectral densities for these profiles.

I. INTRODUCTION

I received my Ph.D. from Mississippi State University in 1983. The topic of my dissertation research was the development of a discrete element model for the prediction of boundary layer flow and heat transfer over rough surfaces. Since that time I have continued to pursue research in this area. Presently, I am, along with my colleague Hugh Coleman, beginning an experimental effort at Mississippi State to study the effects of surface roughness on boundary layer flow and heat transfer which is funded by the Air Force Office of Scientific Research.

Surface roughness can have a large influence on the skin friction and heat transfer on turbine blades and vanes. Increases in heat transfer rates of 100% have been reported for rough turbine blades as apposed to otherwise identical smooth blades [1]. Therefore, the Propulsion Laboratory at Wright-Patterson, AFB has had an interest in the prediction of boundary layer flow and heat transfer over rough surfaces. Because of this mutual interest, I was selected to spend the 1986 Summer Faculty Fellowship at the Components Branch of the Propulsion Laboratory.

Two sub-programs were considered during my summer program: 1. A preliminary study to determine the feasibility of using a circular Couette flow with a high Prandtl number fluid to experimentally investigate the effects of surface roughness on the adiabatic wall temperature, and 2. A program to measure and classify the actual surface roughness on turbine blades and vanes. This report discusses the results of these two sub-programs in the following two sections. Also recommendations for further work are given.

II. ADIABATIC WALL TEMPERATURE

The adiabatic wall temperature is the equilibrium temperature which a surface must attain to have zero heat transfer. It is the natural reference temperature for heat transfer in flows in which viscous heating is important such as high speed gas flows and moderate

speed flows of high Prandtl number fluids. It is the natural reference because it is the delimiting temperature; surface temperatures above the adiabatic wall temperature result in heat transfer into the fluid while those below this temperature result in heat transfer into the surface.

Traditionally the adiabatic wall temperature has been defined in terms of a "recovery" factor, R, as

$$T_{aw} = T_e + R \frac{U_e^2}{2C_p} \qquad (1)$$

The recovery factor for boundary layer flow is usually given in terms of the Prandtl number as $R = \sqrt{Pr}$ for laminar flows and $R = \sqrt[3]{Pr}$ for turbulent flows. These formulae are generally only accurate for fluids with Pr near 1 ($0.5 \leq Pr \leq 10$). However, for laminar plane Couette flow $R = Pr$ for all cases.

The use of the term "recovery" temperature causes some confusion. In the boundary layer the adiabatic wall temperature is a result of the balance of viscous work and thermal diffusion and advection not a direct recovery of kinetic energy in a stagnation process. For $Pr > 1$ the recovery temperature is greater than the stagnation temperature.

Recent calculations by Hodge, Taylor and Coleman [2] indicate that surface roughness can have significant effects on the adiabatic wall temperature. To the author's knowledge, no systematic experimental investigation of this phenomenon has ever been conducted.

Objectives

A set of simple experiments was envisioned. The test bed would be a circular Couette flow with a high Prandtl number fluid. This arrangement would allow a large number of surfaces and Reynolds numbers to be tested more rapidly and cheaply than supersonic wind tunnel test.

The objectives of this sub-program were to develop a computer model for turbulent rough wall Couette flow and to use this model to determine the conditions and fluid properties which would be required to make the Couette flow a useful test bed for the investigation of surface roughness effects on adiabatic wall temperature.

The development of this model, the results of the investigation and conclusions are given below.

Development of the Computer Model

To model circular Couette flow with rough surfaces, the discrete element approach of Taylor, Coleman and Hodge [3] is used. Discrete element equations are derived for circular Couette flow with a rough surface. These equations are then solved numerically. Comparisons with exact laminar solutions and data for turbulent flow are given for partial verification.

The physical system considered is shown in Figure 1. The inner cylinder of radius r_i rotates with angular velocity w_i. The outer cylinder of radius r_0 is taken to be stationary. It is assumed that the outer cylinder is the one with a rough surface.

By performing a torque balance on an annular control volume which is assumed to contain roughness elements, the following differential momentum equation is obtained.

$$0 = \frac{d}{dr}\left(\beta\hat{\mu}r^3\frac{dw}{dr}\right) - \frac{\rho C_D d r^3 r_0 w^2}{2L1} \qquad (2)$$

where β = the blockage factor; $\hat{\mu} = \mu + \mu_t$, the total apparent turbulent viscosity; $w = V/r$, the angular velocity of the fluid; ρ = the fluid density; C_D = the local element drag coefficient; d = the local element diameter; L = the circumferential element spacing; and l = the longitudinal element spacing.

By performing a similar energy balance, the following differential energy equation is obtained

$$0 = \frac{d}{dr}\left(\beta\hat{K}r\frac{dt}{dr}\right) + \beta\hat{\mu}r\left(r\frac{dw}{dr}\right)^2 + \frac{\rho C_D d r_0 r w^3}{2L1}$$

$$+ \frac{\pi K Nu_d r_0 (t_R - t)}{L1} \qquad (3)$$

where $\hat{K} = K + K_t$, the total apparent turbulent conductivity; t = the fluid temperature; Nu_d = the local element Nusselt number; and t_R = the roughness element temperature.

The blockage factor can be shown to be

$$\beta = 1 - \frac{\pi}{4}\frac{d^2}{Ll}\frac{r_o}{r} \tag{4}$$

where the term r_o/r accounts for the curvature of the base wall.

The boundary conditions for equations (1) and (2) are

$$w = w_i, \quad t = t_i \quad \text{at} \quad r = r_i \tag{4a}$$

$$w = 0, \quad t = t_o \quad \text{at} \quad r = r_o \tag{4b}$$

A mixing length model for μ_t such as that given by Taylor et al. [3] is used for turbulent closure, but for the Couette flow the wake or core region is ignored. Following Kays [4] a turbulent Prandtl number (Pr = 1) is used to determine K_t.

The effective rough wall shear stress, τ_{eff}, can be determined by integrating equation (1) and recalling that for circular Couette flow the torque must be constant. The result is

$$\tau_{eff} = \beta_o\tau + \int_{r_i}^{r_o} \frac{\rho C_D dr\, 3v^2}{2Llr_o} \tag{5}$$

Likewise the effective rough wall heat flux, q_{eff}, is determined from equation (2)

$$q_{eff} = \beta_o q_o - \int_{r_i}^{r_o} \frac{\pi K Nu_d (t_R - t)}{Ll} \tag{6}$$

Equations (1) and (2) are easily solved by the implicit finite difference scheme discussed in Taylor et al. [3]. The equations are first linearized. Then the derivatives are approximated using three point Lagrangian interpolating polynomials. This results in a system of linear algebraic equations with a tridiagonal coefficient matrix for which the solutions may be simply written down. The solution of the original nonlinear equations is obtained by repeating this procedure in an iterative fashion. The reader is referred to Taylor et al. [3] for details.

To impose the adiabatic condition on the outer cylinder for a smooth wall the temperature boundary condition in (4b) is changed to $dt/dr = 0$ at $r = r_o$. However, for rough surfaces this is not appropriate. The effective rough wall heat flux of equation (6) must

be forced to zero. For this case a temperature is assumed for the rough wall and it is also assumed that the element temperature is equal to this temperature (valid for element Biot numbers <0.1). The equations are solved for this boundary condition and the value of q_{eff} is computed. The value of the wall temperature is then perturbed as small amount and the solution is repeated. A secant method iterative procedure is carried out until the wall/element temperature is found which gives $q_{eff} \approx 0$.

In order to verify the computer code several test cases were run. For laminar smooth wall cases exact integrals of equations (1) and (2) can be obtained. Figures 2 and 3 show comparisons of the exact solutions and computed values for nondimensional angular velocity, $(w - w_i)/(w_o - w_i)$, and temperature, $(t - t_i)/(t_o - t_i)$, profiles for laminar flow. Figure 4 shows a comparison of the exact and computed adiabatic wall temperature for laminar flow. Inspection of the figures reveals that the agreement is exact. This test is not severe and exact agreement is expected. However, it does demonstrate that the main solution technique and code are functional. Also shown in Figure 4 is the plane Couette flow solution for comparison.

Schlichting [5] reports moment coefficient data for turbulent smooth wall circular Couette flow, where the moment coefficient is defined as

$$C_M = \frac{4\tau_i}{\rho r_i^2 w_i^2} \qquad (7)$$

Figure 5 shows a comparison of these data with the computed values of C_M. Inspection of the figure reveals that the agreement is excellent. This agreement demonstrates that the turbulence model is functional for the smooth wall case. Also shown in the figure are computed values for two rough cases. The roughness elements used are a uniform array (L = 1) of hemispheres with a radius of k = 0.025 inches. The inner and outer cylinder radii are r_o = 2.0 inches and r_i = 2.25 inches. The Taylor number (after G. I. Taylor) is defined as $T_a = \rho r_i w_i (r_o - r_i) \sqrt{(r_o - r_i)/r_i}/\mu$. For constant r_o and r_i T_a reduces to a weighted Reynolds number.

Results

If the circular Couette flow is to be an appropriate test bed, the adiabatic wall temperature difference between the smooth and rough wall cases should be greater than 5°F. In order to determine the conditions needed to meet this requirement, three cases were considered: 1. a hydraulic oil (Pr = 450), 2. water (Pr = 5), and 3. a light oil (Pr = 100).

Figure 6 shows the results of T_{aw} computations for the hydraulic oil. The Couette geometry is r_o = 2.25 inches and r_i = 2.0 inches; the roughness is a uniform array of hemispheres of radius k = 0.025 inches with spacing of L/k = 4. The fluid properties are $\mu = 6.9 \times 10^{-4}$ slugs/ft-s, ρ = 1.68 slugs/ft^3, k = 0.016 ft-lb/s-ft-F, and Cp = 10,500 ft-lb/slug-F. Inspection of the figure reveals the adiabatic wall temperature is 2°F higher for the rough case at the maximum rotation considered of 8000 rpm. In addition, the computations show that the power required to drive the rough wall flow at 8000 rpm is approximately 12 horsepower.

Figure 7 shows the results of T_{aw} computations for water. The Couette and roughness geometry are the same as those considered above, and the properties of water are taken to be constant at 80°F. Inspection of the figure reveals a curious switch. The rough wall T_{aw} is now predicted to be less than the smooth wall case. At the maximum rotation considered of 20,000 rpm the rough wall T_{aw} is 0.4°F less than the smooth value. The calculations show that approximately 65 horsepower is required to drive the rough wall flow at 20,000 rpm.

Figure 8 shows computations for a light oil in the same geometry with the same roughness as above. The properties of the fluid are $\mu = 1.5 \times 10^{-4}$ slugs/ft-s, ρ = 1.47 slugs/ft^3, k = 0.017 ft-lb/s-ft-F, and Cp = 12,300 ft-lb/slug-F. Inspection of the figure reveals that T_{aw} for the rough surface is approximately 0.3°F above the smooth wall case at the maximum rotation considered of 8000 rpm. About 9 horsepower are required to drive the rough wall flow at 800 rpm.

The switch from enhanced adiabatic wall temperature for the oils to decreased adiabatic wall temperature for the water is explained by the influence of the roughness on shear and heat diffusion. The roughness increases shear, shear rate and heat diffusion. For the very

viscous fluids, such as the oils (high Pr), the viscous work is increased more than the heat diffusion and the balance is tipped in favor of higher adiabatic wall temperature. For less viscous fluids, such as water, the balance is tipped in favor of heat diffusion and the adiabatic wall temperature decreases.

Conclusions

From the results presented and discussed above, it is suggested that the high Prandtl number Couette flow is not an appropriate test bed for investigations of the effect of the influence of surface roughness on turbulent adiabatic wall temperature.

Significant differences between smooth and rough wall adiabatic wall temperatures could be obtained at very high rotational speeds. However, the huge power requirements and the resulting huge cooling load required to maintain a constant uniform inner cylinder temperature make such a case impractical. For more viscous higher Prandtl number fluids such as glycerin (Pr = 80,000), it would be difficult to obtain turbulent flow.

III. CHARACTERIZATION OF ACTUAL SURFACE ROUGHNESS

Traditional measures of surface roughness such as root mean height or average height are not satisfactory for the characterization of the influence of surface roughness on boundary layer flow and heat transfer. A more appropriate means of characterization is to use the complete statistical description of the surface.

To obtain a complete statistical description of a rough surface, a sample of the surface profile is obtained. This profile is then digitized and the height distribution density function, autocorrelation function and power spectral density function are calculated. These statistical functions can then be used to compute all the other statistical properties such as peak distributions, peak densities, etc. as discussed by Nayak [6].

The objectives of this sub-program, the measurement techniques used, the statistical computations, and typical results of these measurements are given below. Also, a scheme for analysis of these data from the standpoint of boundary layer modeling is outlined.

Objectives

The objectives of this sub-program were: 1. to develop techniques to use the surface profilometer which is in the AFWAL propulsion lab for measurement of the surface statistics, 2. to collect surface profile data on typical turbine blades, and 3. to analyze these data from the standpoint of modeling boundary layer flow and heat transfer.

Measurement Techniques

Surface profiles were collected using the Rank-Hobson-Taylor Surtronic III surface profilometer. This stylus profilometer is capable, by setting the appropriate amplification ratio, of resolving surface detail between 0.1 micron and 100 microns. The major function of this instrument used in this program is the x-y recording ability. Figure 9, shows a typical section of a recorded surface record. The instrument also has some internal computational ability and can compute the average roughness height, the maximum peak to valley height, the average of the 5 largest peak to valley heights, the number of peaks per unit profile length, and other similar parameters. The parameters which are computed directly by the instrument were developed mainly for surface to surface contact analysis and are of little direct use for boundary analysis.

Once the profile records such as Figure 9 are obtained they are digitized and conditioned for statistical computations. The profiles can be digitized with an x-y digitizing device. The record is placed on the digitizer bed, the profile is followed by hand and selected points are entered into the computer as x-y coordinates. The x-spacing of these points should be at most 1/5 of the major structural scale of the roughness. The x-y coordinates are then converted into the physical coordinate system of the record by using the proper scaling ratios. Since the turbine blades are curved, the profilometer will pickup some of the trends of the sutstrate. These trends must be removed for proper statistical analysis of the roughness. This trend removal is easily accomplished by least squares fitting of a polynomial to the x-y data. This trend curve is then used as the new origin of the y-coordinate.

Statistical Computations

Once the surface profile has been digitized and conditioned as discussed above the following statistical functions are computed: 1. height distribution density, 2. autocorrelation and 3. power spectral density.

The height distribution is computed by dividing the range of surface heights into regions or bins and counting the number of occurances in each bin. The number of bins is set based on the total number of points, N, in the record by $N_{bin} = N/20$. To decide in which bin a given point belongs the following algorithm is followed

$$\text{Compute:} \quad C = \frac{y_{max} - y_{min}}{N_{bin}} \tag{7a}$$

$$\text{Compute:} \quad L = \frac{y_i - y_{min}}{C} \tag{7b}$$

The bin, index ℓ, in which y_i falls is the integer ℓ = max integer < L. The number of points, n_ℓ, in each bin is determined by adding 1 to n_ℓ each time ℓ is encountered. The probability density histogram is then computed by $p_\ell = n_\ell/N$ for $\ell = 0, 1, \ldots N_{bin} - 1$.

The autocorrelation is computed by evaluating the integral

$$R_X(\tau) = \lim_{L\to\infty} \int_{-L/2}^{L/2} y(x)y(x + \tau)dx \tag{8}$$

The finite sample approximation for equation 8 is

$$R_{xj} = \frac{1}{N-j} \sum_{i=1}^{N-j} y_i y_{i+j} \qquad ;j = 4,1,\ldots m \tag{9}$$

where m = N/10. The limit on m is required to avoid digital instabilities in a finite sample.

The power spectral density Psd is defined as the cosine transformation or $R_X(\tau)$

$$Psd(f) = 4 \int_0^\infty R_X(\tau) \cos(2\pi f\tau)d\tau \tag{10}$$

The finite sample approximation or equation 10 is

$$Psd_k = 2\Delta x\left(R_0 + 2\sum_{j=1}^{m-1} R_{xj}\cos\frac{\pi kj}{m} + R_m\cos\pi k\right) \qquad (11)$$

where $k = 0, 1, \ldots m$

and the frequency corresponding to k is

$$f = \frac{k}{2m\Delta x} \qquad (12)$$

To reduce the effects of leakage which result from the digital approximation in equation (11), the Psd is computed with a Hamm window as follows

$$\tilde{Psd}_o = \frac{1}{2} Pds_o + \frac{1}{2} Psd_1$$

$$\tilde{Psd}_k = \frac{1}{4} Psd_{k-1} + \frac{1}{2} Psd_k + \frac{1}{4} Psd_{k+1}, \quad k=1,\ldots m-1 \qquad (13)$$

$$\tilde{Psd}_m = \frac{1}{2} Psd_{m-1} + \frac{1}{2} Psd_m$$

A detailed discussion of these statistical computations can be found in the book by Otnes and Enochson [7].

Results

Profiles have been obtained for several test specimens at Wright-Patterson AFB and statistical analysis of these data have been computed. Also, a large number of profiles were obtained on first and second stage turbine blades for the F-100 and TF-39 engines. These profiles were obtained in the last week of the summer program and statistical computations were not completed.

A sample of the profile record and the statistical computations for sample turbine blade T-N4 is shown in Figure 9. From the figure it is seen that the height distribution is very nearly Gaussian. Also, the autocorrelation function shows an exponential like decay with a correlation length of 5 mils. Figure 11 shows similar information for turbine vane V-002. Inspection of the figure reveals that the height distribution is not Gaussian for this record. Also, the

autocorrelation and power spectrum show a tendency for periodicity. The designations T-N4 and V-002 are the author's own numbering system for the sample specimens and have no other significance.

The results in Figures 9 and 10 are given to demonstrate the appropriate statistical analysis. The following section describes briefly how these statistics might be used to predict the effect of the rough surface on the flow and heat transfer.

A much larger sample has been obtained for F-100 and TF-39 first and second stage turbine blades. A summary of these data are given in Table 1. Figure 11 demonstrates the approximate location of the traces and provides a key for the trace numbers and number of cycles symbols in Table 1. All traces were taken near the mid-span of the blade and the direction of the traces were from the base toward the tip. All traces were approximately 1/2 inch long.

Inspection of Table 1 and the graphical summary in Figure 12 reveals interesting qualitative information. Figures 12(a)-(c) show some indication of the aging process on F-100 1st stage turbine blades. The figures reveal that the roughness location on these blades is on the suction side near the leading edge. Also, the roughness seems to increase rapidly with the number of cycles. The other traces in Figures 12(a)-(c) show a much smaller increase in the roughness with age, Figure 12(d) shows the results for an F-100 1st stage blade with shower head cooling. This blade is approximately twice as old as the one in 12(c); however, it has approximately the same average roughness. Traces 42-46 in Table 1 give some idea of the scatter in the roughness from blade to blade. For the 5, 1800-2000 cycle, 1st stage F-100 turbine blades with shower head cooling, the average roughness varies from 5.2 μm to 9.8 μm with a mean value of 7.3 μm. Figures 12(e) and (f) show the distributions of average roughness for two 1st stage TF-39 blades of unknown cycles. The pattern of the roughness on these blades is very different from that on the F-100 blades. The largest roughness on these blades is located on the pressure side at the mid-chord and the trailing edge.

Proposed Boundary Layer Computational Procedure

The proposed surface roughness computational procedure is the discrete element method of Taylor, Coleman and Hodge [3]. This method is discussed in the first subprogram of this report (Section II) and is discussed in detail in reference [3]. Only a brief summary is given here. This method includes the effects of the local element drag and heat transfer directly in the boundary layer equations. A set of discrete element equations for a general surface is given in Figure 13. The terms α_x and α_y are blockage terms and represent the average fraction of a surface which is blocked to fluid flow. The terms C_D and Nu_d are the local element drag coefficient and Nusselt number. The term d is the local element diameter at a given y-location. The terms N_x and N_z are the numbers of elements in an x-trace and a z-trace as shown in Figure 14. The products of the solutions of these equations are velocity and temperature profiles and surface measureables such as total surface skin friction and Stanton numbers.

The equations in Figure 13 can be used to predict the heat transfer and fluid mechanics of a boundary layer over a general rough surface if some method exist to determine the distribution of element diameters at each y-location as shown in Figure 15. Patir [8] has presented a method which, given the height distribution function and autocorrelation function, can generate a statistically equivalent distribution of diameters as shown in Figure 15.

The proposed procedure is:

1. collect the statistical information on the rough surface - traces - height distributions - autocorrelations functions
2. use methods such as Patir's [8] to obtain the statistically equivalent distribution of element diameters
3. use the discrete element method to predict the total skin friction and heat transfer of the boundary layer.

Conclusions and Recommendations

Measurement techniques have been developed and demonstrated for the complete statistical description of a rough surface. A procedure has been introduced by which this information might be used to predict the effects of surface roughness on boundary layer and heat transfer.

Measurements of surface roughness profiles and parameters have been made for several F-100 and TF-39 engine blades. The summary of these measurements indicates that inservice blades can be much rougher than new blades.

Because of time constraints in the ten week program the profile data for the F-100 and TF-39 blades could not be reduced. Also, profiles on only a limited number of blades could be obtained. Therefore, it is recommended that a follow on mini-grant program be funded to complete the reduction of the data from the summer program and to collect and reduce profile data on a larger sample of blades.

IV. ACKNOWLEDGMENTS

The support of the Air Force office of Scientific Research is gratefully acknowledged. The organization and conduct of the 1986 Summer Faculty Research Program by Universal Energy systems was excellent. A special thanks goes to Dr. Richard B. Rivir for his support and encouragement during this program. Also, thanks go to Mr. Weslie Thomas of Kelly AFB, TX for his assistance in obtaining engine blades for the measurements.

V. REFERENCES

1. Rivir, Richard B., Personal Communication.

2. Hodge, B. K., Taylor, R. P. and Coleman, H. W., "An Investigation of Surface Roughness Effects on Adiabatic Wall Temperature," AIAA-85-1657.

3. Taylor, R. P., Coleman, H. W. and Hodge, B. K., "A Discrete Element Prediction Approach for Turbulent Flow Over Rough Surfaces," Mississippi State University, TFD-84-1.

4. Kays, W. M., Convective Heat and Mass Transfer, McGraw-Hill, 1966.

5. Schlichting, H., Boundary-Layer Theory, McGraw-Hill, 1968.

6. Nayak, P. R., "Random Process Model of Rough Surfaces," J. of Lubrication Tech., July 1971.

7. Otens, R. K. and Enochson, L., Digital Time Series Analysis, John Wiley and Sons, 1969.

8. Patir, N., "A Numerical Procedure for Random Generation of Rough Surfaces," Wear, Vol. 47, pp. 263-277, 1978.

Table 1
Summary of surface roughness measurements

	Engine	Stage	Age	Trace Location	Ra μm	Rt μm	Rtm μm
1.	F100	1st	A-R	2	1.3	9.7	7.6
2.	F100	1st	B-R	2	1.0	---	5.6
3.	F100	1st	C-R	2	0.9	6.3	4.8
4.	F100	1st	D-R	2	2.0	15.2	10.2
5.	F100	1st	T-R	2	1.4	8.8	7.2
6.	F100	2nd	C R	2	1.1	8.4	6.6
7.	F100	1st	E	1	2.7	19.9	---
8.	F100	1st	E	2	1.7	15.8	9.4
9	F100	1st	E	1	7.0	60.0	44.0
10.	F100	1st	E	2	1.4	9.3	7.0
11.	F100	1st	E	3	1.7	15.3	9.7
12.	F100	1st	E	5	1.8	17.8	12.7
13	F100	1st	E	6	2.2	18.2	14.7
14	F100	1st-SH	E	1	7.2	57.0	40.0
15	F100	1st-SH	E	2	3.3	24.0	13.0
16	F100	1st-SH	E	3	2.9	19.2	10.3
17	F100	1st SH	E	4	2.5	23.2	13.7
18	F100	1st-SH	E	5	2.9	19.7	16.7
19	F100	1st-SH	E	6	2.6	23.0	13.7
20	F100	1st	A	1	3.8	24.3	18.3
21	F100	1st	A	2	2.5	16.9	14.6
22	F100	1st	A	3	2.4	16.1	13.0
23	F100	1st	A	4	2.8	17.7	13.8
24	F100	1st	A	5	2.5	18.4	12.1
25	F100	1st	A	6	2.1	15.1	12.4
26	F100	1st	B	1	4.3	41.4	25.8
27	F100	1st	B	2	2.4	16.6	11.1
28	F100	1st	B	3	3.4	16.8	10.8
29	F100	1st	B	4	2.6	19.0	14.8
30	F100	1st	B	5	2.5	16.6	17.3
31	F100	1st	B	6	4.3	31.2	30.0
32	F100	1st	D	1	9.2	82.2	55.3
33	F100	1st	D	2	2.8	19.1	17.1
34	F100	1st	D	3	3.0	17.9	16.0
35	F100	1st	D	4	3.3	33.8	20.3
36	F100	1st	D	5	4.2	32.5	25.6
37	F100	1st	D	6	4.3	32.4	23.3
38	F100	2nd	C	1	3.2	23.1	16.7
39	F100	2nd	C	1	3.4	24.2	15.0
40	F100	2nd	C	1	3.7	42.7	23.2
41	F100	2nd	T-R	1	2.1	13.8	11.4
42	F100	1st-SH	E	1	5.4	40.8	32.2
43	F100	1st-SH	E	1	5.2	39.5	28.1
44	F100	1st-SH	E	1	7.6	70.0	35.9
45	F100	1st-SH	E	1	8.7	83.5	39.1
46	F100	1st-SH	E	1	9.8	105.7	51.7
47	TF39	1st	NEW	1	1.5	12.9	7.4
48	TF39	1st	T	1	4.4	36.8	27.6
49	TF39	1st	T	2	3.6	30.7	20.5
50	TF39	1st	T	4	2.5	14.7	11.8
51	TF39	1st	T	5	4.3	33.5	26.4
52	TF39	1st	T	6	4.7	33.9	27.7
53	TF39	1st	T	1	6.4	52.2	30.7
54	TF39	1st	T	2	2.7	18.2	12.4
55	TF39	1st	T	4	3.8	29.7	23.9
56	TF39	1st	T	5	12.7	106.5	65.9
57	TF39	1st	T	6	11.5	103.1	68.5
58	TF39	2nd	NEW	1	1.7	13.0	9.8
59	TF39	2nd	A	1	5.0	37.8	26.1

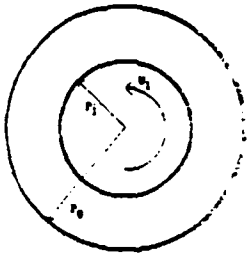


Figure 1. Schematic of circular couette flow

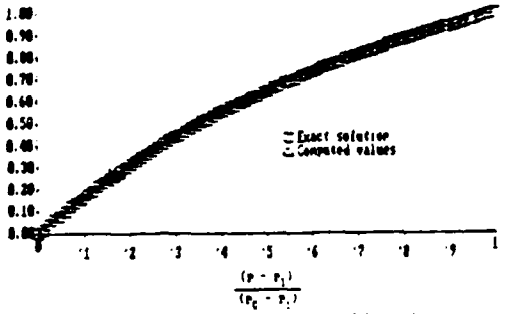


Figure 2. Nondimensional angular velocity profiles -- laminar

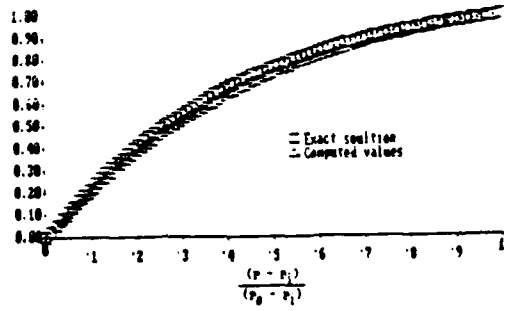


Figure 3. Nondimensional temperature profiles -- laminar

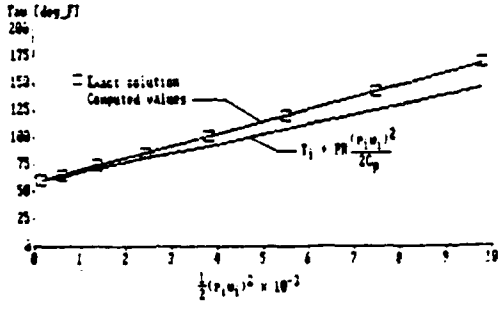


Figure 4. Comparison of laminar adiabatic wall temperature results

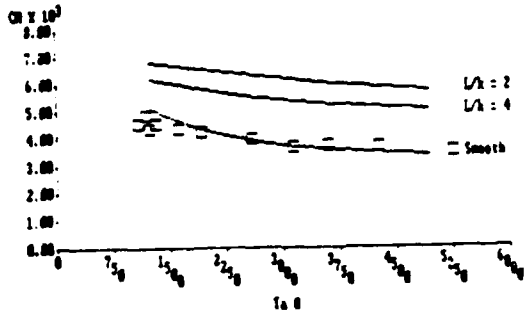


Figure 5. Computed moment coefficient and smooth data of G.I. Taylor

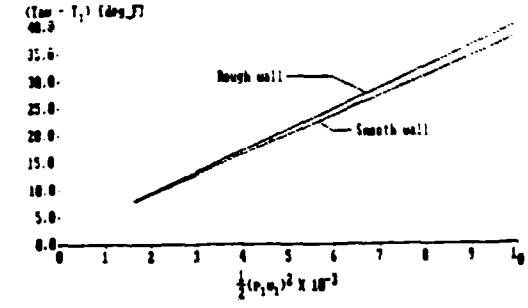


Figure 6. Adiabatic wall temperature for hydraulic oil

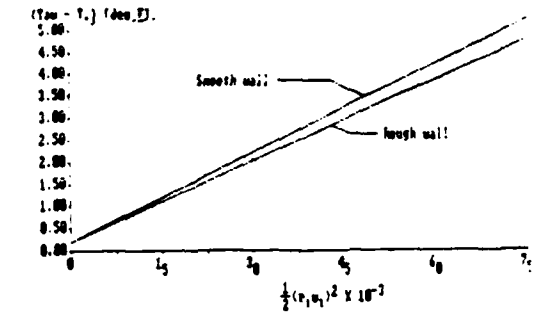


Figure 7. Adiabatic wall temperature for water

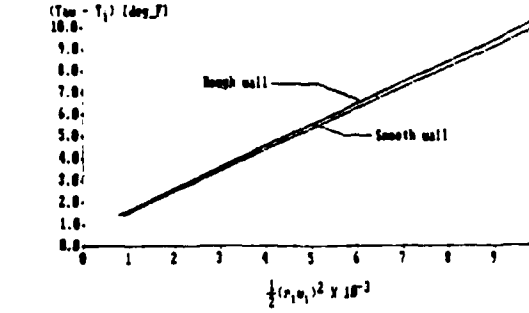


Figure 8. Adiabatic wall temperature for light oil

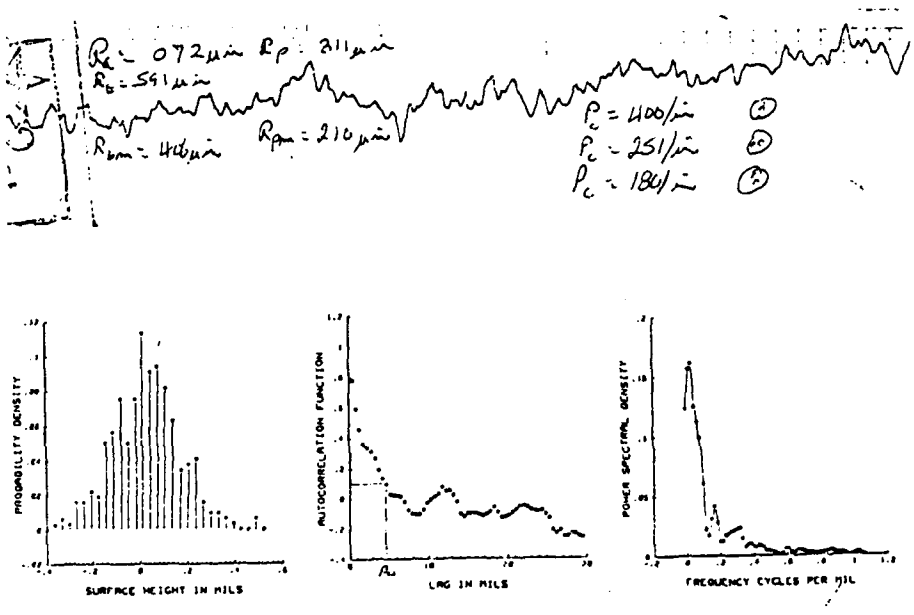


Figure 9. Profile trace and statistical data for blade T-N4

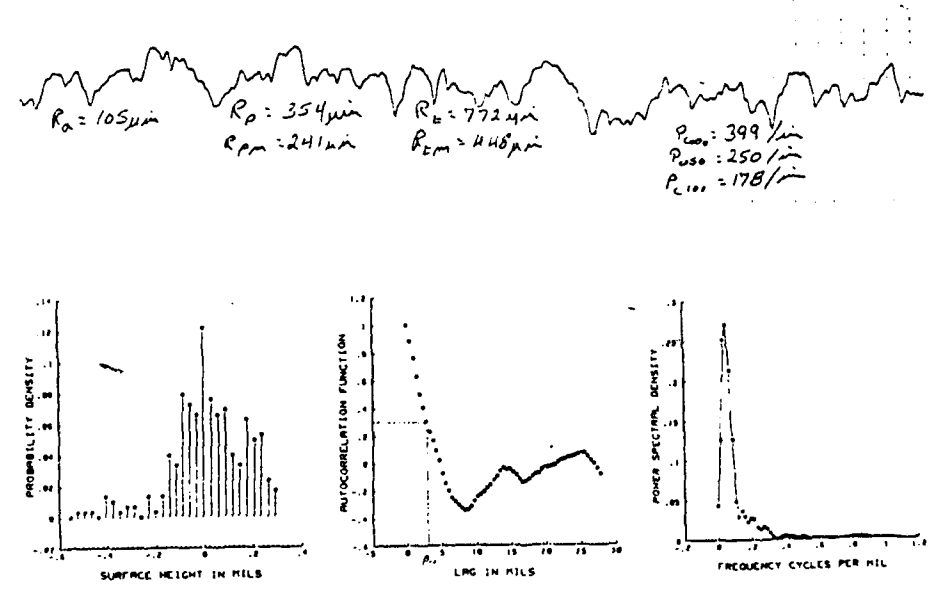
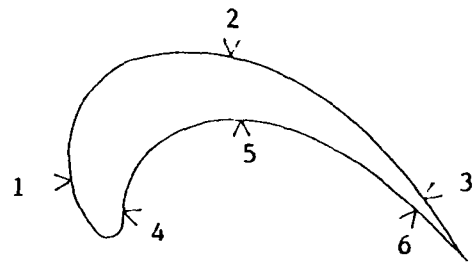


Figure 10. Profile trace and statistical data for vane V-002

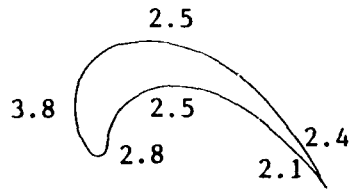


a) Approximate trace location

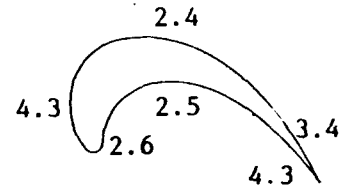
SYMBOL	CYCLES	SYMBOL	CYCLES	SYMBOL	CYCLES
◊	0-NEW	B	1200-1400	C	2600-3000
?	UNKNOWN	B	1400-1600	D	3000-3200
A	0-200	B	1600-1800	D	3200-3400
A	200-400	B	1800-2000	D	3400-3600
A	400-600	C	2000-2200	D	3600-3800
A	600-800	C	2200-2400	D	3800-4000
A	800-1000	C	2400-2600		
a	1000-1200	C	2600-2800		

b) Key for number of cycle symbols

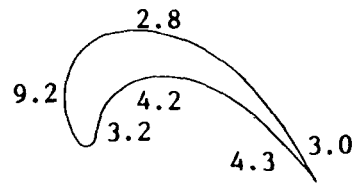
Figure 11. Nomenclature used in Table 1



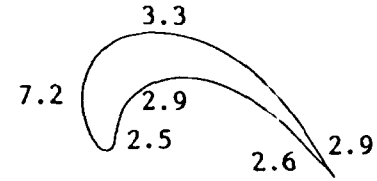
a) F-100 200-400 cycles



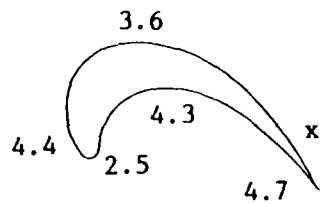
b) F-100 600-800 cycles



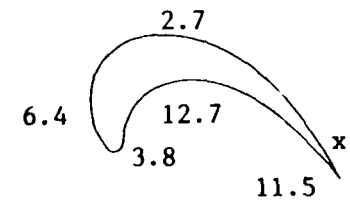
c) F-100 800-1000 cycles



d) F-100 1600-1800 cycles - shower head



c) TF-39 unbroken cycles



f) TF-39 unbroken cycles

Figure 12. Average roughness summary. Numbers indicate average roughness in μm

Continuity:

$$\frac{\partial}{\partial x}\left[\rho(1-\alpha_x)U\right] + \frac{\partial}{\partial y}\left[\rho(1-\alpha_y)V\right] = 0$$

x-Momentum:

$$(1-\alpha_x)\rho U\frac{\partial U}{\partial x} + (1-\alpha_y)\rho V\frac{\partial U}{\partial y} = \frac{\partial}{\partial x}\left[(1-\alpha_x)P\right] + \frac{\partial}{\partial y}\left[(1-\alpha_y)\left(\mu\frac{\partial U}{\partial y} - \rho\overline{u'v'}\right)\right] - \frac{\rho U^2 N_x(0)}{2L_xL_z}\sum_{i=1}^{N_z(y)} C_{Di}d_i(y)$$

Energy:

$$(1-\alpha_x)\rho U\frac{\partial H}{\partial x} + (1-\alpha_y)\rho V\frac{\partial H}{\partial y} = \frac{\partial}{\partial y}\left[(1-\alpha_y)\left(\frac{K}{C_p}\frac{\partial H}{\partial y} - \rho\overline{v'h'}\right)\right] + U\frac{\partial}{\partial x}\left[(1-\alpha_x)P\right] + (1-\alpha_y)\overline{\Phi} - \frac{\rho U^3 N_x(0)}{2L_xL_z}\sum_{i=1}^{N_z(y)} C_{Di}d_i(y) + \frac{\pi K N_x(0)\,(T_R - T)}{L_xL_z}\sum_{i=1}^{N_z(y)} Nu_{di}$$

$$\alpha_x = \alpha_y = \frac{\pi N_x(0)}{4L_xL_z}\sum_{i=1}^{N_z(y)} d_i^2(y)$$

Figure 13. Discrete element boundary layer equations

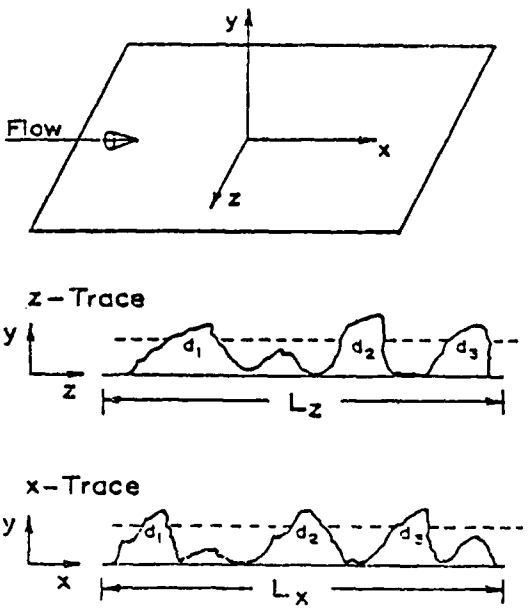


Figure 14. Profilometer trace nomenclature

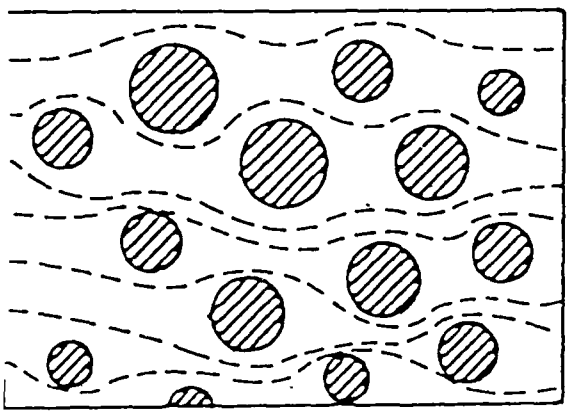


Figure 15. Statistical reproduction of element diameter distribution

1986 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT SUMMER SUPPORT PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by

Universal Energy Systems, Inc.

FINAL REPORT

Atmospheric Modeling for Operational Tactical Decision Aid

Prepared by:	Ken Tomiyama
Academic Rank:	Assistant Professor
Department and	Electrical Engineering
University:	Pennsylvania State University
Research Location:	Air Force Wright Aeronautical Laboratory
	Wright-Patterson Air Force Base, Ohio
USAF Researcher:	Roger L. Cranos
Date:	September 10, 1986
Contract Number:	F49620-85-C-0013

ATMOSPHERIC MODELING FOR OPERATIONAL TACTICAL DECISION AID

by

Ken Tomiyama

ABSTRACT

The Tactical Decision Aid (TDA) is an integrated target/atmosphere/sensor model that is used to estimate target acquisition ranges for infrared sensors. It employs an extensive 8000 plus-line computer code, LOWTRAN, to evaluate the atmospheric extinction of infrared signals for various climatological conditions. The Operational TDA (OTDA) is a simplified version of the TDA housed on a hand-held computer. It is intended for field use. Since LOWTRAN is too voluminous to be employed for the OTDA, pre-computed extinction data tables are currently in use. The process of manual input of data from the tables to the OTDA is cumbersome and is prone to erroneous readings. Therefore, it is proposed to develop compact atmospheric extinction models which may be coded onto hand-held computers. Various types of atmospheric extinction, which are significant for the TDA application, were considered and simple models were developed based on the LOWTRAN computation.

Acknowledgements

I would like to express my gratitude to the Air Force Systems Command and the Air Force Office of Scientific Research for sponsoring my appointment as a Summer Faculty Research Fellow. I would also like to acknowledge the Air Force Wright Aeronautical Laboratory (AFWAL), especially the Avionics Laboratory (AFWAL/AARI), for providing me with the opportunity to spend a fruitful summer at the laboratory. I would also like to thank members of the laboratory staff for their help in countless ways; first and most importantly, Roger L. Cranos and Don Tomlinson for the guidance necessary for my research, Jim Stewart for approving my stay at the Laboratory, Jan Servaites for putting up with noisy office mates, Bill Lanich and Ron Gruenzel for enlightening discussions, Jeff (J.D.) Sweet for his patience in answering numerous silly questions concerning the computer system, Frank Jenks for his assistance in running LOWTRAN6, Major Mike Abel and Major Bob Hughes for advise on meteorological data, Bill (Willie) Martin and Dan Zimmerman for introducing me to a splendid local cuisine, Second Lieutenant Pat Marshall for sharing his IBM compatible PC, and all other members for making my stay at the Laboratory successful and enjoyable.

I. Introduction

I received my Ph.D. from the Department of System Science at the University of California, Los Angeles. Then, I joined the University of Texas at El Paso, and compiled extensive research experiences in the areas of atmospheric extinction modeling and target detectability of infrared sensors. I moved to The Pennsylvania State University, and continued with research on various system modeling problems.

The project sought at the AFWAL/AARI originally involved systematic error budgeting of the Targeting Systems Characterization Facility. The combination of my system scientific background and experience with infrared sensor studies resulted in my initial selection. However, discussions with scientists at AFWAL/AARI resulted in the assignment to develop compact models for atmospheric extinction in relation to the Operational Tactical Decision Aid. The new project is more directly related to my background.

II. Objectives of the Research Effort

The overall project at AFWAL/AARI-3 that I was involved with was the study of the Tactical Decision Aid (TDA) which evaluates target detection characteristics of infrared sensors, and is used in tactical decision making. It employs an extensive 8000 plus-line computer code, LOWTRAN, (Kneizys 1980, Kneizys 1983) to evaluate the atmospheric extinction of infrared signals for various climatological conditions. The operational TDA (OTDA) is a compact TDA, which is housed on a hand-held computer, and is intended for field use. There exists a

need for compact analytical extinction models for use in the OTDA to replace the error-prone manual reading of the pre-computed extinction tables. My individual objectives are to:

1. Identify various extinction mechanisms which are active in the spectral region of interest, 8 - 12 (μm), for the TDA application.
2. Develop compact analytical models for all active extinction mechanisms, based on the LOWTRAN computation.
3. Evaluate the accuracy of the developed models.

The code, LOWTRAN, will be extensively studied for accomplishing these objectives.

III. Background

The Operational TDA (OTDA) cannot employ LOWTRAN for extinction computation due to its size. Instead, pre-computed extinction data tables are currently used. However, it is inconvenient to carry the printed tables the to field, and use of the tables is prone to erroneous readings. It is preferred to have an extinction computation program as a part of the OTDA. Thus, the development of such a program was proposed. As an initial step towards this goal, compact analytical models for various extinction mechanisms will be developed here based on the LOWTRAN computations.

It is noted that the quantity to be modeled is an average transmittance over 830 - 1250 (cm^{-1}), corresponding to 8 - 12 (μm) band, which is the primary band of sensitivity for infrared sensors. It was assumed that optical path between the

sensor and target can be considered horizontal and that the paths are located below 2 km altitude. It was decided to use the altitude of 300 (m), which is the altitude of the sensor test cite at AFWAL/AARI, as a representative altitude.

IV. Atmospheric Extinction Models

IV.1 Introduction

Infrared radiation passing through the atmosphere loses its intensity as a result of interactions with atmospheric constituents. This process is characterized by the extinction coefficient k, or the atmospheric transmittance t, which are defined as,

$$t = \exp\{-k\} = \frac{I(\text{emitted})}{I(\text{received})}, \qquad (1)$$

where I(emitted) and I(received) are the emitted and received intensities of electromagnetic wave, respectively.

The extinction coefficient k consists of contributions from many extinction mechanisms including; molecular resonant absorptions, molecular continuum absorptions, molecular scattering, aerosol (including fog) absorption and scattering, and rain absorption. The most recent version of LOWTRAN, LOWTRAN6, also includes cirrus cloud extinction.

LOWTRAN computes LOW resolution TRANsmittance, called band transmittance, which is basically an averaged transmittance over a small wavenumber interval. It adopts a basic assumption of the law of superposition, where the total extinction is the sum of individual contributions. This assumption enables us to

deal with each extinction mechanism separately.

As LOWTRAN suggests, some of the extinction mechanisms are inactive in the wavenumber region of interest, 830 - 1250 (cm^{-1}). Only the extinction due to the following need to be considered over this spectral region; water vapor, uniformly-mixed gasses, ozone, water vapor continuum, aerosol, and rain.

The OTDA requires the evaluation of the average transmittance over the 8-12 (μm) wavelength band. Therefore, the extinction models will be developed to represent the relationships between the average transmittance and various climatological conditions, including the optical path length.

IV.2 Molecular Resonant absorptions

In LOWTRAN, resonant absorptions due to the three gasses of concern are computed utilizing two intermediate quantities; called an absorber amount U and an equivalent absorber amount x, and two empirical transmittance models; one for ozone and another for both water vapor and uniformly-mixed gasses.

Transmittance profiles for the three absorbers over 830 - 1250 (cm^{-1}) band were generated, using LOWTRAN6, at 5 (cm^{-1}) intervals for various atmospheric conditions. Then the resulting profiles were averaged and stored together with the atmospheric conditions. Finally, analytical models were developed based on the LOWTRAN expressions and were curve-fitted, in an optimal manner, to the relationships of the average transmittances versus atmospheric conditions.

IV.2.1 Water Vapor

The transmittance expression used in LOWTRAN for a hori-

zontal path with homogeneous meteorological conditions of pressure P (mbar), temperature T (K), relative humidity RH (%), and path length R (km) at wavenumber v (cm^{-1}) is as follows.

$$t = f(x), \tag{2-a}$$

$$x = C(v)\ PN^{a}\ TN^{b}\ U, \tag{2-b}$$

$$PN = P/P_o, \quad TN = T_o/T, \tag{2-c}$$

$$U = 0.1\ WH\ R, \quad WH = 0.01\ RH\ F(T_o/T), \tag{2-d}$$

where f(.), C(v), a, b, PN, TN, P_o, T_o, WH, and F(.) are the empirical transmittance function, spectral parameter, absorber parameters (a = 0.9, b = 0.45), normalized pressure, normalized temperature, standard pressure (1013.25 mbar), standard temperature (273.15 K), water vapor density (g/m^3), and an empirical function for saturated water vapor density (g/m^3) at temperature T, respectively. The values of C(v) for each absorber, are stored in LOWTRAN at 5 (cm^{-1}) intervals over absorption bands where the absorption is non-trivial.

In earlier efforts to model the molecular resonant absorption (Gruenzel 1978, Pierluissi 1979), it was found that the following double exponential function agrees excellently with LOWTRAN empirical transmittance functions.

$$t = \exp\{-10^{\,a_o + a_1 x}\}, \tag{3-a}$$

$$x = \log C(v) + n \log(PN) + m \log(TN) + \log(U), \tag{3-b}$$

where a_o, a_1, n, and m are model parameters to be chosen.

This function was selected as our model since it agrees well with the band transmittance, which is a weighted average of transmittances. It is noted that the spectral parameter C(v) in this expression may be eliminated in our model since only an averaged transmittance is to be calculated. As a result, the model can be simplified to

$$t = \exp\{-10^{a_0 + a_1 \log(PN) + a_2 \log(TN) + a_3 \log(U)}\}, \quad (4)$$

or

$$t = \exp\{- A_0 PN^{a_1} TN^{a_2} U^{a_3}\}, \quad (5)$$

where a_0, a_1, a_2, a_3, and $A_0 = 10^{**}a_0$ are the model parameters.

For the optimal determination of the model parameters, we take the double logarithm of Eq. (4). This linearizes the model in terms of the unknown parameters as follows:

$$\log\{-\ln(t)\} = a_0 + a_1 \log(PN) + a_2 \log(TN) + a_3 \log(U). \quad (6)$$

Linear regression techniques may now be utilized to obtain the optimal parameter values.

IV.2.2 Uniformly-Mixed Gasses

The absorber in question here is a mixture of various atmospheric gaseous molecules whose density profiles are relatively unperturbed, except for the pressure and temperature dependencies. Therefore, the corresponding absorber amount U is a function of the pressure, temperature, and the path length only. The transmittance expression for this absorber is the same as that for the water vapor given in Eq. (2), except that

the pressure and temperature dependencies within the absorber amount U is integrated into those appearing in x. This leads to the following LOWTRAN model.

$$t = f(x), \tag{7-a}$$

$$x = C(v)\ PN^{a}\ TN^{b}\ U, \tag{7-b}$$

$$U = R, \tag{7-c}$$

where the absorber parameters a and b have values 1.75 and 1.375, respectively.

Thus, an appropriate model is again Eq. (4) or (5) with the expression for U being replaced by the path length R.

IV.2.3 Ozone

The transmittance expression for ozone is the same as that for the water vapor, except for the absorber parameter values, a = 0.4 and b = 0.2, and the expression for the absorber amount U. With the ozone density WO in g/m^3, U is given by

$$U = 46.667\ WO\ R. \tag{8}$$

Therefore, the appropriate model expression is again given by Eqs. (4) or (5) together with the absorber amount expression in Eq. (8).

IV.3 Water Vapor Continuum Absorption

The LOWTRAN6 expression for the water vapor continuum absorption consists of self and foreign components. The expression for a homogeneous path is given by,

$$t = \exp\{-v\ \tanh(hcv/2kT)\ [Rs\ Cs + Rf\ Cf]\ WH\ R\}, \tag{9}$$

where hc/k = 1.43879 (K/cm^{-1}), Rs and Rf are self (water vapor versus total air at standard condition) and foreign (all other molecular species versus total air) number density ratios, and Cs and Cf (1/(cm^{-1} mol/cm^2)) are wavenumber dependent parameters for self and foreign components, respectively.

The temperature dependence of the self component Cs is taken into account through a factor Kp defined by,

$$Kp = \begin{cases} 1, & T < 260, \\ (296 - T) / (296 - 260), & 260 < T < 296, \\ 0, & 296 < T, \end{cases} \qquad (10)$$

as

$$Cs = (1 - Kp)\, Cs_1 + Kp\, Cs_2. \qquad (11)$$

The parameters Cs_1, Cs_2, and Cf are stored in LOWTRAN at 10 (cm^{-1}) wavenumber intervals over regions where the water vapor continuum absorption is non-trivial.

It is noted that the number density ratio Rs is linearly dependent on the water vapor concentration WH. On the other hand, the sum of the two densities, water vapor and all others, is linearly dependent on the product PN*TN since it is the air density. Therefore, our model needs to carry linear dependencies on PN*TN and WH in an additive fashion. It is also noted that the wavenumber dependent coefficient in Eq. (9) can be imbedded into Cs and Cf. By combining all of these observations, we obtain the following expression:

$$t = e^{-[\{q(Cs_1'+Cf') + Kp(Cs_2'-Cs_1')\}\, WH + r\, Cf'PN*TN]\, WH\, R}, \qquad (12)$$

where q and r are wavenumber independent constants, and Cs_1', Cs_2', and Cf' are scaled wavenumber dependent parameters.

For our model, the averages of Cs_1', Cs_2', and Cf' over 830 - 1250 (cm^{-1}) region are computed from the LOWTRAN data, and then those parameters in Eq. (12) are replaced with respective averages. Then the expression is simplified by combining the constants as,

$$t = \exp\{ -C_0 [PN\ TN + (C_1\ Kp + C_2)\ WH]\ WH\ R\}, \qquad (13)$$

where C_0, C_1, and C_2 are the final model parameters.

IV.4 Aerosol Extinction

Aerosols are active over 830 - 1250 (cm^{-1}) region in both absorption and scattering. Since LOWTRAN has models for the extinction, we will consider the modeling of the extinction, instead of the absorption and scattering individually.

The transmittance due to aerosols is again given by an exponential law as,

$$t = \exp\{-X\ H\ R\}, \qquad (14)$$

where X is the aerosol extinction profile which is dependent on the type of aerosol, the relative humidity RH and the wavelength . H is the aerosol density profile which represents the visibility and the height dependencies.

We first consider the aerosol extinction profile X. There are ten aerosol types used in LOWTRAN. Due to our assumption that applications of the OTDA is limited to horizontal paths below 2 km altitude, we only need to consider four humidity

dependent aerosols; RURAL, URBAN, MARITIME, TROPOSPHERE, and two humidity independent ones; FOG1 and FOG2. LOWTRAN stores four extinction profiles X( ) for each humidity dependent aerosol corresponding to the relative humidities of 0, 70, 80, and 99 (%), and one each for FOG1 and FOG2. These profiles are first averaged to eliminate wavelength dependence. Then the humidity dependencies in four aerosols are modeled using the following empirical relationship which is suggested in Shetlle 1979, based on the observation by Hanel (Hanel 1976),

$$X = c_1 (1 - RH/100)^{c_2}, \tag{15}$$

where c_1 and c_2 are model parameters. Noting that this relationship represents a straight line in log-log scale, optimal values for these parameters were obtained using the linear regression technique. The same model is also used for the two humidity independent models, FOG1 and FOG2, by setting c_2 to 0 to eliminate humidity dependence.

Next, the visibility dependent aerosol density profiles H(.) are studied. In the first 2 km height, H(.) is represented by three empirical functions of the visibility at 0, 1, and 2 (km) altitudes. These three profiles are fitted by the inverse relationship,

$$H(VIS) = d_1 VIS^{-1} + d_2, \tag{16}$$

which is used in LOWTRAN for interpolation of H(.). Then, using the assumption that a typical altitude at which the OTDA is applied is 300 (m), the weighted average of the two profiles

at 0 and 1 (km) heights is adopted as our model.

Finally, Eqs. (14), (15) and (16) are combined to form the following aerosol model,

$$t = \exp\{-(VIS^{-1} + d_2')\, c_1'(1 - RH/100)^{c_2} R\}, \qquad (17)$$

where d_1 is imbedded into c_1' by factoring it out to reduce the number of parameters. It is noted that d_2' is independent of the aerosol type.

IV.5 Rain Model

Because the rain extinction model used in LOWTRAN is a simple analytic function of the rain rate RR (mm/hr) and the range R (km), we can adopt it with a slight modification. After combining some constants to minimize the number of parameters, the model becomes as follows:

$$t = \exp\{-0.3647\, RR^{0.63} R\}. \qquad (18)$$

V. Summary of Numerical Results

The model equations derived in the previous section are summarized in Table 1 together with the obtained optimum parameter values. The input variables are listed in Table 2 together with the default values for some of them. A simple FORTRAN program, called CTRAN, was coded to check the accuracy of the obtained model. Results of this preliminary error analysis are summarized in Table 3. Figure 1 illustrates relative humidity dependence of four aerosol extinctions and their models given in Eq. (15).

Table 1.

Extinction Models for the OTDA

(1) Molecular resonant absorption

$$t = \exp\{- A_o (P/P_o)^{a_1} (T_o/T)^{a_2} U^{a_3}\}$$

Water Vapor: $A_o = 0.0850$, $a_1 = 0.4981$, $a_2 = 0.2989$, $a_3 = 0.5582$, $U = 0.1\ WH\ R$, $WH = 0.01\ RH\ F(T_o/T)$

Uniformly-mixed Gasses: $A_o = 0.0118$, $a_1 = 1.0792$, $a_2 = 0.8488$, $a_3 = 0.6178$, $U = R$

Ozone: $A_o = 0.0076$, $a_1 = 0.3091$, $a_2 = 0.1541$, $a_3 = 0.7498$, $U = 46.667\ WO\ R$

,where $F(s) = s \exp\{18.9766 - 14.9595\ s - 2.43882\ s^2\}$.

(2) Water Vapor Continuum Absorption

$$t = \exp\{ -C_o [(P/P_o)(T_o/T) + (C_1 Kp + C_2) WH] WH U\}$$

$$Kp = \begin{cases} 1, & T < 260 \\ (296 - T)/(296 - 260), & 260 < T < 296 \\ 0, & 296 < T \end{cases}$$

$C_o = 1.655E\text{-}03$ $C_1 = 0.5693$ $C_2 = 0.5437$

(3) Aerosol Extinction

$$t = \exp\{-(VIS^{-1} - 0.005183)\ c_1'(1 - RH/100)^{c_2} R\}$$

#	Model	c_1'	c_2	Default VIS(km)
1	RURAL	0.3670	-0.02877	23
2	URBAN	0.3119	-0.08499	5
3	OCEAN	0.4013	-0.3417	23
4	TROPOSPHERIC	0.08054	-0.04621	50
5	FOG1	4.487	0	0.2
6	FOG2	1.309	0	0.5

(4) Rain Extinction

$$t = \exp\{-0.3647\ RR^{0.63} R\}$$

Table 2.

Input Variables for Extinction Models

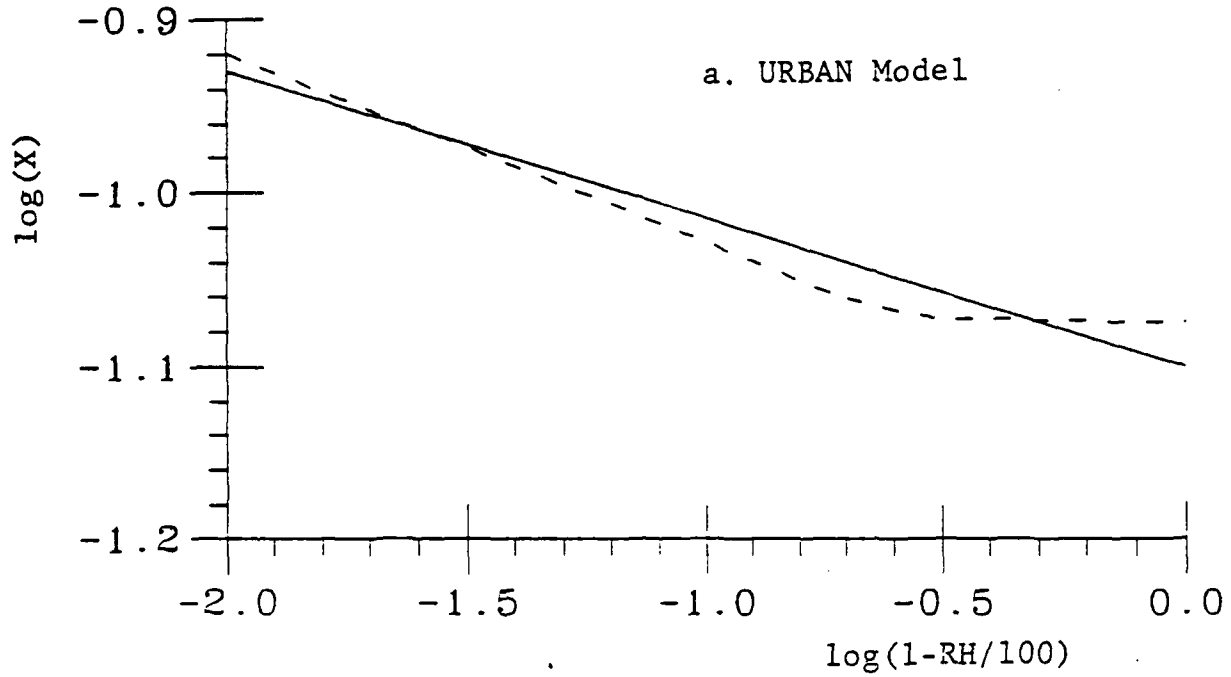
Variable	Notation (Units)	Default
Pressure	P (mbar)	None
Temperature	T (°C)	None
Relative Humidity	RH (%)	None
Ozone Density	WO (G/m^3)	6.0E-05
Visibility	VIS (km)	*
Aerosol Model	I (integer)	0
Rain Rate	RR (mm/h)	0.0
Range	R (km)	None

* See Table 1 for model dependent default values.

Table 3.

Preliminary Error Analysis of Extinction Models

Model		R.M.S. Error
Water Vapor		0.0038
Uniformly-Mixed Gasses		0.0002
Ozone		0.0002
Water Vapor Continuum		0.0056
Aerosol	RURAL	0.0126
	URBAN	0.0427
	OCEAN	0.0172
	TROPOSPHERIC	0.0017
	FOG1	0.0011
	FOG2	0.0183



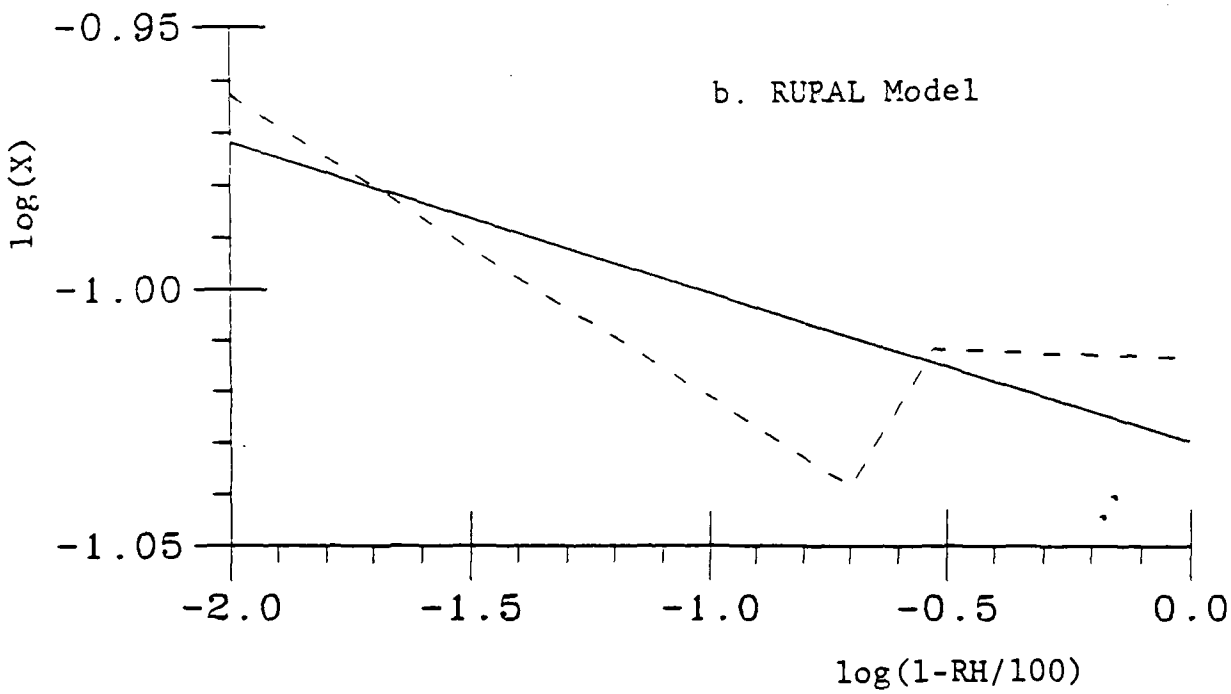
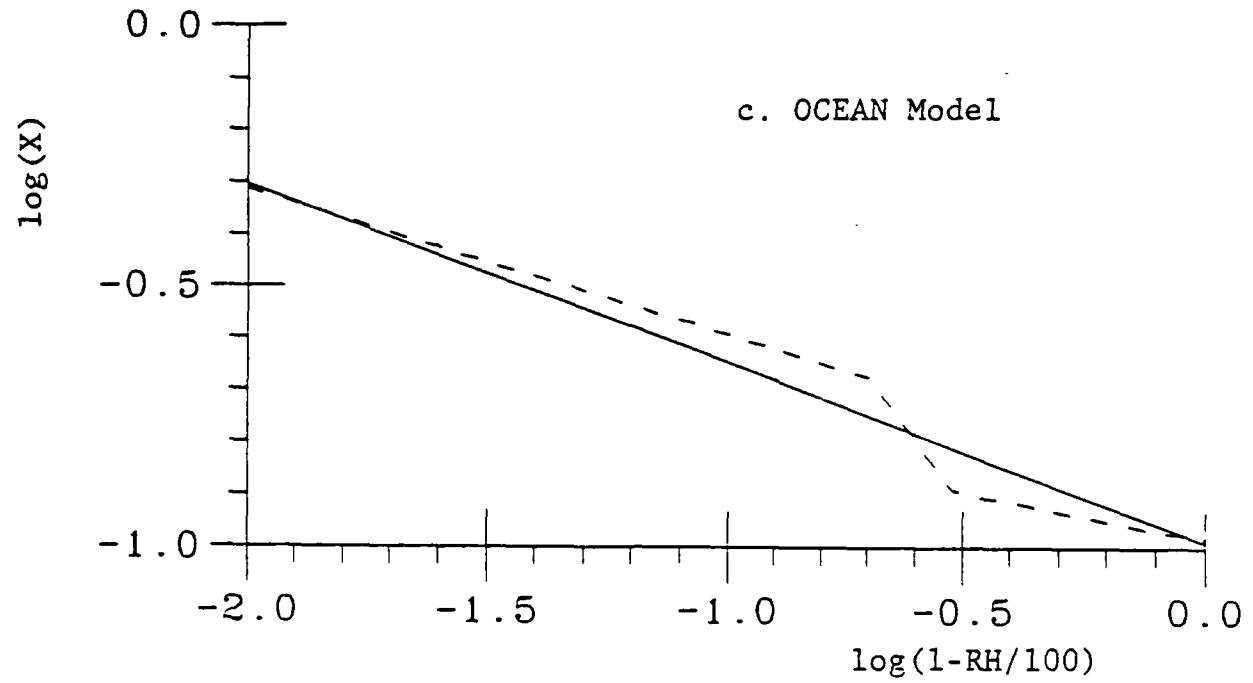


Figure 1. Relative humidity (RH) dependent aerosol extinction profile X. The dashed and solid lines are for the original model for LOWTRAN and the derived model for the OTDA, respectively.



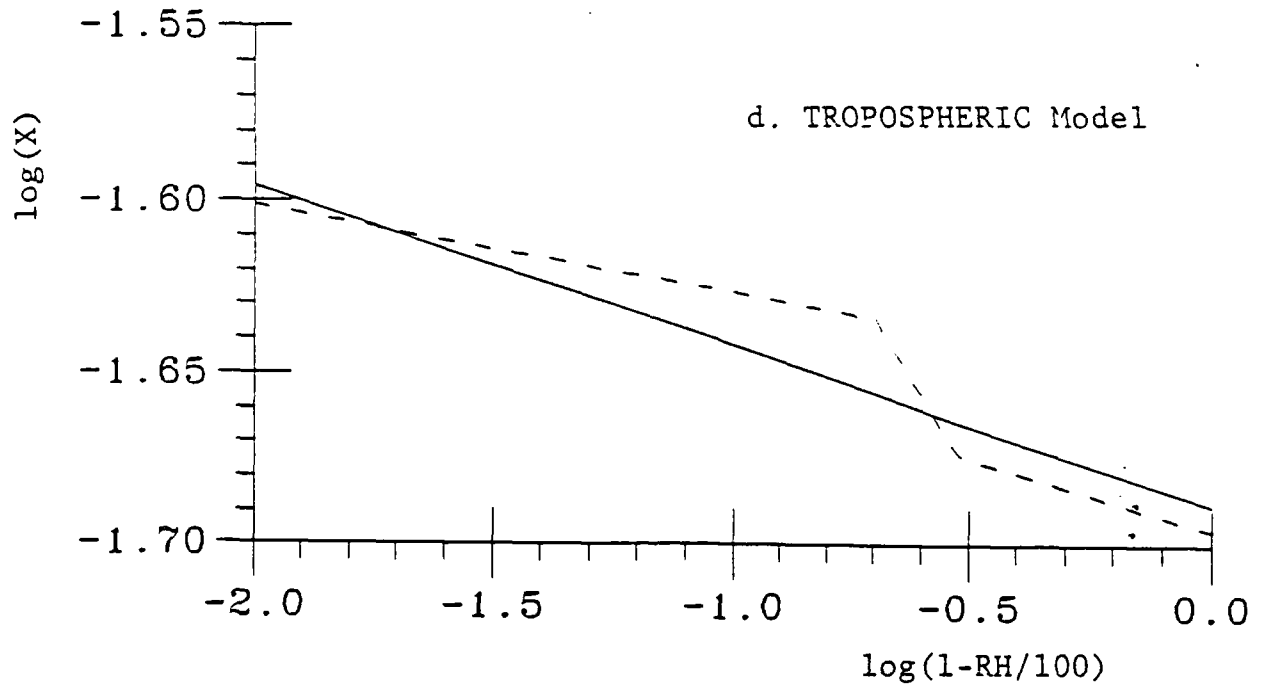


Figure 1. Continued.

VI. Conclusions

Simple models for atmospheric extinctions due to various atmospheric absorption mechanisms are developed for the use in the Operational Tactical Decision Aid. The developed models are; for three molecular resonant absorptions due to water vapor, uniformly-mixed gasses, and ozone; for water vapor continuum absorption; for aerosol extinction; and for rain extinction. All of those absorption mechanisms are active in the 8 - 12 (μm) band. Preliminary error analysis showed good overall accuracy and excellent accuracy for molecular resonant absorption models.

VII. Recommendations

1. Although the developed models generally agreed well with LOWTRAN computations, they need to be fully tested against all conceivable atmospheric conditions, including a range of possible optical path lengths. This may results in modifications of the obtained models.
2. After thorough testing, the developed models should be implemented on the OTDA to replace extinction data tables.
3. One of the perpetual question associated with the capability of a program like the TDA is the sensitivity of the output against the perturbation in the input variables. Regardless of the accuracy of the computation, the output will be in error if the input is disturbed. Therefore, in order for the TDA to be truly effective, it is highly recommended to perform a sensitivity analysis.

REFERENCES

Kneizys, F.X., E.P. Shettle, W.O.Gallery, J.H. Chetwynd, Jr., L.W. Abreu, J.E.A. Selby, R.W. Fenn, and R.A. McClatchey. (1980) "Atmospheric Transmittance/Radiance: Computer Code LOWTRAN 5," Report AFGL-TR-80-0067, Air Force Geophysics Laboratory, Hanscom Air Force Base, Mass.

Kneizys, F.X., E.P. Shettle, W.O.Gallery, J.H. Chetwynd, Jr., L.W. Abreu, J.E.A. Selby, S.A. Clough, and R.W. Fenn. (1983) "Atmospheric Transmittance/Radiance: Computer Code LOWTRAN 6," Report AFGL-TR-83-0187, Air Force Geophysics Laboratory, Hanscom Air Force Base, Mass.

Gruenzel, R.R. (1978) "Mathematical Expressions for Molecular Absorption in LOWTRAN 3B," Applied Optics, 17, 2591-2593.

Pierluissi, J.H., K. Tomiyama, and R.G. Gomez. (1979) "Analysis of LOWTRAN Transmission Functions," Applied Optics, 18, 1607-1612.

Shettle, E.P., and R.W. Fenn. (1979) "Models for the Aerosols of the Lower Atmosphere and the Effects of Humidity Variations on Their Optical Properties," Report AFGL-TR-79-0214, Air Force Geophysics Laboratory, Hanscom Air Force Base, Mass.

Hanel, G. (1976) "The Properties of Atmospheric Aerosol Particles as Functions of the Relative Humidity at Thermodynamic Equilibrium with the Surrounding Moist Air," Advances in Geophysics, Ed. by H.E. Landsberg and J. Van Mieghem, Academic Press, New York, 19, 73-188.

1986 USAF-UES Summer Faculty Research Program/

Graduate Student Summer Support

Sponsored by the

Air Force Office of Scientific Research

Conducted by the

Universal Energy Systems, Inc.

Final Report

Vigilance behavior of military personnel: A study of individual differences

Prepared by: Phillip D. Tomporowski

Academic Rank: Assistant Professor

Department and University: Department of Psychology, University of Alabama

Research Location: Air Force Human Resources Laboratory, Brooks AFB, Texas

Date: September 15, 1986

Contract No.: F49620-85-C-0013

Vigilance Behavior of Military Personnel: A Study of Individual Differences

by

Phillip D. Tomporowski

Abstract

The sustained attention and cognitive abilities of 829 military personnel were assessed during a 3-hr test period. Testing included a 60-min visual vigilance task and seven tests of information processing. Subjects were assigned randomly to complete one of four vigilance tasks that differed in cognitive complexity. All subjects completed the seven information processing tests. Tests were administered via computer-controlled systems. In addition, subjects completed the Armed Forces Vocational Aptitude Battery (ASVAB). Initial data analysis revealed the vigilance abilities of observers were differentially influenced by the complexity of the task. Individual differences in vigilance behavior will be evaluated by correlating each subject's vigilance performance to measures of information processing abilities and general aptitude.

Acknowledgments

I would like to thank the Air Force Systems Command, Air Force Office of Scientific Research for sponsorship of my research. The Human Resources Laboratory provided a unique opportunity to conduct research that would be extremely difficult to perform in an academic setting. Further, the support and guidance provided by personnel at the Human Resources Laboratory was first rate. I would like to acknowledge the administrative assistance provided by Dr. Jeffrey Kantor and Dr. Raymond Christal and the technical support provided by Drs. Dan Woltz and Bill Tirre.

I. Introduction

At the University of Alabama my research focuses on vigilance behavior; that is, the ability of observers to maintain levels of sustained attention over long periods of time. While almost all people evidence a decrement in vigilance behavior over time, some people are better than others in maintaining levels of sustained attention. I am attempting to isolate basic cognitive processes believed to underlie sustained attention and to explain why individuals differ in their watchkeeping ability.

The analysis of basic cognitive abilities of personnel has been of interest to the Air Force for some time. The Learning Abilities Measurement Program (LAMP) within the Human Performance Laboratory at Brooks AFB evaluates the cognitive abilities of military personnel. My assignment at the Human Performance Laboratory was to assess the relation between information processing abilities of military personnel and their performance on vigilance tasks.

II. Objectives of the Research Effort

The objective of the research project was to evaluate the ability of military personnel to maintain levels of sustained attention to vigilance tasks that differ in complexity. At this time there is no general consensus

regarding the effects of altering task complexity on vigilance behavior. Depending upon the methodology employed, it is possible to reduce or exacerbate the vigilance decrement. In the present research, the cognitive complexity of the vigilance tasks was varied. Cognitive complexity refers to the degree to which observers are required to employ mental strategies or other decision making processes to perform the task.

It was predicted that the sustained attention of watchkeepers would vary as a function of task complexity as well as the information processing ability of the observer. That is, as the level of task complexity increased, the task would place a greater cognitive demand upon the watchkeeper. Individuals with superior information processing abilities were expected to meet the demands of the more complex vigilance task and they would perform the vigilance test more efficiently than observers with poor information processing abilities.

III. Method

To test these predictions four vigilance tests were developed. Each test presented visual background events (nontarget stimuli) and target stimuli, successively, at a high rate. The density of target stimuli (i.e., the probability of the presentation of a target) was 5 percent

for all tests. Presentation of all stimulus events were computer controlled (Terak-1850). Stimulus displays were presented on a CRT and subjects entered responses on a computer key board. The tests varied with respect to the type of background and target stimuli presented. One test was a sensory discrimination task in which the background events were 4.0 cm squares and the target events were 4.5 cm squares. The stimuli for the second test were digits from 1 to 9 presented randomly. The target event was the number 7 but only when it was preceded by the number 4. All other numbers constituted background events. The third task also presented digits; however, the digits were presented (from 1 to 9) serially and repetitively. On some occasions a number was missing (e.g., 1,2,3,..5,6,...). The target event for this test was the missing number while all other numbers constituted background events. The fourth task involved the presentation of pairs of digits. Target events were those pairs of digits that sum to a value of 11. All other digit pairs constituted background events.

Each test session consisted of two practice tests and the main vigilance test. Each section of the test was preceded by written instructions which appeared on the CRT screen. During the first 3-min practice session stimulus

events were presented on the CRT for 150 ms followed by a 2500 ms interstimulus interval (ISI). Target stimuli were presented on 25 percent of all stimulus events. Target stimuli appeared randomly during the practice session with the constraint that prohibited two successive target events. A key press during the 2650 ms period between the onset of the target stimulus and the onset of the next stimulus event was considered a "hit" while failure to respond constituted a "miss". A key press following any background event constituted a "false alarm". The initital practice session was designed to ensure the subject understood task requirements. Verification of their ability to discriminate background and target events were obtained by calculating a measure of perceptual sensitivity, P(A), which is obtained via Signal Detection Analysis (Green & Swets, 1966). A P(A) score of .80 was considered necessary before the subject was able to begin the second 3-min practice test. Task conditions for the second practice were identical to those the subject encountered in the main vigilance task. Stimuli were presented for 150 ms with a 1000 ms ISI. Target events occurred randomly on 5 percent of the total number of events. Each subject's hit and false alarm rate was calculated as well as their perceptual sensitivity P(A).

The third part of the test session was the main vigilance task. It was 60 minutes in length. Events were presented for 150 ms with a 1000 ms ISI. Target events were presented in a random fashion with the limitation that target events could not occur on two successive event presentations and at least one target cue would occur during a 60-sec period. Measures of each subject's hit rate, miss rate, false alarm rate, and perceptual sensitivity P(A) were calculated for each 10-min block and for the total 60-min test period. The average response time required of subjects to detect target events was also recorded for each block and for the entire task.

Cognitive Test Battery

Following the completion of the vigilance task, each subject performed seven tests of cognitive functioning. These tests were computer generated and presented to subjects in random order. The test battery was given following the subject's completion of the 60-min vigilance task and a 5-min rest period. The test battery included:

1) Simple reaction time. Each subject was asked to depress a key as rapidly as possible upon the onset of a visual stimulus on the computer screen. Subjects received 6 practice trials; 3 trials required responses with the

right forefinger and three trials with the left forefinger. The test was composed of 100 trials in blocks of 20 trials each. Cue onset varied randomly between 1.0 and 3.0. Subjects responded to cue signals with their right forefinger on 60 trials and with their left forefinger on 40 trials. Mean and median reaction times were calculated for left and right hand responses.

2) Posner Physical Match. Subjects were instructed to examine two letters presented simultaneously and to determine whether they were identical (e.g. AA) or different (e.g. Aa). If letter pairs were the same, the subject was instructed to press a key with their forefinger of their right hand, if the letters differed, a key press with the left hand was required. Subjects were asked to respond as quickly as possible. Three practice trials and 200 test trials were administered. Feedback was given following each response in the form of a "correct" or "incorrect" statement presented for 1 sec on the computer screen. A 500-ms interval was programmed between the offset of the feedback statement and the onset of the next test trial. Mean and median response times were calculated for all responses and for correct responses. The percent correct was also obtained.

3) Posner Name Match. Subjects were instructed to

examine two letters presented simultaneously and to determine whether they had the same name (e.g. AA or Aa) or different names (e.g. AB or Ab). Subjects were given 10 practice trials and 128 test trials. The method of stimulus presentation, response requirements, feedback, and data summarization was identical to the Posner Physical Match task.

4) Memory Span Test. A series of numbers were presented sequentially on the CRT from left to right. Each item was presented for 500 ms and separated by 250 ms (e.g. 8 ___ ___ ___ ___; ___ 4 ___ ___ ___, etc). List length varied in the number of items presented (5, 7, and 9 items). Each subject received a block of 8, 5-item lists, and a block of 10, 7 and 9-item lists. Immediately following the presentation of each list, subject were tested for their memory of each item of the list. On half of the tests probe items were presented starting from the first position of the list (e.g., 8 ___ ___ ___ ___; ___ 7 ___ ___; etc) and half presented from the last position of the list (e.g. ___ ___ ___ ___ 3; ___ ___ 2 ___; etc). The probe items were correct on half of the presentations. The subject's task was to depress a response key with the forefinger of their right hand if they considered the probe to be correct and to depress a

key with the forefinger of their left hand if the probe was considered incorrect. Subjects were given feedback in the form of a 250 ms statement of "correct" or "incorrect" on the CRT immediately following each response. Probe items were presented 500 ms following the offset of the feedback statement. The percent correct and mean and median reaction times for each list length was calculated for each subject.

5) Missing Digit Task. A series of 8 digits were presented successively, 250 ms each. The subject was instructed to detect and depress the key for the digit (1 - 9) not presented. A 250 ms feedback statement was given following each response. Lists were separated by a 2-sec interval and the onset of each list was cued by a 500 ms visual signal. The percent correct and mean and median response times were calculated for each subject.

6) ABC Test. Subjects were instructed to perform a two-stage, self-paced task that demanded the participant to study three numeric expressions (e.g. Y = 8 - 5, B = 4 + 5, H = 7 - 2) and then to answer three questions (e.g. Y = ?, B = ?, H = ?). The numeric expressions varied with respect to the number of computational steps required to solve the equation during the study phase of the test; items were grouped according to three levels of

computational difficulty (e.g., G = 9; G = 4 + 5; G = (8 -2) + 3). Subjects received 18 groups of numeric expressions at each level of difficulty. Measures of subject performance included mean and median response time for the study stage and test stage at each level of difficulty. Also, the mean percent correct during the test stage was calculated.

7) Remote Association. Fifty pairs of words were presented simultaneously to subjects. Participants were instructed to determine whether the words had meanings that were similar or different. Subjects depressed a response key with the forefinger of their right hand following the presentation of word pairs considered similar and they depressed a response key with the forefinger of their left hand when the word pair was considered different. Feedback in the form of a 250 ms statement of "correct" or "incorrect" followed each response. The percent correct and mean and median response terms were calculated for each subject.

Test of General Cognitive Ability

All subjects were required to complete the Armed Service Vocational Aptitude Battery (ASVAB) prior to induction into the military or soon thereafter. This paper and pencil aptitude test consisted of 10 separate

subtests: word knowledge (WK); mathematics knowledge (MK); paragraph comprehension (PC); numerical operations facility (NO); mechanical comprehension (MC); general science knowledge (GS); arithmetic reasoning skill (AR); coding speed (CS); automobile and shop knowledge (AI); and electronics information (EI). Results from these 10 tests were used to create four aptitude factors; general (G); administrative (A); mechanical (M); and electronic (E). The factors were derived via Conversion Tables for ASVAB (DOD 1304.12W - July 1986).

IV. Preliminary Results.

Data were collected from 829 Air Force enlistees in their sixth day of Basic Training. At this time only descriptive statistics of the subjects' performance on the four vigilance tasks are available. All other data are in the process of being collated and converted into computer compatible data sets.

Figures 1 through 4 provide graphic representations of the subjects' performance on the four vigilance tasks. A decrease in the number of targets detected across time was noted during the performance of the sensory discrimination task, digit identification task, and missing digit task (Figures 1, 2, and 3) an increase in the number of target stimuli detected during the

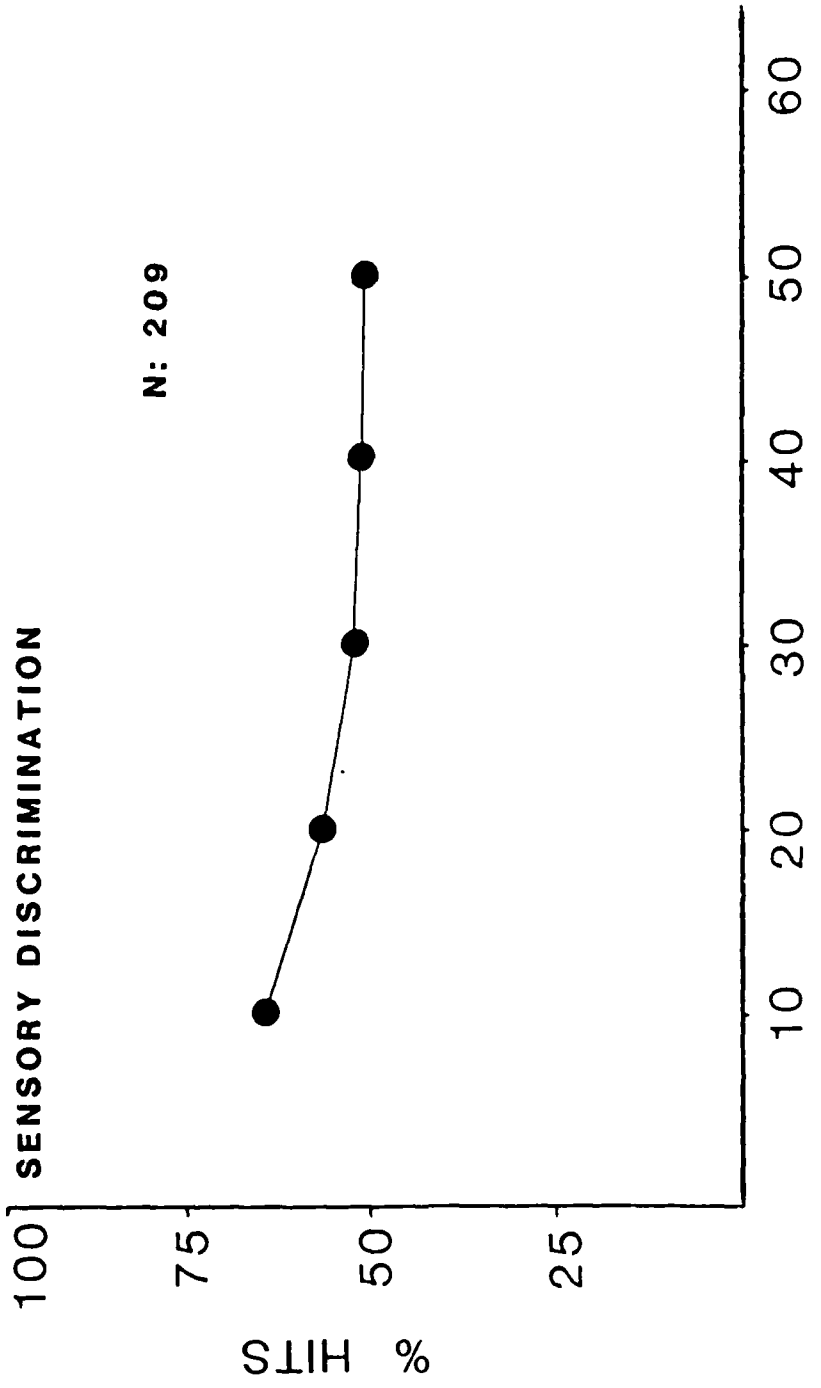


Figure 1. Percent correct detections of target stimuli during a 60-min test period

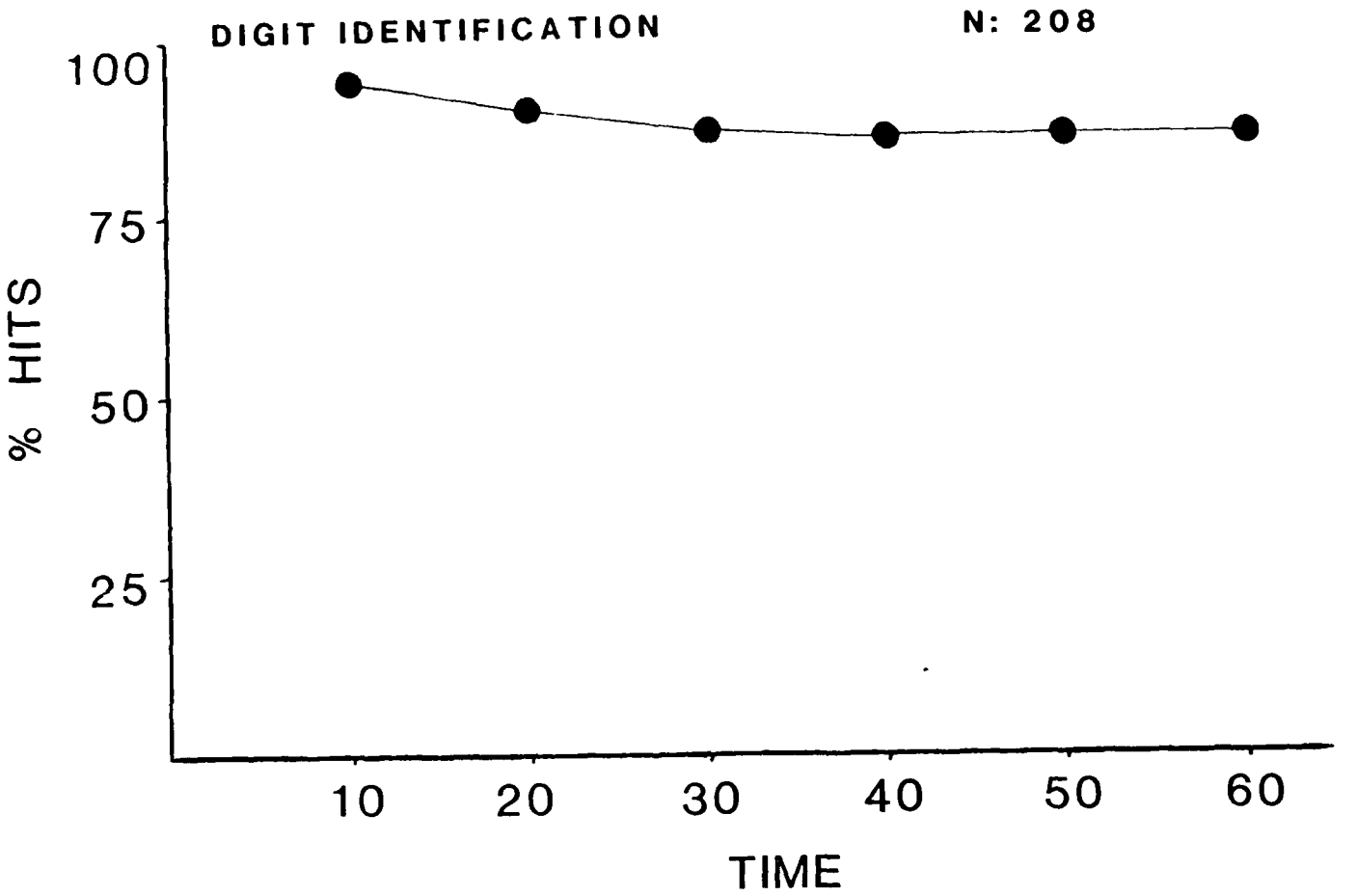


Figure 2. Percent correct detections of target stimuli during a 60-min test period

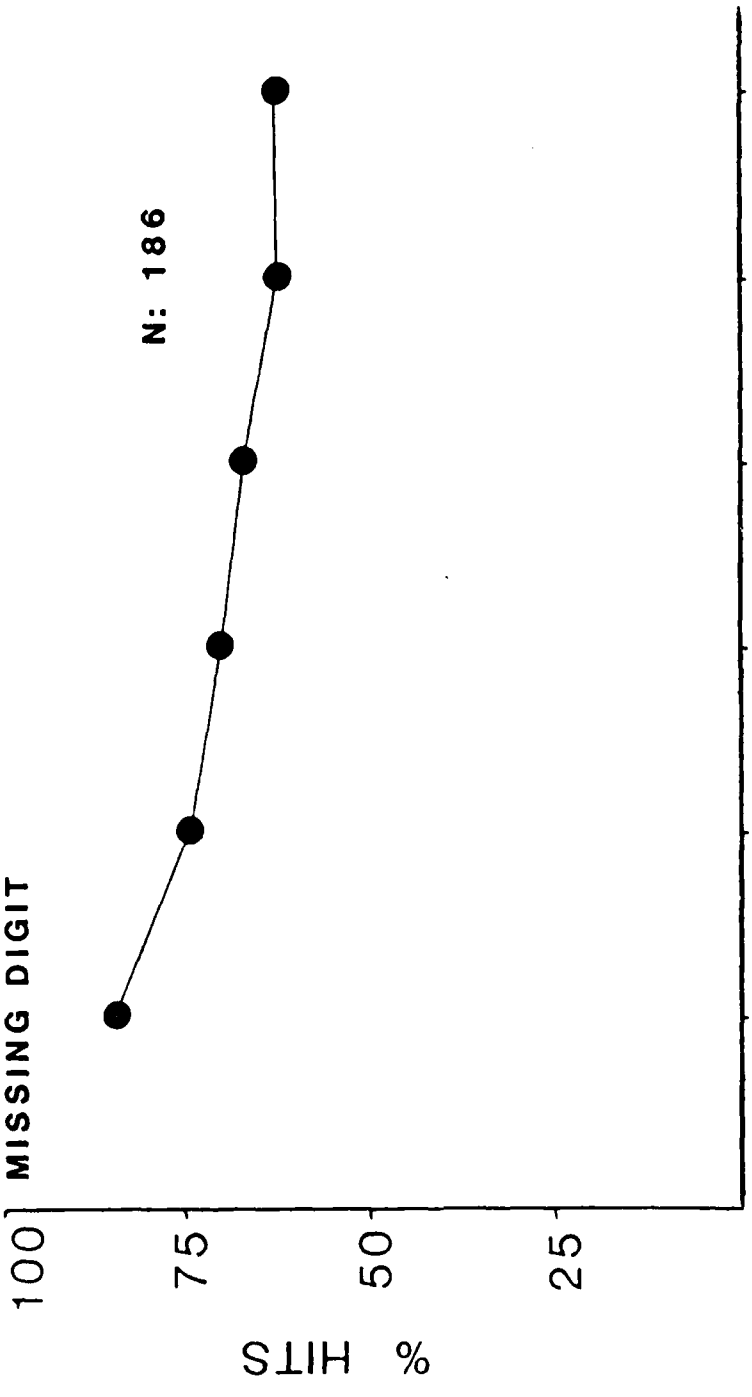


Figure 3. Percent correct detections of a missing digit during a 60-min test period

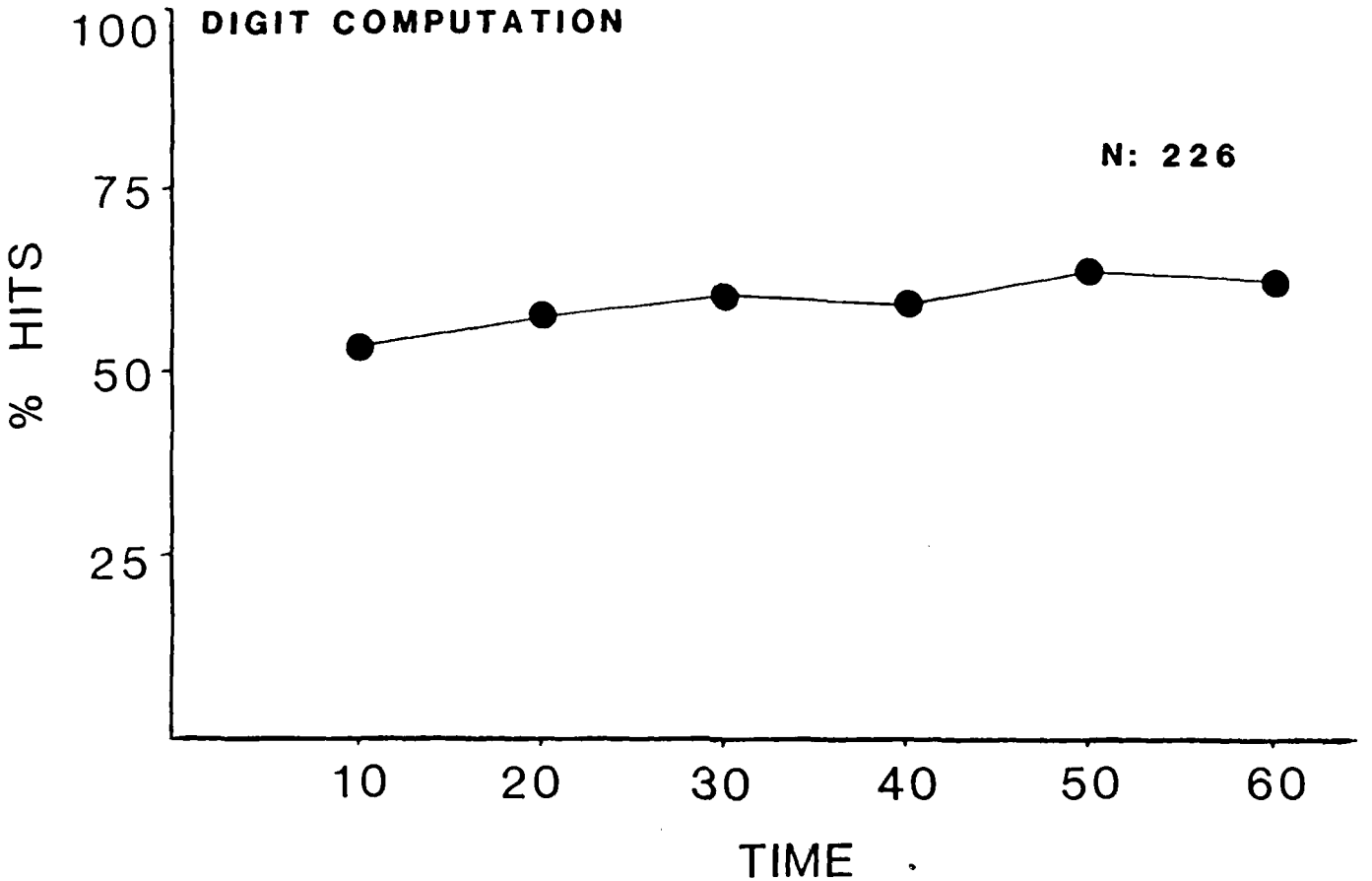


Figure 4. Percent correct detection of digit pairs that sum to a specified value during a 60-min test period.

watchkeeping period was observed of subjects who perform the vigilance task that demanded them to sum digit pairs (Figure 4).

V. Discussion.

The performance of subjects on the four vigilance tasks follow predictions made by Davies and Parasuraman (1981) and Warm (1984). On simple, repetitive vigilance tasks the number of missed targets increase with time; however, vigilance task that are more complex and demand greater allocations of attention result in performance that remains stable, or in some cases improves.

Whether individual differences in the performance of the four vigilance tasks is related to specific cognitive ability remains to be determined. The statistical analyses planned will provide insight into this question. Subject performance on vigilance tasks will be correlated with their ASVAB factor scores and with their performance on each of the seven tests which compose the cognitive test battery.

VI. Recommendations.

1. A thorough evaluation of individual differences in watchkeeping ability should be made. The present study provides a starting point for such an evaluation. It is suggested that each of the four vigilance tasks be

systematically evaluated. Each task can be varied with respect to: A) background event rate, B) density of target stimuli, and C) duration of the task.

2. A series of watch keeping tasks should be developed that employ observational requirements that simulate actual military job responsibilities. The ecological validity of laboratory research protocols is of paramount interest.

References

1. Green, D. L., and Swets, J. A. Signal Detection Theory and Psychophysics, New York: John Wiley, 1966.
2. Davies, D. R., and Parasuraman, R. The Psychology of Vigilance, New York: Academic Press, 1981.
3. Warm, J. S. Sustained Attention in Human Performance, New York: John Wiley, 1984.

1986 USAF-UES Summer Faculty Research Program/
Graduate Student Summer Support Program

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

Final Report

An Investigation of Unsteady Vorticity Production by a Pitching Airfoil

Prepared By:	T.R. Troutt
Academic Rank:	Associate Professor
Department & University:	Department of Mechanical Engineering Washington State University
Research Location:	Frank J. Seiler Research Laboratory, Aerospace Mechanics Directorate, United States Air Force Academy
USAF Researcher:	Major John M. Walker
Date:	September 2, 1986
Contract No:	F49620-85-C-0013

An Investigation of Unsteady Vorticity Production by a Pitching Airfoil

by

T.R. Troutt
J.A. Albertson

Abstract

This research investigation concentrated on developing insight into the unsteady aerodynamics produced by pitching airfoils in uniform airstreams. The specific experimental situation addressed involved a two-dimensional NACA 0015 airfoil which was pitched at constant rates through angles of attack from 0-60 degrees. The experimental results analyzed included high speed motion picture flow visualizations and simultaneous pressure measurement from eighteen chord locations distributed over the surface of the airfoil. The research techniques employed involved both the development of simple analytical procedures for computation of the vorticity generation rates by the surface of the pitching airfoil and digital image techniques for enhancing and quantifying interpretations of the visualization results. The analytical results demonstrate that the nondimensional parameter $\alpha^+ = \dot{\alpha}c/U$, where $\dot{\alpha}$ is the airfoil pitching rate, c is the airfoil chord length, and U is the free stream velocity, should be the primary parameter for predicting the relative influence of pitching rate and free stream velocity on the unsteady aerodynamics around an airfoil. This result agrees closely with experimental findings. The results from the image analysis techniques show that the initiation of the dynamic stall vortex on the airfoil top surface corresponds closely to a halt in the increase in lift as a function of attack angle curve. The subsequent fast growth of the dynamic stall vortex is found to accompany a rapid decrease in the lift on the airfoil.

I. Introduction

My previous related research interests have been in the areas of turbulent free shear flows and reattaching separated flows. The dominant proportion of my efforts in these areas has involved the acquisition and analyses of experimental data with the objective of developing improved understanding of the role of quasi deterministic flow structures in the development of these turbulent flows. Representative publications associated with these previous efforts are given in references 1, 2, 3, 4, 5.

The unsteady aerodynamic research pursued by graduate student, Ms. Julie Albertson, and myself this summer was a natural extension of my previous experience involving separated and turbulent flows. The specific area studied this summer involved the analysis of unsteady separation phenomena created by pitching airfoils in a uniform airstream. The motivation behind this work involves the interest of the U.S. Air Force in developing improved understanding of the aerodynamics produced by unsteady maneuvering of high performance aircraft.

II. Objectives of the Research Effort

The primary goal of this research program was to improve understanding of dynamic stall phenomena associated with airfoils undergoing rapid accelerations or pitching motions. This goal was pursued through two avenues involving analyses of experimental data acquired previously by the aeromechanics directorate at the Seiler Research Laboratory. The two research avenues followed were: 1. The development of a simple analytical technique to understand and quantify the vorticity generation

rates from the unsteady surface motions and instantaneous surface pressure gradients; 2. The development and application of digital image analysis techniques to enhance both qualitative and quantitative interpretations of flow visualization results.

One goal which was proposed in my original objectives statement was to perform additional focussed experiments to supplement and extend previous studies in this area. Because of some delays in getting a new wind tunnel facility and associated peripherals into working status no additional experiments were performed. It is hoped at this time that additional experiments employing the new wind tunnel facility at the Seiler Research Laboratory could be performed at a somewhat later time with the support of the mini grant associated with the USAF-UES Summer Faculty Research Program. A grant application concerning this work is in preparation.

III. Approach

III.1 Vorticity Generation Analysis

The first avenue of investigation concerned the application of Morton's (6) vorticity generation analysis to the problem of vorticity production by a pitching airfoil.

Morton's analysis shows that the vorticity generation in a homogeneous fluid may only take place at rigid boundaries and that the generating mechanisms are the tangential surface pressure gradient and the tangential surface acceleration of the boundary.

This deduction can be mathematically represented by the equation

$$\frac{d\Gamma_z}{dt} = \rho^{-1} \left(\frac{\partial p}{\partial s}\right)_o + \frac{dU_s}{dt} \qquad (1)$$

where $d\Gamma_z/dt$ is the flux of z direction vorticity from the surface (the z axis is tangential to the surface and normal to the s direction), ρ is the fluid density, $\left(\frac{\partial p}{\partial s}\right)_o$ is the tangential pressure gradient along the rigid surface in a direction normal to the z axis, and dU_s/dt is the tangential acceleration of the rigid surface. This equation explicitly separates the effects of the fluid dynamics through the pressure gradient from the effects of the rigid surface acceleration on the vorticity generation rate at the surface. Although the surface pressure gradient is dependent upon the flow and cannot easily be obtained for turbulent flows, except from experiments, the surface acceleration term can be easily calculated for specific experimental situations.

For the case of an airfoil pitching about a fixed pivot the surface acceleration term can be given by the following relation

$$\frac{dU_s}{dt} = r_o \left[\dot{\alpha}^2 \cos(\theta + \xi) + \ddot{\alpha} \sin(\theta + \xi)\right] \qquad (2)$$

where r_o is the radial distance from the pivot point to the specified surface position, $\dot{\alpha}$ is the airfoil rotation rate, θ is the angle between the chord axis and r_o, ξ is the angle of the local surface tangent and $\ddot{\alpha}$ is the rotational acceleration. A diagram showing the various geometrical quantities and the employed coordinate system is given in figure 1.

For the specific airfoil experiments evaluated thus far, $\dot{\alpha}$ was maintained constant during the measurements so equation (2) reduces simply to

$$\frac{dU_s}{dt} = r_o \dot{\alpha}^2 \cos(\theta + \xi) \qquad (3)$$

The surface acceleration term for a specific airfoil geometry pitching about at a fixed pivot at a constant rotation rate is thus solely

dependent on $\dot{\alpha}^2$. Competition between the pressure gradient term and the surface acceleration term with respect to vorticity generation for constant density flow can therefore be scaled by $\dot{\alpha}^2 c^2/U^2$ or $\dot{\alpha}c/U$ where c is the chord length of the airfoil and U is the free stream velocity. This nondimensional parameter has in fact been noted previously by Walker and Helin (7) as the primary controlling factor in their pitching airfoil experiments. This parameter was assigned the symbol α^+ by Walker and Helin (7) and will therefore be referred to as such throughout the rest of this report.

Figure 2 shows the computed values of the vorticity flux density from the top surface of a NACA 0015 airfoil as a function of chord position for various values of the parameter α^+. For these curves the hinge position was located on the airfoil chord at 0.25 c from the leading edge. This hinge position was used for the bulk of the pitching airfoil experiments conducted at the Seiler Laboratory. The results show that both positive and negative vorticity is produced by the top surface of the airfoil with a zero vorticity flux located approximately at the hinge chord location because the sum of the two angles $(\theta + \xi)$ goes through 90° in this vicinity. Since the NACA 0015 is a symmetric airfoil the bottom surface vorticity production is a negative of the top surface generation rate, and the total instantaneous vorticity generation produced by the tangential surface acceleration of the pitching airfoil will be identically zero at all times. In addition one can also conclude from Morton's (6) discussion that the total vorticity production from the tangential surface pressure gradient will also identically cancel at all times such that no net vorticity is created by the body.

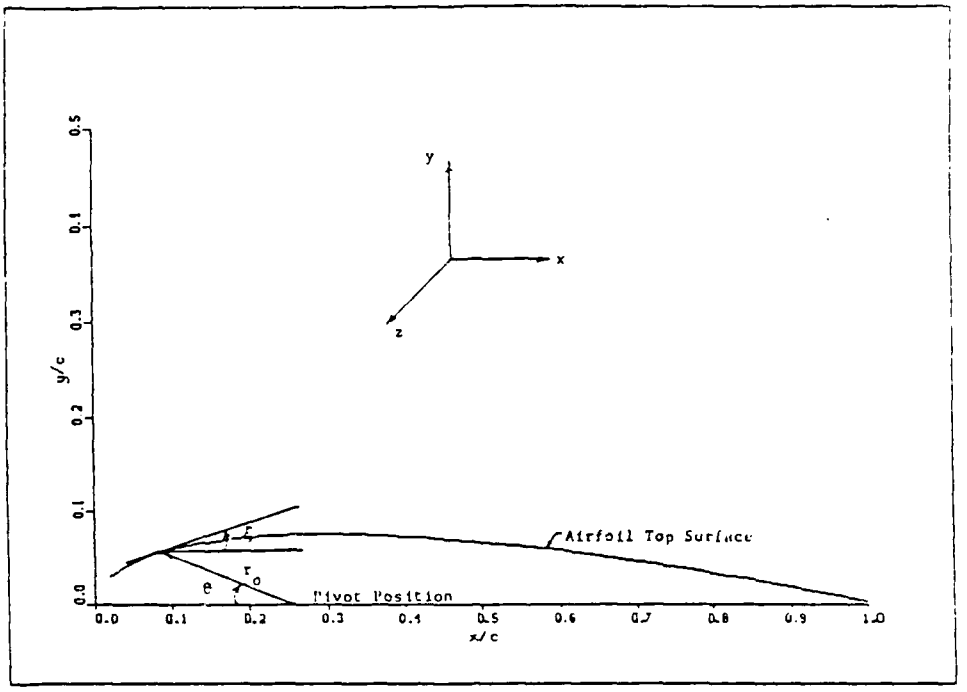


Figure 1. Schematic of Airfoil Geometry and Coordinate System.

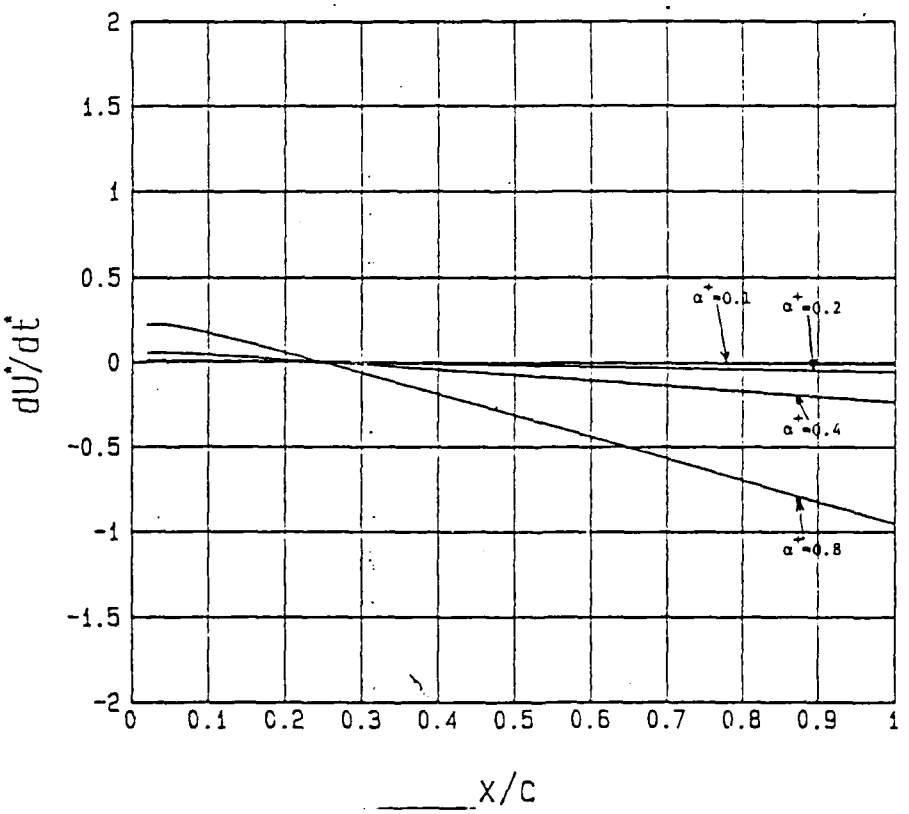


Figure 2. Tangential Surface Acceleration Curves for Pitching Airfoil at Various α^+ Values with Pivot at x_h = 0.25 c.
$U^* = U_s/U$, $t^* = tU/c$.

Vorticity generation from the tangential surface acceleration for the upper surface pitching about various hinge positions is shown in figure 3. As the hinge position is moved down the chord the relative amounts of positive and negative vorticity generation change with the zero flux position corresponding closely to the pivot location.

The determination of the vorticity generation by the tangential surface pressure gradients is considerably more complex than the surface acceleration analysis since the surface pressure gradient depends directly on the flow character. To evaluate this generation term point surface pressure measurements from previous experimental results provided by Major J.M. Walker, Chief of the Mechanics Directorate at the Seiler Research Laboratory were analyzed. The raw surface pressure data was obtained from 18 pressure transducers located at various downstream position along the airfoil surface.

Typical surface pressure measurements are shown for a value of $\alpha^+ = 0.2$ at various attack angles in figures 4a-d. Both the instantaneous coefficient of pressure curve and its tangential surface gradient are shown in each figure. The surface pressure gradients at all angles of attack are positive over most of the airfoil top surface. According to equation (1) these positive gradients will produce positive z direction vorticity. This positive vorticity production will tend to induce reversed or separated flow close to the airfoil. This is, of course, the situation that arises in steady aerodynamics as the attack angle nears the stall angle, and the separation region spreads over the airfoil surface. For pitching airfoils this result is avoided by apparently preventing the spread of the separation region over the entire upper airfoil surface.

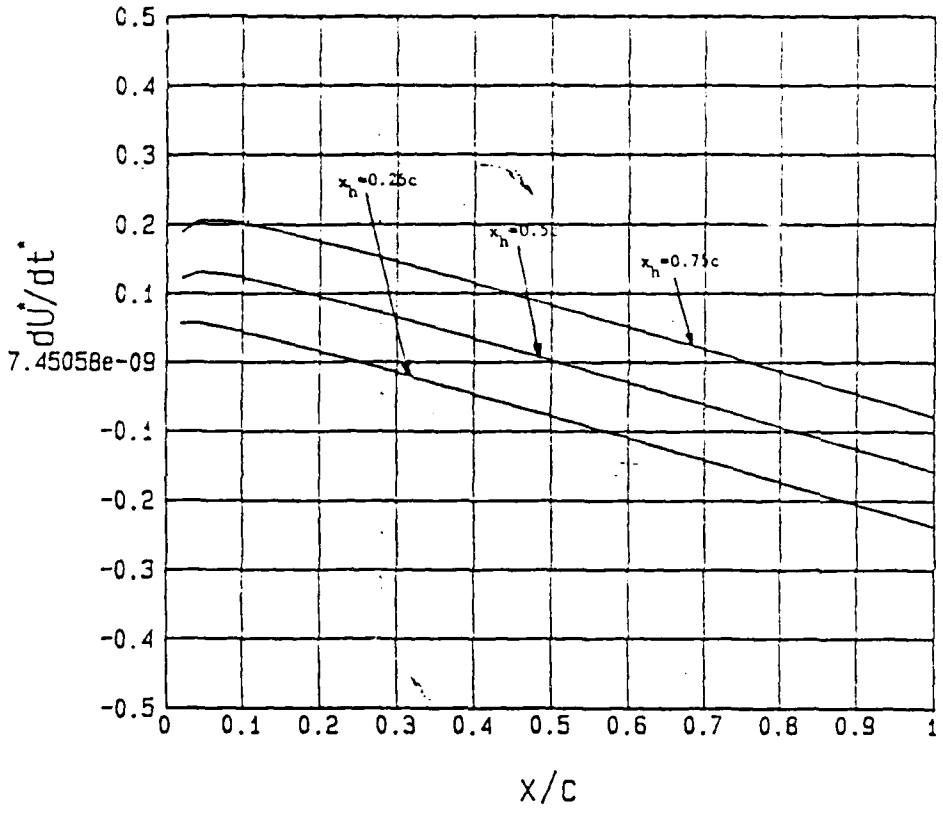
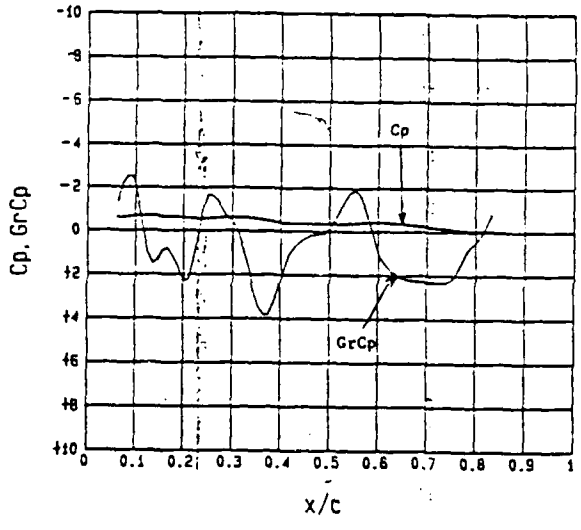
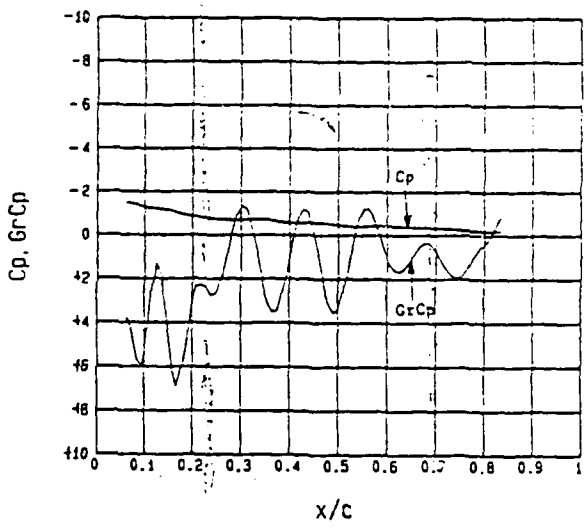


Figure 3. Tangential Surface Acceleration Curves for Pitching Airfoil with Various Pivot Locations, $\alpha^+ = 0.2$.

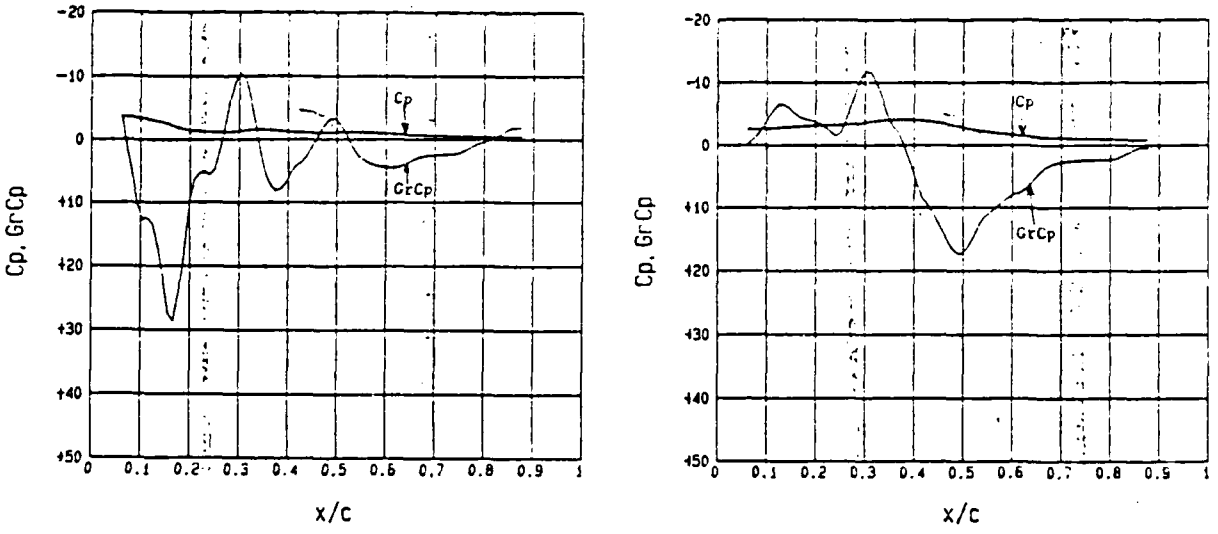


a. $\alpha = 0°$



b. $\alpha = 10°$

Figure 4. Tangential Surface Pressure Gradients for Pitching Airfoil, $\alpha^+ = 0.2$.



c. $\alpha = 20°$

d. $\alpha = 40°$

Figure 4. Tangential Surface Pressure Gradients for Pitching Airfoil, $\alpha^+ = 0.2$.

A comparison of the two vorticity generation terms can now be made. The negative z vorticity produced by the surface acceleration term downstream of the pivot location will tend to counteract the positive vorticity produced by the surface pressure gradients. However, for cases involving relatively low α^+ values (0.1 - 0.4) it would not appear that the surface acceleration values are of comparable magnitude to the levels of pressure gradient encountered even at moderately low angles of attack. The large influence of the airfoil pitching motion on the maximum coefficient of lift observed by Walker, Helin and Chou (8) can thus apparently not be explained easily through a simple superposition of vorticity generation terms. Since the pressure gradient term is a product of the fluid dynamics of the system small changes in the boundary conditions must be reflected as large effects in the flow through the nonlinearity of high Reynolds number flow. It therefore appears that well designed experiments are needed which better separate

the individual vorticity generation terms so that their separate effects can be better understood.

III.2 Image Analysis of Flow Visualization Experiments

The influence of airfoil pitching motion on the development of stall formation and the relationship between dynamic stall vortex growth and lift and pressure drag over the airfoil were studied using digital image analysis of flow visualization results previously obtained. The subject of the flow visualization was also a two dimensional NACA 0015 airfoil driven at constant pitching rates over attack angles from zero to 60 degrees in a uniform air stream. The digial image analysis techniques were performed using an International Imaging System 570. Photographs obtained from high-speed motion film taken at α^+ values of 0.1 and 0.2 were studied over angles of attack between the formation of the stall vortex and its eventual detachment from the airfoil surface.

The major objective of the image analysis technique was to clearly determine the stall vortex boundary from each photograph. Once this vortex boundary was defined the vortex growth and position could then be quantified as a function of attack angle.

The first step in the image analysis procedure employed an imaging technique which enhanced the contrasts between light and dark areas of the smoke streakline flow over the airfoil. The resulting image from this technique was then used to determine the leading and upper edges of the stall vortex at each attack angle.

Determination of the trailing edge of the stall vortex was somewhat more difficult since this edge merged with the thin turbulent separation region covering the remaining portion of the airfoil. This difficulty

was overcome by tracking the trailing edge of the dynamic stall vortex backwards in time from its eventual departure from the airfoil surface. Enhanced flow visualization pictures from four attack angles are displayed in figure 5 for an $\alpha^+ = 0.1$. The development of the dynamic stall vortex region with increasing attack angle is readily apparent in these figures.

The vortex area growth as a function of attack angle was quantified using an image analysis program to color the vortex region. A 6 x 6 pixel grid was then placed over the artificially colored area. The number of grid squares associated with the colored region were then computed to produce a reasonably accurate vortex area estimate. For each α^+ value the area measurement was initiated when the dynamic stall vortex was first discernable on the top surface of the airfoil. The area measurement was then continued until the vortex became clearly detached from the airfoil surface or the maximum attack angle was reached.

The relationship between the vortex location and the pitching motion of the airfoil was determined from measurements of the vortex center on an axis perpendicular to the chord. The vortex center was defined to be halfway between the leading and trailing edges of the stall vortex at each angle of attack and was obtained from the digitally enhanced images used in the vortex growth measurements. The vortex growth and location data obtained were then plotted as a function of angle of attack.

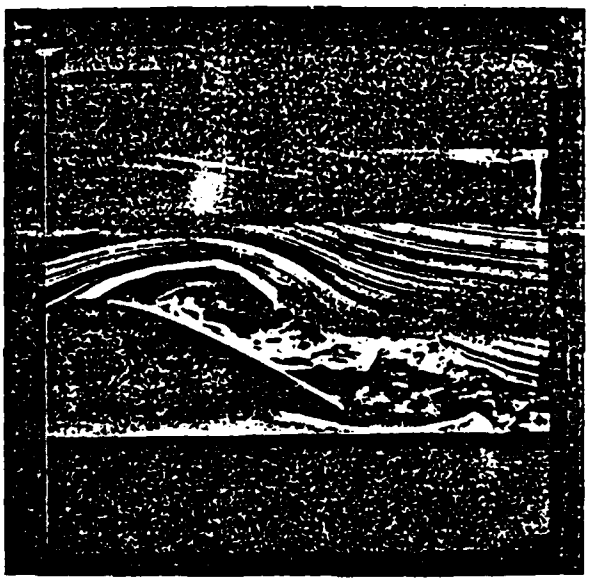
Vortex growth as a function of α plots at constant α^+ values but differing $\dot{\alpha}$ and b values are shown in figures 6a and b, for α^+ values of 0.1 and 0.2 respectively. The curves for $\alpha^+ = 0.1$ in figure 6a indicate similar growth rates at constant α^+ for each vortex if allowance is made



α = 24.7



α = 25.3



α = 28.7



α = 29.3

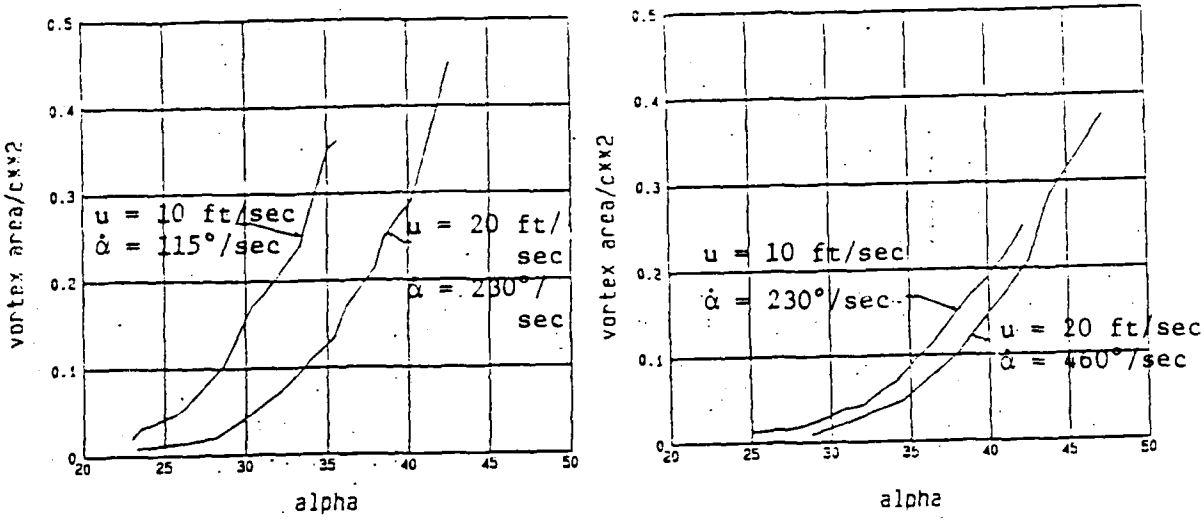
Figure 5. Enhancement of dynamic stall vortex boundaries. α^{+} = 0.1, u = 10^{ft}/sec.

for the slightly different vortex initiation angles. The $\alpha^+ = 0.2$ curves in figure 6b also show similar growth rates which appear slightly slower than the $\alpha^+ = 0.1$ curves.

Figure 7 shows experimentally measured lift coefficients as a function of α for an $\alpha^+ = 0.2$ from a similar airfoil. These data were furnished by W.D. Siuru from analyses of experiments performed by the Seiler Laboratory Aeromechanics group. The lift coefficient curves indicate that a general reduction in slope occurs near an attack angle of 24° which corresponds closely to the observation of the vortex initiation discussed previously. The curve acquires a negative slope near an attack angle of 38° which also corresponds closely to the initiation of the fast vortex growth previously mentioned.

Figure 8, also from Siuru's analyses, displays the experimentally measured pressure drag coefficient as a function of α for an $\alpha^+ = 0.2$. The drag coefficient increases up to approximately 40° and then begins a steady decrease. The beginning of this decrease appears to correspond closely to the detachment of the stall vortex from the airfoil surface shown previously.

Plots of the stall vortex center location as a function of attack angle are shown in figures 9a and b for $\alpha^+ = 0.1$ and 0.2 respectively. The similarity of the curves at constant values of α^+ indicates again the importance of α^+ as a parameter for this unsteady flow. In addition, the attack angle at which the stall vortex center moves into the $x/c = 0.25$ position coincides with the angle at which the stall vortex growth rate undergoes a rapid increase for both α^+ values of 0.1 and 0.2. This result indicates that a possible connection between the pivot location and the dynamic stall vortex behavior may exist. Further experiments are needed to clarify this possibility.



a. $\alpha^+ = 0.1$

b. $\alpha^+ = 0.2$

Figure 6. Dynamic stall vortex area versus angle of attack.

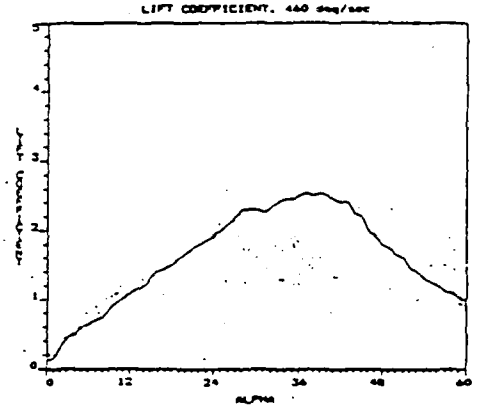


Figure 7. Lift coefficient versus angle of attack.
$\alpha^+ = 0.2$, u = 20 ft/sec

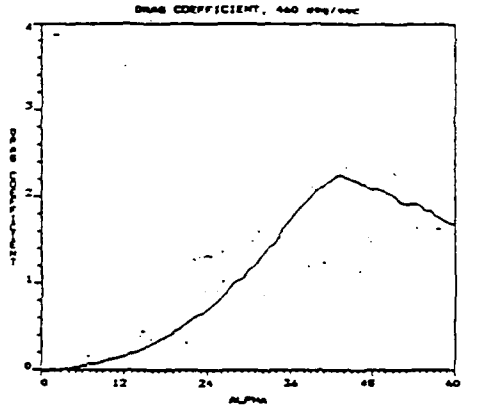
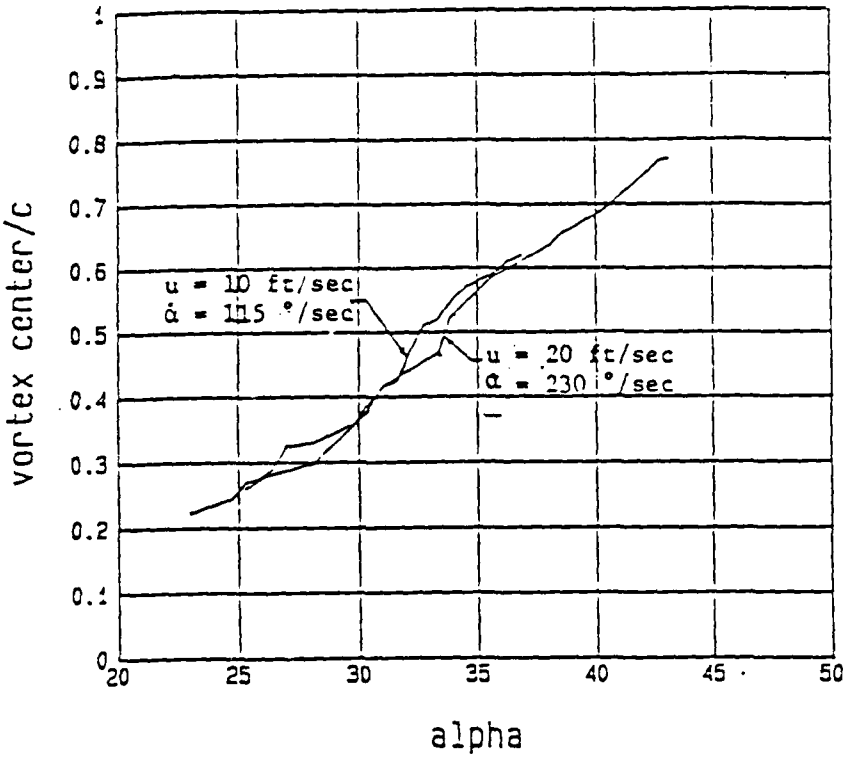
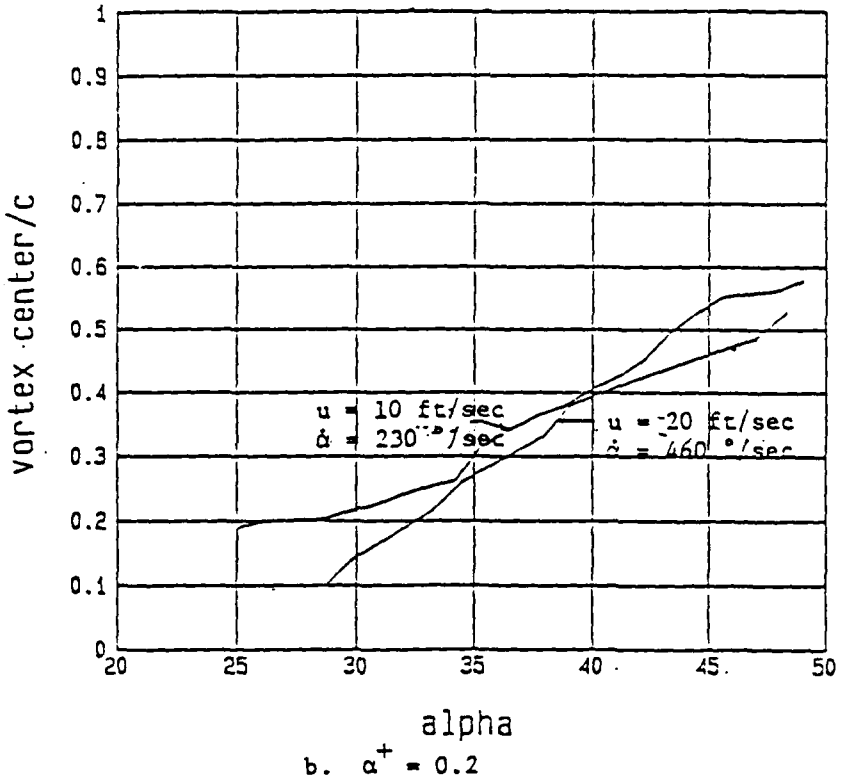


Figure 8. Drag coefficient versus angle of attack.
$\alpha^+ = 0.2$, u = 20 ft/sec



a. $\alpha^+ = 0.1$



b. $\alpha^+ = 0.2$

Figure 9. Dynamic stall vortex position along the airfoil versus angle of attack.

IV. Conclusions

The following conclusions from this study apply to a NACA 0015 airfoil pitching at constant rates.

1. The nondimensional parameter α^+ was shown to be associated with the ratio of vorticity generation rates by unsteady surface motions and surface pressure gradients.
2. The maximum value of lift obtained by the pitching airfoil occurs near the angle of attack at which the dynamic stall vortex becomes apparent.
3. The rapid growth of the dynamic stall vortex with increasing attack angle was found to coincide with decreasing values of lift.
4. The departure of the dynamic stall vortex from the airfoil surface was found to coincide with a decrease in the pressure drag.

V. Recommendations

This study has shown that important insights into the nature of unsteady vorticity production from pitching airfoils can be developed through the use of simultaneous flow visualization and point sensor measurements of flow variables. In addition the use of digital image analyses techniques was shown to enhance and clarify interpretations from the visual information. It is recommended that the analysis techniques employed here be used on certain additional focussed experimental situations to further examine the effects of unsteady vorticity production on airfoils. Three experimental studies which appear to have

high potential for further analyses using the technique employed in this research are:

1. A study of airfoils subjected to translational accelerations such that the effects of vorticity generation from surface pressure gradients and unsteady surface motions can be more explicitly separated.
2. A study of airfoils for which the pivot location is changed to determine if the surface distribution of the vorticity generation from the unsteady acceleration term is an important factor.
3. A study of airfoils pitching at low α^+ values to determine how small values of unsteady motion are able to cause large changes in the aerodynamic forces on the airfoils.

VI. References

1. Browand, F.K. and Troutt, T.R., "A Note on Spanwise Structure in the Two-Dimensional Mixing Layer," J. of Fluid Mech., 97, 1980, 771-781.

2. Troutt, T.R. and McLaughlin, D.K., "Experiments on the Flow and Acoustic Properties of a Moderate Reynolds Number Supersonic Jet," J. of Fluid Mech., 116, 1982, 123-156.

3. Troutt, T.R., Scheelke, B. and Norman, T.R., "Organized Structures in a Reattaching Separated Flow Field," J. of Fluid Mech., 143, 1984, 413-427.

4. Browand, F.K. and Troutt, T.R., "The Turbulent Mixing Layer: Geometry of Large Vortices," J. of Fluid Mech., 158, 1985, 487-504.

5. Bhattacharjee, S., Scheelke, B. and Troutt, T.R., "Modifications of Vortex Interactions in a Reattaching Separated Flow," AIAA J., 24, 4, 1986, 623-692.

6. Morton, B.R., "The Generation and Decay of Vorticity," Geophys. Astrophy. Fluid Dynamics, 28, 1984, 277-308.

7. Helin, H.E. and Walker, J.M., "Interrelated Effects of Pitch Rate and Pivot Point on Airfoil Dynamic Stall," AIAA Paper 85-0130, The Aerospace Sciences Conference, Reno, NV, Jan. 1985.

8. Walker, J., Helin, H., and Chou, D., "Unsteady Surface Pressure Measurements on a Pitching Airfoil," AIAA Paper 85-0532, AIAA Shear Flow Control Conference, Boulder, CO, March 1985.

1986 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT SUMMER SUPPORT PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

A Study of Turbulence Models for Predicting Aero-Optic Interactions

Prepared by: C. Randall Truman

Academic Rank: Assistant Professor

Department and University: Department of Mechanical Engineering, University of New Mexico

Research Location: Air Force Weapons Laboratory/ARDF, Kirtland AFB, New Mexico

USAF Researcher: Dr. Bruce Masson

Date: September 30, 1986

Contract No.: F49620-85-C-0013

A Study of Turbulence Models for Predicting Aero-Optic Interactions

ABSTRACT

Turbulence models for the prediction of aero-optic interactions were surveyed. It was found that the prediction of index-of-refraction flucuations which degrade the quality of a coherent beam traversing the field is not feasible at present. Thus a series of predictions and measurements is proposed to determine the relationships between index-of-refraction flucuations and turbulence quantities which can be predicted and measured. Preliminary predictions and laser-Doppler velocimetry measurements of an axisymmetric jet were made.

The implementation of a sequence of successively more complex turbulence models is proposed for comparison with measurements of velocity, density, or temperature flucuations for a model shear layer. These results will be compared with measurements of optical quality of the shear layer using laser interferometry or similar techniques. The predictions should culminate in a second-order turbulence model which can accurately predict the complex turbulent flows of interest for chemical lasers. The validation of the turbulence models by measurements for a model shear layer is necessary before reliable predictions of more complex flows can be made.

I. INTRODUCTION

My research interests include turbulence modeling for complex three-dimensional flows. My dissertation research at Arizona State University was a prediction of source flow between rotating disks using a mixing-length turbulence model which modeled streamline divergence and anisotropy due to rotation. The prediction of this flow has since been extended to include a two-equation turbulence model. My research in turbulence modeling for complex turbulent flows at the University of New Mexico has been supported by Sandia National Laboratories, Albuquerque. Most recently, I have been studying mixing-length turbulence models for supersonic flow using a parabolized Navier-Stokes algorithm. The sensitivity of algebraic turbulence models to starting conditions, grid spacing, and strong crossflow which may lead to separation is being studied. This study is now examining "half-equation" models which can introduce history effects into algebraic turbulence models.

The present study of the turbulence modeling requirements for predicting aero-optic interactions is a natural extension of my earlier work. The subject was chosen through discussions with the AFWL/ARDF staff, with recognition of the limitations of the ten-week appointment upon experimental investigations or code development. This survey of turbulence models available to predict the effects of turbulent flucuations upon a coherent beam

traversing the turbulent field was seen as necessary to begin a program within ARDF to develop the capability to predict such flows. Preliminary work on the measurement and prediction of a model shear layer was done by the two graduate students, Kyle Ross and Tim Clark, who worked with me this summer.

II. RESEARCH OBJECTIVES

The propagation of a coherent beam of light through turbulent media is of critical importance to the performance of chemical lasers (Skifstad 1974; Fuhs 1981; Sutton 1985). Turbulent density and/or temperature fluctuations induce index-of-refraction flucuations which can significantly degrade beam quality (Bogdanoff 1984). Several regions of turbulent flow can occur along the beam train, including the chemically-reacting flow within the laser cavity, along beam tubes within the laser, at mirrors which require a cooling or purge flow, and at the aerodynamic window through which the beam exits to the surroundings. Exterior flows between the laser source and target will also affect beam quality. These may include the flow over the dome of a ground-based laser, flow in the atmosphere near the earth's surface which is affected by structures, terrain, or heat sources, and turbulence within the atmosphere.

The objectives of this research were:

1. to survey turbulence models to determine suitable models for predicting index-of-refraction flucuations in shear flows;

2. to determine the level of computational effort and turbulence modeling required to predict large-scale turbulent structure in shear flows; and

3. to begin an experimental and numerical study of a simple flow from which computed and measured length scales can be determined.

The study was limited to shear flows because they appear to have the greatest influence upon beam quality. The extension to other flows should be straightforward once the turbulence modeling for aero-optic interactions in shear layers is established. Chemically-reacting flows were not considered in detail; such flows would be modeled in a manner similar to that used for combustion problems.

The first objective required an examination of the origin of index-of-refraction flucuations, and thus a determination of the level of turbulence modeling required to predict such flucuations or similar turbulent quantities. The second objective was a result of a hypothesis that the primary optical degradation occurs as a result of large-scale turbulence structure. The deterministic nature of large-scale structure and the selection of appropriate length scales were examined with

respect to the computational method and turbulence model which should be employed. The last objective involved a free shear layer which was planned by AFWL/ARDF before the summer research began. An axisymmetric jet of air or helium will be injected into quiescent air in the laboratory, and its evolution measured by a laser-Doppler velocimeter(LDV). The LDV system has a three-dimensional traversing capability which can be employed to map the flow field. The efforts of two graduate students, Kyle Ross and Tim Clark, on the experimental and computational aspects, respectively, will be briefly discussed.

The result of this study is intended to be a set of recommendations for future work which the Air Force Weapons Laboratory/ARDF should undertake to develop a predictive capability for aero-optic interactions. Because of the necessity of verifying turbulence models with experimental data, the nature and scope of the measurements required to validate the proposed turbulence models was an integral part of this study.

III. COMPUTATIONAL METHODS AND TURBULENCE MODELS

The existence of regular, or coherent, structure within turbulent shear flows is now well established (Brown & Roshko 1974; Roshko 1976; Cantwell 1981). The possibility as well as the necessity of predicting such structure in order to accurately model aero-optic interactions was a

primary consideration in this work. While it is no doubt expensive to model a turbulent flow at a level which can predict deterministic structure, it may be necessary if the density or concentration flucuations within the coherent structure dominate the refractive-index flucuations which degrade the coherence of a beam passing through the field.

The choice of turbulence model has a direct impact upon the computational method employed. The use of time-mean or ensemble-average transport equations and a Reynolds-stress turbulence closure will lead to the use of conventional numerical techniques (Reynolds 1976). The solution method is determined primarily by whether the flow can be modeled as parabolic, hyperbolic or elliptic. However, for solutions which compute part of the turbulent flucuations directly, more expensive solution methods must be employed. While this choice has a direct impact upon the computation cost, the discussion below will center on the requirements and expected results of the turbulence modeling.

Time-mean or ensemble-average equations. The historic treatment of turbulence has been a stochastic model with the flow variables computed as time-mean or ensemble-average quantities and the turbulent flucuations treated only in a statistical sense (Reynolds 1976). The resulting Reynolds stresses must be modeled in some kind of turbulence closure. Such models are tractable for nearly

all flows, and if sufficient spatial resolution can be achieved, adequate prediction of mass, momentum and energy terms is obtained. Because of the averaging, turbulent flucuations may not be modeled adequately, particularly in flows in which deterministic turbulent structure is significant.

The Boussinesq approximation that the Reynolds stress is proportional to an eddy viscosity requires modeling the eddy viscosity with algebraic relations or transport equations for such quantities as turbulent kinetic energy and its rate of dissipation. Algebraic models are frequently used for efficient solution, with constants adjusted to compute different classes of flow. Two-equation models demonstrate somewhat greater versatility than algebraic models, but include several additional adustable constants.

Second-order models compute directly the Reynolds stresses and other second-order moments of turbulent flucuations (Reynolds 1976; Bradshaw 1978). Transport equations for each second-order moment must be derived; higher-order moments which appear in these equations must then be approximated. Such models are viewed by many as the highest level of closure which is practical for most problems given the present generation of computers.

The modeling of the Reynolds stresses depends upon the assumption of single- or two-point closure (Bradshaw 1978),

in which the Reynolds stress is approximated in terms of products of turbulent flucuations at a single point in space or between two adjacent points, respectively. The former, which is usually chosen for expediency, assumes a single scale for the turbulence; its limitations are apparent for flows in which several scales of turbulence are important. The latter results in a spectral model, which includes all scales of turbulence. While the single-point closure cannot be universally valid, it is suprisingly accurate for a broad range of flows (Launder 1985).

A useful approximation is a "two-scale" model, in which only two length (or time) scales are modeled (Hanjalic, et al. 1980). This model has been effective for flows in which turbulent kinetic energy is produced and dissipated at significantly different scales. Such models should be considered for flows in which large-scale structures are important because turbulence energy is produced at the scale of the large coherent structures, but is dissipated at the smallest, or Kolmogorov scales (Tennekes & Lumley 1972). Several promising two-scale models have been developed for two-equation models (Fabris & Harsha 1981; Birch 1983; Wilcox 1986).

The limitation of the eddy-viscosity closure is that the assumptions required to formulate each model results in modeling constants which must be adjusted for each new

class of flow. The models discussed below compute the turbulence directly with the objective of avoiding modeling constants with their resulting limitations on generality and versatility.

Large-eddy simulation (LES). The purpose of large-eddy simulation is to compute the large-scale structure directly (Reynolds 1976; Ferziger 1985). The equations are filtered according to the grid spacing rather than time- or ensemble-averaged. Turbulent eddies which are smaller than the grid spacing are modeled so that the turbulence closure resides in the subgrid-scale (SGS) model. The behavior of the small eddies is presumably more universal (Reynolds 1976). The SGS model can be as simple as an eddy viscosity-type model or as sophisticated as a spectral model. The prediction of the large-scale, deterministic structure eliminates the problem of modeling disparate scales within a single turbulence closure. The primary limitations of LES are computing enough grid points, specifying boundary conditions, and developing SGS models for flows other than simple shear layers. The latter may be the most severe restriction; specifying an SGS model may be as difficult as (and no more universal than) specifying a Reynolds-stress closure. Moreover, in chemically-reacting flows, the chemical reactions occur at the molecular level and may not be accurately modeled by the SGS model (Ferziger 1985). The large computation times for

LES predictions will become less significant as parallel-processing machines become available.

Direct numerical simulation. With the smallest scale in the flow resolved by an extremely small grid spacing, direct numerical simulation requires no turbulence closure (Rogallo & Moin 1984). However, the difficulty and expense of this calculation makes it feasible only as a "numerical experiment." The results of the computation can be used to compute turbulent statistics and other quantities, and to develop information which can be used to enhance lower-order turbulence models. The large number of grid points required to resolve even simple flows results in computation times on the order of 10 to 100 CRAY CPU hours. Nonetheless, direct numerical simulations have been attempted for a variety of flows, including a chemically reacting turbulent mixing layer (Riley, et al. 1986).

IV. SECOND-ORDER TURBULENCE MODELS

A second-order or Reynolds-stress transport model is proposed as the model which should be developed to predict the turbulent flows of interest in chemical lasers. Although such a model cannot predict coherent structure directly, it contains sufficient physics to adequately model most flows. Lumley (1981) argues that because the organized structures usually contain only a small part of the turbulent energy they play a secondary role; second-

order modeling, which ignores the coherent structures, is thus "satisfactory" for most flows. Lumley (1985), in fact, claims to have successfully modeled a variable-density mixing layer with a second-order model. Several groups of investigators have carefully examined second-order models to ensure that the relevant flow physics is correctly modeled.

Launder and coworkers have constructed a robust Reynolds-stress model (Launder, et al. 1975) which has been extended to multiple time scales (Hanjalic, et al. 1980), variable-density flows (Ha Minh, et al. 1982; Vandromme & Kollmann 1982), and the transport of a scalar quantity (Elgobashi & Launder 1983) such as temperature. This is perhaps the most widely used Reynolds-stress transport model.

Second-order models have been rigorously defined by Lumley (1978) and colleagues, who have stressed realizability and invariance to obtain a nearly universal model. Shih and Lumley (1986) examined the near-wall region to derive a realizable model. Lumley's earlier experimental work is used to describe proper measurement techniques for comparison of data with predictions.

Donaldson and coworkers have derived second-order models in a manner similar to Lumley. Donaldson and Sandri (1982) studied the possibility of predicting large-scale structure within the second-order closure; Donaldson and

Varma (1976) described the prediction of turbulent reacting flows.

Each of these second-order models appears to have about the same capability to compute complex turbulent flows. Much of the work of Lumley and Donaldson has been concerned with atmospheric modeling. The work of Launder is more engineering oriented, and thus somewhat more accessible and understandable. The three models should be carefully examined with regard to the turbulent flows of interest for chemical lasers before one is chosen. For flows with chemical reaction, transport equations for the many second-order moments must be developed with appropriate approximations and modeling of the higher-order moments. A thorough discussion of chemically-reacting turbulent flows is given by Libby & Williams (1976).

V. LENGTH SCALES FOR AERO-OPTIC INTERACTION

Since it is not feasible to compute directly the index-of-refraction flucuations which determine the degradation of a coherent beam of light passing through a turbulent flow field, it is necessary to relate the quantities which can be computed to measurements. This comparison should be carried out for a simple model shear flow which is well defined yet retains the characteristics typical of the flows of interest. An axisymmetric jet of helium into quiescent air (with density ratio of 1:7) is proposed as

such a model flow which has large-scale structure in its turbulent mixing layer. The variable density due to mixing creates an index of refraction variation which can be computed and measured. Preliminary work in measuring and computing this flow was carried out at AFWL/ARDF this summer by Kyle Ross and Tim Clark, respectively. The geometry of the flow is simple and the mean-flow development of the jet can be predicted with a parabolic thin-shear-layer model. Comparison between measurements and predictions should be made throughout the development of the turbulence model, beginning with a simple algebraic eddy viscosity model, and gradually including more sophisticated turbulence models.

The relationship between turbulent flucuations of velocity, temperature, or density to refractive-index flucuations can probably be expressed most conveniently in terms of length scales. In addition, some type of length scale can characterize the effects of the turbulence upon the beam. Length scales computed by a time-mean or ensemble-average method will be single-point correlations since no spectral information is available. Measurements of time and length scales for a model shear flow may be single- or two-point correlations. The former is desirable for comparison with predictions and the latter for comparison with measurements of optical quality. Having both scales would eliminate dependence on the Taylor

hypothesis, which may not be valid for flows with coherent structure (Tennekes & Lumley 1972). Measured scales can be determined from correlations of the flucuating quantities. The two-point LDV probe of Fraser, et al. appears promising for measuring two-point correlations of velocity which can yield a length scale.

Several measurement techniques to determine optical quality are available, including laser interferometry, Rayleigh scattering (Gouldin & Halthore 1986), and laser shadow (Grandke 1986) techniques. Gouldin & Halthore report density measurements for turbulent, premixed flames, and Grandke describes a crossed beam method which allows three-dimensional spatial resolution. The development of measurement techniques must move in parallel with the predictions, so that the second-order predictions can be compared with detailed and comprehensive measurements.

The goal of this study is to be able to provide an optical designer with information about the index-of-refraction flucuations across a turbulent flow field. This prediction would be derived from predictions of mean-flow quantities and second-order moments of relevant flucuations. The turbulence model would be calibrated against a model shear flow, so that reliable predictions of complex flows can be made. One result of the parallel computational and experimental studies would be a better understanding of the importance of coherent structure in

index-of-refraction flucuations. The free shear layer is probably a better choice for this examination than a wall-bounded shear layer.

Several important questions arise about the proposed measurements. One is the effect of the varying index of refraction upon LDV measurements. Secondly, the axisymmetric geometry requires that Abel transforms be employed to decode the measurements when the flow is viewed normal to the streamwise flow. Finally, the presence of coherent structures in the jet mixing layer may require conditional sampling of the measurements.

VI. RECOMMENDATIONS

The following recommendations are made in order for AFWL/ARDF to develop the capability to predict the effects of aero-optic interaction upon beam quality for the various turbulent shear layers which arise in chemical lasers:

1. Define a simple shear layer for which meaningful measurements and predictions can be made. Such a model flow is essential to the development of reliable turbulence models and measurement techniques. The variable-density axisymmetric jet is proposed. This constitutes a severe test of the prediction scheme (Launder 1985). Consideration must also be given to the effects of axisymmetry upon measurements within this flow,

particularly when looking through the cross-section of the jet to determine optical quality.

2. Develop measurement techniques for the model flow to determine mean-flow quantities such as velocity and density or concentration, as well as fluctuations of these variables. Data-reduction techniques must be developed to determine characteristic scales from these measurements.

3. Build a sequence of turbulence models for the model flow beginning with simple eddy-viscosity ideas. The progression to higher-order turbulence models must be systematic and gradual to allow full understanding of the benefits and limitations of each model, and to allow comparison with the measurements. Parallel development of the predictions and the measurements is appropriate, since the higher-order models will require more detailed and sophisticated measurements. The progression from algebraic eddy viscosity, to turbulent transport equations for turbulent kinetic energy and its dissipation rate, and to a Reynolds-stress model will allow implementation of such effects as variable density and chemical reactions. The end result should be a reliable turbulence model whose behavior is well understood for at least one model flow. The extension to other shear layers is then possible with favorable results more probable.

ACKNOWLEDGMENTS

The support of this work by the Air Force Systems Command and the Air Force Office of Scientific Research through the Summer Faculty Research Program is gratefully acknowledged. I wish to thank the staff of ARDF at the Air Force Weapons Laboratory for their assistance during this project. Dr. Bruce Masson, Capt. Richard Charles, Capt. Nan Founds, and Capt. Marty Trout all provided helpful information and stimulating discussion which helped me understand the problems of aero-optic interactions.

REFERENCES

Birch, S.F. 1984 Multi-Length Scale Turbulence Models, AFOSR-TR-84-0249.

Bogdanoff, D.W. 1984 The Optical Quality of Shear Layers: Prediction and Improvement Thereof. AIAA J, 22, 58-64.

Bradshaw, P. 1978 Introduction. In Turbulence, ed., P. Bradshaw, Springer-Verlag, 2nd ed., 1-44.

Brown, G.L. & Roshko, A. 1974 On Density Effects and Large Structure in Turbulent Mixing Layers. J Fluid Mech, 64, 775-816.

Cantwell, B.J. 1981 Organized Motion in Turbulent Flow. Ann Rev Fluid Mech, 13, 457-515.

Donaldson, C.duP. & Sandri, G. 1982 On the Inclusion of Information on Eddy Structure in Second-Order-Closure Models of Turbulent Flows. In Fluid Dynamics of Jets with Applications to V/STOL, AGARD CP-308, 25-1 to -14.

Donaldson, C.duP. & Varma, A.K. 1976 Remarks on the Construction of a Second-Order Closure Description of Turbulent Reacting Flows. Combustion Science and Technology, 13, 55-78.

Elgobashi, S.E. & Launder, B.E. 1983 Turbulent Time Scales and the Dissipation Rate of Temperature Variance in the Thermal Mixing Layer. Phys Fluids, 26, 2415-2419.

Fabris, G. & Harsha, P.T. 1981 Multiple-Scale Turbulence Modeling of Free Turbulent Flows. ASME Paper 81-FE-20.

Ferziger, J.H. 1985 Large Eddy Simulation: Its Role in Turbulence Research. In Theoretical Approaches to Turbulence, eds., D.L. Dwoyer, et al., Springer-Verlag, 51-72.

Fraser, R., Pack, C.J. & Santavicca, D.A. 1986 An LDV System for Turbulence Length Scale Measurements. Exper Fluids, 4, 150-152.

Fuhs, A.E. 1981 Overview of Aero-Optical Phenomena. In Wavefront Distortions in Power Optics, ed., C.A. Klein, Proceedings of SPIE, 293, 36-55.

Gouldin, F.C. & Halthore, R.N. 1986 Rayleigh Scattering for Density Measurements in Premixed Flames. Exper Fluids, 4, 269-278.

Grandke, T. A Crossed Beam Laser Shadow Technique for Spatially Resolved Measurements in Turbulent Flow. Exper Fluids, 4, 289-295.

Ha Minh, H., Launder, B.E. & MacInnes, J. 1982 The Turbulence Modelling of Variable Density Flows: A Mixed-Weighted Decomposition. In Turbulent Shear Flows 3, eds., L.J.S. Bradbury, et al., Springer-Verlag, 291-308.

Hanjalic, K., Launder, B.E. & Schiestel, R. 1980 Multiple-Time-Scale Concepts in Turbulent Transport Modelling. In Turbulent Shear Flows 2, eds., L.J.S. Bradbury, et al., Springer-Verlag, 36-49.

Launder, B.E. 1985 Progress and Prospects in Phenomenological Turbulence Models. In Theoretical Approaches to Turbulence, eds., D.L. Dwoyer, et al., Springer-Verlag, 155-186.

Launder, B.E., Reece, G.J. & Rodi, W. 1975 Progress in the Development of a Reynolds-Stress Turbulence Closure. J Fluid Mech, 68, 537-566.

Libby, P.A. & Williams, F.A. 1976 Turbulent Flows Involving Chemical Reactions. Ann Rev Fluid Mech, 8, 351-376.

Lumley, J.L. 1978 Computational Modeling of Turbulent Flows. Adv Appl Mech, 18, ed., C.-S. Yih, Academic Press, 123-176.

Lumley, J.L. 1981 Coherent Structures in Turbulence. In Transition and Turbulence, ed., R.E. Meyer, Academic Press, 215-242.

Lumley, J.L. 1985 Strange Attractors, Coherent Structures and Statistical Approaches. In Theoretical Approaches to Turbulence, eds., D.L.Dwoyer, et al., Springer-Verlag, 359-363.

Reynolds, W.C. 1976 Computation of Turbulent Flows. Ann Rev Fluid Mech, 8, 183-208.

Riley, J.J., Metcalfe, R.W. & Orszag, S.A. 1986 Direct Numerical Simulations of Chemically Reacting Turbulent Mixing Layers. Phys Fluids, 29, 406-412.

Rogallo, R.S. & Moin, P. 1984 Numerical Simulation of Turbulent Flows. Ann Rev Fluid Mech, 16, 99-137.

Roshko, A. 1976 Structure of Turbulent Shear Flows: A New Look. AIAA J, 14, 1349-1357.

Shih, T.-H. & Lumley, J.L. 1986 Second-Order Modeling of Near-Wall Turbulence. Phys Fluids, 29, 971-975.

Skifstad, J.G. 1974 Gas-dynamic Phenomena in Chemical Lasers. Acta Astron, 1, 781-812.

Sutton, G.W. 1985 Aero-Optical Foundations and Applications. AIAA J, 23, 1525-1537.

Tennekes, H. & Lumley, J.L. 1972 A First Course in Turbulence. MIT Press.

Vandromme, D. & Kollmann, W. 1982 Second Order Closure for Variable Density Free Shear Layer. In Turbulent Shear Flows 3, eds., L.J.S. Bradbury, et al., Springer-Verlag, 275-290.

Wilcox, D.C. 1986 Multiscale Model for Turbulent Flows. AIAA Paper AIAA-86-0029.

1986 USAF-UES Summer Faculty Research Program/

Graduate Student Summer Support Program

Sponsored by the

Air Force Office of Scientific Research

Conducted by

Universal Energy Systems, Inc.

Final Report

Analysis and Fate of Organic Components of Aqueous Film Forming Foams

Prepared by:	Roy M. Ventullo
Academic Rank:	Associate Professor
Department and	Department of Biology
University:	University of Dayton
Research Location:	Environics Laboratory, AFESC/RDVW
	Tyndall AFB, Panama City, Fl 32403
USAF Research:	Dr. Jim Spain
Date:	September 15, 1986
Contract No.:	F49620-85-C-0013

Analysis and Fate of Organic Components of Aqueous Film Forming Foams

by

Roy M. Ventullo

Abstract

Analytical methods were developed to determine the biodegradation of specific organic components of aqueous film forming foam (AFFF) in laboratory scale microcosms. To this end, surfactant and non surfactant organic components were separated , identified, and quantitated by high performance liquid chromatography (HPLC), gas chromatography - mass spectrometry (GC/MS), and specific analytical methods for surfactants. Reversed phase HPLC was used extensively in attempts to separate and detect both fluoro- and hydrocarbon surfactants. Several of the methods were used to conduct initial biodegradation studies using soil and water from the base fire training facility. Biodegradation of the ethylene glycol monobutyl ether solvent and the anionic surfactants was noted. Further laboratory (analytical and biodegradation) and field work is required to establish the rate and extent of biodegradation of AFFF components.

Acknowledgments

I would like to thank the Air Force Systems Command and the Air Force Office of Scientific Research for sponsorship of my research. The staff at the Environics Laboratory, AFESC/RDVW, Tyndall AFB provided a stimulating setting to conduct research. Special thanks to Dr. Jim Spain, Shirley Nishimo, and Mike Henley for guidance and technical help as well as easing my transition into a new laboratory. I also appreciate the help of Jim Cornette, Doug Downey, Wink Zachritz, and Msgt. Stork for help in planning and carrying out various aspects of this study.

I. Introduction

I received my Ph.D. in Microbiology from the University of Georgia studying the uptake of hydrocarbons by Acinetobacter. As a post doctoral fellow at the University of Calgary I studied the heterotrophic and biodegradative activity of microbial communities in lotic environments. My most recent research interest has been the biodegradation of xenobiotic chemicals in surface and subsurface (groundwater) environments. Specifically, we have been working on acclimation to and biodegradation of substituted benzene compounds and detergent chemicals in aquatic systems.

The Air Force, specifically the Air Force Engineering Services Center, Tyndall AFB is continually investigating the environmental impacts associated with day-to-day operations. One major operation is firefighting training which occurs at more than 100 sites. The Air Force is currently interested in treating the wastewater generated from firefighting training exercises. These wastewaters contain residual fuel and aqueous film forming foam (AFFF) constituents. Several investigations have shown that the AFFF components have adverse effects upon both sewage treatment systems and aquatic environments (1,2,3). This has lead to a curtailment of fire training activities at several Air Force and Navy facilities.

The research problem at Tyndall AFB involves the potential of on-site systems as an effective method for treatment of these wastewaters. Although several studies have examined the degradation of jet fuel, few studies have examined the biodegradation of AFFF (4,5,6,7) and all except one (7), used non-specific methods such as COD, BOD and TOC to assess the fate of AFFF. These parameters do not measure AFFF directly. Since the residual jet fuel can contribute significantly to these measurements, AFFF degradation was likely overestimated.

II. Objectives of the Research Effort

The overall objective of the Aqueous Film Forming Foams project is to evaluate the effectiveness of on-site biological treatment to remove AFFF containing wastewaters. Although attempts have been made to analyze the fate of AFFF's, there are no known analytical procedures for measuring AFFF components. This makes the process of evaluating treatment options extremely difficult.

The specific goals of my project were:

1) development of methods to separate and quantitate the organic components of AFFF's

2) use these methods to determine the rate of biodegradation of these components in laboratory scale microcosms.

III. AFFF Characterization

Ansul Co. Ansulite and 3M FC-206/203 are the AFFF's currently being used by the Air Force. After initial review of the information available on the AFFF's, we began laboratory work using FC-206 as a model AFFF because information about it's components and possible methods of analysis were available from 3M (Table 1). We decided to use High Performance Liquid Chromatography (HPLC) for the analysis of the surfactant components of FC-206. HPLC is a powerful tool for the separation and quantitation of complex mixtures of organic materials. It is not limited by sample volatility or thermal stability of the components.

Reverse phase liquid chromatography has been used to separate the components of AFFF (7). After a brief review of the literature on surfactant analysis that was available, we modified the methods used in the above report. The Waters HPLC system used in this study consisted of a 2 pump gradient system, gradient controller, and recording integrator. Peaks were detected with a variable wavelength UV/VIS detector, (Waters), a conductivity detector (Dionex), or differential refractometer (Perkin Elmer). Chromatograms shown in the Figures were produced using a 30 cm, 10 u particle size, uBondpak C_{18} reverse phase column with an acetonitrile (ACN)/reagent grade Millipore Q water solvent system (30:70 v/v).

The chromatograms shown in Figure 1 were obtained after 6 weeks of methods development (varying gradients, columns, solvents, detectors, wavelengths, etc.). The three detectors were connected in series so A thru C represent chromatograms of the same sample. The UV peaks with retention times (RT) of 0 - 3.5 min are those materials which are not retained on the column. These peaks likely represent inorganic salts and the urea component of AFFF. This is supported by the numerous peaks seen in the conductivity detector for the same time period. RT of the conductivity peaks correspond to peaks 0.4 min earlier in the UV chromatogram because of the great distance from the column to the conductivity detector. (ie.. peak 1.61 corresponds to peak 1.23 on the UV scan). The dip seen in Figure 1B is due to water in the 200 uL sample, which has a lower refractive index than the solvent. Any peak recorded by the other detectors during and just after this dip are going thru the column in the void volume.

The peaks at RT 4.67 and 4.85 are probably the anionic and nonionic hydrocarbon surfactants. Both these peaks have absorption maxima between 260 and 278 nm (Figure 2 shows chromatograms at 254 nm) in addition to some absorption at 215. The second peak also gave a response in the conductivity detector indicating that it was an ionic species.

Discussions with a 3M chemist as well as a review of surfactant chemistry lead us to believe that the anionic surfactant is a linear alkyl sulfonate (LAS) and the nonionic surfactant an alkyl phenol ethoxylate (APE). These surfactants are generally found commercially as a mixture of alkyl chain lengths rather than pure compounds. Both surfactant types are aromatic (Table 2) and the LAS would also be ionic. These identifications are tentative and further characterization of these peaks as well as comparison with purified standards by GC/MS or HPLC/MS is required.

Analysis of the fluorosurfactant components (FS), FM-4197 and FM-3820 has proven to be more difficult, although some characterization was done this summer (Table 2). A limited literature search indicated that little HPLC has been done with these compounds. We are currently using a computer search to conduct a more extensive review.

The 3M Corporation and Ansul Co. were kind enough to send samples of the two flourosurfactants found in their products. We chose to work on the 3M products. Although neither of the FS showed absorption in the UV range (data not shown), both were detectable by refractive index detection (Figure 3). In Figure 1, the RI peak at 4.45 min elutes before the peaks tentatively identified as the hydrocarbon surfactants. This is also seen in Figure 2. We have tentatively assigned the peak detected by RI which elutes before the hydrocarbon UV peaks as the FS component of FC-206.

We also used GC/MS to examine the flourosurfactants (FS) from both 3M and Ansul. A Hewlett Packard Model 5987 GC/MS equipped with a silicone capillary column was operated with a temperature gradient from 60 to 350 C. Figure 4 shows some typical chromatograms and mass spectra of components detected in Ansul FC A and B and 3M 4197 and 3820. Figure Ai/Aiv and Figure Ci/Ciii are chromatograms of Ansul and 3M fluorosurfactants, respectively. The predominant peak in the chromatograms is the ethylene glycol monobutyl ether (EGMBE), the solvent the FS were dissolved in. This was detected and confirmed by MS in Figure Bi, Biv, and Bv. Peak 9.11 was also confirmed to be EGMBE. A second glycol was also detected in chromatograms of FC-206 (Figure Bii and Biii). Figures Aii and Av are the mass spectra of a predominant peak in Ansul FS A and B. Figures Cii and Civ are the spectra for the predominant peak in FS 3820 and 4197. The spectra contain all the m/e fragments expected from a perfluoroalkyl moiety. A spectrum for perfluorodecane is shown in Figure Aiii for comparison.

Many of the small peaks in both chromatograms were identified as containing perflouroalkyl groups. However, none of the peaks were identified as the parent fluoro-surfactant. This was not unexpected since the parent compound is not very volatile. We hypothesize that the perfluoroalkyl peaks we see are thermal degradation products of the FS.

The use of chemical or thermal destruction and subsequent GC/MS analysis as a method for identification and quantitation of the FS components in HPLC needs to be pursued.

Using the information above, calibration curves were estabished for FC-206 for concentrations ranging from 200 to 10,000 ppm AFFF. Two RI peaks were measured in the HPLC chromatograms (Figure 5a). The peak with RT = 1.57 is an unkown component; the peak with RT = 4.00 is the tentative FC peak. The two hydrocarbon surfactant peaks were also quantitated (Figure 5b). All showed good linearity over the range tested. These peaks were subsequently used to determine biodegradation of AAAF.

The Methylene Blue Active Substances (MBAS) assay was used to detect the concentration of anionic surfactants in FC-206 (8,9). A calibration curve for AFFF is shown in Figure 6a.

Another important organic component in the AFFF, the EGMBE solvent (Table 1), was quantitated by gas chromatography. A Perkin Elmer GC equipped with a flame ionization detector and a Carbowax 20-M (terminated with terphthlic acid) packed column was run isothermally at 180 C. A calibration curve for EGMBE in FC-206 is shown in Figure 6b.

IV. Biodegradation of FC-206

The source of microbial communities used in the biodegradation experiments were water samples (Panama City activated sludge, aeration lagoon water, the

oil-water seperator, the earthen pit at the fire training facility) and soil (from outside and inside the pit). Ten grams of soil or 10 ml of water from each of these environments were added to three 500 ml flasks containing 100 ml distilled water, 20,000 ppm FC-206 AFFF and 100 ug/L phosphorus ($NaPO_4$) and nitrogen (NH_3NO_3). All flasks were incubated at 30 C on a rotary shaker. A fourth flask was prepared in a similar manner and autoclaved to serve as a sterile control. No chemical or microbiological changes were noted in the controls during the experiment. The biodegradation of AFFF components was followed during the last two weeks of the project.

In all cases there was visible bacterial growth in the flasks, indicating that the components of AFFF can serve as a source of carbon for growth. On the average the numbers increased from 10^6 to 5 x 10^7 cells/ml. EGMBE decreased to below detectable limits in all flasks by the end of two weeks. The MBAS only showed a decrease at the last sampling time (2 weeks). The greatest reduction (3%) was noted with the 2 soil derived communities. The UV detected LAS peak decreased by 2%, indicating that the MBAS decrease was mostly due to LAS degradation. No reduction in the FS peaks were noted. These results suggest that some specific components of AFFF are biodegraded and that further long-term experiments are needed to determine the biodegradability of other components. Our initial results indicate that on-site bacteria can be used to treat AFFF wastewaters.

V. Recommendations

1. The analytical work that was begun this summer should be continued. Confirmation of the RI and UV peaks as surfactants needs to be carried out. Further, use of methods to seperate anionic and nonionic components before HPLC (8,9) is strongly suggested.

2. The MBAS and CTAS standard methods should be used to measure the anionic and nonionic surfactants in present studies. These direct measurements for the surfactant components of AFFF can be made interference free (9).

3. The initial lab-scale aerobic biodegradation experiments were encouraging. These should be continued and expanded using the MBAS and CTAS methods as well as the improved HPLC method above for the surfactants and GC method for the EGMBE component.

4. Anaerobic biodegradation studies should be initiated. These can be bench scale microcosms using sulfate reducing, denitrifying, and methanogenic conditions. Thses can be conpared to the aerobic degftradation rates.

References

1. Fink, P.T., Air Force Technical Report, CEEDO-TR78-45, "Aqueous Film Forming Foam Treatability", Civil and Environmental Engineering Development Office, Tyndall AFB, FL, 1978.

2. Wang, E.H., Air Force Technical Report, CEEDO-TR77-7, "Bioassay of Air Force FC-206 Fire Fighting Foam", Civil and Environmental Engineering Development Office, Tyndall AFB, FL, 1977.

3. Chan, D.B., Civil Engineering Laboratory Technical Memorandum No. M-54-78-06, "Disposal of Wastewater Containing Aqueous Film Forming Foam (AFFF)", The U.S. Navy Civil Engineering Laboratory, Port Hueneme, CA, 1978.

4. Thomas, J.F., and LeFebvre, E.E., Report No. EHL(K) 74-3, "Biodgradability and Toxicity of FC-200, Aqueous Film Forming Foam", USAF Environmental Health Laboratory, Kelly AFB, TX, 1974.

5. LeFebvre, E.E., and Thomas, J.F., "Biodegradability and Toxicity of AER-O-WATER 3 and AER-O-WATER 6 Aqueous Film Forming Foam" USAF Environmental Health Laboratory, Kelly AFB, TX, 1974.

6. LeFebvre, E.E., and Inman, R.C., Report No. EHL(K) 74-26, "Biodgradability and Toxicity of FC-206, Aqueous Film Forming Foam", USAF Environmental Health Laboratory, Kelly AFB, TX, 1974.

7. Bass, C.M., "The Fate and Effects of Aqueous Film Forming Foam (AFFF) as It Passes Through Sand, Soil, or Activated Carbon Biofilter with Appendices on High Performance Liquid Chromatography and the Beckman Microtox Toxicity Monitor Applications to AFFF. Ph.D. Dissertation, University of Oklahoma, Norman, OK, 1982.

8. American Public Health Association. Standard Methods for the Examination of Water and Wastewater, 15th ed., 1983.

9. Milwidsky, B.M., and Gabriel, D.M., Detergent Analysis: A Handbook for cost-effective quality control. Halstead Press (John Wiley and Sons, Inc.), New York, 1982.

Table 1. Components of 3M AFFF FC-206.

Ingredient	% (v/v)
Diethylene glycol monobutyl ether	15
Water	76
Fluoroalkyl Surfactants (2)	5
Hydrocarbon Surfactants (2)	5
Urea	4
Tolyltriazole	<0.1

Table 2. Characteristics of Individual Surfactants of FC-206.

Component Name	Characteristics
Hydrocarbon 1	Anionic, MBAS +, UV detectable, (215/254/270 nm), probably LAS, biodegradation reported
Hydrocarbon 2	Nonionic, CTAS +, UV detectable, probably alkyl phenol ethoxylate, biodegradation reported (slow)
Fluorocarbon 3820	Anionic, MBAS +, No UV, RI +, conductivity maybe (?), ID perfluoro group in GC/MS, degradation not likely
Flourocarbon 4197	Zwitterionic, MBAS +, No UV, RI +, conductivity (?), ID perfluoro group in GC/MS, degradation not likely

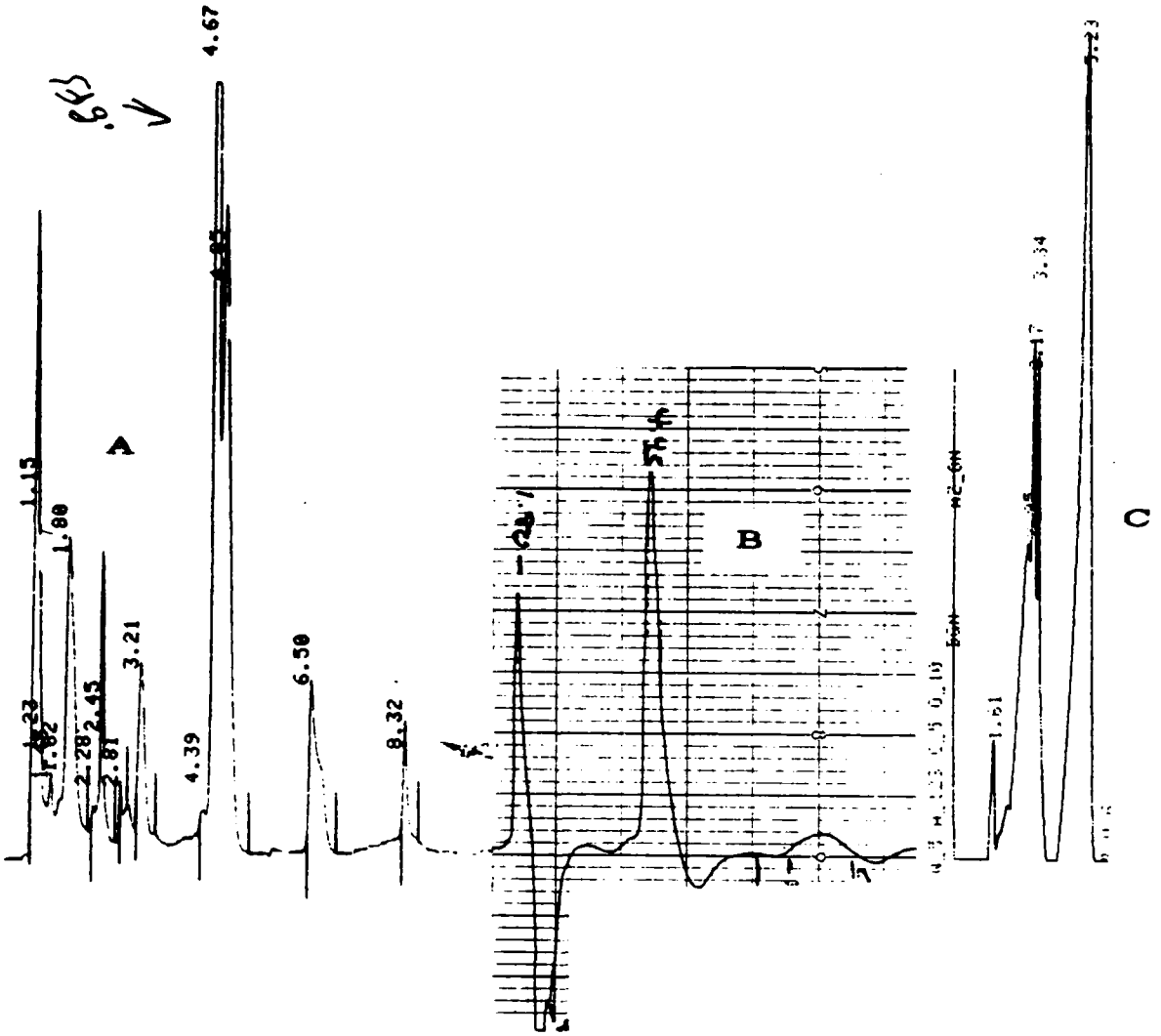


Figure 1. Reverse phase HPLC of 20,000 ppm (v/v) 3M FC-206. Isocratic conditions acetonitrile/water (30:70 v/v). (A) UV detection (215 nm, 0.2 OD full scale). (B) Refractive Index detection (range 200). (C) Conductivity detection (100 Si full scale).

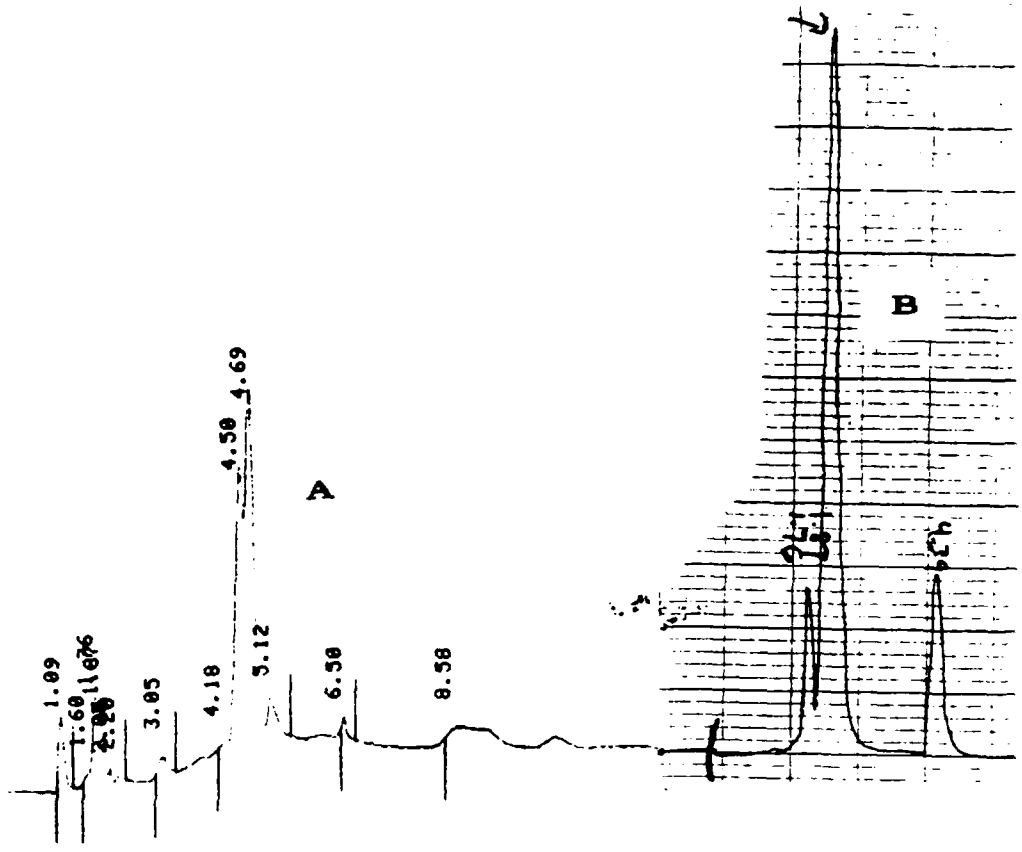


Figure 2. HPLC of 20,000 ppm (v/v) FC-206. (A) UV detection (254 nm, 0.2 OD full scale). (B) Refractive Index detection (range 200).

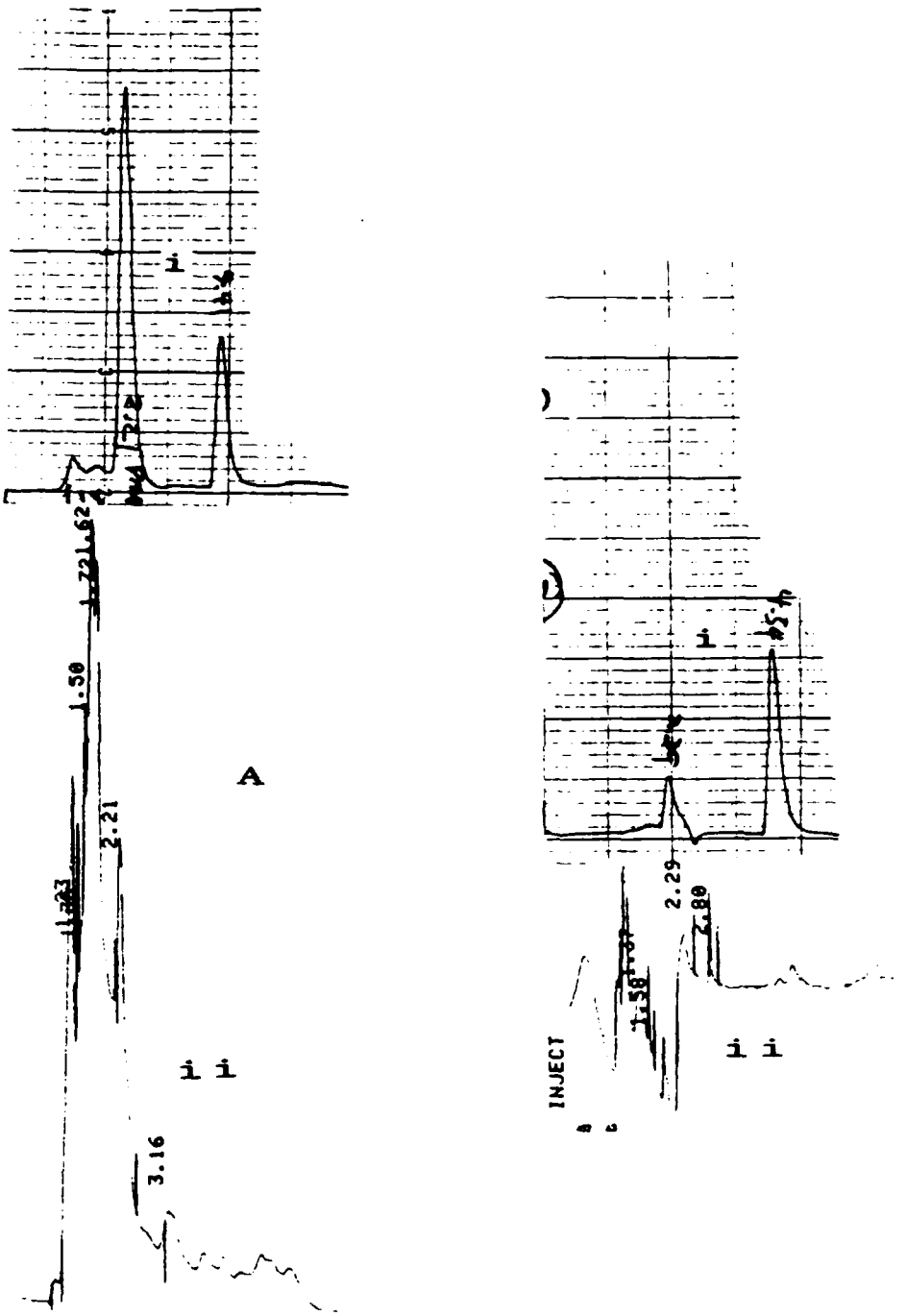


Figure 3. Comparison of RI and UV detection of FS FM-4197 and FM-3820.
(A) (LEFT) FM-4197, i. RI detection (range 500) ii. UV detection (215 nm). (B) (RIGHT) FM-3820, i. RI detection (range 500) ii. UV detection (215 nm).

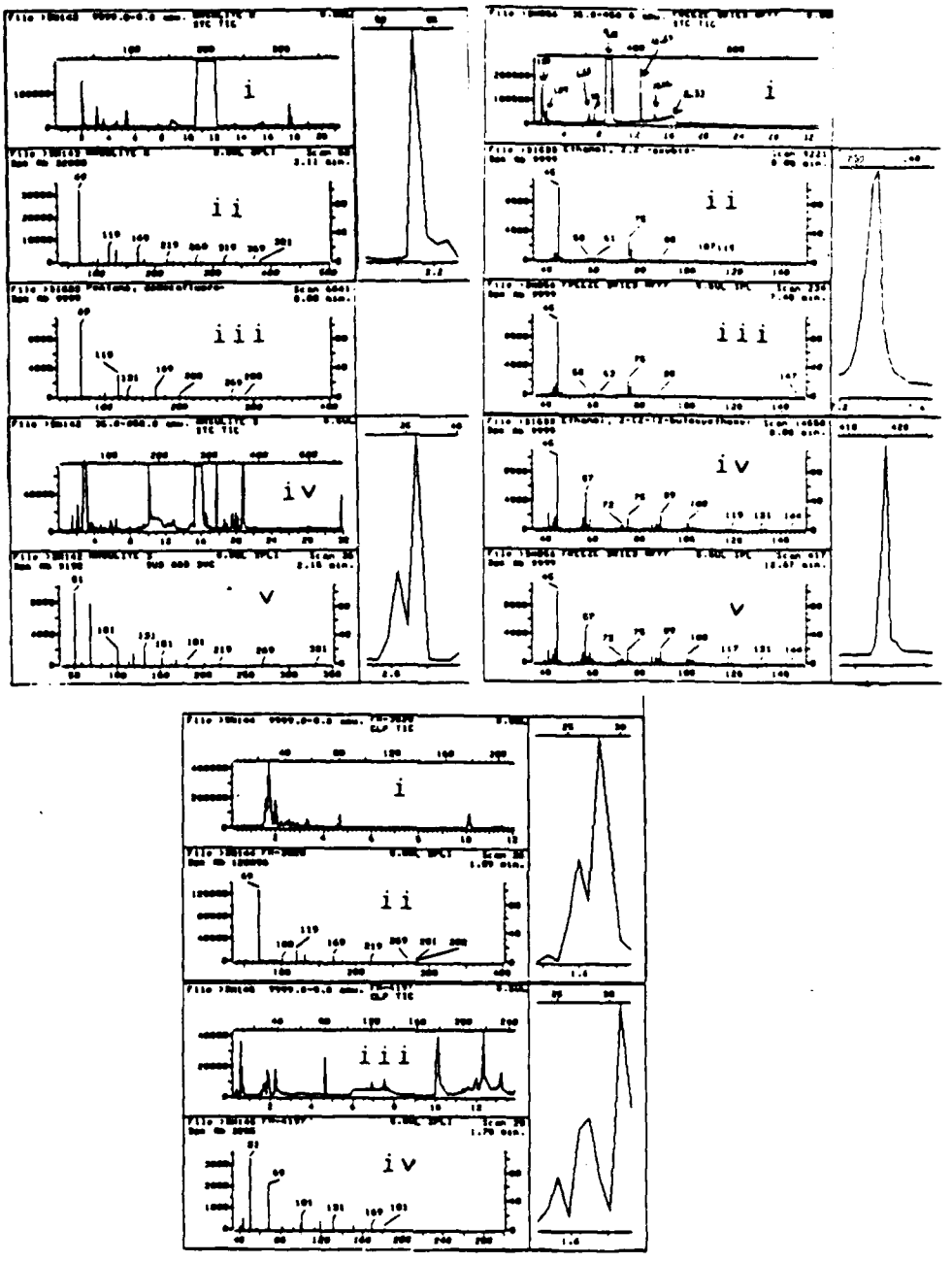


Figure 4. Capillary gas chromatagraphic separation and detection by mass spectrometry of AFFF components.
(A) Ansulite FS A and B components, i. GC separation of A, ii. mass spectra of peak with RT 3.11 min, iii. mass spectra of perfluorodecane standard, iv. GC separation of B, v. mass spectra of peak with RT 2.16.
(B) Freeze dried 3M FC-206 AFFF, i. GC separation of AFFF, ii. mass spec of 2-2'-oxybis-ethanol standard, iii. mass spectra of peak with RT 7.40 min, iv. Mass spectra of ethylene glycol monobutyl ether, v. mass spectra of peak with RT 12.67 min.
(C) 3M FS 3820 and 4197, i. GC separation of 3820, ii. mass spectra of peak with RT 1.59 min, iii. GC separation of 4197, iv. mass spectra of peak with RT 1.78 min.

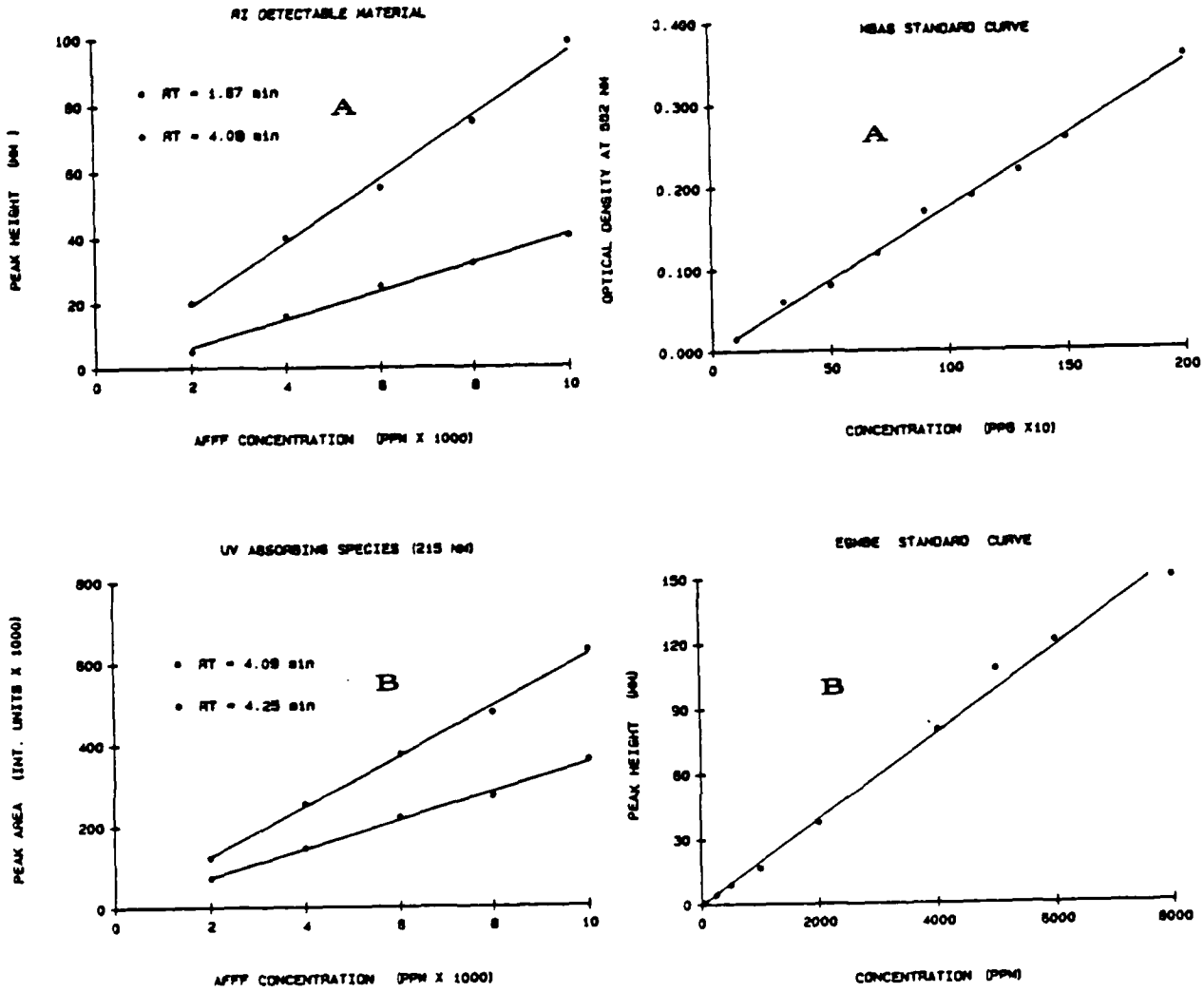


Figure 5. (LEFT) Calibration curves for HPLC peaks tentatively identified as surfactants in 3M FC-206. (A). Refractive Index detectable materials. (B). UV absorbing peaks. Similar curves are obtained at 254 nm.

Figure 6. (RIGHT) Calibration curves for (A) Methylene Blue Active Substances (MBAS) and (B) ethylene glycol monobutyl ether solvent in 3M FC-206 AFFF.

1986 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/
GRADUATE STUDENT SUMMER RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

Comprehension and Cohesion of Text

Prepared by:	Doris J. Walker-Dalhouse
Academic Rank:	Associate Professor of Reading
Department and University:	Department of Administration, Curriculum, and Instructional Media Jackson State University
Research Location:	Human Resources Laboratory Intelligence Division Air Force Systems Command Brooks Air Force Base, Texas
USAF Researcher:	Lt. Col. Hugh Burns

Date:	July 30, 1986
Contract No:	F49620-85-C-0013

Comprehension and Cohesion of Text

by

Doris J. Walker-Dalhouse

ABSTRACT

The issues of discourse analysis, schema and reading comprehension were examined and relevant literature read. An annotated bibliography was compiled on research relative to the underlying theoretical foundation of cognitive text processing and reading comprehension. Walter Kintsch's Model of Discourse Analysis was analyzed in relative depth to ascertain its implications for research in Reading and applicability to the field of Artificial Intelligence. Air Force technical training manuals were examined to informally assess the tecnical demands made upon the reader. Preliminary thought was given to the relevance of the Kintsch model to the understanding of technical prose.

ACKNOWLEDGEMENTS

Gratitude is extended to the Air Force Systems command, and the Air Force Office of Scientific Research for sponsorship of the Summer Faculty Research Program administered by Universal Energy Systems. I would like to thank the Human Resources Laboratory, especially Colonel Gene A. Berry, Division Chief of the Brooks/AFB Training Systems Division and the Intelligence Division Staff, who provided both an intellectually stimulating and cooperative environment for research generation.

Special thanks are extended to Lt. Col. Hugh Burns, with whom I worked most closely, and Dr. Philip Gillis for their assistance in the crystallization of my research plan and guidance relative to future implementation. Additional, thanks are extended to Dr. William Tirre, Learning Aptitude Measurement Program, and Dr. Bob Pokorny, Basic Skills Program, for their technical assistance in the sharing of expertise and resources which proved invaluable in the formation of my research objectives.

I. INTRODUCTION

Artifical Intelligence (AI) is the study of intelligent behavior and its application to the development of machines that perform tasks or that aid in the performance of tasks normally requiring human intelligence. The Field of AI has a strong relationship to reading and writing along with computer science, because a truly intelligent computer must be capable of reading input and writing output. Balajthy, 1985)

The AI branch under the leadership of Lt. Col. Hugh Burns, Jr. is attempting to build computers capable of teaching future Air Force job training skills by using intelligent tutoring systems. The knowledge based systems will integrate rules about student learning, instruction and specific job fields within instructional tutors. Thus, the AI unit's focus is on the application of artifical intelligence to instructional delivery and support of future AF training needs.

Having completed graduate training in Reading Education and developed an interest in the nature of reading comprehension development in mature readers and the concomitant instructional implications, I find the area of artifical intelligence to be a furtile one for advancing research in the area of reading.

Comprehension of written text or the operation of the brain in processing language is of major interest in the implementation of the Air Force's goals. I envision my schemata, background knowledge, as both drawing from and adding to the research about the effect of semantic and syntactic factors on the comprehensibility of print materials.

II. OBJECTIVES

The objectives for the specified research period were formulated during the site visit and subsequently expanded at the onset of the research period. They were as follows:

1) To investigate semantic avenues of readability and propositional analysis of text.

2) To review literature pertaining to discourse analysis, schema theory, and linguistic theory.

3) To design a test of reading comprehension as a measure of cognitive processing of technical text.

III. PROCEDURE

In order to develop a knowledge base for an understanding of the theoretical basis of discourse analysis, schema, and reading comprehension, a bibliographic search was made of literature pertaining to the subject from 1972-1986. A select number of the materials found were read and annotated in preparation for formulating a follow-on research proposal. Summaries of the main ideas of findings from twenty of those articles are presented in the following section.

IV. RESULTS

Richgels, D.J. Schema theory, Linguistic Theory and Representation of Reading Comprehension. Jounral of Educational Research, 76 (1982), 54-62.

Reading comprehension is designed as the result of successful interaction of reader with a text, the influence of schema theory and linguistic theory on its development are discussed. Shank's, Frederiksen's, and Kintsch's models of comprehension are explained with emphasis placed on their value to reading education. Kintsch's strategy-based model is postulated as the most effective in assimilating psychological and linguistic theories which will help promote an understanding of how readers comprehend.

Bransford, J.D. and M.K. Johnson. Contextual Prerequisites for Understanding: Some Investigations of Comprehension and Recall. Journal of Verbal Learning and Verbal Behavior, 11 (1972), 717-726.

A review of research indicating the importance of relevant contextual knowledge as a prerequisite for comprehending prose passages is presented. The objective of the present research was to manipulate the availability of prior knowledge to determine its influence on S's ability to understand and recall linguistic materials.

Experiment 1. found that comprehension rating and recall of ideas were higher for subjects in the Context Before group (ones who saw the appropriate context picture before learning the passage), as opposed to the Context After, and No Context groups. In experiments II, III and IV, S's were grouped into Topic Before, Topic After and No Topic groups. Comprehension and recall were greater in the Topic Before condition than in the Topic After and No Topic groups. These results indicate that in order for prior knowledge to aid comprehension, it must become an activated schema.

Britton, B. and S.M. Glynn. Effects of Text Structure on Cognitive Capacity during Reading. Journal of Educational Psychology, 76 (1982),51-61.

Report of findings from three studies investigating the demand of text processing on learners' cognitive capacity using a secondary' task technique. Experiments 1 and 2 required reading technical texts or rewritten versions of standard English texts with structural simplifications. Simplified vocabulary and syntax were found to require less cognitive capacity to process with syntactical factors primarily responsible for the observed differences. The use of signal words in Experiment 3 indicated the importance of ideas and their relationships required less cognitive capacity to process than texts without signal words.

Rasool, J.M. and J.M. Royer. Assessment of Reading Comprehension Using the Sentence Verification Technique. Evidence from Narrative and Descriptive Texts. Journal of Educational Research, 79 (1986), 180-184.

The Sentence Verfication Technique (SVT), created by Royer, Hastings and Hook (1979), was used with narrative and descriptive texts to determine the technique's validity as a measure of reading comprehension for third grade students designated by their teachers as good, medium, and poor readers. An ancillary purpose of the study was to examine the relationship between SVT performance and student exposure to and knowledge relative to various text types.

Results revealed the SVT's sensitivity to teacher defined differences in reading competence and a positive correlation with student comprehension of printed text.

Bowman, M. A Comparison of Content Schemata and Textual Schemata or the Process of Parachuting. Reading World, 21 (1981), 14-22.

A discussion of Anderson, Pichert, and Shirey's (1979) contention that content schemata (the reader's existing or prior knowledge) is more significant to the reader than textual schemata (knowledge of discourse-level convention of text). Relevant research supporting either content or textual schemata as being more instrumental in reading and discourse recall is presented. It is postulated that the more useful schemata is determined by the reader-material match than by one's schemata set. The author believes that it would be more valuable to examine the interaction of content schemata with textual schemata.

Richgels, D.J. Dispelling the Mystery about Comprehension: Kintsch's Model and Implications, Reading Horizons, 24 (1984), 72-78.

Presents Walter Kintsch's model of comprehension and information on linguistic and schema theory in an attempt to enable teachers to be more consistent in their teaching, and more confident of their intuitive knowledge of comprehension. Implications of Walter Kintsch's model of comprehension for readers and teachers for comprehension development are presented. Emphasis for students consists of expanding their repertoire of concepts, paraphrasing sentences, identifying main ideas, and inferring text throughout the comprehension process. Instructional applications for teachers focus on readability and text analysis.

Reinking, D. and R.Schreiner. The Effects of Computer-Mediated Text on Measures of Reading Comprehension and Reading Behavior. Reading Research Quarterly, 20 (1985), 536-552.

A report of research designed to study the effect of computer mediated text manipulators on the reading comprehension of good and poor intermediate grade readers. After exposing the subjects to four experimental conditions using variations of computer or print forms, a comprehension test of six expository passages (3 low difficulty and 3 high difficulty) was administered. Results indicate that reading comprehension can be influenced by computer textual manipulations controlled more by the reader's exposure to textual manipulations as opposed to the availability of textual manipulations for his selection.

Goldman, S.R., Hogaboam, T.W., Bell, L.C. and C. Perfetti. Short-Term Retention of Discourse during Reading. Jounal of Educational Psychology, 72 (1980), 647-655.

Presentation of research into the recoding process during oral and silent reading and individual differences in this process for 8, 9, and 12 year old children. A probe memory task was used to assess the availability of recently read discourse in working memory. In Experiment 1, specifically designed to determine the effect of sentence boundaries and length of discourse segment, a positive relationship was found between these two factors. Experiment 2 was used to determine the limits of working-memory capacity and the role of the sentence boundary as a cue for recoding. Both were found to be related to individual differences among readers. Whereas, length and sentence boundary affected skilled and older readers, performance was consistently low for less skilled readers.

Royer, J. M. and G. Cable. Facilitated Learning in Connected Discourse. Journal of Educational Psychology, 67 (1975), 116-123.

The research tested the prediction that nonspecific facilitated learning of a second prose passage occurs when an initial passage read by subjects contained concrete referents designed to increase the comprehension of a second passage designated as difficult to understand. Twelve groups of two hundred forty S's read two passage before being asked to recall the information obtained from the second passage. Experiment 2 tested the prediction that facilitated learning of a second passage would occur only when subjects receive a concrete initial passage and an abstract second passage. The results of Exp. 2 supported the original hypotheses.

Carrell, P.L. Facilitating ESL Reading by Teaching Text Structure. TESOL Quarterly, 19 (1985), 727-752.

Study examined the issue of interaction of text structure with the reader's formal, rhetorical schemata in affecting reading comprehension. The major question asked was, "Can English, as a second language, reading comprehension be facilitated by explicity teaching text structure?" ESL students were trained to recognize top level rhetorical organization of texts, and to use text type to increase their reading comprehension as measured by the amount of the original text recalled. Results of qualitative data analysis found that training facilitated recall of supporting details in addition to major topics and subtopics.

Dooling, D.J. and R. Lachman. Effects of Comprehension on Retention of Prose 1. Journal of Experimental Psychology, 88 (1971), 216-222.

The purpose of the study was to determine if knowledge of a passage's theme facilitates retention of it's words using free recall or paced binary recognition. Comprehension of metaphorical passages was manipulated by prior presentation or nonpresentation of a short title representative of the theme or main idea. Experiment 1 showed that the experimental manipulation of surrogate processes facilitated comprehension. Experiment II designed a paced binary-recognition word test following the presentation of material in three forms: a) random words; b) random phrases, and c) prose. Superior recognition was limited to those words in the binary-recognition test which were directly relevant to the theme of stories. Abstract representation of the main idea of metaphorical text is believed to be useful as a mnemonic device in the retention of prose.

Duffy, T.M. and P. Kabance. Testing a Readable Writing Approach to Text Revision. Journal of Educational Psychology, 74 (1983), 703-746.

Comprehensibility is identified as the one dimension of text usability which has received little attention because of the difficulty in defining and measuring it objectively in a learning context. The author's contend that readability formulas only focus on the writing style of prose while ignoring graphics, graphic-prose coordination, format or organization. The present research was a test of the effectiveness of using readable writing revisions as a technique for producing more comprehensible text. In three of the four experiments performed, comprehension was not facilitated by the simplication of vocabulary or sentence structure. Vocabulary simplication was found to strongly affect the performance of low ability readers required to complete a memory task or reading-to-learn task.

Anderson, R.C., Pichert, J.W., and L.L. Shirey. Effects of the Reader's Schema at Different Points in Time. Journal of Educational Psychology, 75 (1983), 271-279.

A distinction is made between textual schemata and content schemata. The purpose of the first of the two research investigations reported was to determine whether a reader's schema has both encoding and retrieval effects, and whether or not they operate independently. Results from Experiment 1 during which subjects took one of two perspectives before reading a passage and then shifted to another perspective after reading, revealed that both have major effects on recall, effects which operate independently. Experiment 2 examine the effect of introducing a new perspective, a lengthy time period after actually reading a passage. Overall results indicate that perspectives assigned prior to, immediately following or considerably long after reading, significantly influence text recall.

Masson, M.E.J. and J.A. Miller. Working Memory and Individual Differences in Comprehension and Memory of Text. Journal of Educational Psychology, 75 (1983), 314-318.

The role of working memory in reading comprehension was investigated as an extension of the related work of Danenan and Carpenter (1980). The research objectives were to determine if working memory as measured by a reading span test was significantly associated with the ability to integrate information from different parts of a test passage to infer ideas, and to determine the influence of working memory on the ability to encode explicitly stated and inferred information into long-term memory. Storing and processing of information in working memory was found to correlate with a) comprehension test scores, b) integration of text information in drawing inferences, and c) long-term memory encoding and recall of explicitly stated and inferred information.

Brooks, L., Dansereau, D.F., Spurlin, J.E. and C.D. Holley. Effects of Headings on Text Processing. Journal of Educational Psychology, 75 (1983), 292-302.

Two studies were conducted to investigate the effects of hierachial, author-provided, intact and embedded outlines (headings) on text learning of unfamiliar expository prose. Embedded headings were found to reliably improve delayed text performance in Experiment 1. Experiment 2 was a test of the use of these headings as processing aids to enhance performance. Results showed a significant and positive effect for instruction in the use of headings. No significant relationship was found between field independence/dependency on the Group Embedded Figures Test and text processing. Both experiments supported the use of embedded headings as processing aids.

Freebody, P. and R. Anderson. Effects of Vocabulary Difficulty and Cohesion and Schema Availability on Reading Comprehension. Reading Research Quarterly, XVIII (1983), 277-294.

The compensation hypothesis, a predictive hypothesis generated from the interactive theory of reading, proposes that when one source of knowledge about the meaning of a text element is inoperative, alternate means of determining meaning are supplied by other knowledge sources. Tests of this hypothesis, were performed in two experiments using sentence verification, free recall and summarization measures. No relationship was found between vocabulary difficulty and text cohesion in Experiment 1 nor between vocabulary difficulty and topic familiarity in Experiment 2. Vocabulary, however, did significantly affect student performance in Experiment 1. The primacy of topic schemata and comprehension was evidenced in Experiment 2 by better recall of the most familiar versions of text passages. Overall results indicated a lack of support for the interactive theory of reading.

Townsend, M.A.R. Flexibility of Schema: A Shifting in Good and Poor Readers. Journal of Reading Behavior, XIV (1982), 169-177.

Study of the ease of good and poor third grade readers in shifting between familiar schemata presented in a listening comprehension task utilizing two short passages describing daily activities. A shift in schemata was necessary for comprehension of the second passage. Good and poor readers in the Cued Shift condition received explicit cues to the passage shift through two repetitions of the title prior to hearing each passage, while those in the Uncued Shift condition were given both passage titles prior to hearing the First passage. Although subjects in the Cued Shift condition showed no difference in recall of the two passages, subjects in the Uncued Shift condition exhibited a decline in second passage recall. Results suggest that schema shifting is equally difficult for good and poor third grade readers.

Townsend, M.A.R. Schema Activation in Memory for Prose. Journal of Reading Behavior, XII (1980), 49-53.

Presentation of research examining the effect of information recall of connected discourse by college students exposed to one of four contexts conditions: appropriate context before reading; appropriate context after reading; no context information, or inappropriate context information. Recall of passage information was significantly lower for subjects who were not given contextual information at the time of acquisition of material. Subjects given inappropriate contextual information recalled less information than subjects given no contextual information.

Carnine, D.W., Kameenui, E.J. and N. Woolfson. Training of Textual Dimensions Related to Text-Based Inferences. Journal of Reading Behavior, XIV (1982), 335-339.

Distinguishes between the text-based and schema-based approaches to comprehending discourse while acknowledging the presence of numerous theoretical frameworks for examining inferences in reading comprehension. Data is provided relative to the use of a systematic instructional practice to examine the effects of specific textual dimensions and precise interventions on a particular type of inference behavior. Systematic instruction utilizing a procedure for identifying the problem statement, rule, and critical information, and which ignored irrelevant or distracting information elicited significantly higher scores on transfer stories.

Marshall, N. and M.D. Glock. Comprehension of Connected Discourse: A Study into the Relationships between the Structure of Text and Information Recalled. Reading Research Quarterly, XIV (1978-79), 11-56.

A historical overview of research pertaining to reading comprehension and definition of related terms is presented. The review provides a perspective for the present study which focused on the effects of variations in the logical network defined by Frederikson (1975) and staging by Grimes (1975) on reading comprehension and recall. Specifically, the aim of the study was to investigate the effects of variation in text structure upon the structure and content of the written recalls of 160 Ivy League and community college students. The following aspects of text structure were examined: the presence or absence of the if-then relationship between clauses, and the comparative or superlative forms of adjectives, and the effects of difference in order of several sentences. These manipulations were found to produce more complete recall by the

Cornell subjects than the Auburn Community College (ACC) subjects. Fredericksen's model of text and memory was supported by the research findings.

Kintsch, W. and J.C. Yarbrough. Role of Rhetorical Structure in Text Comprehension. Journal of Educational Psychology, 74 (1982), 828-834.

The experiments examined the role of text comprehension strategies utilizing rhetorical cues. Four rhetorical forms: classification, illustration, comparison and contrast, and procedural description were used to construct one structurally simple text and one structurally complex text for each form. Rhetorical form of texts were found to strongly affect the subjects' ability to answer topic and main idea questions. While good versions of the text produced better performance than shown on poor versions, performance on a cloze test was identical for both versions. Statistical significance was found for complexity with more correct answers given after simple texts were read than after reading complex texts.

V. RECOMMENDATIONS

It is recommended that the literature and research preparation begun during the 1986 Summer research period be continued in a follow-on Air Force Mini Grant. The purpose of the proposed research would be to investigate the cohesion and comprehensibility of Air Force technical manuals in one of four training areas. Semantic theories of text cohesion should be further examined to provide the foundation for technical discourse, and the resulting instructional implications for AF staff and reading educators. Such information will be useful in rearranging underlying meaning of text to make them more readable.

This research will be of significance in providing Artificial Intelligence personnel with additional data for the development of intelligent computer assisted instructional (ICAI) sytems of educational significance to AF personnel and the larger academic community.

REFERENCES

Anderson, R.C., Pichert, J.W. and L.L. Shirey. Effects of the Reader's Schema at Different Points in Time. Journal of Educational Psychology, 75 (1983) 271-279.

Bowman, M.A. Comparison of Content Schemata and Textual Schemata or the Process of Parachuting. Reading World, 21 (Oct. 1981) 14-22.

Bransford, J.D. and M.K. Johnson. Contextual Prerequisites for Understanding: Some Investigations of Comprehension and Recall. Journal of Verbal Learning and Verbal Behavior, 11 (1972) 717-726.

Britton, B. and S.M. Glynn. Effects of Text Structure on Cognitive Capacity During Reading. Journal of Educational Psychology, 76 (1982) 51-61.

Brooks, L., Dansereau, D.F. Spurlin, J.E. and C.D. Holley. Effects of Headings on Text Processing. Journal of Educational Psychology. 75 (1983) 292-302.

Carnine, D.W., Kameenui, E.J. and N. Woolfson. Training of Textual Dimensions Related to Text-Based Inferences. Journal of Reading Behavior XIV (1982) 335-339.

Carrell, P.L. Facilitating ESL Reading by Teaching Text Structure. TESOL Quarterly 19 (1985) 727-752.

Dooling, D.J. and R. Lachman. Effects of Comprehension on Retention of Prose 1, Journal of Experimental Psychology, 88, (1971) 216-222.

Duffy, T.M. and P. Kabance. Testing a Readable Writing Approach to Text Revision. Journal of Educational Psychology, 74 (1983), 703-746.

Freebody, P. and R. Anderson. Effects of Vocabulary Difficulty and Cohesion and Schema Availability on Reading Comprehension. Reading Research Quarterly XVIII (1983) 277-294.

Goldman, S.R., Hogaboam, T.W., Bell, L.C. and C. Perfetti. Short-term Retention of Discourse During Reading. Journal of Educational Psychology, 72 (1980) 647-655.

Kintsch, W. and J.C. Yarbrough. Role of Rhetorical Structure in Text Comprehension. Journal of Educational Psychology, 74 (1982) 828-824.

Marshall, N. and M.D. Glock. Comprehension of Connected Discourse: A study into the Relationships between the Structure of Text and Information Recalled. Reading Research Quarterly, XIV (1978-79) 11-56.

Masson, M.E.J. and J.A. Miller. Working Memory and Individual Differences in Comprehension and Memory of Text. Journal of Educational Psychology. 75 (1983) 314-318.

Rasool, J.m. and J.M. Royer. Assessment of Reading Comprehension Using the Sentence Verification Technique: Evidence from Narrative and Descriptive Texts. Journal of Educational Research, 79, (1986) 180-184.

Reinking, D. and R. Schreiner. The Effects of Computer-Mediated Text on Measures of Reading Comprehension and Reading Behavior. Reading Research Quarterly, 20 (1985) 536-552.

Richgels, D.J. Dispelling the Mystery about Comprehension: Kintsch's Model and Implications. Reading Horizons, 24 (1984) 72-78.

Richgels, D.J. Schema Theory, Linguistic Theory, and Representation of Reading Comprehension. Journal of Educational Research, 76 (1982) 54-62.

Royer, J.M. and G. Cable. Facilitated Learning in Connected Discourse. Journal of Educational Psychology, 67 (1975) 116-123.

Townsend, M.A.R. Flexibility of Schema: A Shifting in Good and Poor Readers. Journal of Reading Behavior, XIV (1982) 169-177.

Townsend, M.A.R. Schema Activation in Memory for Prose. Journal of Reading Behavior, XII (1980) 49-53.

1986 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT SUMMER SUPPORT PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

Effect of Hyperoxia on the Permeability of the Blood-Brain Barrier in the Rat

Prepared by:	Donald W. Welch, Ph.D
Academic Rank:	Research Scientist
Department and	Hyperbaric Laboratory
University:	Texas A&M University
Research Location:	Clinical Investigative Facility
	Wilford Hall Medical Center, Lackland AFB
USAF Researcher:	William Ehler, D.V.M., GM-14, DAF
Date:	September 22, 1986
Contract No:	F49620-85-C-0013

Effect of Hyperoxia on the Permeability of the Blood-Brain Barrier in the Rat

by

Donald W. Welch, Ph.D

ABSTRACT

The overall objective of this project was to examine the effect of increased partial pressures of oxygen on the permeability of the blood-brain barrier (BBB).

Sterile surgical techniques were developed to chronically implant silastic catheters in the jugular vein in rats. With daily flushing these catheters remained patent for 8-14 days. Blood-brain barrier integrity was evaluated under blind conditions using sodium flourscein as the permeability marker. Tissue was examined at ten different levels of the central nervous system. The tissue was examined with a dissecting scope while simultaneously being exposed to ultraviolet light in a specially constructed light box.

Rats were exposed to hyperbaric oxygen at 2.0, 2.5 and 3.0 atmospheres absolute (ATA). The maximum exposure employed was 3.0 ATA for a total of 210 minutes on 100% oxygen. Even under these severe conditions the BBB of the rat could only be termed as "minimally leaky". It was decided that these studies should be expanded to rabbits and guinea pigs to compare our results with those previously reported. Rabbits and guinea pigs exposed to 100% oxygen at 3.0 ATA for a total of 90 minutes showed a dramatic increase in BBB permeability when compared to controls. These preliminary observations indicate a significant difference in the susceptibility of the blood-brain barriers of these species to hyperoxic conditions.

Acknowledgments

I would like to thank the Air Force Systems Command and the Air Force Office of Scientific Research for sponsorship of my research. I would also like to thank several members of the staff of the Clinical Investigative Facility at Wilford Hall Medical Center, Col. Marion J. Stansell, Lt. Col. John H. Cissik, Lt. Col. George C. Bell, and most especially William Ehler for his daily support and assistance. In addition, I would like to thank the technical staff of the facility for their constant support and help.

I want to thank William P. Fife, Director of the Texas A&M University Hyperbaric Laboratory, Dr. Ralph Storts and Ms. Debbie Young for their continual enthusiasm and dedication to this project.

Finally, I would like to thank my wife, Julia, for her willingness and patience in surviving an extremely hectic summer, in order that I might make the most of this great opportunity. I would also like to thank my two lovely daughters, Anna and Deborah, for constant love and devotion to their Daddy.

I. INTRODUCTION

I received my Ph.D from Texas A&M University studying the potential of hyperbaric oxygen (HBO) therapy to alter the course of Experimental Allergic Encephalomyelitis (EAE), which is the best characterized animal model system for multiple sclerosis (MS). Our studies indicated that clinical symptom onset could be delayed if HBO therapy began simultaneously with injection of the sensitizing protein and was continued on a daily schedule. These studies expanded to include investigating blood-brain barrier (BBB) alterations which occur in conjunction with the disease process. Previous studies showed that a breakdown of the BBB occurs prior to clinical symptom onset in the disease. Our on-going research has focused on the nature of the breakdown and with a major concentration on the mechanism of the breakdown.

In addition, I have been very active in several other research projects which deal with the subject of drug distribution under hyperbaric conditions. These include investigations which have evaluated HBO and various drugs on spinal cord injuries and the examination of several types of cardioactive drugs under hyperbaric conditions.

The Clinical Investigative Facility at Wilford Hall Medical Center currently houses the only U.S. Air Force hyperbaric unit dedicated to animal research. When combined with their excellent surgical and animal care facilities, this laboratory provided a research environment which is most certainly one-of-a-kind in the USAF if not in the entire scientific community. Because of this unique combination I was assigned to work on the effect of HBO on the BBB at

the Clinical Investigative Facility at Wilford Hall Medical Center (Lackland AFB, San Antonio, Texas).

II. OBJECTIVES OF THE RESEARCH EFFORT:

The overall objective of this project was to examine and characterize, on a preliminary basis, the effect of increased partial pressures of oxygen on the permeability of the blood-brain barrier.

My individual objectives were:

1. Establish the surgical procedures required to chronically implant catheters in the rat, in order that multiple samplings and injections could be performed.

2. Establish the histological techniques necessary for a quick and accurate evaluation of BBB permeability changes utilizing circulating dyes as indicators.

3. Establish with the above techniques a dose/response relationship between oxygen partial pressure and BBB alterations in the rat.

The following additional objective was established during the course of the project.

1. To examine, on a preliminary basis, the differences in sensitivity to HBO between the rat, the rabbit and the guinea pig. Blood-brain barrier permeability changes were used as the indicator of sensitivity.

III. SURGICAL PROCEDURES

At least 48 hours prior to surgery, silastic catheters were constructed as follows. A 2.5 inch sleeve of 0.03" x 0.065" silastic tubing was soaked in xylenes for approximately one hour. This served to expand the tubing to facilitate the insertion of the smaller bore tubing. A 5.5 in. section of silastic tubing (0.02" x 0.037") was

threaded into the larger bore tubing so that one end was flush. This was accomplished by threading a piece of 1-0 braided silk through the large bore tubing, tying it to the end of the small tubing and working the smaller bore tubing through the shaft of the larger tube. These catheters were allowed to dry for several hours. Following drying, a ring made with silastic medical adhesive was positioned approximately one inch from the free end of the small bore tubing. This ring was fashioned so that the tubing could be easily inserted into a small opening in the vessel. The ring was molded in a cone shape with the pointed end positioned to enter the vessel first. The entire diameter of the cone was approximately the same as the outside diameter of the large bore tubing (0.067"). The silastic adhesive was allowed to cure for at least 24 hours before being used. The end of the catheter inserted in the vessel was trimmed to a short bevel to facilitate vessel entry.

All instruments, catheters and other supplies were autoclaved prior to surgery. The surgical procedure described can only be considered "clean" surgery because of the necessity of dorsally exteriorzing the catheter between the shoulder blades of the animal. However, extreme care was taken to maintain a sterile environment in the area of the incision and at the point of entry of the catheter into the vascular system.

Because of its proximity to the surface and its relatively large diameter, the jugular vein was chosen as the vessel for catheter implantation.

Animals were anesthetized with Nembutal or acepromazine-ketamine and the area under the jaw and neck, and the scapula area were

shaved free of hair. The clipped areas were thoroughly cleaned with betadine, and an incision (right of midline) was made from the area of the thoracic inlet towards the right jaw for approximately one inch. This incision only penetrated the skin. Just prior to this procedure a small incision was made over the scapula of the animal. A subcutaneous tunnel was then formed with curved hemostats, and the catheter passed through to the area of the vein. The jugular vein was gently isolated just anterior to its entry into the thoracic cavity and two lengths of 4-0 silk were positioned under the vein.

The catheter was prepared for vessel entry by inserting a 22g tubing adapter into the scapula end, and filling the catheter with sterile saline. The cranial end of the vessel was ligated with one of the lengths of 4-0 silk. The ends of the tie were used at a later time to help secure the catheter. A loose knot was made around the caudal end of the vessel. With slight tension on the vein a small perpendicular cut was made and then enlarged with a small longitudinal cut extending caudally from the first cut. The catheter was inserted and the caudal suture tied just above the silastic adhesive ring. Care was taken to insure that the vessel was not occluded by the caudal tie. The catheter was then further secured by the ends of the rostral tie. This also secured the two sections of the vessel. The wound was closed with a non-absorbable suture. A heparin lock was placed in the tubing adaptor and the entire unit secured with several sutures along the midline of the back extending caudally. At least one suture was taken at the point of catheter entry to insure that the opening remained as small as possible.

Post-operative care included a daily application of bacteriocidal salve at the point of catheter entry and a daily flushing

using sterile saline. It was not necessary to use heparinized saline because clotting is not a problem with silastic tubing.

Using this technique catheters remained patent for 8-14 days.

IV. HISTOLOGY TECHNIQUES

Because of the desire to use a large number of animals to do a survey of a variety of different oxygen partial pressures as well as an exhaustive series of controls it was necessary to establish techniques which would allow the BBB to be quickly and effectively evaluated for permeability alterations. This portion of the project and the subsequent evaluation of experimental animals was completed under the supervision of Dr. Ralph Storts, D.V.M. (Texas A&M University, College of Veterinary Medicine, Dept. of Pathology).

In an effort to obtain a comprehensive picture of central nervous system (CNS) changes, ten levels of tissue were examined in each animal. Following perfusion of each animal with buffered formalin, the brains were extracted and cut into eight cross-sectional segments. The spinal cord was examined at the thoracic and cervical levels.

After sectioning, the tissue was examined under a dissecting scope for vital dye staining. Tissue stained with sodium flourscein was examined in a specially constructed ultraviolet (UV) light box which allowed sections to be viewed with a dissecting scope while being exposed to UV light. All tissue was evaluated under blind conditions, one to three hours after staining.

V. EXPERIMENTAL EXPOSURES

The use of vital dyes as BBB markers has long been established in the literature (1, 2). In these investigations the

vital dyes trypan blue and Evans blue were compared to determine which one provides the best staining under the proposed experimental conditions. In addition, the micromolecular dye, sodium flourscein, was examined as a potential marker for permeability changes.

Preliminary studies comparing the vital dyes indicated that trypan blue provided the best staining. However, additional studies indicated the superior nature of sodium flourscein and it was therefore used in the majority of the animals as an indicator of BBB integrity.

Experimental animals were exposed to hyperoxic environments at 2.0, 2.5 and 3.0 atmospheres absolute (ATA). Dye injections were done by one of three methods:

1. Animals were exposed to HBO for the specified time, anesthetized immediately after exposure, injected with the appropriate dye via the femoral vein and sacrificed by cardiac perfusion with 4% buffered formalin.

2. Animals were anesthetized, given the appropriate dye, exposed to oxygen and then immediately upon removal from the chamber reanesthetized (if necessary) and sacrificed by intracardiac perfusion.

3. Several of the animals were anesthetized, injected with dye, given the appropriate HBO exposure, injected with dye again and then sacrificed by intracardiac perfusion.

In all cases the dye was allowed to circulate for a minimum of 15 minutes.

Hyperbaric oxygen exposures consisted of placing the animals in a steel hyperbaric chamber, compressing the chamber to approximately 0.15 ATA, flushing the chamber to 100% oxygen and then compressing the chamber to the appropriate depth with oxygen. Compression rates were

approximately 0.33 ATA per minute. Carbon dioxide was controlled with an internal scrubber system. The percentage of carbon dioxide was maintained below 0.03%.

The first series of animals were exposed to 100% oxygen at 2.0 ATA (1520 mmHg) for 30 and 60 minutes. All of the animals evaluated within this series were injected with either trypan blue or Evans blue (1.5 cc, 2% solution) immediately following HBO exposure. Upon examination none of the tissue sections showed any penetration of the dye.

Following the negative results of the animals at 2.0 ATA the decision was made to begin using sodium flourscein (NaFl) because it is a micromolecular dye whereas trypan blue and Evans blue are macromolecular dyes.

The next group of animals was exposed to 100% oxygen at 2.5 ATA (1900 mmHg) for a total of 90 minutes. All of the animals evaluated at this depth were injected with sodium flourscein (10% solution) at a dose of 1.0 milligram per kilogram body weight. Animals were injected with dye immediately following oxygen exposure. Upon microscopic examination no significant increase in BBB permeability was noted.

The last group of experimental animals was exposed to 100% oxygen at 3.0 ATA (2280 mmHg) for a total of 90, 180 and 210 minutes. Groups exposed for a total of 90 and 180 minutes received NaFl injections immediately following HBO exposure. Tissue examination revealed that the BBB was not significantly compromised.

In the event that the BBB was possibly breaking down and repairing itself before the dye could be injected following the exposure, the animals exposed for a total of 210 minutes were injected

with dye (NaFl) both before and after the dive. Tissue examination revealed no significant increase in BBB permeability.

A final group of rats was exposed to hyperbaric air at 3.0 ATA (480 mmHg oxygen, 1800 mmHg nitrogen) for a total of 90 minutes. The animals were immediately decompressed to the surface following this exposure, injected with NaFl which was allowed to circulate for a total of 15 minutes and then sacrificed by intracardiac perfusion. This group showed a minimum amount of BBB leakage but the increase in permeability was not uniform among all of the exposed animals. These animals were compared to a group of 1.0 ATA air controls. No permeability changes were noted in the control group.

VI. SPECIES DIFFERENCES

Past observations by two independent research groups (Ehler, unpublished observations; Chambi, unpublished observations), indicate that the BBB in the rabbit is much more susceptible to oxygen induced permeability changes than those which we observed in the rat. Rats have been used for numerous BBB studies, but not in studies focused on the role of oxygen. The majority of studies utilizing rats as the experimental model system have addressed permeability changes due to mechanical factors such as bubbles or emboli. For this reason we felt it necessary to examine the possibility of significant species differences in BBB susceptibility to increased partial pressures of oxygen. Because of the limited time and funds available for a study of this nature these findings are very preliminary in nature.

In the series of studies utilizing rabbits, a total of three rabbits were used. Two of the rabbits were exposed to 100% oxygen at a pressure of 3.0 ATA for a total of 90 minutes. These animals were

immediately decompressed and injected with dye. Because past studies used vital dyes one animal was injected with trypan blue and the other was injected with NaFl. Following a 15 minute circulation time the animals were anesthetized and sacrificed by intracardiac perfusion. The third animal which remained at 1.0 ATA and served as the control for the group was injected with NaFl and perfused. The same ten levels of the CNS as those viewed in the rat were examined for BBB permeability changes.

The control animal showed no increase in permeability. The animal which was injected with trypan blue showed a very minimal penetration of the dye. However, the animal which was injected with NaFl showed a dramatic increase in permeability. There was a clear indication of increased diffusion in the cerebrum and included the cortex, neostriatum, thalmus and periaqueductal area. In addition, the cerebellar cortex had comparable staining in this area also.

Guinea pigs were exposed to the identical oxygen exposure as described for rabbits and the results indicated a definite increase in BBB permeability, with areas of CNS similar to those of the rabbits being compromised.

Even though these studies must be considered preliminary because of the limited number of animals and the single oxygen exposure used, the results strongly suggest a significant difference in the susceptability of the blood-brain barriers between these species.

VII. RECOMMENDATIONS

1. Survey of species differences

a. The preliminary results reported in this study indicate a significant difference in species susceptibility to increased partial pressures of oxygen. The next step is to expand on

these observations and define the magnitude of the differences in susceptibility. It is extremely important to determine the exposure pressure and time at which BBB permeability is altered in each species and to determine the limits and scope of the "breaking point". The definition of these parameters would enable us to possibly understand the discrepancies which are present in the literature. Only after these parameters are established can such basic questions begin to be asked as 'What factors or components make the BBB an effective barrier?'. These investigations should address such issues as the effects of long-term low dose exposure or multiple, short-term high dose exposures. In an effort to maintain the integrity of the BBB does the brain "adapt" to conditions of hyperoxia? Is the "adaption" short or long?

2. The exact nature of the increase in permeability must be examinied in the susceptible species. A variety of physiologic conditions, structural components and biochemical constiuents combine to form the BBB. Which of these areas are vulnerable to hyperoxic conditions and which are not? Can they truly be seperated? Which areas are susceptible to mechanical damage? The characteristics of the increase in permeability must be documented (size limits, charge requirements, etc.).

3. After the above question is answered the next question which should be addressed is, "What components or structural differences enable the rat BBB to remain intact when exposed to levels of oxygen that trigger a significant increase in permeability in other laboratory species". Previous investigations have reported after exposure to 85% oxygen for seven days rats have a 50% increase in

pulmonary superoxide dismutase (SOD) and became "tolerant" to 100% oxygen (4, 5). "Tolerant" rats survived more than four days of exposure to 100% oxygen while control rats died within 72 hours. Guinea pigs, hamsters and mice did not develop oxygen tolerance under similar conditions and did not have as large a change in pulmonary SOD as did the rat. This report supports the theory that SOD provides a defense against increased levels of superoxide and the ability to increase production of SOD is a significant factor in this defense. Whether the ability to significantly increase SOD levels in CNS tissue or within the cells comprising the BBB is present as an effective defense mechanism in rats and absent in other species is a question that only research will answer. Very basic questions as to the mechanisms of resistance and susceptibility must be asked. Only by answering these basic questions can we begin to effectively address the issue as to how the barrier can be manipulated to achieve a specific objective.

4. The question of whether or not the BBB can be modified by drugs is becoming more important everyday. Can the barrier be modified by over-the-counter drugs or does a family of drugs exist that triggers a modification of the barrier? These questions must be answered before we can responsibily extend our expansion into space and before we can begin to take significant steps to increase the operational effectiveness of the weakest link in flight and space operations-man.

5. The question must be asked as to whether or not the barrier is modified by increased partial pressures of other gases. The gas which must first be considered is carbon dioxide.

6. The question must be asked as to what the consequences

and differences are between mechanical alteration (bubbles, emboli, etc.) and biochemical alteration (SOD levels, etc.). What are the conseqences when the two are combined as they are on a regular basis in the treatment of decompression sickness and arterial gas embolism.

7. One of the most critical questions which must be answered is which animal model most closely resembles humans. What are the critical factor of the resemblance? Should the choice of models change under certain conditions? What are the conditions or criteria for change?

REFERENCES

1. Chryssanthou, C., M. Springer, and S. Lipschitz, "Blood-brain and blood-lung barrier alteration by dysbaric exposure," Undersea Biomed. Res. 4 (1977) 117-129.
2. Goldman, E.E., "Vitalfarbung am Zentrulnervensystem," Berlin 1913 Eimer.
3. Crapo, J.D. and D.F. Tierney, "Superoxide dismutase and pulmonary oxygen toxicity," Amer. J. Physiol. 226 (1974) 1401-1407.
4. Crapo, J.D., "The role of superoxide dismutases in development of pulmonary oxygen tolerance," Enviro. Health Perspect. 16 (1976) 182-183.

1986 USAF-UES SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

Multimodal Information Exchange for Individual and Group Problem Solving

Prepared by: A. Rodney Wellens

Academic Rank: Associate Professor

Department and University: Department of Psychology, University of Miami

Research Location: Armstrong Aerospace Medical Research Laboratory, Human Engineering Division, Technology Development Branch

USAF Researcher: Michael D. McNeese

Date: 19 August 1986

Contract No: F49620-85-0013

Multimodal Information Exchange for Individual and Group Problem Solving

by

A. Rodney Wellens

ABSTRACT

USAF C^3 systems represent the organizational equivalent of neural networks for Air Force operations. Recent research efforts within the Human Engineering Division of AAMRL have attempted to improve USAF C^3 systems by developing design guidelines based upon the performance of humans and machines as information processors and decision makers. The task undertaken by the present investigator was to explore the ways in which studies of advanced multichanneled telecommunication systems might be incorporated within the context of an established C^3 Operator Performance Engineering (COPE) project.

A three tiered analysis of the problem was conducted involving appropriate literature reviews, observations of ongoing experiments and visits to available field sites. An on-site computerized search of the artificial intelligence and social psychology literatures was conducted to analyize person-machine interactions within advanced information processing systems. An observational study was also conducted in conjunction with an ongoing team problem solving experiment to determine the role of verbal and nonverbal communication factors in clustered and segregated team configurations. Finally, visits were made to local teleconferening facilities to discover the benefits and drawbacks of a variety of computer, audio and video system hardware configurations.

Results have been reported in three separate reports written by the investigator that are on file in AAMRL/HEC. Recommendations are given regarding future COPE experiments based on both theoretical and applied concerns.

Acknowledgments

I would like to thank the Air Force Systems Command and the Air Force Office of Scientific Research for sponsoring this research effort. I would also like to express my thanks to the members of the Technology Development Branch, Human Engineering Division, Armstrong Aerospace Medical Research Laboratory for their support and guidance. In particular, I am grateful to Mr. Michael D. McNeese for sharing his insights into artifical and human intelligence and to Mr. Donald Monk for allowing me access to COPE facilities and providing me with valuable research contacts. I would like to thank Mr. Walter Sommers for issueing the original invitation for me to visit Wright Patterson Air Force Base and giving me the freedom to explore a wide variety of research areas. Special thanks are given to Mr. Tracy Vogler, a fellow summer visitor at AAMRL, for his careful observation of the videotapes described in this report and his all around assistance in the preparation of auxillary reports.

Finally, I would like to thank my wife, Lynda, and my children Marci and Michael, for their willingness to relocate. Their understanding, patience and support throughout the summer was much appreciated.

I. Introduction

The investigator received his graduate training as an experimental social psychologist at Vanderbilt University where he helped to quantify and compare several models of cognitive balance. Upon receiving his Ph.D., he moved to the University of Miami where he established an interactive television laboratory for the study of social interaction. The laboratory provided a method of carefully studying the verbal and nonverbal behavior patterns of individuals engaged in face-to-face and television mediated interactions. In addition to testing several models of homeostatic regulatory mechanisims governing social interaction, the investigator soon discovered that his research results had relevance to media studies dealing with the design of video teleconferencing systems. He spent the summer of 1983 at Stanford University where he participated in a NASA study investigating the role of humans and intelligent machines within the context of a proposed space station. During the summer of 1985 the investigator conducted an extensive computerized search of the multimedia telecommunications literature and formulated a generic psychological distancing model of media effects. He also expanded his laboratory to include computer and audio teleconferencing facilities. When the Air Force Summer Faculty Research Program was announced, the investigator telephoned Dr. Billy Welch to seek his advice regarding a relevant research program. Dr. Welch suggested the Human Engineering Division of the Armstrong Aerospace Medical Research Laboratory. The investigator soon discovered through Mr. Walter Summers that the Technology Development Branch had several ongoing projects related to man-machine interactions and remote decision making. A consideration of telecommunication factors appeared directly relevant to ongoing concerns for improving Air Force command, control and communication systems. After an exchange of reprints, it was decided that the investigator's primary work effort would be within the C^3 Operator Performance Engineering Project (COPE).

II. Objectives of the Research Effort

After a brief pre-summer visit a preliminary statement of objectives was formulated that read:

> "Review relevant theoretical and empirical literatures dealing with telecommunication systems, man-machine interfaces and team problem solving tasks. Develop a research paradigm for studying information media effects upon group problem solving processes and performance."

Upon returning to Miami, these objectives were translated into a series of concrete tasks to be accomplished. A tentative timetable was developed. The project sheet shown in Figure 1 was sent to the Technology Development Branch for approval. It was accepted with only minor revisions (i.e., computer facility familiarization was eliminated because of adequate vendor support for programming and the investigator bringing his own personal computer for word processing). It was emphasized that the investigator try to accomplish first that which could only be accomplished on site (e.g., formulating ideas with colleagues and, where possible, collecting relevant empirical data).

III. Literature Reviews and Conceptual Papers

The investigator completed a literature review of the teleconferencing literature prior to his arrival at AAMRL (Wellens, 1986). The review included the formulation of a "psychological distancing" model of electronic media that orders interactive telecommunication devices along a "close-far" continuum according to their information carrying capacity or bandwidth. Figure 2 shows a

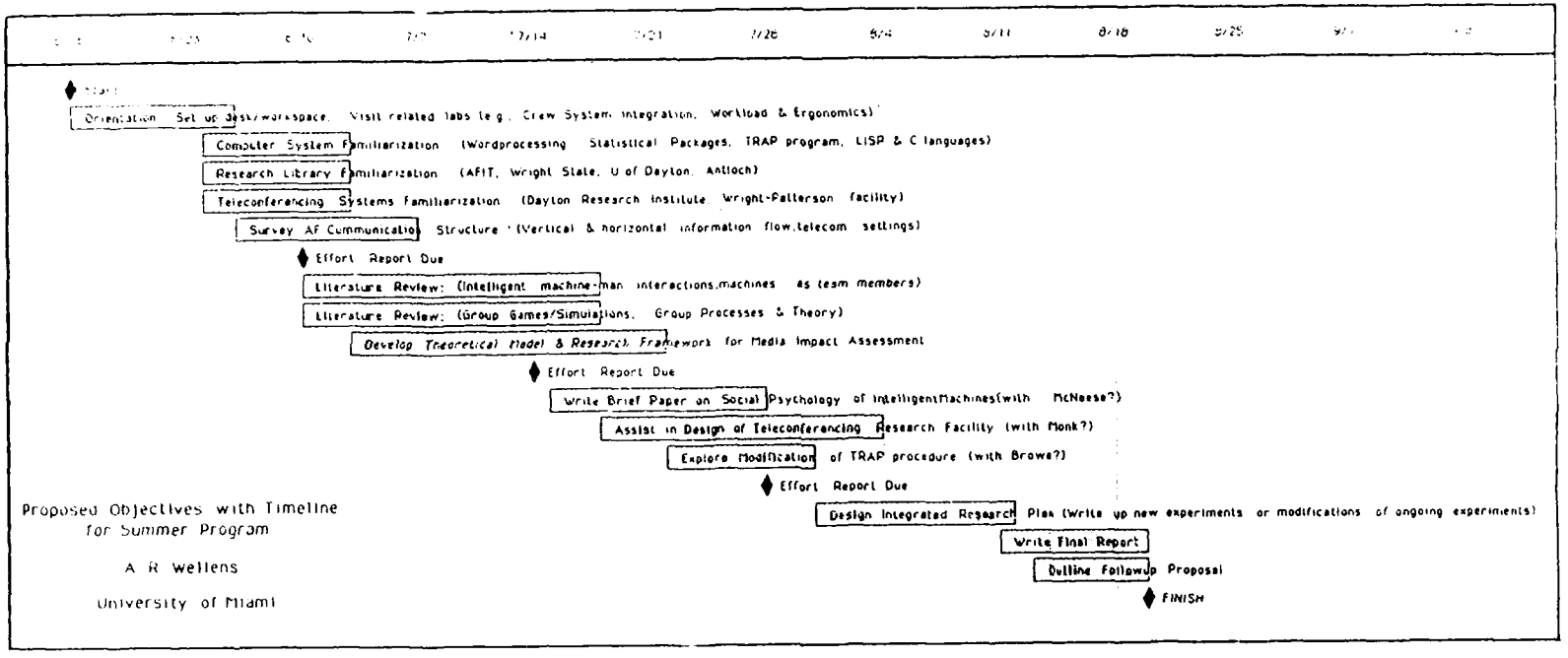


Figure 1. Proposed objectives with timeline submitted for summer study.

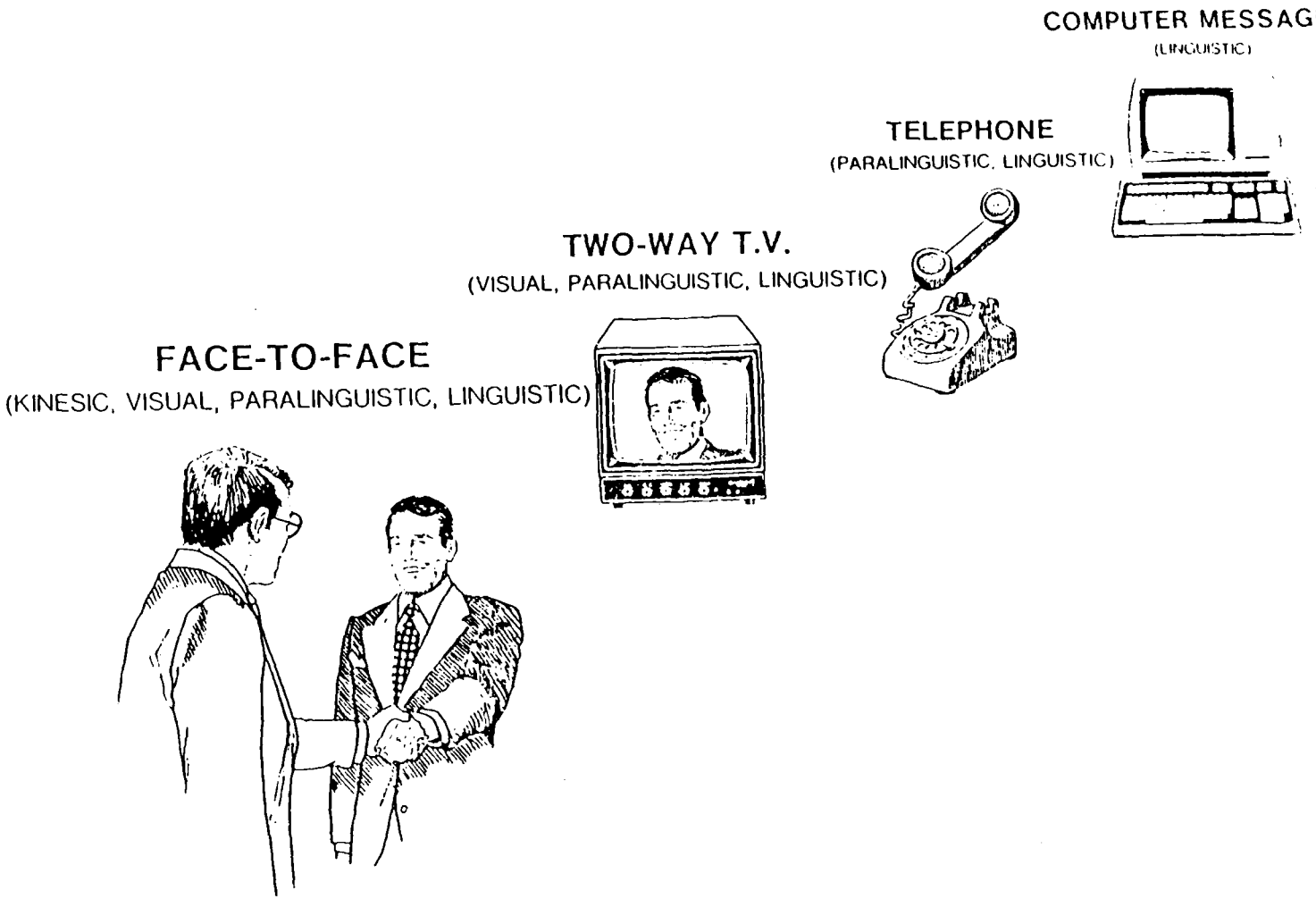


Figure 2. Psychological distancing model used by Wellens (1986) for comparing electronic media.

graphic representation of the model. Table 1 summarizes research findings resulting from a review of over 50 multimedia studies.

In order to familiarize himself with issues surrounding man-machine interfaces and the role of intelligent machines in future decision making networks, a computerized search of the behavioral, social science and computer science databases, accessable through the AFWAL library, was conducted. The diagram presented in Figure 3 shows the topic areas and key words used to guide the search.

The area of focus was defined as the "Social Psychology of Intelligent Machines" and was developed into a paper recently submitted for publication by the investigator and his summer USAF research contact, Mr. Michael D. McNeese. The paper, written for a lay audience, takes the reader on a brief tour of current thought regarding the nature of artificial intelligence and outlines the need for developing knowledge pathways for connecting synthetic and organic intelligence. The paper also outlines the dangers inherent in depending too heavily upon smart machines as social partners.

IV. Observational Study and Pilot Project Report

An unforeseen opportunity arose to collect a limited amount of behavioral observation data in support of the overall goal of developing a research paradigm to study media effects upon group decision making. Shortly after the investigator's arrival, he was introduced to the Team Resource Allocation Problem (TRAP) developed by Dr. Clifford Brown (Brown & Leupp, 1985). An experiment was underway to investigate the impact of group versus isolated viewing of information displays, alphanumeric versus graphic display formats and high versus low time stress upon team performance (McNeese & Brown, under review). The

Table 1
Representative Findings Reported in Sampled Comparative Media Studies (from Wellens, 1986)

Interpersonal Relations

1. Perceptions of others are generally more favorable in face-to-face and "close" (video) media than "far" (audeo and computer) media.
2. "Close" media (face-to-face and video) are seen as more pleasant and friendly than "far" (audio) media.
3. Video provides a greater feeling of social contact than audio.
4. Face-to-face meetings are preferred over video, audio or computer conferencing for situations involving interpersonal relations or conflict resolution.
5. Video is perceived as more effective than audio for forming impressions of others.
6. Users are more confident in their perceptions of others via video than via audio, but not necessaily more accurate.

Social Influence and Group Processes

7. Persuasion (acceptance and retention of ideas) decreases with inappropriately "close" media.
8. Obedience (submission to power and authority) is higher face-to-face than mediated situations.
9. Individuals are perceived as more persuasive and trustworth in audio than face-to-face or video.
10. Coalition formation and "we" feelings are more likely to develop within face-to-face and video groups than between groups linked by audio.
11. Cooperative responses decrease as the communication media linking interactants decreases in communication richness.
12. Role differentiation is less pronounced in mediated settings than in face-to-face settings. "Lean" media produce more even contribution rates from group members than "rich" media.

Information Exchange and Media Effectiveness

13. Simple exchange of information is as effective using audio and video as face-to-face.
14. Informationally "rich" media are preferred when the complexity of information to be exchanged increases.
15. Verbal communication rates are higher and solutions reached faster in media employing a voice channel rather than written text.
16. Simultaneous verbalizations and interrruptions are higher in face-to-face than audio-only interactions.
17. Video interactions produce more interruptions and speech disfluencies than audio, but fewer than face-to-face.

The Social Psychology of Intelligent Machines

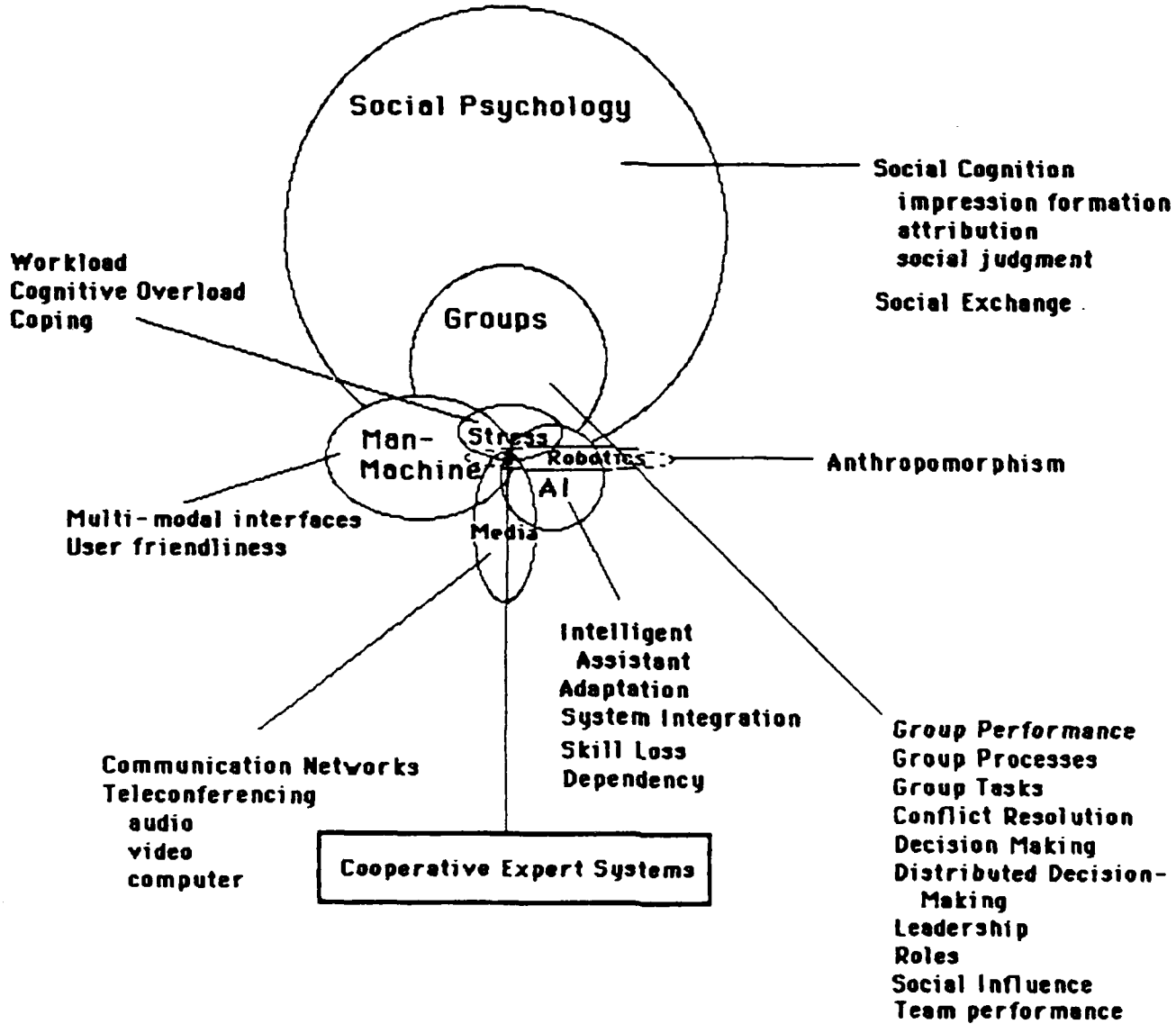


Figure 3. Research domains and key words used to guide computerized search of the literature relevant to "The Social Psychology of Intelligent Machines.

present investigator suggested that all remaining teams be videotaped while performing their tasks. This recommendation was made in order to test the utility of employing direct behavioral observational techniques for assessing group decision making processes in future studies. The recommendation was immediately implemented with the help of MSgt. Danny Bridges and Mr. Bill McGovern.

Eight three-person teams were videotaped while they were engaged in the Team Resource Allocation Problem (TRAP). An experimental behavioral rating system that drew upon Bales (1950) interaction process analysis technique was developed. A summer high school apprentice assigned to AAMRL, Mr. Tracy Vogler, used the rating method to record a subset of team members verbal and nonverbal behaviors across all experimental conditions.

The results of this pilot study have been reported in a paper entitled, "Groups in TRAP: An Exploratory Analysis of Behavioral Dynamics," written by the present investigator in collaboration with Mr. Vogler. Preliminary findings suggested that of the nine verbal and nonverbal behaviors sampled, verbal commands and gesturing were the best indicators of group functioning and responded directly to two of the three independent variables manipulated. Verbal commands occurred at a faster rate under conditions of high as opposed to low time stress. Directed commands from emergent leaders appeared to covary positively with team performance under conditions of high time stress and negatively under conditions of low time stress. Gesturing increased dramatically as a function of group versus isolated viewing of displays, but was negatively related to team performance. Other behaviors that were less clearly related to experimental manipulations were nonetheless helpful in defining team operating climates.

The observational study proved to be a useful exercise for sharpening the investigator's thinking about behavioral manifestations of group problem solving activities. It was concluded that direct observation of group decision making activities formed an important adjunct to the study of group performance outcomes. It was recommended that behavioral observation be continued and expanded in future team research. The behavioral rating system used in the current study should be refined with greater attention paid to subcategories of verbal behavior.

V. Teleconferencing Site Visits and COPE Opportunity Paper

In order to further develop a sense of context within which to make recommendations regarding future telecommunication research, arrangements were made to visit three different teleconferencing facilities located at Wright Patterson Air Force Base (WPAFB). Each site had its own unique combination of equipment that provided its users with different kinds of multichanneled information.

The first site was located within the F-16 SPO and consisted of an off-the-shelf slow scan television and graphics system used to augment group-to-group audioteleconferencing between the Aeronautical Systems Division at WPAFB and the Fort Worth Division of General Dynamics. The Interand Discon 1000 system has had frequent use averaging two conferences per week over the first seven months of its existence at WPAFB. The system has been used to support meetings ranging from normal briefings to engineering trouble shooting and budget discussions. The system's $125K price tag has been justified in realized savings of TDY costs of about the same amount over the last seven months. The system is viewed by its users as a success.

The second site was located within the Foreign Technology Division (FTD) at WPAFB and consisted of a totally unique system that had been developed under a DARPA contract for demonstrating limited motion teleconferencing at very low bandwidths (19.2 Kbps). It was supplemented by a two-way audio link and a Shared Graphic Work Space that was interfaced with a PDP 11/23. These system components have been descibed elsewhere by Kelly (1982) and White (1983). The system is used to link FTD at WPAFB with another government site in Maryland. Several system bugs, including poor audio quality and slow transmission of hard copy, combined with frequent staff turnover at its remote site, has led to only limited success of this system. The system is currently undergoing a hardware and software upgrade with two additional sites being added to the network shortly.

The third facility visited was housed within the Human Resources Laboratory at WPAFB. The Advanced Technology Multimedia Communications Device (ATMC) was developed by the University of Dayton Research Institute and has been described elsewhere (Hoffman & McKinney, 1983). It is a modularized system that provides freeze frame video, color graphics with cursor pointing and audio teleconferencing between laboratory sites at WPAFB and Brooks AFB. System use is currently limited to audio teleconferencing. The sophisticated graphics system is used primarily in support of local production of overhead transparencies.

It was instructive to note that the degree of success enjoyed by the systems examined appeared to be inversely proportional to system cost and complexity. It was recommended by the investigator that a laboratory facility capable of emulating the rich variety of systems encountered would be a useful tool for uncovering system attributes that contributed to task success and user satisfaction.

A report written by the present investigator entitled, "Potentials for Telecommunication Research within the C^3 Operator Performance Engineering Project," provides more details concerning the systems outlined above together with a description of the scope of current research efforts within COPE. Current efforts range from experiments assessing video display technologies, integrated workstation designs and intelligent machine assistants. A series of recommendations were made regarding the kinds of research efforts that could potentially fall within the boundaries of COPE. These included tasks that focused upon C^3 system information acquisition and transmission, situation assessment, and initiation and monitoring of action plans. An integration of field and laboratory based studies was recommended that balances a need for engineering guidelines against an equally valid need to develop theoretical models of system operations and performance.

VI. Recommendations

The three papers authored by the investigator and referenced in the body of this report contain detailed recommendations regarding future research dealing with team decision making and multimodal information systems. The recommendations listed below serve to highlight a few representative categories:

1. Continue and expand the use of behavioral observation techniques for the study of group decision making. Observation of team communication patterns can provide insight into problem solving and decision making processes that are inaccessible within single person problem solving tasks or group tasks that focus only on performance outcomes.

2. Upgrade current videotaping facilities to allow for multicamera, unobstructed viewing of individual team members in group and isolated conditions. Computerize tape handling and behavioral scoring procedures for more efficient data collection and processing.

3. Refine behavioral rating categories to include more detail regarding verbal activities. Talking was the dominant behavior observed during team problem solving and reflected both information exchange and attempts to coordinate activities.

4. Expand the range of COPE activities to include a wider variety of laboratory tasks. Develop gaming situations that focus upon information integration and situation assessment.

5. Survey Air Force C^3 system organization and facilities to develop a comprehensive model of system architecture. Include a consideration of upper level decision making as well as lower level support operations.

6. Encourage the development of alternative theoretical orientations for testing hypotheses regarding human cognitive and social functioning within team situations.

7. Develop a flexible media laboratory to test the impact of new telecommunication technologies upon C^3 system functioning.

8. Integrate research efforts dealing with artificial intelligence, communication and group problem solving.

References

Bales, R. F. (1950). Interaction process analysis: A method for the study of small groups. Cambridge, MA: Addison-Wesley.

Brown, C. E., & Leupp, D. G. (1985). Team performance with large and small screen displays. AAMRL Tech report no. TR-85-033.

Hoffman, C. J., & McKinney, W. A. (1983). ATMC, the multi-media teleconferencing system. In L. Parker & C. Olgren (eds.), The Teleconferencing Resource Book: A Guide to Applications and Planning, pp. 215-222. New York: Elsevier Science Publishers.

Kelly, C. W. (1982). Video teleconferencing for cirsis applications. Signal, (July, 1982), pp. 27-36.

McNeese, M.D, & Brown, C. E. (under review). Large group displays and team performance: An evaluation and projection of guidelines, research and technologies. (AAMRL Tech report prepared 15 April 1986).

Wellens, A. R. (1986). Use of a psychological distancing model to assess differences in telecommunication media. In L. Parker & C. Olgren (eds.) Teleconferencing and Electronic Media, Vol. V., pp. 347-361. Madison Wisconsin: Center for Interactive Programs, University of Wisconsin.

White, P. (1983). Evaluation of a five node videoconference system and shared graphics workspace, U. S. Department of Denfense, Washington, D. C. Report prepared for DARPA, No. ARPA4/F1605/83225-WH.) London: EIU Informatics.

Auxillary Reports

(On file with Mr. Walter Summers and Mr. Michael McNeese, AAMRL/HEC)

Wellens, A. R. Potentials for telecommunication research within the C^3 Operator Performance Engineering Project. Prepared 18 August 1986. 12 pages text, 26 pages appendices, 5 references.

Wellens, A. R., & McNeese, M. D. The social psychology of intelligent machines. Submitted for publication in R. Kurzweil's forthcoming book, "The Age of Intelligent Machines." Approved for release 4 August 1986, ASD 86 1613; 7 pages text, 22 references.

Wellens, A. R., & Vogler, T. J. Groups in TRAP: An exploratory analysis of behavioral dynamics. Prepared 15 August 1986. 15 pages text, 46 pages appendices, 5 references.

1986 USAF-UES Summer Faculty Research Program/
Graduate Student Summer Support Program

Sponsored by the
Air Force Office of Scientific Research
Conducted by the
Universal Energy Systems, Inc.
Final Report

Preliminary Study of an Optical Implementation of the Conjugate Gradient Algorithm

Prepared by:	Stephen T. Welstead
Academic Rank:	Assistant Professor
Department and University:	Mathematics and Statistics University of Alabama in Huntsville
Research Location:	Rome Air Development Center, OCTS, Griffiss Air Force Base, New York
USAF Research:	Dr. Vincent Vannicola
Date:	August 15, 1986
Contract No.:	F49620-85-C-0013

Preliminary Study of an Optical Implementation of the Conjugate Gradient Algorithm

by

Stephen T. Welstead

ABSTRACT

An analysis was done on an acousto-optic signal processor in order to make recommendations about possible improvements in the algorithm being used. The optical nature of the processor necessitates looking at data in an analog, rather than digital, fashion. The minimization problem, which the processor is solving, is formulated in an analog way. This leads to an operator equation in Hilbert space, rather than the usual matrix equation. Operator theoretic versions of steepest descent and conjugate gradient algorithms are discussed. Block diagrams are given for these algorithms, along with recommendations for possible optical implementations.

Acknowledgments

I would like to thank the Air Force Systems Command and the Air Force Office of Scientific Research for sponsorship of my research. Also, I am indebted to the people of the Rome Air Development Center, Griffiss AFB, for providing the opportunity for this research. In particular, I would like to thank Dr. Vincent Vannicola of RADC for guiding my efforts.

I. Introduction

My Ph.D. dissertation topic at Purdue University concerned the numerical solution of integral equations. After several years of working in that area, I decided to branch out into other areas, and began looking at optical signal processing and image processing. Through the Center for Applied Optics at the University of Alabama in Huntsville I was able to do funded research in optics in 1985.

The research problem I selected at the Rome Air Development Center, Griffiss AFB, NY, concerns the investigation of different mathematical algorithms for implementation on an acousto-optic signal processor. Because of the analog nature of the optics involved, the problem to be solved turns out to be an integral equation. The combination of optics and integral equations makes this problem particularly well suited to my background.

II. Objectives of the Research Effort

The objective of the research effort is to improve the performance of an acousto-optic signal processor (already in experimental operation) by implementing a more efficient mathematical algorithm. The system in operation now uses a Howells-Applebaum least mean square (LMS) algorithm. It was felt that performance could be improved if a more powerful algorithm, such as the conjugate gradient algorithm, were implemented. My individual objectives were:

1. Familiarize myself with the acousto-optic processor now in operation in order to fully understand its implementation of the LMS algorithm.
2. Formulate the mathematical problem to be solved, keeping in mind the special nature of the optical processing involved.
3. Study the conjugate gradient and related algorithms and investigate the feasibility of implementing these algorithms in an optical system similar to the one now in operation.
4. Construct a block diagram for the conjugate gradient algorithm and make recommendations about possible optical

implementations.

III. Analog Formulation of the Minimization Problem for the Optical System

We consider the signal processing problem of cancelling noise from a main signal. The receiving configuration consists of a main antenna and N omni-directional side antennas. An acousto-optic processor for such a system has been proposed and implemented by Vannicola and Penn [VP1,VP2]. A similar system has been considered by Vander Lugt [V1].

We denote by $n_1(t), n_2(t), \ldots, n_N(t)$ the signals received at the side antennas at time t, and by $s(t) + n(t)$ the main signal plus noise received at the main antenna. Each side channel signal $n_i(t)$ is input through an acousto-optic device which produces a continuum of delayed signals $n_i(t-x)$, where the delay x ranges from $\emptyset$ to a value R which depends on the acousto-optic device (R is typically in the range 5-50 microseconds). We call x a 'spatial' variable here, since it represents position across the acousto-optic device, although it can also be thought of as another time variable.

Our problem is to form a weighted combination of the delayed secondary signals $n_i(t-x)$ in such a way that the result is a good estimate of the noise n(t) received at the main channel. The continuous nature of the delays necessitates that we look at this problem in an analog way, rather than the usual discrete formulation involving matrices and vectors. Thus, we define a Hilbert space H consisting of the set of complex vector valued functions $\underline{h}(x) = (h_1(x), \ldots, h_N(x))$ defined on the real interval $[\emptyset, R]$ with inner product

$$(\underline{h}, \underline{g}) = \sum_{i=1}^{N} \int_{\emptyset}^{R} h_i(x) \overline{g_i(x)}\, dx$$

where for a complex variable z, $\overline{z}$ denotes its complex conjugate.

Define the functions of two variables $f_i(x,t) = n_i(t-x)$, $i = 1,\ldots,N$, and let $\underline{f}(x,t)$ be the vector valued function whose i^{th} component is $f_i(x,t)$. Our problem, then, is to determine a vector valued weight function $\underline{w}(x) = (w_1(x),\ldots,w_N(x))$ so that the scalar function

$$y(t) = (\underline{f}(.,t),\underline{w}) \tag{3.1}$$

is a good approximation of the noise $n(t)$.

The output of the system is the "error" signal $e(t)$, which is the main signal plus noise minus the estimated noise:

$$e(t) = s(t) + n(t) - y(t).$$

The quantity we wish to minimize is

$$E(|e(t)|^2) = \int_0^\infty |e(t)|^2 \, dt. \tag{3.2}$$

E can be thought of as an expected value over time, although for the purposes of our minimization problem we have omitted any reference to a probability density function. (One can also think of (3.2) as an "energy" integral.) Setting $d(t) = s(t) + n(t)$, we find

$$\begin{aligned} |e(t)|^2 &= e(t)\overline{e(t)} \\ &= (d(t) - y(t))(\overline{d(t) - y(t)}) \\ &= |d(t)|^2 - y(t)\overline{d(t)} - d(t)\overline{y(t)} + |y(t)|^2. \end{aligned}$$

Then

$$\begin{aligned} E(y(t)d(t)) &= \int_0^\infty y(t)\overline{d(t)} \, dt \\ &= \int_0^\infty \overline{d(t)} \left\{ \sum_{j=1}^N \int_0^R f_j(x,t)\overline{w_j(x)} \, dx \right\} dt \\ &= \sum_{j=1}^N \int_0^R \left\{ \int_0^\infty f_j(x,t)\overline{d(t)} \, dt \right\} \overline{w_j(x)} \, dx \end{aligned}$$

$$= (\underline{b},\underline{w})$$

where

$$\underline{b}(x) = \begin{pmatrix} \int_0^{\infty} f_1(x,t)\overline{d(t)}\,dt \\ \vdots \\ \int_0^{\infty} f_N(x,t)\overline{d(t)}\,dt \end{pmatrix}.$$

Similarly,

$$E(\overline{y(t)}d(t)) = \overline{E(y(t)d(t))}$$
$$= \overline{(\underline{b},\underline{w})}$$
$$= (\underline{w},\underline{b}).$$

Also, using (3.1), we find

$$E(|y(t)|^2) = \int_0^{\infty} y(t)\overline{y(t)}\,dt$$
$$= \int_0^{\infty} \left\{ \sum_{j=1}^{N} \int_0^{R} f_j(x,t)\overline{w_j(x)}\,dx \right\} \overline{(\underline{f}(\cdot,t),\underline{w})}\,dt$$
$$= \sum_{j=1}^{N} \int_0^{R} \overline{w_j(x)} \left\{ \int_0^{\infty} f_j(x,t)\overline{(\underline{f}(\cdot,t),\underline{w})}\,dt \right\} dx$$
$$= (A\underline{w},\underline{w}),$$

where A is an operator mapping H to H, defined by

$$\underline{Aw}(x) = \begin{pmatrix} \int_0^{\infty} f_1(x,t)\,\overline{(\underline{f}(\cdot,t),\underline{w})}\,dt \\ \vdots \\ \int_0^{\infty} f_N(x,t)\,\overline{(\underline{f}(\cdot,t),\underline{w})}\,dt \end{pmatrix}.$$

The j^{th} component of $\underline{Aw}$ can be rewritten as

$$\int_0^{\infty} f_j(x,t) \sum_{k=1}^{N} \int_0^{R} \overline{f_k(s,t)} w_k(s)\,ds\,dt =$$

$$= \sum_{k=1}^{N} \int_0^{R} w_k(s) \int_0^{\infty} f_j(x,t)\overline{f_k(s,t)}\,dt\,ds$$

$$= (\underline{w}, \underline{A}_j(x,\cdot)),$$

where $\underline{A}_j(x,s)$ is a vector valued function of two spatial variables whose k^{th} component is given by

$$A_{jk}(x,s) = \int_0^{\infty} f_k(s,t)\overline{f_j(x,t)}\,dt$$

for $k = 1,\ldots,N$.

The operator A can be thought of as an analog "outer product". It corresponds to the cross-correlation (or covariance) matrix in the usual discrete formulation of the problem.

It is straightforward to show that for functions $\underline{w}$, $\underline{u}$ in H we have

$$(\underline{Aw},\underline{u}) = (\underline{w},\underline{Au})$$

so that A is a self adjoint operator on H. Also, one can show

$$(A\underline{u},\underline{u}) = \int_{\emptyset}^{\infty} |(\underline{u},\underline{f}(\cdot,t))|^2 \, dt. \qquad (3.3)$$

The expression on the right side of (3.3) is $> \emptyset$ for $\underline{u} \neq \emptyset$, thus A is a positive operator on H.

Our minimization problem can thus be reformulated as the problem of minimizing the functional F defined by

$$F(\underline{w}) = E(|d(t)|^2) - (\underline{b},\underline{w}) - (\underline{w},\underline{b}) + (A\underline{w},\underline{w}) \qquad (3.4)$$

The right side of (3.4) is just $E(|e(t)|^2)$ (see (3.2)). The quantity $E(|d(t)|^2)$ is independent of $\underline{w}$ and is of no consequence in the minimization problem. Since A is a positive operator, the problem of minimizing (3.4) is equivalent to solving the operator equation

$$A\underline{w} = \underline{b} \qquad (3.5)$$

([M, Theorem 2.1]). Equation (3.5) can also be written as a coupled system of integral equations:

$$\sum_{k=1}^{N} \int_{\emptyset}^{R} w_k(s)\overline{A_{jk}(x,s)} \, ds = b_j(x), \quad j = 1,\ldots,N \qquad (3.6)$$

IV. The Existing Architecture

Before examining any new algorithms for the solution of (3.5), let us first look at what the existing optical system (reported in [VP1,VP2]) is doing. Figure 1 is a simplified diagram of this system, showing one side channel only.

This architecture is implementing the LMS algorithm, which is an approximate version of the method of steepest descent. Assume, for the moment, that we are receiving signal samples at discrete time intervals $t = i\Delta t$ for some fixed Δt. Consider the following iterative scheme for determining the weight function w(x) (one side channel only):

$$w_{i+1}(x) = w_i(x) + a_i p_i(x). \qquad (4.1)$$

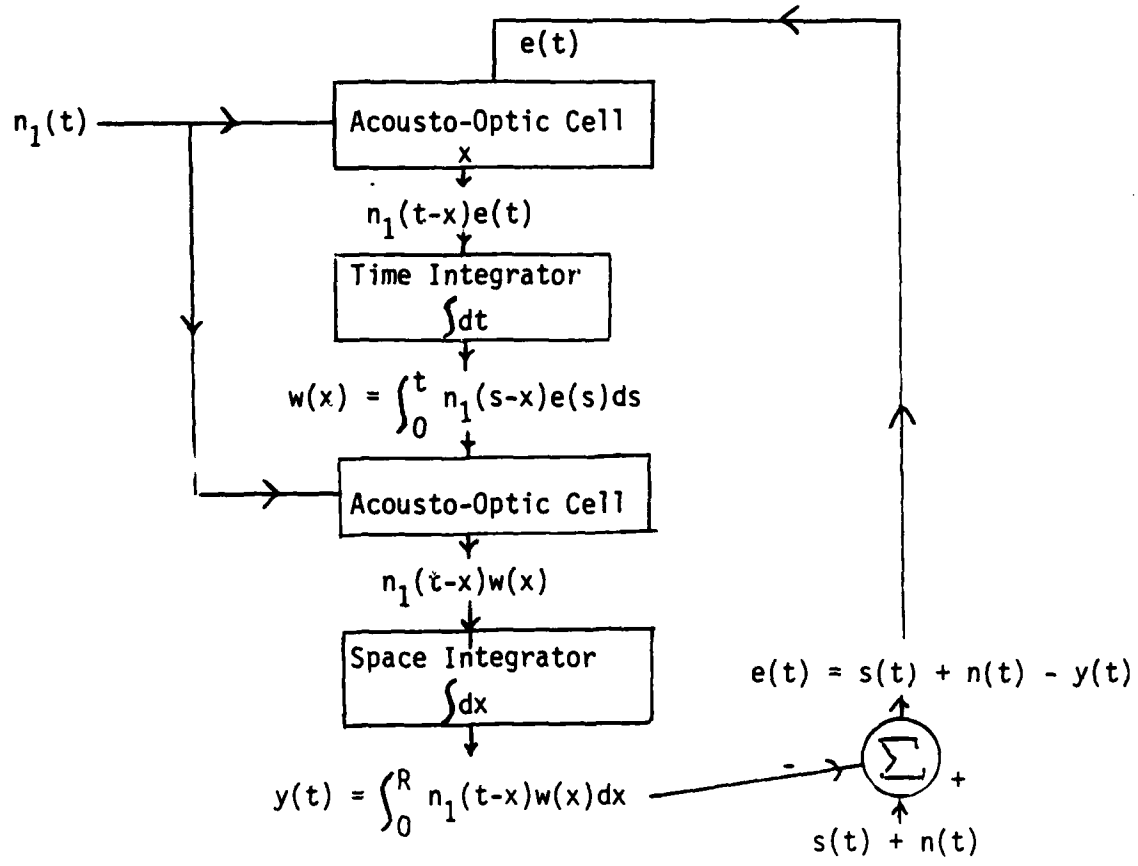


Figure 1
BLOCK DIAGRAM FOR CURRENT SYSTEM

The functions $p_i(x)$ are direction 'vectors', a_i are scalars, and $w_i(x)$ is the i^{th} iterative approximation of $w(x)$. The method of steepest descent uses, for $p_i(x)$ the negative gradient of the functional F, defined by (3.4), evaluated at $w_i(x)$. This gradient is $-2r_i(x)$ (cf., [L]) where

$$r_i(x) = b(x) - Aw_i(x) \tag{4.2}$$

is the i^{th} residual.

The LMS method uses the gradient of $|e_i|^2$, instead of the gradient of $F(w) = E(|e_i|^2)$, where $e_i = e(i\Delta t)$. This gradient can be computed as $-2e_i n_1(i\Delta t-x)$. The algorithm thus becomes

$$w_{i+1}(x) = w_i(x) + a_i e_i n_1(i\Delta t-x) \tag{4.3}$$

where we have absorbed the factor 2 as part of a_i Notice that now we do not have to compute the operator A, as is necessary in determining the residual defined by (4.2).

The iterative scheme given by (4.3) can be easily solved by observing that

$$\sum_{i=\emptyset}^{K-1} w_{i+1}(x) - w_i(x) = w_K(x) - w_\emptyset(x).$$

Taking $w_\emptyset(x) = \emptyset$, we obtain from (4.3)

$$w_K(x) = \sum_{i=\emptyset}^{K-1} a_i e_i n_1(i\Delta t - x). \tag{4.4}$$

We now make the assumption that the step size a_i (or "beam steering signal" [MM]) has been incorporated into the signal n_1 Letting $\Delta t \to \emptyset$, we get the analog version of (4.4), namely

$$w(x) = \int_\emptyset^t e(s) n_1(s-x)\, ds. \tag{4.5}$$

This is the first integral which appears in Figure 1. It can also be interpreted as a correlation of n_1 with the "error" signal $e(t)$.

The advantage of this implementation is its simplicity. There is no iteration loop, rather, the iteration scheme has been solved directly, and an expression for the solution implemented. Also, the problem of computing the operator A has been avoided completely.

The disadvantage is that it may not produce an accurate solution for $w(x)$. The method of steepest descent typically can have very slow convergence, and one would expect this LMS method to be even slower.

V. The Conjugate Gradient Algorithm

We now consider another iterative method for solving equation (3.5). We return to the iteration equation (4.1), which we now write for the case of multiple side channels, so that $\underline{w}_i$

(x)and $\underline{p}_i(x)$ are elements of the Hilbert space H:

$$\underline{w}_{i+1}(x) = \underline{w}_i(x) + a_i\, \underline{p}_i(x).$$

We now choose the direction vectors $\underline{p}_i(x)$ to be a set of linearly independent, A-orthogonal vectors, ie., the $\underline{p}_i(x)$ are such that

$$(\underline{p}_i, A\underline{p}_j) = \emptyset \qquad \text{for } i \neq j. \tag{5.1}$$

The scalar a_i is chosen at each step of the iteration process to minimize the value of $F(\underline{w}_i)$. The iteration method is then said to be a conjugate directions method (the use of the word 'conjugate' here comes from the fact that vectors satisfying (5.1) are said to be A-conjugate).

If one chooses as the vectors $\underline{p}_i(x)$ the A-orthogonalized residuals $\underline{r}_i(x) = \underline{b}(x) - A\underline{w}_i(x)$ then one obtains the conjugate gradient method. This method can be summarized by the following iteration scheme (cf.,[L]):

$$\begin{aligned}
\underline{w}_{i+1}(x) &= \underline{w}_i(x) + a_i\, \underline{p}_i(x) \\
\underline{p}_{i+1}(x) &= \underline{r}_{i+1}(x) - c_i\, \underline{p}_i(x) \\
a_i &= (\underline{r}_i, \underline{p}_i)/(\underline{p}_i, A\underline{p}_i) \\
c_i &= (\underline{r}_i, A\underline{p}_i)/(\underline{p}_i, A\underline{p}_i) \\
\underline{r}_i(x) &= \underline{b}(x) - A\underline{w}_i(x).
\end{aligned} \tag{5.2}$$

The value of a_i comes from minimizing $F(\underline{w}_i)$, and the value of c_i comes from A-orthogonalizing the vectors $\underline{r}_i(x)$.(It is a nontrivial property of this method that one need only A-orthogonalize $\underline{p}_{i+1}(x)$ with respect to $\underline{p}_i(x)$, and not all the preceding $\underline{p}_j(x)$'s, to obtain a complete A-orthogonal set (cf. [L]).)

What is the motivation for considering conjugate direction methods? One reason is the following fact. Suppose we have some weight value $\underline{w}_\emptyset(x)$ and we compute a new value $\underline{w}_1(x)$ from

$$\underline{w}_1(x) = \underline{w}_\emptyset(x) + a_\emptyset \underline{p}_\emptyset(x)$$

where $\underline{p}_\emptyset(x)$ is any nonzero direction vector and $a_\emptyset$ is chosen to minimize $F(\underline{w}_1)$ (so $a_\emptyset$ is given by the expression in (5.2) with $i=\emptyset$). Now suppose $\underline{w}^*(x)$ is the true solution of (3.5). Then the correct direction to go in, in order to reach exactly $\underline{w}^*$ on the next step, is always going to be A-orthogonal to the previous direction vector used (in this case, $\underline{p}_\emptyset(x)$). To see this, note that the direction from $\underline{w}_1$ to $\underline{w}^*$ is $\underline{w}^* - \underline{w}_1$ and

$$\begin{aligned}
(\underline{p}_\emptyset, A(\underline{w}^* - \underline{w}_1)) &= (\underline{p}_\emptyset, \underline{b} - A\underline{w}_1) \\
&= (\underline{p}_\emptyset, \underline{b} - A(\underline{w}_\emptyset + a_\emptyset \underline{p}_\emptyset)) \\
&= (\underline{p}_\emptyset, \underline{b} - A\underline{w}_\emptyset) - a_\emptyset(\underline{p}_\emptyset, A\underline{p}_\emptyset) \\
&\qquad\qquad (a_\emptyset \text{ is real}) \\
&= (\underline{p}_\emptyset, \underline{r}_\emptyset) - a_\emptyset(\underline{p}_\emptyset, A\underline{p}_\emptyset) \\
&= \emptyset \qquad \text{from (5.2).}
\end{aligned}$$

In the discrete case, when A is a finite dimensional matrix, there are only finitely many directions which are A-orthogonal to a given vector. Conjugate direction methods search through this finite list until exactly the right direction vector is found. They are thus guaranteed to converge to the exact solution (ignoring roundoff errors) in a finite number of steps. In contrast, if the method of steepest descent does not obtain the exact solution in one step, then it will always take infinitely many steps to reach the exact solution ([LS]).

We are not dealing with a finite dimensional matrix, but rather with an "infinite dimensional" operator A, so the finite step advantage mentioned above is not, in general, applicable to our situation (there are, however, cases when finite convergence is attained for an operator A (cf. [LS]). However, the conjugate gradient method will always converge more rapidly than the method of steepest descent (see [D] for estimates on the rate of convergence).

VI. Block Diagrams

In this section, block diagrams for two iterative methods are presented. These diagrams are constructed with optical implementation in mind (eg., there is no storage of data or previously computed results). Boxes labeled "compute A", "compute b", etc., are indicated in detail in separate diagrams.

Figure 2 shows a diagram for the method of steepest descent. This method is included here because it is simpler to implement than the conjugate gradient algorithm, yet it contains most of the computational difficulties (computing A, $\underline{b}$, inner products, and inverting scalars). If this method can be implemented optically, then it would be relatively straightforward to modify the resulting system for the conjugate gradient algorithm.

Figure 6 shows the diagram for the conjugate gradient algorithm. As one can see, it contains all of the computations required by steepest descent, plus additional computations required for obtaining the vectors $\underline{p}_i(x)$. Since we have not assumed the possibility of storing previously computed results, we must compute both $\underline{r}_i$ and $\underline{r}_{i+1}$ in each iteration loop. Each of these residual computations requires the computation of A and $\underline{b}$ (see Figures 3-4).

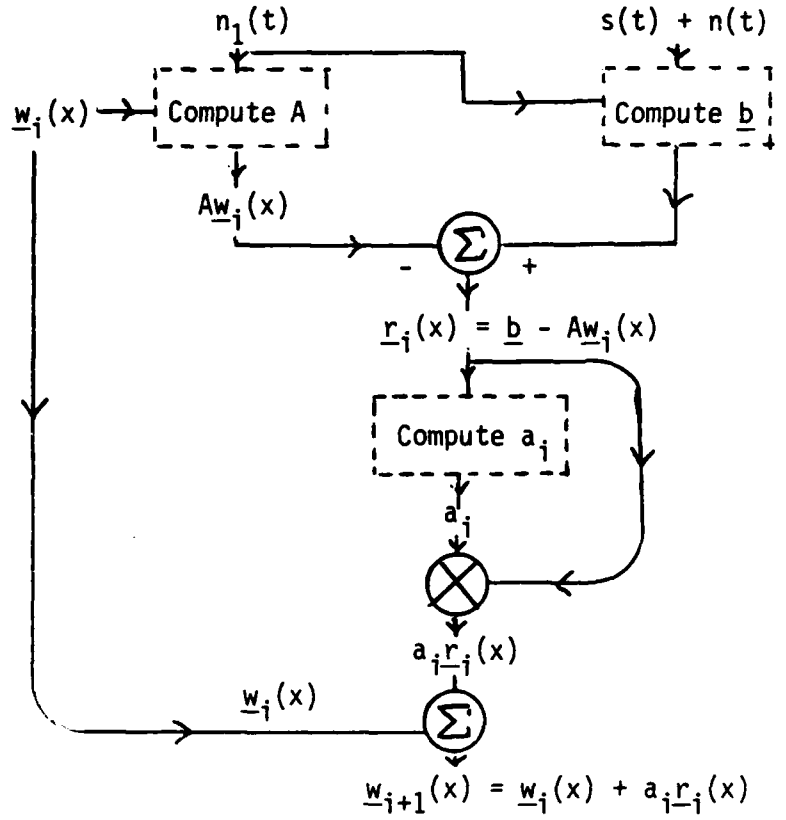


Figure 2. THE METHOD OF STEEPEST DESCENT

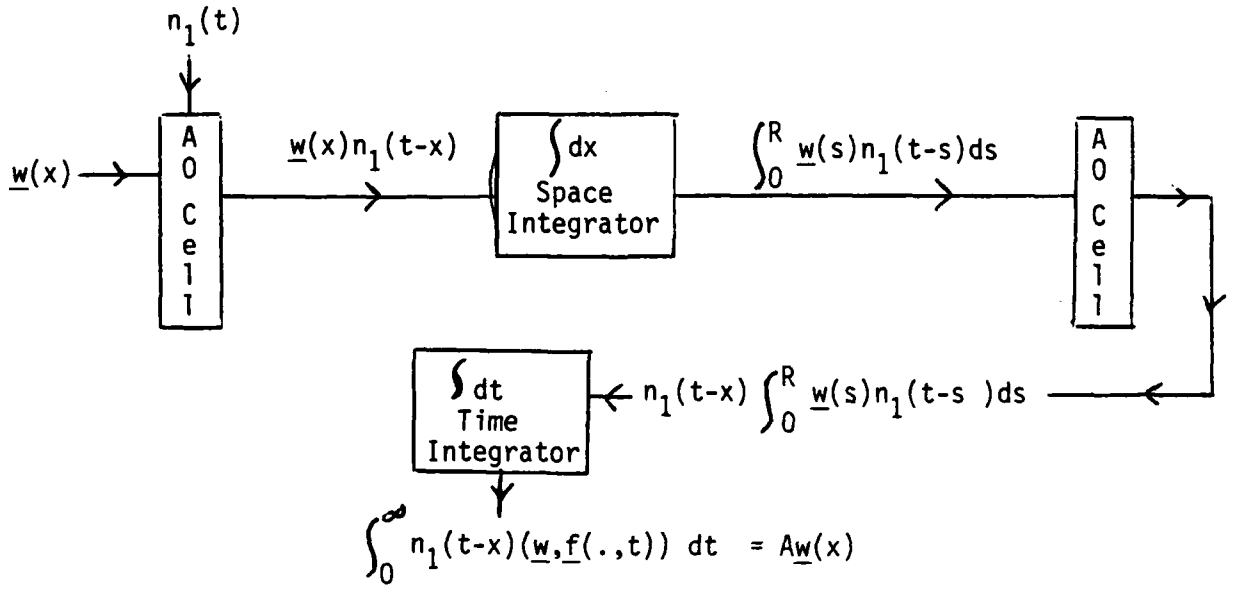


Figure 3. COMPUTING $A\underline{w}(x)$

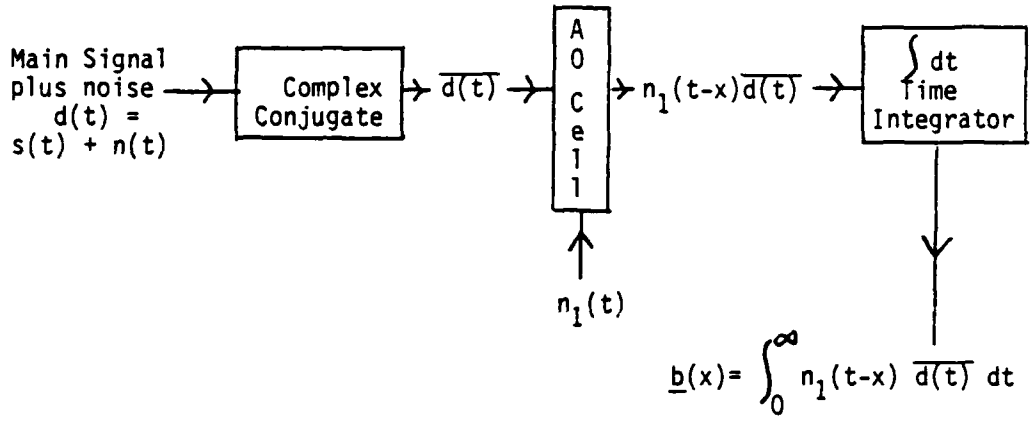


Figure 4. COMPUTING $\underline{b}(x)$

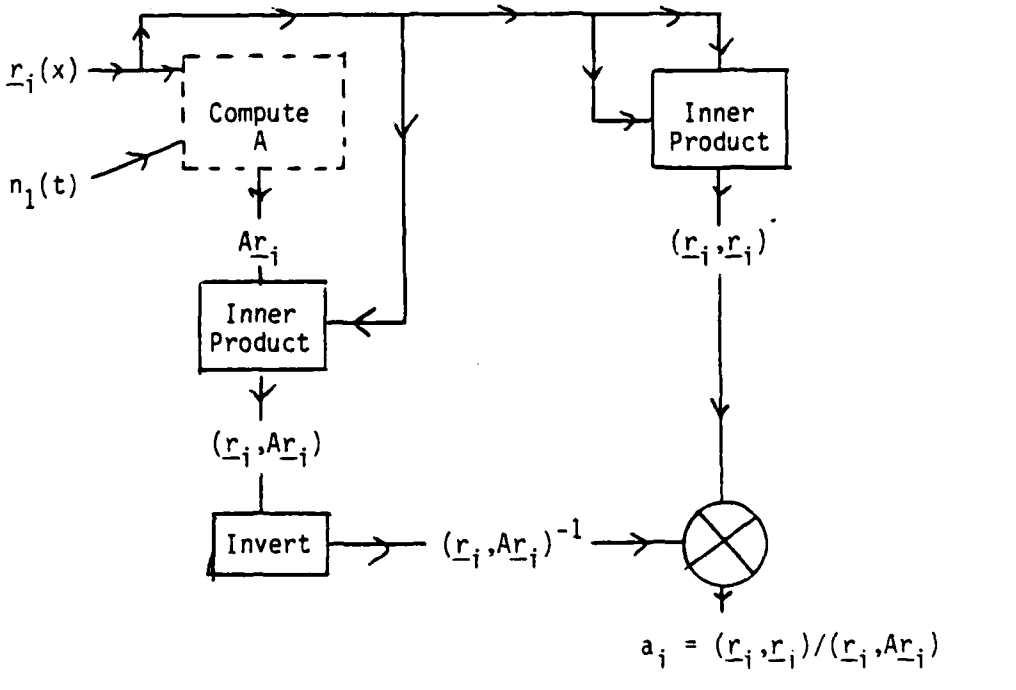


Figure 5. COMPUTING a_i

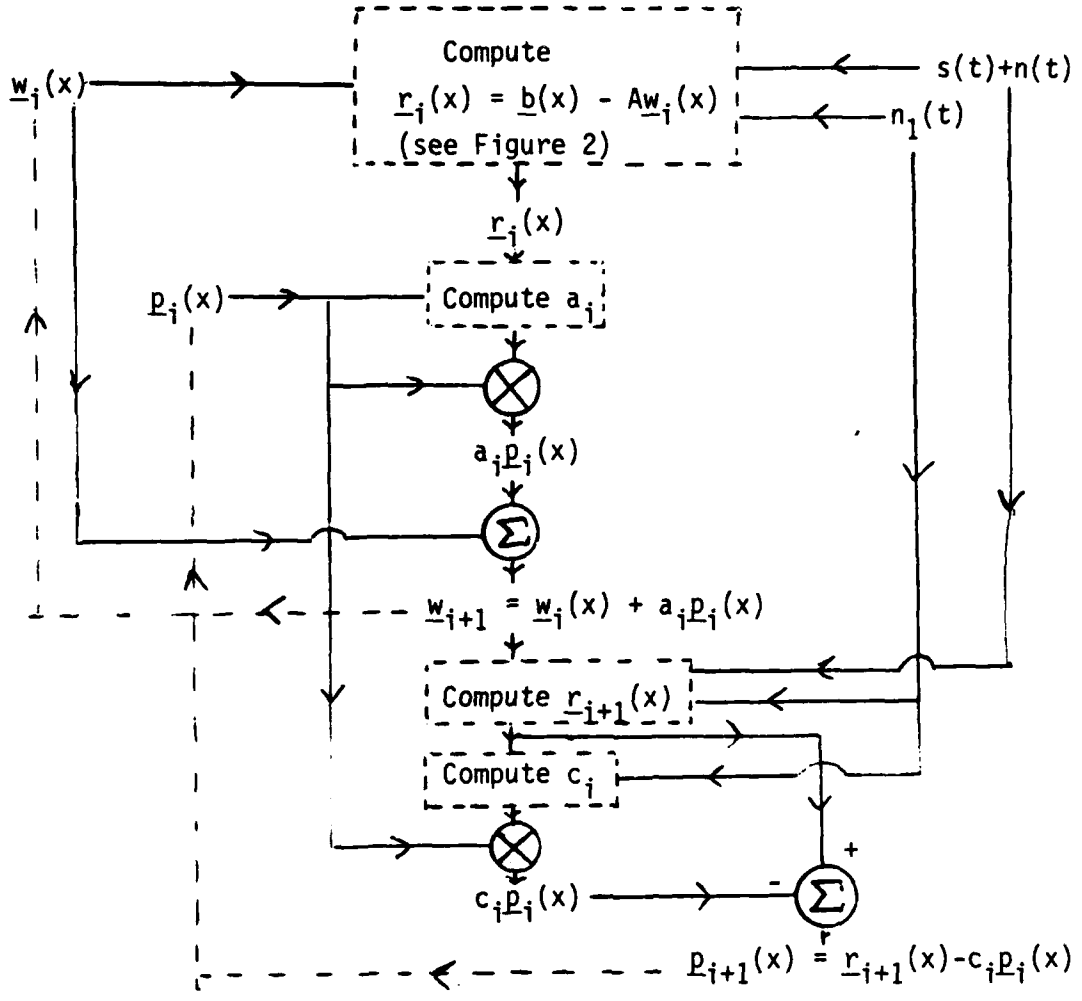


Figure 6. THE CONJUGATE GRADIENT ALGORITHM

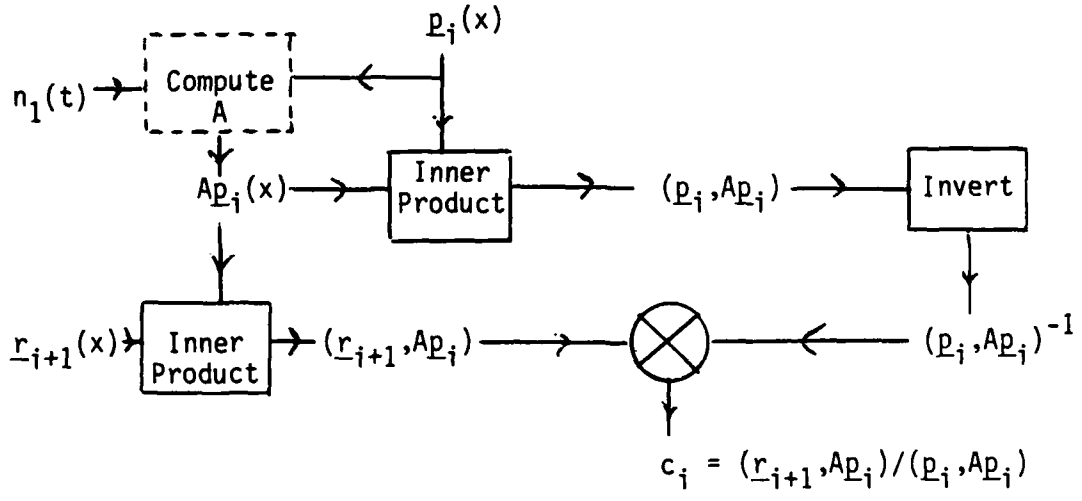


Figure 7. COMPUTING c_i

VII. Recommendations

In order to achieve an optical implementation of either of the iterative methods discussed in the previous section, we must first be able to implement the individual operations shown in the block diagrams. The major operations are:

1. Computation of the operator A (Figure 3). This requires two acousto-optic cells and space and time integration. In the current system, space integration is done with a lens, and time integration with a liquid crystal light valve (LCLV). In fact, all the operations needed to compute A are done in the current system, so this should present no problem. Since A essentially represents an "outer product", reference [A], which discusses optical computation of outer products, may be helpful.

2. Computation of $\underline{b}$ (Figure 4). This requires another acousto-optic cell, and a time integration (LCLV).

3. Inner Products. Each inner product requires complex conjugation, pointwise multiplication, space integration, and summation. Also, both vectors will be represented as light, so an acousto-optic cell (which has one electronic input) may not be appropriate. An efficient optical means will have to be found to compute these inner products.

4. Inverting Scalars. This may be the hardest operation to implement optically. It may require its own iterative loop.

The iterative loops involved in both the steepest descent and conjugate gradient algorithms are a major departure from the existing optical system. There are two alternatives to approaching the implementation of these loops, both having to do with the idea that the weights are supposed to be slowly varying with time.

The first approach would be to consider taking in data in blocks, rather than continuously, and doing a set number of iterations on each block of data to determine the weights. The value of the weights would be updated as each block of data comes in. This would be a true implementation of the algorithms as outlined above.

A second approach, which requires further analysis but would be easier to implement, would be to do just one iteration as part of the existing loop. Since the weights are assumed "constant" in time, this would have the effect of many iterations as new data is continuously brought in and sent through the system. Also, it would be "adaptive", just as the present system is, in that changes in the data should eventually be reflected in changes in the weights. This implementation would not too different from the existing system, but further analysis is needed to determine if the algorithms are still valid when the iterations are slowly varying in time.

My recommendation is that further analysis of the second approach mentioned above be carried out. It should be compared to the first approach, ie., the standard implementation of the algorithms. A computer simulation study comparing both should be done. If an optical implementation seems feasible, it should be carried out for the steepest descent method first, since most of the computational difficulties are encountered there. The conjugate gradient algorithm can be implemeted as a straightforward modification of steepest descent.

References

[A] Athale, R. A., and Lee, J. N., "Optical Processing Using Outer Product Concepts", Proc. IEEE, 72, No. 7, 1984, p. 931.

[D] Daniel, J. W., "The Conjugate Gradient Method for Linear and Nonlinear Operator Equations", SIAM J. Numer. Anal., Vol. 4, No. 1, 1967, p. 10.

[L] Luenberger, D. G., Optimization by Vector Space Methods, Wiley, 1969.

[LS] Lasdon, L. S., Mitter, S. K., Waren, A. D., "The Conjugate Gradient Method for Optimal Control Problems", IEEE Trans. on Auot. Control, AC-12, No. 2, 1967, p. 132.

[M] Mikhlin, S. G., The Problem of the Minimum of a Quadratic Functional, Holden-Day, INc., 1965.

[MM] Monzingo, R. A., and Miller, T. W., Introduction to Adaptive Arrays, Wiley, 1980.

[V1] Vander Lugt, A., "Adaptive Optical Processor", Applied Optics, Vol. 21, No. 22, p. 4005.

[VP1] Vannicola, V. C., and Penn, W. A., "Acousto-Optic Adaptive Processing", GOMAC Digest of Papers, Vol X, 1984.

[VP2] _____, and Lowry, M. F., "Recent Improvements in the Acousto-Optic Adaptive Processor", GOMAC Digest of Papers, Vol. XI, 1985.

1986 USAF-UES SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

UNIVERSAL ENERGY SYSTEMS, INC.

FINAL REPORT

ASSESSING COGNITIVE SKILLS THROUGH A SUPERVISORY CONTROL SIMULATION

Prepared by:	Shih-sung Wen
Academic Rank:	Professor of Psychology
Department & University:	Department of Psychology Jackson State University
Research Location:	School of Aerospace Medicine/NGN Brooks AFB, San Antonio, TX
USAF Researcher:	John C. Patterson
Date:	September 9, 1986
Contract No:	F49620-85-C0013

ASSESSING COGNITIVE SKILLS THROUGH A SUPERVISORY CONTROL SIMULATION

by

SHIH-SUNG WEN

ABSTRACT

The present study intended to investigate the ability of the NASA-Ames Supervisory Control Simulation (SCS) to measure general intelligence and visuospatial skills of Air Force crewmen. Subjects performed computer-displayed SCS tasks which require attention, concentration, estimates of time, memory, visuo-motor coordination, judgment, and strategies. The data collection was not completed due to the time limit in the summer research program. A continuation of the research through a Mini Grant was recommended.

ACKNOWLEDGEMENTS

This research was supported by the Air Force Systems Command, Air Force Office of Scientific Research, and the School of Aerospace Medicine, Psychology Division at Brooks AFB, San Antonio, TX.

Thanks are due to Dr. John C. Patterson for providing adequate research facilities, subjects, and valuable advice. Thanks are also due to Dr. Bryce Hartman for his support and encouragement.

I. Introduction

Recent advancement in avionic technology has altered a great deal of the traditional man-machine relationship in cockpit. The highly sophisticated computers and intelligent software systems in the aircraft are able to assume control and monitoring of a routine flight, handle repetitive tasks, and automatically diagnose and correct mechanical errors. The achivement in automation has drastically reduced psychomotor or physical task demands on piloting (Hart & Sheridan, 1984). Cognitive demands, on the other hand, have greatly increased and become a major concern for a pilot.

Today, the computer-automation and multifunction display systems constitute the key element of the cockpit environment. The versatility of hardwares and softwares provides a pilot with instant messages for actions or reactions. Various options can be displayed for a pilot to judge, choose, or reject. This type of interactive decision-making processes inevitably requires such cognitive processes as attention, perception, comprehension, reasoning, judgement, memory, learning, and problem-solving (Hennessy, 1984).

Currently, few instruments are available to adequately measure a wide range of the aforementioned cognitive processes (Hart & Sheridan, 1984; Wen, 1985; Wierwille, Rahimi, & Casali, 1985). In their effort to develop a comprehensive workload assessment instrument, Hart and her associates (1984) introduced a Supervisory Control Simulation System (SCS). The system combines workload-sensitive multiple tasks, includes

time pressures on performance, varies task difficulty levels, requires decision-makings in forming or choosing task strategies, allows dynamic interactions to occur between the subject and tasks, and obtains subjective ratings of mental workload. The SCS has initially proven to be successful in differentiating Type A and Type B behaviors (Hart, et al, 1985). Other capabilities of the SCS to measure cognitive processes, however, remain to be attested. In his recent review of workload assessment techniques, Wen (1985) suggested that the SCS could measure some important cognitive processes in the future flight.

II. Purpose of the Research

The present research attempted to investigate some selected cognitive correlates of the SCS performance of Air Force flight personnel. The criterion measures are general intellectual abilities (as measured by the Wechsler Adult Intelligence Scale) and visuospatial skills (as measured by the Halstead Category Test).

III. Method

Subjects

Subjects were Air Force crewmen referred to the psychological division of the School of Aerospace Medicine, Brooks AFB, San Antonio, TX.

Tests

Wechsler Adult Intelligence Scale (WAIS). The purpose of the WAIS (Wechsler, 1958) is to measure an adult's intellectual skills in terms of the purposefulness, rationality, and

adaptability of behavior. It consists of eleven subtests, six of whcih are classified as verbal (Information, Comprehension, Arithmetic, Similarity, Digit Span, and Vocabulary) and five are classified as performance (Digit Symbol, Picture Completion, Picture Arrangement, Block Design, and Object Assembly). It has been the most frequently administered individual adult intelligence scale.

Halstead Category Test (HCT). The test was developed by Halstead (1947) to measure an individual's capacity to deduce general principles from the presented specific items. The test consists of seven subtests, each with its own underlying principle. When shown a slide, the subject must deduce the principle represented by the slide. The test is appropriate to measure an individual's conceptual ability, visuospatial skills, attention, concentration, memory, and problem-solving strategies (Golden, et al, 1981).

Instrument

The SCS system is a real time, interactive, multifunction display system of tasks in which certain activities take place that need to be monitored. The simulation system was programmed on an 'enhanced' IBM PC/AT hardware with a 12-inch Princeton graphic color monitor. A visual mouse system is plugged into a RS232 port to communicate with the computer. The program packages include POP.EXE, SCALE.RAT, PROFILE, and MASTER.RAT.

Simulation Task

The SCS is performed through a game-like graphic display

task consisting of five different boxes of task elements (represented by the symbol of *, +, -, #, =) and of thirteen function keys to manipulate boxes and task elements (See Figure 1 below).

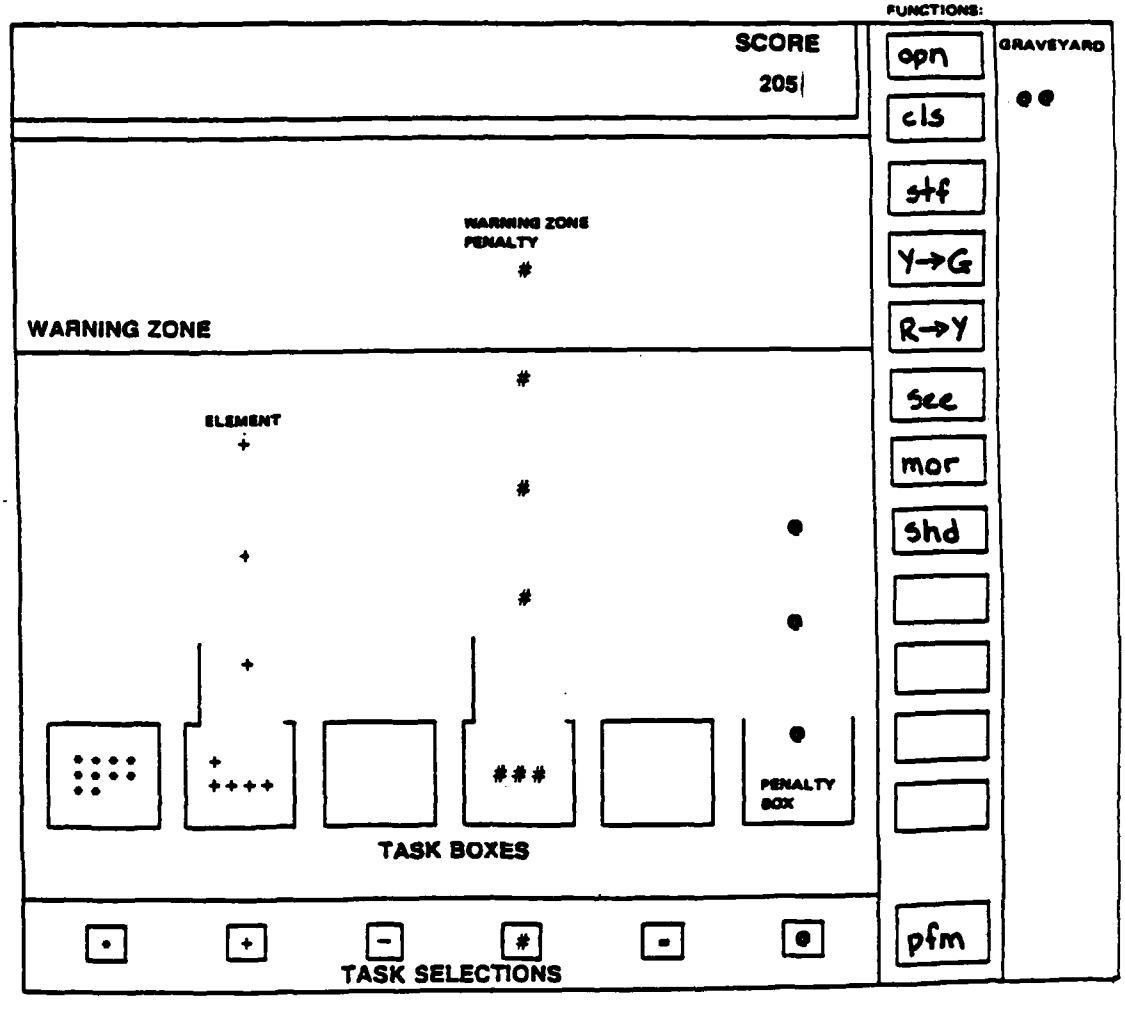


Figure 1: SCS task elements and functions on the CRT.

The purpose of the simulation is to measure an operator's cognitive ability to select, monitor and perform rule-guided semi-automatic tasks by achieving the maximum score with minimum penalty. For example, when a task key and a function key (e.g. OPEN) are sequentially pointed and clicked (hense called selecting) by an arrow cursor with a mouse control device, the box will open for task elements to exit in order. The up-moving task elements soon need to be removed by selecting PFM (Performance) key before they reach the warning zone and finally enter the graveyard. When an element is being performed before entering the warning zone the operator scores. On the other hand, when an element passes through the warning zone and enters the graveyard, points will be automatically deducted as a penalty. Task elements remain in the box will mill around with increasing speed. Score and penalty weights can be predetermined through the scenario generator (POPGEN).

Four scenarios were designed for the present study -- one (FUNFLY) for practices and three (FLYA, FLYB, and FLYC) for formal assessment. Scenario FLYA presents five tasks in sequence (staggered), while scenario FLYB presents five tasks simultaneously (massed). The initial rate and the acceleration rate of milling and traveling of task elements in both FLYA and FLYB are identical across tasks. FLYC mixes task schedules of FLYA and FLYB while the acceleration rate for task elements varies from task to task.

Workload Rating Scale (WRS)

A ten-item single-dimension rating scale was designed to estimate the task workload actually experienced by each subject. The scale includes the rating of overall workload, task difficulty, time pressure, performance level, mental and sensory efforts, physical effort, frustration level, stress level, fatigue, and activity type (skill based vs knowledge based). The scale items are to be presented on the CRT immediately after completion of each task scenario. The subject uses the mouse system to mark his/her responses to each item along a line of continuum. The computer program in the package automatically digitizes it for data analysis.

Procedure

The WAIS and the HCT were administered to each subject prior to SCS tasks. For the SCS performance each subject was briefed on the purpose and the nature of the task followed by demonstrations and practices. After a 30-minute practice session with FUNFLY scenario each subject performed scenarios FLYA, FLYB, and FLYC at 400 seconds each in sequence. Upon completion of each scenario the WRS was presented on the CRT for rating. Individual data on simulations and ratings were automatically recorded and saved in the hard disk and were retrieved later for analysis.

Result

Due to the time limitation on the summer faculty research program, only five out of fifty subjects completed

the required assessment. The data collection, however, would be continued at the focal point by the trained researcher.

Further Continuation of the Research

It is strongly recommended that data collection, data analysis, and report writing be continued through a Mini Grant Program. To this end, a proposal will be submitted to the UES for a favorable consideration.

Reference

Golden, C. J., Osmon, D. C., Moses, J. A., & Berg, R. A. (1981). Interpretation of the Halstead-Reitan neuropsychological test battery. New York: Brune & Stratton.

Hart, S. G., & Sheridan, T. B. (1984). Pilot workload, performance, and aircraft control automation. In R. T. Hennessy (Ed), Human factors considerations in high performance aircraft, AGARD-CP-371.

Hart, S. G., & Battiste, V., & Lester, P. T. (1984). POPCORN: A supervisory control simulation for workload and performance research. 20th Annual Conference on Mannual Control, Sunnyvale, CA.

Hart, S. G., Battiste, V., Chesney, M., Ward, M., & McElroy, M. (1985). Comparison of workload, performance, and physiological measures: Type A personalities vs Type B. Unpublished manuscript.

Hennessy, R. T. (1984). Human factor considerations in high performance air craft. AGARD-CP-371.

Wen, S. S. (1985). POPCORN as a tool for future cognitive workload assessment: A conceptual analysis. 1985 USAF-UES Summer Faculty Research Program, Final Report. Universal Energy System, Inc., Dayton, OH.

Wierwille, W. W., Rahimi, M., & Casali, J. G. (1985). Evaluation of 16 measures of mental workload using a simulated flight task emphasizing mediational activity. Human Factors, 27 (5), 489-502.

1986 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT SUMMER SUPPORT PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

UNIVERSAL ENERGY SYSTEMS, INC.

FINAL REPORT

HYPERBARIC (3ATA) OXYGEN 100% THERAPY AS AN ADJUVANT IN THE TREATMENT OF RESUSCITATED (BROOKE FORMULA) GUINEA PIGS' BURN (3^0, 50 BSA) SHOCK.

Prepared by:	Stanley J. Whidden, M.D., Ph.D.
Academic Rank:	Chief Researcher
Department and University:	Baromedical Research Institute Jo Ellen Smith Medical Center
Research Location:	USAF School of Aerospace Medicine Hyperbaric Medicine Division Brooks AFB, Texas 78235-5301
USAF Research:	MICHAEL W. LISCHAK, Major, USAF, MC JOHN W. FANTON, Major, USAF, BSC
Date:	September 16, 1986
Contract No:	F49620-85-C-0013-760

HYPERBARIC (3ATA) OXYGEN 100% THERAPY AS AN ADJUVANT IN THE TREATMENT OF RESUSCITATED (BROOKE FORMULA) GUINEA PIG BURN (3^0, 50 BSA) SHOCK.

by

Stanley J. Whidden, M.D., Ph.D.

ABSTRACT

Twenty four (24) male guinea pigs (400± 35 kg) with indewelling arterial and venous catheters and indwelling thermistors were temporarily anesthetized with Metofane and scalded through a template in 100°C water to produce a full skin thickness burn over 40% of the BSA. These animals were treated with the Brooke resuscitation fluid formula isobaricly or hyperbaricly O_2 (100% 6 ATA). All animals showed a post burn (PB) decrease 2°C±.3 in body temperature and 50% drop in cardiac output. The burn produced a hemoconcentration at 1/2 hour PB which returned to preburn by 6 hours in both groups. Plasma lactate levels rose PB in all samples but were higher in the isobaric group. Glucose plasma levels increased sharply 30 minutes PB but returned at three hours and remained there for the rest of the experiment. Blood pH was acidotic in all groups with blood O_2 levels being higher in the hyperbaric O_2 group. Heart rate and blood pressure both dropped PB at 1 hour but recovered by 6 hours in both groups. Histopathery samples are still being studies and as of this time there are no conclusive observations. These data trend to indicate that there was some improve-ment on the hemodynamic and metabolic changes during burn shock in the guinea pigs treated with hyperbaric O_2 (3 ATA 100%) and Brooke

Formula as compared to isobaric (1 ATA 100%) and Brooke Formula. More work needs to be done to delineate these observations by the addition of more central control groups.

ACKNOWLEDGEMENTS

I would like to thank the Air Force Systems Command and the Air Force Office of Scientific Research for sponsorship of my research. In a meaningful effort, scientific, clinical research must be conducted in an intellectually stimulating and supporting environment. The most successful scientific studies are indeed a team effort and this work is no less the case. Brooks AFB, USAF School of Aerospace Medicine, Hyperbaric Medicine Division (SAM/HM) provided this environment. I would like to thank the following members of the SAM Staff: Colonel John E. Touhey, Major Wilbur T. Workman, Major Michael W. Lischak, and Captain William E. Drew for clinical professional help and the following for technical assistance: SSG Bruce Beecher, USA; SSG Thomas P. Okey, USA; SSgt Wayne R. Knight, USAF; Sgt Berger S. Nuusa, USAF; Sgt Andrew D. Woodrow, USAF; Amn Lois M. Cheesebrough, USAF.

The Veterinary Research Support Section deserves a special thanks because without them this work could not have been done. They are: Major John W. Fanton, TSgt Steven Dusgh, MSgt David Cushing, Sgt Carole Wilson, Ms Carol Pyka, Mr Val Corothers, Mr Charles Teeters. The strong literary support provided by the USAF School of Aerospace Medicine, Strughold Aeromedical Library, Brooks AFB deserves thanks to Bonnie Fridley, Joe Franzello and Dewey A. Goff, Jr. Lt Colonel Steve Hubbard of the Brooke Army Medical Center, Surgical Research Unit provided histopathology support and are indeed as well deserved of "thank you very much".

A special thanks goes to Ms Starr Foster of Rothe Developers, Inc. for her painstaking techniques in Biochemical preparations.

The laborious task of typing the manuscript was performed by Mrs Linda Tondre, who has the distinct honor of serving as secretary for the Clinical Investigations Branch of the Hyperbaric Medicine Division.

I. Introduction:

I received my Ph.D. from the Auburn University majoring in Clinical Physiology with minors in Anatomy, Biochemistry, and Pharmacoloy. I later was a NIH post-doctoral fellow at the Louisiana State University Medical Center at New Orleans Physiology and Surgery Departments. Here I studied purturbation of fluid resuscitation formulas in burn shock (50% BSA) metabolic and hemodynamic. My medical degree was completed at Universidad Autonoma de Civdad Juarez, Mexico in 1984 and later that year I further completed Diving/Hyperbaric Physicians Tenth Course by the Undersea Medical Society (UMS)/N.O.A.A.)

The research problem at the USAF School of Aerospace Medicine, Hyperbaric Medicine Division, Brooks AFB, was to investigate and determine in a controlled statistical study of the effectiveness of hyperbaric (3ATA) oxygen (100%) therapy (HBO) as an adjuvant in the treatment of resuscitated (isotonic saline) burn (50% BSA 3^{o}) shock using the guinea pig shock model. The problem under investigation at Brooks AFB was, therefore, very similar to the problems I had studied for my Ph.D. post-doctoral work, and my M.D. degree. Because of the similarity I was assigned to work on this project.

II. Objectives of the Research Effort

The overall salient axis of this investigation is to determine if there is a difference in the natural resuscitated course of metabolic and hemodynamic burn shock perturbation when HBO therapy is used in an adjunvant therapy.

1. The primary flank objective of this investigation is to determine if there is a difference in the natural resuscitation course in burned skin would healing when HBO therapy is used as an adjuvant therapy.

2. The secondary flank objective of this investigation is to determine if there is a difference in the natural resuscitation course histopathology when HBO therapy is used as an adjuvant therapy.

III. Review of Literature

On any future modern battlefield where U.S. forces are to be employed in hostile action, burn trauma shall present as a significant percentage of the treated medical patients. If hyperbaric oxygen (HBO) therapy proves to be efficacious in this controlled animal study, then it would enable us to look at attempting a small pilot study, which in turn could lead to this method of becoming a more accepted mode of adjuvant threatment therapy in the Department of Defense.

Over two decades ago, R. Adams Cowley[8] felt that hypoxia was a dominant feature of the pathophysiology of shock and started hyperbaric oxygen investigations into hemorrhagic shock. With the support of the U.S. Army Medical Research and Development Command, he established a Clinic Shock Unit at the University of Maryland and began studies into clinical shock biochemistry and HBO. He studied hemorrhagic, tumbling injury, and bacteremic shock when HBO was used. This initiated a debate in the leterature in which

many papers were published revolving around hamorrhagic shock models. Most of the more recent papers show no significant differences in the cardiovascular effects of hemorrhage when exposed to HBO and fluid resuscitation[15,16].

In over fifty (50) articles written on the topics of burns or shock, none were found which specifically dealt with hyperbaric oxygen therapy used in conjunction with resuscitative measures (i.e., intravenous fluid therapy) for shock produced from thermal burns. C.H. Wells, et al[56], came close to attempting to understand this problem by studying post-burn plasma extravasition and cardiac output. In unresuscitated animals with HBO (2 & 3 ATA) it was found that all animals underwent pronounced decreases in cardiac output after injury. The decline in cardiac output was slightly but significantly ($p=0.05$) less in the HBO than in untreated animals. Clearly, the literature reflects that hyperbaric 3 ATA oxygen (100%) will not in and by itself reduce lactic acid or trigger recovery of cardiac output, or correct completely the stagnant hypoxia caused by shock hemorrhage. Clearly, the focus has been upon the whole body response to injury rather than on the effect of therapy upon the cell or organelles. If one considers such factors as neurohumoral agents, endocrinology, prostaglandins-thromboxanes, and shock toxins or depressant factor, then shock is merely not just a hypoxic state. However, HBO may meanfully inreract to stop the cascade of the sum total of these events that progresses cellularly even though the host has stabilized clinically.

The history of the treatment of burn shock is marked by a curious dialecticism. Beginning more than four decades ago, an extensive debate raged in the literature over the optimum regimen of fluid resuscitation[36]. Contro-

versy focused on the volume of fluid infused, the sodium content of that fluid, the role of additional colloid, blood products, and artificial plasma expanders. Rates of infusion were varied as were the routes chosen. Survival differences between groups of patients constituted the initial comparison of regimens[28, 47]. With increasing technological sophistication, these comparisons came to be expressed in terms of cardiovascular[38, 39], cardiopulmonary[14, 52], and renal[34] functions. No attempt was made to delineate the metabolic consequences of these fluid choices, or the role of hyperbaric oxygen (HBO). The focus continued to be on the host's (or patient's) response to the injury, rather than on the effect of that therapy on the host.

Against this lack of critical information has come the concept that burn shock is no longer a clinical problem, that a variety of different fluids, in a number of different regimens, constitutes adequate therapy[43]. This experiment was performed to question this assumption by investigating the metabolic consequences of defined fluid lactate resuscitation schedules with and without HBO in a controlled laboratory animal environment.

The guinea pig was selected as the experimental animal because of its proven value in burn research. Previous experimentation[40, 1] has delineated its cardiovascular responses to burn injury both with and without fluid resuscitation[11, 12, 19, 60], the metabolism of glucose[62, 64, 65], lactate, free fatty acids,[44, 45, 46] and glycerol[48] during burn shock. The burn produced in this model has been studied anatomically at both the gross and light microscopic level and histochemically. Additionally, the guinea pigs' small size

allows it to be used in small animal chambers. Finally, the metabolic and physiologic similarities between man and guinea pig[54, 66, 1] make it an appropriate choice.

IV. EXPERIMENTAL METHOD/APPROACH:

Twenty four (24) male guinea pigs, weighing 400 ± 50 grams, were utilized for the experiment. They were implanted with vascular catheters and recording probes during a sterile surgical procedure on day 1, and were recovered from anesthesia overnight. The following day they were reanesthetized, the burns were performed, and the guinea pigs was immediately placed in the hyperbaric chamber for experimental therapy and data recording. Only two guinea pigs were utilized each day, one each in morning and afternoon sessions. Only one guinea pig was in the chamber during "dives". SURGICAL PROCEDURE: On day 1 the guinea pigs were anesthetized with sodium pentobarbital (35 mg/kg) IP and with an injection of Lidocaine HCl (10 mg subcutaneous) in the sternohyoid area. All hair was closely clipped and a polyethlene catheter (PE50) was surgically implanted in the left carotid artery for drawing blood samples and measuring blood pressure and heart rate. A thermistor (Edwards 2.5F) was placed in the right carotid artery and advanced into the arch of the aorta for the determination of cardiac output by thermal-dilution and for the measurement of body temperature[18, 26]. An injectable catheter (Edwards Laboratories, 2.5Fr injectate) was advanced through the right jugular vein approximately to the entrance of the right atrium. Sulfathiazole was applied topically before the incision is closed. The catheters was externalized through the skin, wrapped around the neck, and held in place with adhesive tape so that they were accessible at the back of the animal's neck. Food was withheld, but water shall be provided ad libitum until the experiment begins the next day.

On day 2 the guinea pigs were anesthetized with the inhalant agent Metofane administered via a face mask/nose cone device. A template[55] was used to produce the burns, one on the animal's ventral and the other on its dorsal surface, to produce a full thickness burn about 40%[20] of the body surface area. The template is a semitubular device in which the guinea pig is placed, and which has a 4 X 6 inch window on the ventral aspect to allow exposure of a specific region of skin. The rest of the body is fully insulated from any thermal effects. A 7-second immersion of the animal on both sides of the "boat" with template in 100°C water produces the burn. Previous studies have confirmed by histologic analysis that a burn between second and third degree, characterized by vesicle formation and some involvement of the superficial dermis, resulted. Full thickness burns of the dermis are anesthetic, or without pain, due to disruption of the cutaneous nerve endings. Each animal was dried thoroughly with a towel immediately after immersion so that the burn process will not continue.

For evaluation of resuscitative measures, the animals were devided into two groups, as follows: Group I received Brooke Formula, Lactated Ringer's solution of 2 cc/kg % burn per 24 hours and isobaric air. and Group II received Brooke Formula and hyperbaric oxygen. The Lactated Ringer's solution was administered at a rate of 1/2 of the calculated dose over the first 8 hours. All solutions were administered at a constant infusion rate. Fluid was administered by means of an IV drip. Ten guinea pigs in each group were required to provide statistical validity. Data acquired from less than this number would be inconclusive.

After the burning procedure was completed, the guinea pigs were allowed to recover from the Metofane anesthetic. They were placed into a small animal chamber and treated by the above method. The implanted catheters extended out of holes so that sampling did not disturb the animals. During the chamber session and data recording, the guinea pigs were conscious.

At the conclusion of the hyperbaric chamber session, each guinea pig was humanely euthanatized with an injection of T-61 (euthanasia agent) via the jugular catheter. Necropsy of the guinea pigs was performed and the burned skin was saved for histologic examination by a pathologist familar with thermal burn effects.

The results of this are cardiovascular parameters including cardiac output, mean arterial pressure and heart rate, and core temperature were determined before immersion at 1, 3, and 6 hours after immersion in all animals. For the cardiac output measurement 0.5ml of ice saline ($0^{o}C.$) will be infused through the venous catheter and valve determined by Edwards Cardiac Output Computer[26]. Two blood samples (0.7ml) were drawn pre-immersion and at 1, 3 and 6 hours post-immersion in all surviving animals. From the blood samples levels are determined of glucose[25], lactate[51], pH _____(bloodgas) and hematocrit.

V. Results:

Because of equipment failure, hardly any complete sets of data exist on any one animal. And more often than not data does not exist in enough quanity to indicate a significant difference occurring across time for any factor within one group or between groups at a particular time. It will require much more time to finish compiling these data, because the lactate assays are as yet to be finished, and the histopathology will not be completed for several months to come.

However, from the raw data several observation can be made. This 50% BSA 3^{o} burn produced a hemoconcentration at 1/2 hour postburn (PoB.) which returned to preburn (PeB) by 6 hours in both groups. All animals showed a PoB decrease $2^{o}C\pm$ 3 in body temperature and 30% drop in cardiac output at 1/2 hour with recovery at about six hours PoB. The plasma lactate levels rose PoB in all samples but appeared higer in the isobaric group. Glucose plasma levels increased sharply 30 minutes PoB but returned at 2 hours and remained there for the rest of the experiment. Blood ptl was acidotic in all groups with blood 0^{2} levels being higher in the hyperbaric oxygen group. Heart rate and blood pressure both dropped PoB at 1 hour, but recovered by 6 hours in both groups. These raw data suggest that there was some improvement on the hemodynamic and metabolic changes during burn shock in the guinea pigs treated with hyperbaric oxygen (3ATA) and Brooke Formula as compared to isobaric 1ATA air and Brooke Formula.

Discussion

The common crucial factor of all shock models is cellular hypoperfusion. Occurring first in the skin, connective tissues, muscle, and gastrointestinal tract, and then in the vital organs of the heart and brain. The result is cellular anoxia and starvation. Under these conditions, the minimal amounts of substrates available for cellular metabolism into the energy pathways are blocked by anoxia. Consequently, it is the cell's energy pathways that the greatest metabolic shock effects occur. In this study, we examine the effect of hyperbaric oxygen in a burn shock model.

These experiments are among the most incredably complex experiments done in shock and trauma pre-clinical research today because it not only involves very frustrating micro-surgical procedure, but the study of hemodynamic and biochemical metabolic changes in which any one of a number of machines could go down at any time and often did. "Murphy's Law" ruled supreme, that is, if anything could go wrong or break, it did.

In a small animal model we developed a technique for surgical implacement if an arterial and venous catheters, as well as a thermistor for cardiac output be thernal-dilution. To our knowledge, we are about the only one of about 5 laboratories that have the ability to do this delicate micro-surgical procedure. It is inexcuseable that an experimental animal should ever feel pain, and through the use of anethesia we provided both relief of pain and suffering. The level of anesthesia was deep enough so that a number of animals actually die from overdose (which is very close to the deepest surgical plane).

Recommendations

1. These experiments are almost finished. At least two more control groups are needed, one at 3ATA on air, and one at surface on oxygen. But because of the limited amount of time this summer, and chamber restriction these could not be done.

In the experiment groups, much work remains to be done. Complete sets of data on any one animal do not exist because of equipment failure. I donated 5 sets of cardiac output catheter sets and clearly this was the minimum needed because by half way through the experiments we only had two working. The Veterinary Research Support Section needs to replace or add some equipment that failed because of stress of diving a large study.

2. Much work remains to be done with this experiment and a larger multilock needs to be used rather than a small animal chamber.

In normal man lipids constitute more than 80% of the stored fuel reserve and plasma fill fatty acids (FFA) are an important fuel. In FFA Assay needs to be established at this facility along with a lactate assays. Lactate is a very important indicator of metabolism.

3. Clearly, I would recommend that a follow-up mini-grant be endorsed to finish this work.

REFERENCES:

1. Adams, H.R., C.R. Baxter, J.L. Parker, and R. C. Senning, Development of Acute Burn Shock in Unresuscitated Guinea Pigs. Circulatory Shock 8:613-625 (1981).

2. Baxter, C.R., Fluid Volume and Electrolyte Changes of the Early Postburn Period. Clinics in Plastic Surgery 1:693-709 (1974).

3. Broder, G., and M.H. Weil, Excess Lactate: An Index of Reversibility of Shock in Human Patients. Science, 143:1457-1459 (1964).

4. Cain, S.M., et al, Tissue Oxygenation During Hemorrhage in Dogs Breathing 1 and 3 Atmospheres of O_2. Journal of Applied Physiology, 22:255-259 (1967).

5. Caldwell, F.T., Hypertonic vs Hypotonic Resuscitation. Journal of Trauma, 19:874 (1979).

6. Caldwell F.T., B.H. Bowser, Critical Evaluation of Hypertonic and Hypotonic Solutions to Resuscitate Severely Burned Children: A Prospective Study. Annals of Surgery. 54:241 (1979).

7. Casali, R.E., B.H. Bowser, V. Smith, J. Enloe, F.T. Caldwell, W.J. Flanigan, Critical Factors in Resuscitation of the Severely Burned Rat: The Relative Merit of Volume, Tonicity, Sodium Load, and Concentration of the Solution Used. Annals of Surgery, 175:138 (1972).

8. Cowley, R.A., et al, Prevention and Treatment of Shock by Hyperbaric Oxygen. Ann N.Y. Acad Sci, 117:673-683 (1965).

9. Cuono, C.B., Early Management of Severe Thermal Injury. Surgical Clinics of North America, 60: 1021-1033 (1980).

10. Evans, E.I., O.J. Purnell, P.W. Robinett, B.A. Batchelor, M.I. Martin, Fluid and Electrolyte Requirements in Severe Burn. Annals of Surgery, 135:804 (1952).

11. Ferguson, J.L., G.F. Merrill, H.I. Miller, J.J. Spitzer, Regional Blood Flow Redistribution During Early Burn Shock in the Guinea Pig. Circulatory Shock, 4:317-326 (1978)

12. Ferguson, J.L., I. Hikawyi-Yevich, and H. I. Miller, Body Fluid Compartment Changes During Burn Shock in the Guinea Pig. Circulatory Shock, 7:457-466, (1980)

13. Fox, C.L., Jr., The Role of Alkaline Sodium Salt Solutions in the Treatment of Severe Burns. Early Treatment of Severe Burns. Annals of New York Academy of Science, 150:823 (1968).

14. Fulton, R.L., R.P. Fischer, Pulmonary Changes Due to Hemorrhagic Shock Resuscitation with Isotonic and Hypertonic Saline. Surgery, 75:881-891 (1974).

15. Goldberger, E., A Primer of Water, Electrolyte, and Acid-Base Syndromes. Lea and Febiger, Philadelphia, 1970.

16. Gross, D.R., et al, Hemodynamic Effects of 10% Dextrose and of Dextram 70 on Hemorrhagic Shock During Exposure to Hyperbaric Air and Hyperbaric Hyperoxia. Aviat Space Environ Med, 55 (12):1118-1128 (1984).

17. Gross, D.R., et al, Hemodynamic Effects of Lactated Ringer's Solution on Hemorrhagic Shock During Exposure to Hyperbaric Air and Hyperbaric Hyperoxia. Aviat Space Environ Med, 54 (8): 710-708 (1983).

18. Guyton, A.C., C.E. Jones, T.G. Coleman, Circulatory Physiology: Cardiac Output and Its Regulation. Philadelphia, London and Toronto: W.B. Sanders Company, (1973), p. 12.

19. Hikawyi-Yevich, I., H.I. Miller, and G.M. Smigh, The Effects of Heparin on Cardiovascular and Metabolic Parameters Following Burn Shock in Guinea Pigs. Circulatory Shock, 7:103-110, (1980).

20. Hong, C.C., R.D. Ediger, R. Raetz, and S. Djurickovic, Measurements of Guinea Pig Body Surface Area. Lab Animal Sciences, 27:474-476 (1977).

21. Issekutz, B., H.I. Miller, P. Paul, and K. Rodahl, Effects of Lactic Acid on Free Fatty Acids and Glucose Oxidation in Dogs. American Journal of Physiology 209, 6:1137-1147 (1965).

22. Itaya, K., And M. Ui, Colorimetric Determination of Free Fatty Acids in Biological Fluids. Journal of Lipid Research, 6:16-20, (1965).

23. Jelenko, C., M.L. Wheeler, et al, Shock and Resuscitation II: Volume Repletion with Minimal Edema Using the "HALFD" Method. Journal of the American College of Emergency Physicians, 7: 326-333 (1978).

24. Jelenko, C., J.B. Williams, et al, Studies in Shock and Resuscitation, I: Use of a Hypertonic, Albumin-containing, Fluid Demand Regimen (HALFD) in Resuscitation. Critical Care Medicine 7, 4:157-167 (1979).

25. Kabaskalian, P., S. Kalliney, and A. Westcott, Enzymatic Blood Glucose Determination by Colorimetry of N, N-Diethylaniline-4 Aminoantipyrine. Clinical Chemistry, 20:606-607, (1974).

26. Lin, Y., C. Dawson, and S. Horvath, The Expression of Cardiac Output in the Albino Rat, Rattus rattus. Comparative Biochemistry and Physiology (B), 33:910, (1970).

27. Lopes, O.U., V. Pontieri, I.T. Velasco, and M. Rocha DeSilva, Hypertonic NaCl Infusions in the Treatment of Hypovalemic Shock, A Surprising Effect. Homeostasis in Injury and Shock, Adv. Physiol. Sci. 26:221, (1980).

28. Markey, K., E. Smallman, R.C. Millican, The Efficacy and Toxicity of Iso-, Hypo, and Hypertonic Sodium Solutions in the Treatment of Burn Shock in Mice. Surgery, 57:698-704, (1965).

29. McCarthy, M.D., and N. Newlin, Range of Efficacy of Sodium Chloride Solutions in Treating Severe Thermal Injury in the Rat. Journal of Laboratory and Clinical Medicine, 41:416-436, (1953).

30. Miller, T.A., et al, Epithelial Burn Injury and Repair. Hyperbaric Oxygen Therapy, 251-257, (1977).

31. Millican, R.C., H. Tabor, and S.M. Rosenthal, Traumatic Shock in Mice. Comparison of Survival Rates Following Therapy. American Journal of Physiology, 170:179-186, (1952).

32. Monafo, W.W., The Treatment of Burn Shock by the Intravenous and Oral Administration of Hypertonic Lactated Saline Solution. Journal of Trauma, 10:575-586, (1970).

33. Monafo, W.W., C. Chuntrasakul, and V.H. Ayvazian, Hypertonic Sodium Solutions in the Treatment of Burn Shock. American Journal of Surgery, 126:778-783 (1973).

34. Monafo, W. W., V. H. Ayvazian, R. Logel, F. Deitz, M. Eve, Renal Function After Thermal Trauma: The Effects of Treatment of Renal Blood Flow and Sodium and Water Excretion. Surgery, 79:342-345, (1976).

35. Monafo, W.W., Volume Replacement in Hemorrhage, Shock and Burns. Advances in Shock Research, 3:47-56, (1980).

36. Morton, J.H., Fluid Replacement in Patients with Large Area, Full and Partial Thickness Burns. Archives of Surgery, 114:247, (1979).

37. Moyer, C.A., H.W. Margoff, W.W. Monafo, Burn Shock and Extravascular Sodium Deficiency-treatment with Ringer's Solution with Lactate. Archives of Surgery, 90:799-811, (1965).

38. Moylan, J. A., J.M. Reckler, and A. D. Mason, Jr., Resuscitation with Hypertonic Lactate Saline in the Thermal Injury. American Journal of Surgery, 125:580-584, (1973).

39. Moylan, J.A., A.D. Mason, Jr., P. W. Rogers, and H. H. Walker, Postburn Shock-A Critical Evaluation of Resuscitation. Journal of Trauma, 13:354-358, (1973).

40. Nicolai, J.P.A., R.J.A. Goris, A Guinea-Pig Model in Burn Research. European Surgical Research, 12:22-29, (1980).

41. Paul, Pavle, B. Issekutz, and H.I. Miller, Interrelationship of Free Fatty Acids and Glucose Metabolism in the Dog. American Journal of Physiology 221, 6:1313-1320, (1966).

42. Peretz, D.I., H.M. Scott, J. Duff, J.B. Dossetor, L.D. Maclean, M. McGregor, The Significance of Lacticacidema in the Shock Syndrome. Chemistry and Metabolism of L- and D-Lactate Acids, Annals of New York Academy of Sciences 119:1133-1141 (1965).

43. Pruitt, B.A., Fluid Resuscitation for Extensively Burned Patients. Journal of Trauma 21:690-692 (1981).

44. Robinson, J.M. H. I. Miller, Free Fatty Acid Kinetics Following Burn Shock in Guinea Pigs. Circulatory Shock, 6: 188(a), (1979).

45. Robinson, K.M., and H. I. Miller, Adipose Tissue and Plasma FFA and the Glycerol Concentrations During Burn Shock in Guinea Pigs. Proceedings of the Society for Experimental Biology and Medicine (December, 1980).

46. Robinson, K.M., and H.I. Miller, Free Fatty Acid Kinetics After Burn Shock in Guinea Pigs. Circulatory Shock, (in press).

47. Rosenthal, S.M., Experimental Chemotherapy of Burns and Shock. III. Effects of Systemic Therapy in Early Mortality. Public Health Reports, 58:513, (1943).

48. Schumer, W., G.S. Moss, and L.M. Nyhus, Metabolism of Lactic Acid in the Macacus Rhesus Monkey in Profound Shock. American Journal of Surgery, 118:200-205, (1969).

49. Schumer, W., The Effects of Shock on the Energy Pathways of the Cell. The Cell in Shock, pp 8-11 Upjohn (1974).

50. Shimazaki, S., T.Yoshioka, N. Tanaka, et al. Body Fluid Changes During Hypertonic Lactated Saline Solution Therapy for Burn Shock. Journal of Trauma, 17:38-43, (1977).

51. Strom, G., Technique for Lactate Determination, Acta Physiologica Scandanavia, 17:440, (1949).

52. Traber, D.L., T.H. Adau, P.D. Thompson, Cardiopulmonary Effects of Fluid Resuscitation Following Burn Wound Sepsis. Circulatory Shock, 6:177, (1979).

53. Wachtel, W.L. G.R. McCahan, W.W. Monafo, Fluid Resuscitation in a Porcine Burn Shock Model. Journal of Surgical Research, 23:405-414, (1977).

54. Wanger, J.E., and P.J. Manning, The Biology of the Guinea Pig, p. 6, Academic Press, New York, (1976).

55. Walker, H.L., and A.D. Mason, A Standard Animal Burn. Journal of Trauma, 8:1049-1057, (1968).

56. Wells, C.H., J.G. Hilton, Effects of Hyperbaric Oxygen on Post-Burn Plasma Extravasation. Hyperbaric Oxygen Therapy, 259-265, (1977).

57. Wilmore, D.W., The Metabolic Management of the Critically Ill. New York, London, Plenum Medical Book Company, 77, p. 14.

58. Wilson, B.J., Stirman: Initial Treatment of Burns. Journal of the American Medical Association, 173: 509-516, (1960).

59. Winer, B.J. Statistical Principles in Experimental Design, p. 298-318, McGraw-Hill, New York, (1962).

60. Wolfe, R.R., H.I. Miller, Cardiovascular and Metabolic Responses in Burn Shock in Guinea Pigs. American Journal of Physiology, 231:892-897, (1976).

61. Wolfe, R.R., H.I. Miller, Burn Shock in Untreated and Saline Resuscitation Guinea Pigs - Development of a Model. Journal of Surgical Research, 21:269-276, (1976).

62. Wolfe, R.R., J.J. Spitzer, H.I. Miller, D. Elahi, Effects of Insulin Infusion on Glucose Kinetics in Normal and Burned Guinea Pigs. Life Sciences, 19:147-156, (1976).

63. Wolfe, R.R., H.I. Miller, and J.J. Spitzer, Glucose and Lactate Kinetics in Burn Shock. American Journal of Physiology, 232:E415-E418, (1977).

64. Wolfe, R.R., J.F. Burke, Effect of Burn Trauma on Glucose Turnover, Oxidation, and Recycling in Guinea Pigs. American Journal of Physiology, 233, E80-E85, (1977).

65. Wolfe, R.R., D. Elahi, J.J. Spitzer, and H.I. Miller, Effect of Burn Injury on Glucose Turnover in Guinea Pigs. Surgery, Gynecology and Obstetrics, 144:359-364, (1977).

66. Wolfe, R.R.,, Review: Acute Versus Chronic Response to Burn Injury. Circulatory Shock, 8:105-115, (1981).

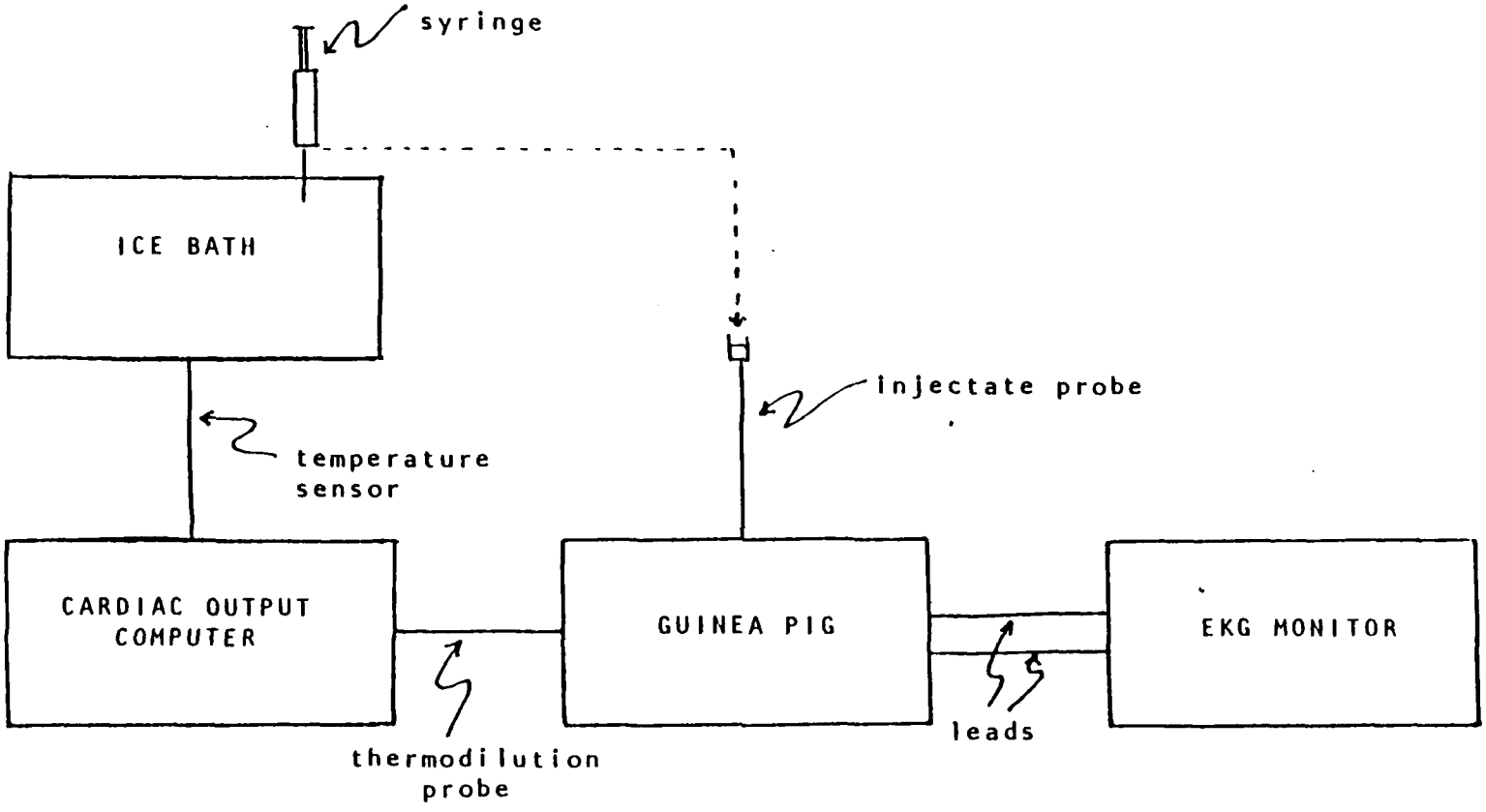
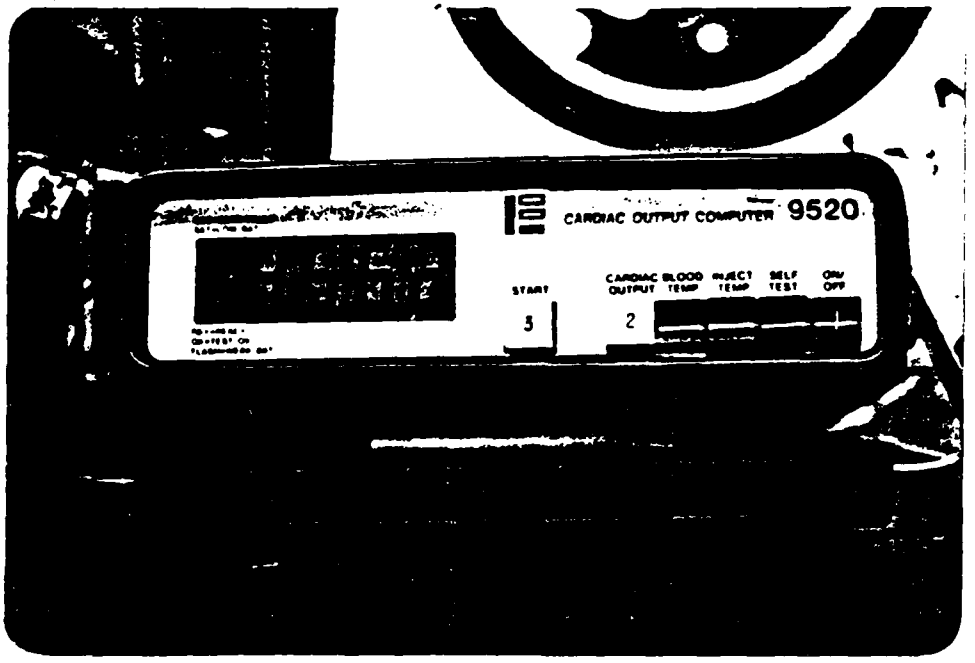


Figure (1) Block Diagram of Setup. The EKG leads are connected at the right and left thoracic cavity of the guinea pig. The thermodilution probe is in the arch of the aorta and the injectate catheter is in the vena cava.



a) From left to right: the chamber, the cardiac output computer, the ice bath for saline, the experimental animal, and the EKG monitor.



b) The cardiac output computer

FIGURE (2) THE EXPERIMENTAL SETUP

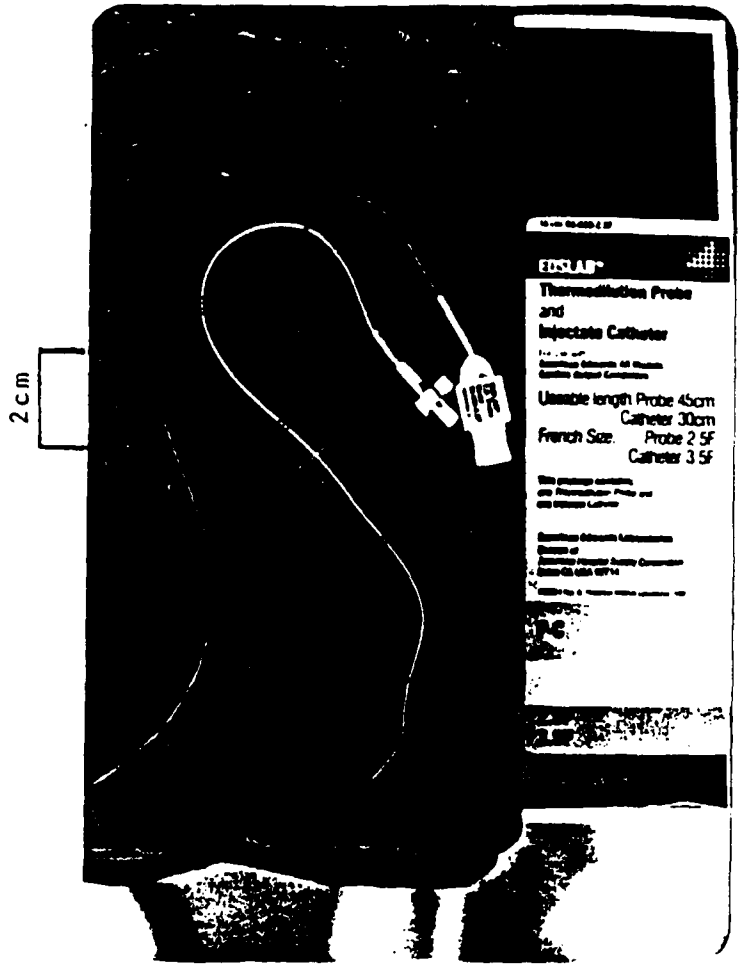


Figure (3) The Thermoldilution Probe and Injectate Catheter
Thermodilution probe - 2.5F, 1mm outside diameter
45cm long
Injectate catheter - 3.5F, 1.75mm outside diameter
30cm long

1986 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/
GRADUATE STUDENT SUMMER SUPPORT PROGRAM

Sponsored by the
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
Conducted by the
Universal Energy Systems. Inc.
FINAL REPORT

Analytical Computer Modeling of the NPN BICFET Device

Prepared by:	Dennis Whitson and W. David Schmidt
Academic Rank:	Full Professor / Graduate Student
Department and University:	Physics Department Indiana University of Pennsylvania
Research Location:	Avionics, Electronics Technology Division, Electronics Research Branch, Device Research Group
USAF Researcher:	Gary McCoy and Chern Huang
Date:	Septemper 24, 1986
Contract No.:	F49620-85-C-0013

Analytical Computer Modeling of the NPN BICFET Device

by

Dennis Whitson and W. David Schmidt

ABSTRACT

A computer program was written for an analytical model of the NPN Bipolar Inversion Channel Field Effect Transistor (BICFET) device. From this analysis a number of conclusions can be drawn: (1) Type of metal used is extremely important; (2) The "Fermi Factor" which determines the electron population in the spike layer could be crucial and the spike layer may have to be grown in the semi-insulator rather than the semiconductor; (3) Collector stretch may be negligible at realistic current density values; (4) For the GaAs/AℓGaAs system a functioning NPN device may be easier to fabricate than a PNP device; (5) There are two independent gain factors: (a) the exponential argument which depends on ψ_{ms} and (b) p_o which depends only on the density of the spike layer dopant.

ACKNOWLEDGMENTS

We would like to thank the Air Force Systems Command and the Air Force Office of Scientific Research for sponsorship of our research. The Avionics Laboratory of Wright Patterson Air Force Base has an environment quite conducive to scientific research. Everyone we interacted with over the summer was extremely helpful and friendly. The general attitude shown at the labs makes for a productive, interesting, and amiable atmosphere. Special thanks go to Gary McCoy amd Chern Huang for giving us the opportunity and the guidance necessary for the research. Also, we would like to thank the members of the computer support team for all their help.

Finally Dennis Whitson would like to thank his wife, Sandy, for her loving support and impressive patience in putting up with our seperation during the summer.

I. Introduction

Recent research interests of Dennis Whitson have been in the area of semiconductor device physics. During the summers of 1981, 82, 83, and 84, he worked at Westinghouse Research & Development labs. Dr. Whitson also worked there one day a week during the school year and spent a sabbatical year (1984-85) at the labs. During this time he was involved mainly with double injection silicon power devices with both injection and MOS gates. The summer of 1985, Dr. Whitson was at GTE labs working with isolation techniques in power intergrated circuits. At that time he decided that he wished to become involved with compound semiconductors and hetereojunction devices. This is an area that avionics/AADR is pursuing and it fit very nicely into his plans.

W. David Schmidt was a M.S. graduate student at Indiana University of Pennsylvania and is now in the Ph.D. graduate program at Dartmouth.

II. Objectives of the Research Effort

The objective of the ten week research effort was to make significant progress in the modeling of two GaAs/AlGaAs devices: the BICFET (bipolar inversion channel field effect transistor) and the DOES (double heterostructure optoelectronic switch).

The three layers of material (e.g. AlGaAs(p)/GaAs(n^+)/GaAs(p)) that form the BICFET (1,2) are also the first three layers of the DOES device (3). It is hoped that both the DOES device and the BICFET can be processed on the same substrate. This would allow both optical switching and electronic signal processing with the same integrated circuit.

The BICFET device (1,2) by itself offers intriguing possibilities:

1. High current gain (~10^5),
2. High current operation (~10^6 A/cm^2),
3. High frequency operation (~400 GH_z). and
4. Different scaling laws than bipolar or MOSFET with no punchthrough.

III. COMPUTER PROGRAM OPTIONS

Since Taylor and Simmons[1] modeled a silicon NPN device we decided to duplicate their work first and then to proceed to modeling GaAs/AlGaAs devices. In the process we also checked a number of their approximations. The computer program therefore has a number of options in it.

Two of the options that can be used or not are "Collector Stretch" and "Fermi Factor". These two options affect the equation used for charge balance, i.e.

$$Qe + Qi + Qc = 0 \tag{1}$$

where Qe = emitter charge, Qi = spike layer charge, and Qc = collector charge. Using the "Fermi Factor" results in

$$Q_i = \frac{-q N_i}{1 + 4\, e^{(\epsilon_p - \Delta E_i)/KT}} \tag{2}$$

where Ni is the charge density per cm^2 of the spike layer, ϵ_p is the acceptor level energy above the valence band in the spike layer, and ΔE_i is calculated from the density of the inversion layer at the interface (semiconductor/spike layer), p_o, using the equation:

$$p_o = N_v\, e^{-\Delta E_i/KT} \tag{3}$$

If the "Fermi Factor" is not used then the denominator in equation(2) is equal to one. When "Collector Stretch" is used we have

$$Q_c = \left(\frac{2q\epsilon_s}{\beta}\right)^{\frac{1}{2}} \left[p_o + \beta \phi_s \left(N_d - \frac{J_c}{q V_s}\right)\right]^{\frac{1}{2}} \tag{4}$$

When "Collector Stretch" is not used $J_c = 0$ in this equation. The inclusion of J_c (Collector current density) in this equation causes the collector depletion area to increase, i.e. to "stretch", whenever the device is conducting current.

Other options include the choice of either calculating certain parameters or using default values. The effective velocities can be calculated

$$V_n = \frac{V_{tn}}{1 + d/\ell_n} \quad ; \quad V_p = \frac{V_{tp}}{1 + d/\ell_p} \tag{5}$$

where d is the thickness of the semi-insulator; V_{tn} and V_{tp} are the thermal velocities in the semi-insulator for electrons and holes, respectively, and

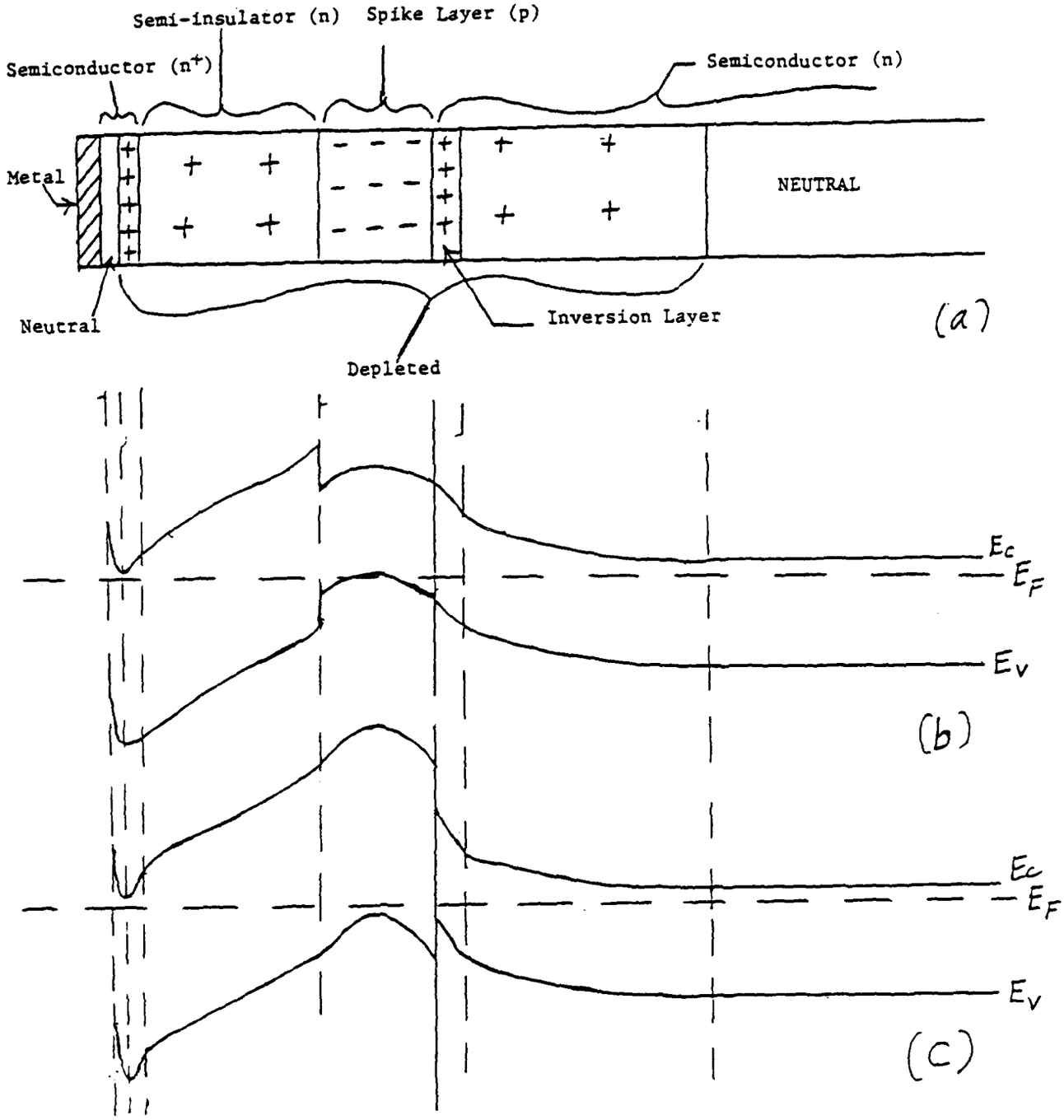


Figure 1. (a) Cross-section of a NPN BICFET. The semi-insulator region is about 200 Å wide and the spike layer is about 20 Å, (b) Equilibrium band diagram with the spike layer in the semiconductor. (c) Spike layer in the semi-insulator.

$$\ell_n = \frac{D_n}{V_{tn}} \quad ; \quad \ell_p = \frac{D_p}{V_{tp}} \tag{6}$$

where D_n and D_p are the diffusion constants in the semi-insulator for electrons and holes, respectively. The default values are:

$$V_n = V_p = V_{tn} = V_{tp} = 10^7 \text{ cm/sec} \tag{7}$$

If Φ_n is calculated we use

$$\Phi_n = K T \ell n\left(\frac{N_c}{N_d}\right) + \Delta E_c + \Psi_{ms} \tag{8}$$

where N_c is the effective density of states of the conduction band, N_d is the donor concentration in the semiconductor, ΔE_c is the difference in conduction bands between the semi-insulator (wide-band gap semiconductor) and the semiconductor, and ψ_{ms} is the difference between the work functions of the metal and the semiconductor. The default value is:

$$\Phi_n = 0.02 \text{ eV} \tag{9}$$

When choosing to calculate ψ_{ms} you also have a choice of the type of metal:

$$\begin{aligned} A\ell &: \psi_{ms} = 0.026\ \ell nN_d - 1.218 \\ n^+\text{Poly}&: \psi_{ms} = 0.026\ \ell nN_d - 1.348 \\ \text{Au}&: \psi_{ms} = 0.026\ \ell nN_d - 0.308 \end{aligned} \tag{10}$$

These values were taken from experimental results[4]. The default value for ψ_{ms} is:

$$\psi_{ms} = -.32 \tag{11}$$

One can also calculate a consistent set of equilibrium values or use the default values:

$$V_i^* = -1.2, \quad \phi_s^* = 0.88 \tag{12}$$

These variables are defined in Figure 2. If the default values are used instead of calculating these values, equation (1) is not satisfied in equilibrium.

Other parameters that must be specified each time the program is run are the spike layer concentration, the semiconductor donor concentration, N_d, the range and number of steps of the collector to emitter voltage, V_{ce}, and the range and number of steps of the source current, J_s. For each value of J_s the full range of values of V_{ce} are stepped through.

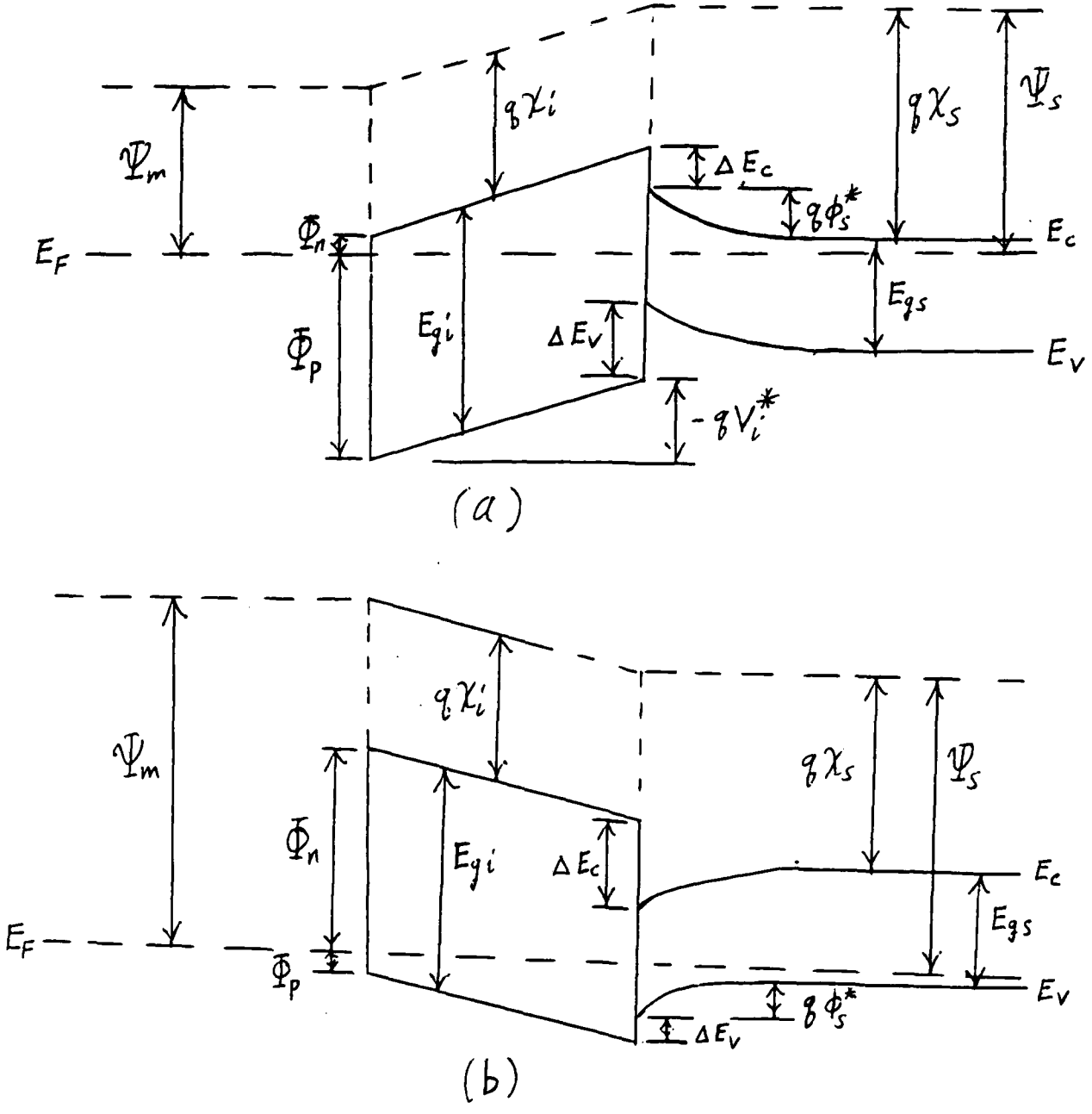


Figure 2: Equilibrium band diagrams assuming that the spike layer is infinitesimal (a) Diagram for NPN device (Figure 1(a)). (b) Diagram for PNP device.

IV APPROXIMATIONS IN THE PHYSICAL MODEL

Certain approximations cannot be checked within the present model. Among these are: (1) treatment of the wide-band gap material as an insulator; (2) assumption that the semi-insulator and the spike layer are totally depleted; (3) neglect of the band shape of the spike layer; and (4) neglect of carrier-carrier scattering and Auger recombination.

Other approximations can be and were checked using the present model: (1) neglect of "collector stretch"; (2) neglect of "Fermi factor"; (3) thermionic emission vs. diffusion; (4) using default value of 0.02 for Φ_n; (5) using default value of -0.32 for ψ_{ms}; (6) assuming that $\Delta E_c = q(\chi_s - \chi_i)$; and (7) using default value of -1.2 for V_i^*.

The treatment of the wide-band gap material as an insulator, i.e., assuming that:

$$V_i = \frac{Q_e}{C_i} \quad ; \quad C_i = \frac{\epsilon_i}{d} \tag{13}$$

does not allow the model to determine whether the semi-insulator is always depleted or not. If the semi-insulator is not totally depleted then there would be a neutral area whose size would change with the applied voltage. This could change the emitter efficiency.

Neglecting the band shape of the spike layer (see Figure 1) could have two effects. If the electric field didn't cross the x-axis at the mid-point of the spike layer then the band shape may be asymmetrical and the effective ΔE_c and ΔE_v could change with semiconductor and semi-insulator donor concentration. Also the curvature as shown in Figure 1(b) would put most of the valence band above the quasi-Fermi level and most of the acceptors would not be occupied by an electron. This would reduce the effective charge in the spike layer and the semi-insulator would probably not be fully depleted.

It is found[5] with p-i-n diodes that carrier-carrier scattering and Auger recombination profoundly reduce the current carrying capacity of the diode. Anything much above 2,000 A/cm^2 is very difficult to attain. Effects of this sort may not allow currents on the order of 10^6 A/cm^2 as hoped for with the BICFET. However, the BICFET has a current path on the order of 300-400 A as opposed to 10-400μm for the p-i-n diode. Thus, with the BICFET these effects may not be as important, but they probably should be considered.

V RESULTS FROM THE COMPUTER MODEL OF THE BICFET

All results of the computer program reported here are for the silicon NPN BICFET. The gain equation for the PNP device is, however, discussed later on. The I-V curve for a device with all the approximations (no collector stretch, no Fermi factor) and all the default values (including spike layer concentration = 1.6×10^{19}) has a response similar to that in Figure 3 except that Jc is constant for larger values of V_{ce} (> 0.5V). Collector stretch is an important factor when:

$$N_d \lesssim \frac{J_c}{q V_s} \tag{14}$$

As seen in Figure 4, the collector stretch causes a negative slope but that the lines become more horizontal as J_c (and J_s) decrease. For the case of collector stretch the results for $N_d = 10^{15}$ are essentially the same as for $N_d = 10^{17}$. For the case of no collector stretch the curves for $N_d = 10^{17}$ show a definite positive slope as seen in Figure 3. However, if J_c is decreased, the curves for collector stretch and no collector stretch are identical for N_d of any value. Thus, if $J_c \leq 10^4$ A/cm^2 collector stretch is probably negligible.

Introducing the Fermi factor into the set of equations that were solved iteratively caused the I-V curves to be distorted as shown in Figure 5. The Fermi factor reduces the effective spike layer charge by a factor of 100. For the equations with collector stretch this produces:

$$J_c \cong \frac{q V_s}{\beta \phi_s} p_o \tag{15}$$

where

$$\phi_s = V_{ce} - V_i + \psi_{ms} \tag{16}$$

For this situation the values of p_o (inversion layer concentration at the interface), V_i and ψ_{ms} are constant. Thus J_c increases as ϕ_s decreases with V_{ce}. However, for the case of no collector stretch the curves are flat for $N_i \geq 5.0 \times 10^{18}$. For $N_i = 5 \times 10^{17}$ (on the order of 100 times less than $N_i = 1.6 \times 10^{19}$) the slope is positive and not negative as it was in the case of collector stretch with the Fermi factor. For the case of no collector stretch the gain increases as N_i is reduced. As shown in Equation (17) the gain increases with a

$$G = \frac{v_n}{v_p} \frac{N_c}{p_o} e^{(\Delta E_v - \Phi_n)/KT} \tag{17}$$

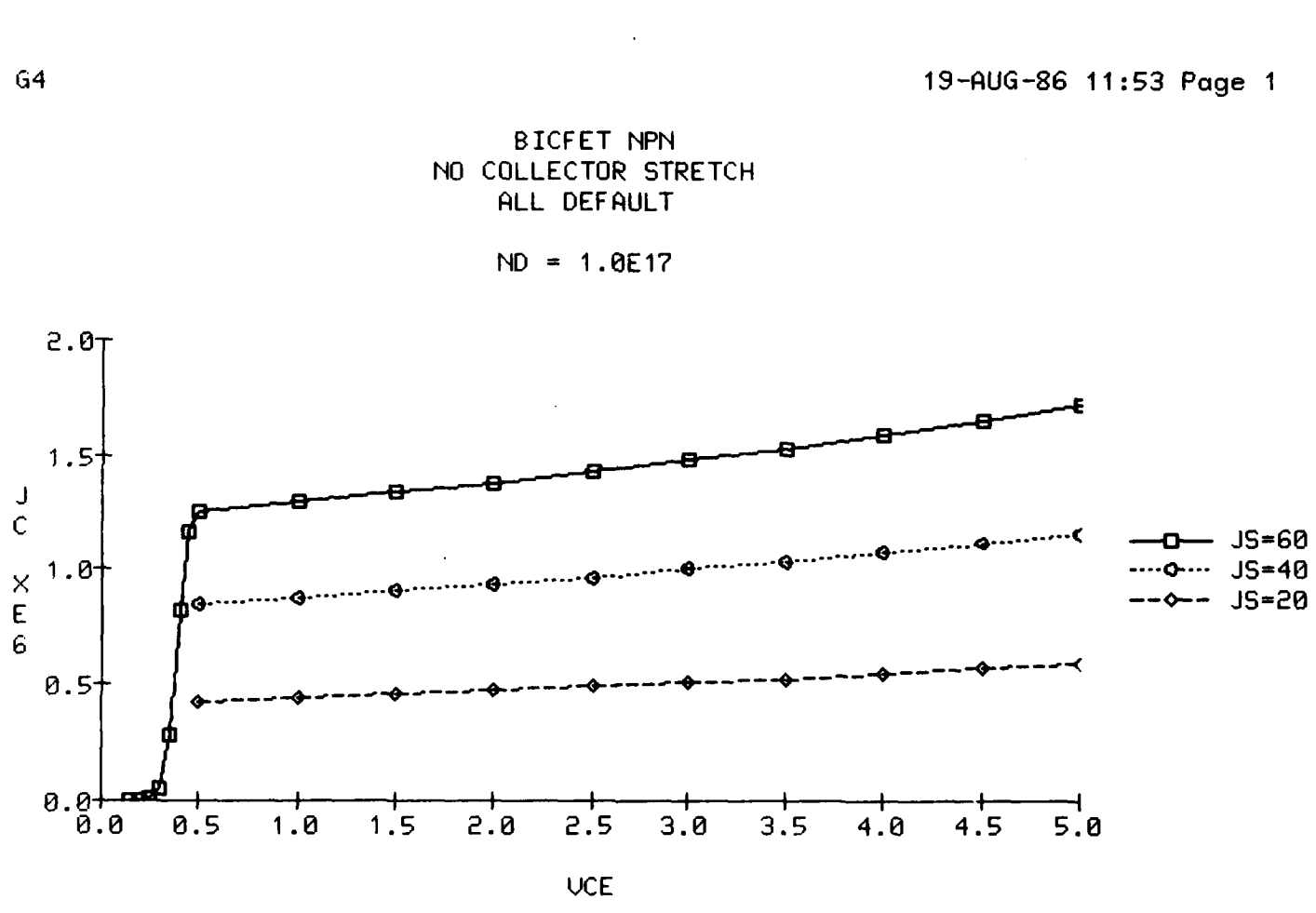


Figure 3: Characteristic curves for a NPN BICFET device with $N_d = 10^{17}$ cm^{-3} and no collector stretch. All default values were used including $N_i = 1.6 \times 10^{19}$ cm^{-3}. The collector current, J_c, has units of 10^6 A/cm^2 while the source current, J_s, is in A/cm^2.

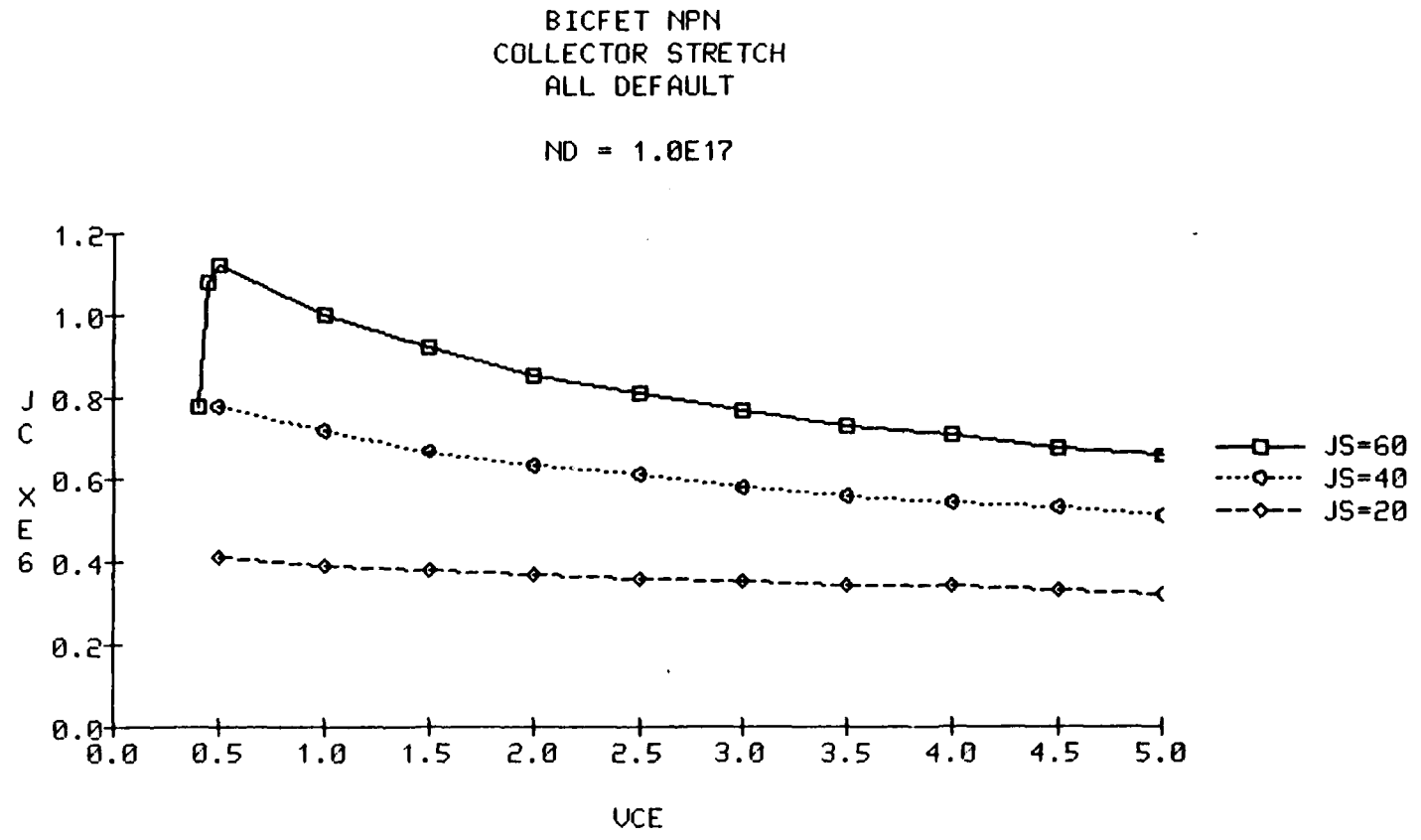


Figure 4: Characteristic curves for a NPN BICFET device with $N_d = 10^{17}$ cm^{-3} and collector stretch. All default values were used including $N_i = 1.6 \times 10^{19}$ cm^{-3}. The collector current, J_c, has units of 10^6 A/cm^2 while the source current, J_s, is in A/cm^2.

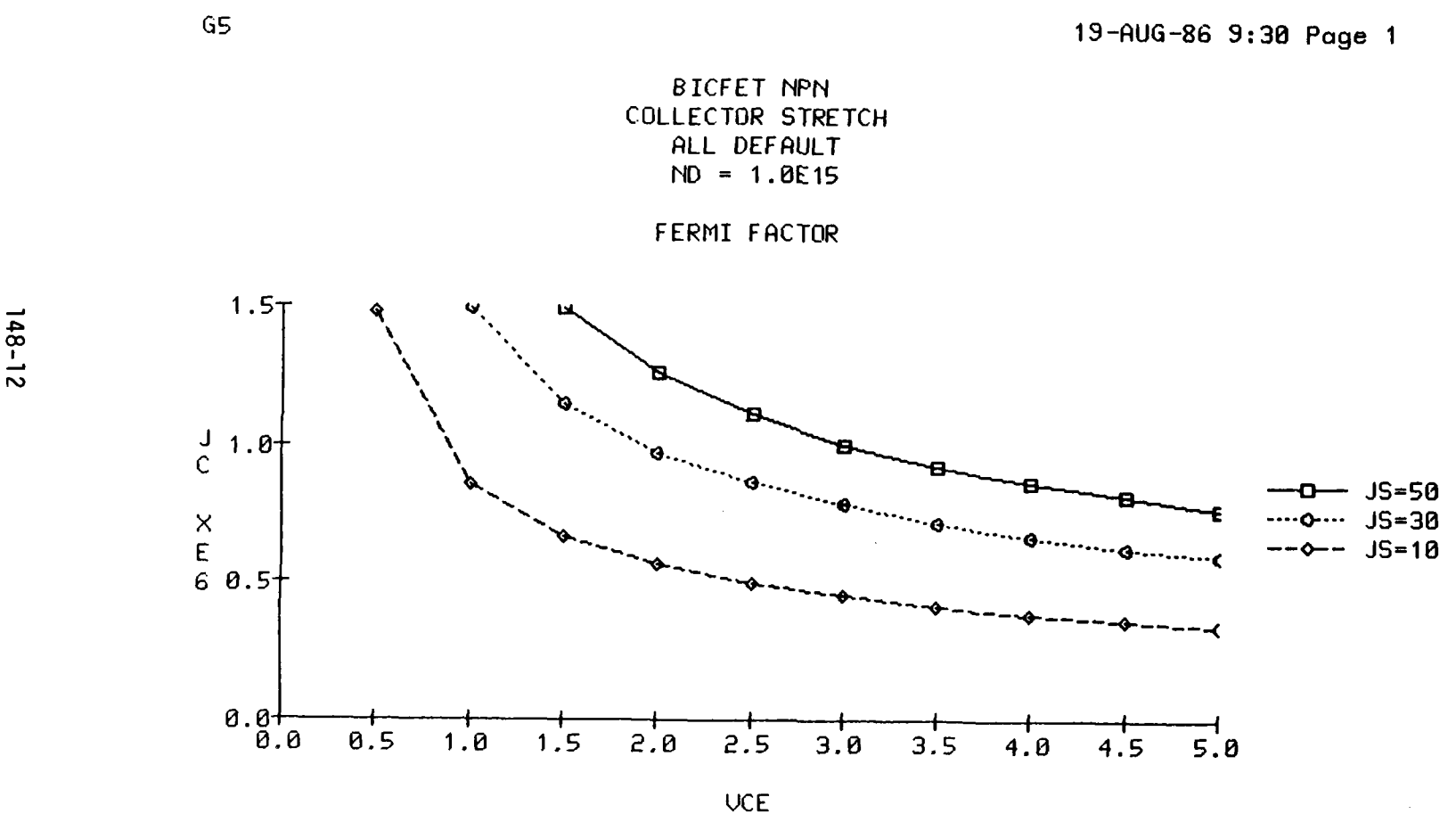


Figure 5: Characteristic curves for a NPN BICFET device with $N_d = 10^{15}$ cm^{-3}, collector stretch and Fermi Factor. All default values were used including $N_i = 1.6 \times 10^{19}$ cm^{-3}. The collector current, J_c, has units of 10^6 A/cm^2 while the source current, J_s, is in A/cm^2.

decrease in p_o. Because of charge balance, the positive inversion charge, p_o, will decrease as the negative spike layer charge ($-qN_i$) decreases. Table I shows this clearly. Note that the product, p_o G, remains constant. This indicates that changing N_i affects only p_o and not the exponential factor.

N_i	p_o	G
5.0 x 10^{19}	64 x 10^{19}	0.21 x 10^4
3.0	23	0.58
1.6	6.4	2.1
1.0	2.5	5.4
0.5	0.62	21
0.1	0.038	350
.05	0.017	800

p_o G = 1.3 x 10^{24}

TABLE I No Collector Stretch, All Defaults, $N_d=10^{15}$, $V_{ce}=1.0$, $J_s=40$

For large values of J_c there will be some distortion due to collector stretch as shown in Figure 5, however, where J_c is low enough (and collector stretch is negligible) substantial gain can be achieved by lowering N_i. There is a limit to the reduction of N_i. The value of N_i must be large enough to deplete the emitter region (the semi-insulator) and to reduce the recombination due to interface states.

When gain is found as a function of J_s, collector stretch is seen to be an important factor for $J_s \geq 20$ A/cm^2, while the Fermi factor gives a dramatically different result for all values of J_s.

The question of thermionic emission vs diffusion in the semi-insulator comes down to one of whether the ratio of effective velocities (V_n/V_p) in the gain equation (Equation 17) is much different than one. For typical silicon values this ratio is about 1.8. This effect is small compared to changes caused by other parameters.

For the discussion of the next three approximations it is useful to consider both NPN and PNP deivces. These approximations are : Φ_n = constant, ψ_{ms} = constant and $\Delta E_c = q(\chi_s - \chi_i)$ where χ_s and χ_i are the electron affinities for the semiconductor and semi-insulator, respectively.

The gain equations are:

$$\text{NPN:} \quad G = \frac{N_c}{p_o} e^{(\Delta E_v - \Phi_n)/KT}$$

$$\text{PNP:} \quad G = \frac{N_v}{n_o} e^{(\Delta E_c - \Phi_p)/KT} \tag{18}$$

where

$$\text{NPN:} \quad \Phi_n = KT \ln \frac{N_c}{N_{ds}} + \Psi_{ms} + q(\chi_s - \chi_i)$$

$$\text{PNP:} \quad \Phi_p = KT \ln \frac{N_v}{N_{AS}} - \Psi_{ms} + \Delta E_v + \Delta E_c - q(\chi_s - \chi_i) \tag{19}$$

where

$$E_c - E_F = KT \ln \frac{N_c}{N_{ds}}$$

$$E_F - E_v = KT \ln \frac{N_v}{N_{AS}} \tag{20}$$

where N_v is the effective density of states for the valence band, n_o is the inversion electron concentration at the interface, N_{ds} is the semiconductor donor concentration, N_{As} is the semiconductor acceptor concentration, and the gain equation for PNP is written down by analogy to that of the NPN.

If we assume that

$$\Delta E_c = q(X_s - X_i) \tag{21}$$

then

$$\text{NPN:} \quad G = \frac{N_{ds}}{p_o} e^{(\Delta E_v - \Delta E_c)/KT} \; e^{-\Psi_{ms}/KT}$$

$$\text{PNP:} \quad G = \frac{N_{AS}}{n_o} e^{(\Delta E_c - \Delta E_v)/KT} \; e^{\Psi_{ms}/KT} \tag{22}$$

For the NPN silicon device modeled in the computer program it was found that certain parameters did not change when ψ_{ms} and N_d were varied. In fact it was found that $p_o \cong 6 \times 10^{19}$, $V_{ce,cut\text{-}in} \cong 0.15$, $V_i \cong -0.08$,

$N_e \cong .12 \times 10^{12}$ cm^{-2}, $N_{cc} \cong 4.68 \times 10^{12}$ cm^{-2} where $N_i = N_e + N_{cc} = 4.8 \times 10^{12}$ cm^{-2}. N_i, N_e and N_{cc} are the charge densities for the spike layer, the emitter and the collector, respectively and V_{ce}, cut-in is the voltage where J_c goes thru zero as V_{ce} is lowered (see Figure 3). However, certain parameters changed dramatically (see Table II) with a change in ψ_{ms}. These results along with those in Table I indicate that the gain for an NPN device can be regarded as having two independent parts: p_o^{-1} and the exponential

Metal	ψ_{ms}	Φ_n	Gain	J_c
Default	-.32	.02	2.1×10^4	$.83 \times 10^6$
Aℓ	-.32	.047	7.5×10^3	$.30 \times 10^6$
n^+Poly	-.45	-.084	1.1×10^6	$.44 \times 10^6$
Au	.59	.96	4.8×10^{-12}	2.3×10^{-10}

Table II $N_{ds} = 10^{15}$ cm^{-3}

factor. In Table I it is seen that p_o changes with N_i but that the exponential factor does not. In the above it is seen that as ψ_{ms} changes the exponential factor changes but that p_o does not.

In order to more clearly define this property the gain equation can be rewritten by taking:

$$\psi_{ms} = 0.026 \ell n\, N_{ds} - B \tag{23}$$

where[4], experimentally:

$$\begin{aligned} A\ell: &\quad B = 1.218 \\ n^+poly: &\quad B = 1.348 \\ Au: &\quad B = 0.308 \end{aligned} \tag{24}$$

Thus, for a NPN device:

$$G = \frac{1}{p_o}\, e^{(\Delta E_v - \Delta E_c)/KT}\, e^{B/KT} \tag{25}$$

where, theoretically:

$$B = kT \ell n N_c + qX_s - \psi_m \tag{26}$$

Now, if we assume:

$$\Delta E_c \doteq q(X_s - X_i) \tag{27}$$

we have:

$$\begin{aligned} \text{NPN:} \quad & G = \frac{N_c}{p_o} e^{(\Delta E_v - \Psi_m + q\chi_i)/KT} \\ \text{PNP:} \quad & G = \frac{N_v}{n_o} e^{(\Psi_m - \Delta E_v - E_{gs} - q\chi_i)/KT} \end{aligned} \tag{28}$$

where

$$\Psi_{ms} = \Psi_m - \Psi_s = \Psi_m - (qX_s + E_c - E_F) \tag{29}$$

and equations (20) have been used. For the case of the GaAs/AlGaAs (semiconductor/semi-insulator) system we will take $qX_i \cong 4.5$ eV. Since we could not find any published values of X_i (electron affinity) for AlGaAs we have taken an average value (see, e.g., reference 6) for some compound semiconductors. From equations (28) it is obvious that one should minimize Ψ_m (metal work function) for a NPN device, but maximize Ψ_m for a PNP device; while ΔE_v and X_i should be maximum (minimum) for a NPN (PNP) device. Also, for a PNP device it would be good to pick a semiconductor with a smaller E_{gs} (energy gap).

For the NPN device let us choose x=0.4 ($Al_x Ga_{1-x}As$ where [7] $\Delta E_v = 0.436X$). Then

$$\Delta E_v + qX_i - \Psi_m \cong 4.67 - \Psi_m \tag{30}$$

and since there are many metals[8] (e.g. Ti: $\Psi_m = 3.92$, Sn: $\Psi_m = 4.51$) with work functions less than 4.67 the NPN has a good chance of being realized. For a PNP device with x=0.2 we have

$$\Psi_m - \Delta E_v - E_{gs} - qX_i \cong \Psi_m - 6.01 \tag{31}$$

Since there are very few metals (Pt: $\Psi_m \cong 6.37$) with a work function larger than 6.0 it would seem that it might be difficult to fabricate a successful PNP device. Thus, it appears that for the GaAs/AlGaAs system the NPN device has a better chance of working than the PNP device.

The final approximation ($V_i^* = -1.2$ and $\phi_s^* = 0.88$ where $\Psi_{ms} = \phi_s^* + V_i^*$) does not affect a conducting device but does affect the equilibrium (superscript asterick indicates equilibrium values) values of the emitter charge density, N_e^*, and the collector charge density,

N_{cc}^*. If the equlibrium values are calculated using a self-consistent (i.e. $N_i = N_e^* + N_{cc}^*$, etc.) iteration loop then p_o^*, N_e^*, and N_{cc}^* are not a function of N_d. (They do change values with the metalization chosen). However, if the default values are used the value of p_o^* is inversly proportional to N_d.

		P_o^*	N_e^*	N_{cc}^*
n^+	Aℓ	~2.5×10^{19}	1.77×10^{12}	3.03×10^{12}
	Poly	~2.2×10^{19}	1.96×10^{12}	2.84×10^{12}
	Au	~5.3×10^{19}	0.45×10^{12}	4.35×10^{12}

(a) Self consistent equilibrium values for $N_d = 10^{15}$, 10^{16}, and 10^{17}.

$N_d = 10^{15}$	2.75×10^{19}	1.77×10^{12}	3.08×10^{12}
$N_d = 10^{16}$	2.75×10^{18}	1.77×10^{12}	1.03×10^{12}
$N_d = 10^{17}$	2.75×10^{17}	1.77×10^{12}	1.12×10^{12}

(b) Default values used.

TABLE III Equilibrium values under various conditions

IV CONCLUSIONS

There are some conclusions that can be drawn from the analysis using this particular model.

1. Type of metal used is extremely important.
2. Experimental values of the electron affinity, X_i, of $Aℓ_x\ Ga_{1-x}As$ are needed as a function of x.
3. The "Fermi factor" may be extremely important.
4. If the "Fermi factor" is important then it may be useful to put the spike layer in the semi-insulator rather than the semiconductor.
5. Collector stretch may be negligible at realistic current density values.
6. For the GaAs/AℓGaAs system a functioning NPN device may be easier to fabricate than a PNP device.
7. There are two independent gain factor: the exponential agrument which depends on Ψ_{ms} and p_o which depends only on N_i.

8. The question as to whether the conduction process in the semi-insulator is analogous to thermionic conduction or diffusion is a relatively small factor.

VII RECOMMENDATIONS

1. Write program for GaAs/AℓGaAs system.
2. Write program for "diodes", i.e. $J_s = 0$ for NPN and PNP devices.
3. Build diodes and compare the experimental results with the above.
4. Measure values of the electron affinity, X_i, for $A\ell_x\ G_{a1-x}As$ as a function of x.
5. Investigate the possibility of experimentally achieving the spike layer in AℓGaAs.
6. Experimentally investigate the ohmic contact to AℓGaAs for both p and n type.
7. Pursue the "Fermi Factor" problem.
8. Build fully functional devices.
9. Investigate the temperature dependence of the BICFET both theoretically and experimentally.
10. Begin both experimental work and theoretical work on the DOES device.

REFERENCES

1. Taylor, G.W. and J. G. Simmons, "The Bipolar Inversion Channel Field-Effect Transistor (BICFET) - A new Field-Effect Solid-State Device: Theory and Structures", IEEE Transactions on Electron Devices, ED-32, 2345-67, Nov. 1985.
2. Taylor, G.W. and J.G. Simmons, "Small Signal Performance of the BICFET", IEEE Transactions on Electron Devices, ED-32, 2368-77, Nov. 1985.
3. Taylor, G.W. J.G. Simmons, A.Y. Cho and R.S. Mand, "A New Double Heterostructure Optoelectronic Switching Device Using Molecular Beam Epitaxy", J. Appl. Phys. 59(2), pp. 596-600, Jan. 15, 1986.
4. Deal, B.E., E.H. Snow, and C.A. Mead, "Barrier Energies in Metal-Silicon Dioxide-Silicon Structures", J. Phys. Chem. Solids, 27, pp. 1873-79, 1966.
5. Sze, S.M. Physics of Semiconductor Devices, 2nd Ed., pg. 122, Wiley-Interscience, 1981.
6. M.P. Shaw, "Properties of Junctions and Barriers", pg.5, Vol. 1, Ed. by William Paul, "Band Theory and Transport Properties", Series Ed. T.S. Moss, "Handbook on Semiconductors".
7. A.J. Hill and P.H. Ladbrooke, Electronics Letters, 22 (#4), 218 (1986).
8. A.F. Ioffe, "Physics of Semiconductors", pg. 230, Academic Press (1960).

1986 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/
GRADUATE STUDENT SUMMER SUPPORT PROGRAM

Sponsored by the
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
Conducted by the
Universal Energy Systems, Inc.
FINAL REPORT

WOMEN IN THE WORKFORCE: OCCUPATIONAL PULMONARY DISORDERS

Prepared by:	Shirley A. Williams
Academic Rank:	Assistant Professor of Biology
Department and University:	Department of Biology Jackson State University
Research Location:	USAF Occupational and Environmental Health Laboratory, USAFOEHL Command Section Brooks Air Force Base
USAF Researcher:	Colonel Bruce J. Poitrast
Date:	August 14, 1986
Contract No:	F49620-85-C-0013

WOMEN IN THE WORKFORCE: OCCUPATIONAL PULMONARY DISORDERS

by

Shirley A. Williams, Ph.D.

ABSTRACT

Although policies, scientific and legal effort have been directed toward limiting the adverse effects of the work environment on the individuals in the workforce, very little attention has been focused on the overall health status of women in the workforce. In addition, comparatively few cohort studies have been devoted to the description of health hazards experienced by women in non traditional as well as traditional occupations. This report describes some pulmonary disorders experienced by women in the workforce, with major emphasis on beryllium lung disease. Because of its elusive nature and degenerative effects on lung tissue as a result of physiological stress, I am submitting an experimental protocol as my mini-grant proposal to define the mechanism(s) of action of beryllium on pulmonary architecture and function.

Acknowledgments

I would like to express my appreciation to the Air Force Systems Command and the Air Force Office of Scientific Research for sponsorship of my research. Special tnanks is extended to Colonel James C. Rock, the entire Command Division Staff, Major William C. Keller, Mrs. Ann Potter, Captain R. Nakasone,and the Aeromedical Library Staff for their support and efforts to make my tenure at the Occupational and Environmental Health Laboratory a productive and meaningful experience. Finally, my deepest appreciation is given to Colonel Bruce J. Poitrast for his guidance of my research and insight to enhance and broaden my career goals and objectives.

I. Introduction

I received my Ph.D. from the University of Alabama at Birmingham where I studied the relationships of xanthurenic acid, pyridoxal phosphate and glucose metabolism in the rabbit. For this study I developed a high performance liquid chromatographic method to measure serum levels of xanthurenic acid in rabbits. Prior to this time, I studied as a research trainee at Oak Ridge National Laboratory, Oak Ridge, Tennessee, at Argonne National Laboratory, Argonne, Illinois, and at the University of Alabama at Birmingham in the Minority Hypertension Training Program.

The United States Air Force Occupational and Environmental Health Laboratory Consulting Division is interested in compiling information on occupational hazards encountered by women in the workforce. Comparatively few cohort studies have been devoted to the description of health hazards experienced by women in the workforce.

The problem under investigation is unlike any problem I have studied or undertaken at any of the research facilities mentioned. However, the opportunity to research and compile information of this nature has served to broaden my perspective of the significance of quality research and investigation of medical and health related problems.

II. Objectives of the Research Effort

The overall objective of studying the occupational hazards experienced by women in the workforce is to provide information to increase the awareness and incite participation of the various work sectors in finding remedies to prevent occupational hazards. This information will be compiled and disseminated to students and personnel of the School of Aerospace Medicine and Occupational and Environmental Health Laboratories.

My individual objectives were as follow:

1. Study the positions held by women in the workforce from 1890-1986.
2. Become familiar with the available sources of information regarding the legal dilemmas and health hazards encountered by women.

3. Study the pulmonary disorders experienced by women in the workforce.
4. Identify and design an experimental protocol to understand the cause of occupational pulmonary disorders.

III. Background

There are only a limited number of epidemiologic studies concerning the occupational hazards experienced by working women. On the whole, studies on traditional female jobs are not engineered or undertaken and cohort studies which investigate a mixed work-population of men and women, tend to exclude women because of small numbers.

This is paradoxical because with the emergence of industrial technology, and the mass production of chemical agents came the awareness that women worked and made essential contributions to the economy. If this support was suddenly cut off, the nations' economy would collapse (1).

History records that women have always participated to varying degrees in the workforce. The number of gainfully employed women in the U.S. increased from 4,005,532 in 1890 to 109,752,116 in 1930. As a result of the industrial revolution and the advent of World War II, women were employed in every sector of the workforce including mining and milling. As World War II escalated, the number of women in the workforce increased. Women were compelled to do so to insure national security and their own economic survival (2).

After World War II, women were virtually eliminated from mining, construction, transportation, and all other types of basic industry (3). Although only traditional female positions were being filled during this time, the number of women in the workforce increased astronomically to a total of 33.3 million during the decade 1960-1970. By 1980, over 45 million working women comprised 42% of the workforce. Fifty-two percent of all women were in the labor force by 1983 and it is estimated that the female workforce approaches 60% today (2,4,5).

The top ten professions for women post WWII were secretary, waitress, nurse, teacher, clerk, seamstress, sales, typist, bookkeeper, and private domestic worker (1,6,7).

The general consensus was that these positions lacked defined career structure and that promotion and prospects were limited--often

"prestige" within a given profession simply mirrored that of the immediate boss (5). In the 1960's women began to move into non-traditional "blue collar" positions such as carpentry, construction and plumbing, as well as the military and other non-traditional professions, such as, law, administration and medicine.

Unfortunately, even though these changes were socially positive, little consideration was given to the work environment, and the ways inwhich women differ from men, if at all, in their response to the work environment.

While Industrial technology was providing us with our present standard of living, it was also using and manufacturing chemicals without any real knowledge of their many actions and invariably their adverse--though unintended side-effects.

As early as the middle thirties, a series of deaths and illness in industry were associated with the use of recently introduced solvents. This focused attention on the rapidly expanding use of solvents in the industrial sector and the dangers associated with their use (8,9). Today we still have mass production of chemical substances with inadequate knowledge of their effects on humans. Other potentially hazardous conditions went unconsidered as well.

Some of these hazards include: stress and emotional illness encountered by managers and administrators; physical maladies suffered from the excessive lifting, bending and carrying found in the health professions; poorly designed work areas and monotonous tasks encountered by workers in clerical and assembly-line positions; physical agents such as noise and vibration in heavy industry or radiation in nuclear medicine and radiographic specialties. In addition, females who moved into predominately male fields found equipment inadequately designed for their size and safety and the non-existence of programs designed for adequate job training.

It is true that men suffer and encounter the same maladies as women in the workplace, however, women's illnesses are usually precipitated after a shorter period of exposure time. A contributing factor to this phenomena may be the ratio of body fat per kg body mass. Since women have a greater proportion of body fat (20-25%) than men (10-15%),

organic solvents which are lipophilic tend to be taken-up in greater quantity and tend to be stored over a longer period of time(10,11).

In spite of contentions of previous years which suggest that there are no differences in the susceptibility in normal women, other than during pregnancy, from men under the same conditions (12), a number of reports conclude that women are indeed more susceptible than men to the effects of exposure to toxic substances, which most often include metals and organic solvents. Ferguson (10) as cited by Messite (13) proposes that where differences in susceptibility have been suggested there is some correlation with the rate of metabolism by liver microsomal enzymes of the parent compounds and the relative toxicity of the parent substance and its metabolites. He suggested that the biotransformation of toxicants is influenced by sex hormones, and that where a metabolite of a chemical is more toxic than the parent compound, then males, with a more rapid rate of liver metabolism, are more likely to be more susceptible. On the other hand, when the parent compound is more toxic than the metabolite, females are more susceptible, since estrogens have a relatively lesser stimulating effect on liver metabolism than androgens. In support of this viewpoint, Hirakawa (14), reported a greater sensitivity of female rabbits than males to the effects of benzene, and the reduction of the relative resistance of males after castration.

Virtually all policies, scientific and legal effort have been directed toward limiting the adverse effects of the work environment on the reproductive capacities of women. Greater attention should be focused on the overall health status of women as well as men in the workforce.

IV. General Legislation

In 1918, the Women's Division, known as Women in Industry Service was established within the Department of Labor to develop policies affecting women's service in war production. The agency noted that available data indicated that a number of women were exposed to occupational conditions which were detrimental to health. Later, in 1920, Congress passed a bill and created the Women's Bureau in the Department of Labor to replace the Women in Industry Service. The Women's Bureau together with the Children's Bureau of the Department of

Labor was the voice of conscience in the industrial sector. The Bureau published maternal and infant mortality rates and continued the work started by the Women in Industry Service Division by formulating standards and policies to promote the welfare of working women (15).

Legislative guidelines formulated and enforced, subsequently to efforts generated by the Women's Bureau including; the Walsh-Healy Public Contract Act of 1936, the Occupational Safety and Health Act of 1970, Title VII of the Civil Rights Act of 1964 (1972), the Toxic Substance Control Act of 1976 and the Pregnancy Discrimination Act of 1978, provide for the general welfare, assure so far as possible every working man as well as woman in the nation safe and healthful working conditions, and preserve human resources (12,16,17). In addition, segregation or classification of an employee in any way which would deprive or tend to deprive any individual of employment opportunities or otnerwise adversely affect his/her status as an employee based on sex is declared illegal. Taken together, the federal statutes outlawed distinctions based on sex when sex is not a bonafide occupational qualification for the job (17-24).

V. An Occupational Hazard - Pulmonary Disorders

Exposure to chemical agents is obviously not confined to industry. Many exposures occur on the road and in the home (25).

There are three main portals of entry for chemical substances into the body--the skin, the gastrointestinal tract and the respiratory tract. Other routes of absorption, generally of less importance, include mucous membranes and open lesions. Absorption through the skin is possible with solvents having a high fat solvent capacity. The quantity of substances reaching the bloodstream via the skin is low compared to the quantity of substances transported through the large surface area of the lungs. In the industrial sense, absorption through the mouth and digestive tract is regarded as an accident rather than a regular exposure hazard (9).

Exposure to organic or non mineral dusts has been known to cause pulmonary disease for several centuries. A disease caused by cotton dust was first described by Ramazzini in his De Morbus Artificum, published in 1713 (26).

Although no inhaled particles are beneficial to the respiratory

system, inhaled air-borne substances in the workplace may or may not be injurious to the respiratory tract. For example, lead, which is capable of causing systemic disease causes no apparent local disease in the respiratory tract or lungs. On the other hand, other contaminants, may act within the respiratory system as primary irritants, sensitizers, fibrogenic agents, sources of ionizing radiation or as chemical carcinogens (27). To further accentuate their degenerative effects, some of the particles or contaminants may also be absorbed and produce systemic disease. In general, when particulates are inhaled, only those which are 5 microns in size or less reach the lungs and are capable of deposition in the airspaces. The larger particles are brushed up by ciliary action and are either coughed up or swallowed. It is also important that lung cells comprising the surfaces in contact with air normally have a rapid turnover or replacement rate, hence partially damaged surface cells are quickly replaced by new and normal cells.

As in other organ systems, however, the capacity for self-protection and repair of injury can be exceeded, and excessive dust or particulate deposition can cause adverse effects within the breathing apparatus. The four basic responses of the respiratory tract to the deposition of inhaled particles are: (1) immunologically induced airway constriction--Type I and Type III Reactions mediated by IgE and IgG immunoglobulins; (2) pharmacologically induced airway constriction--resulting from the release of serotonin and histamine; (3) acute irritation and reflex broncho-constriction associated with irritant gases and fumes, e.g., chlorine, ammonia, ozone, oxides of nitrogen; and (4) non-specific response to dust, excluding toxic properties and propensity of particles to generate an immune response.

Depending on the intrinsic chemical and physical nature of inhaled particles, the chemicals absorbed onto their surfaces, the duration of exposure and susceptibility of the host, the biologic response may be acute or chronic (28,29).

Acute respiratory symptoms of the upper respiratory tract are often due to regional inflammation, often perceived as irritation. With nasal and paranasal sinus irritation, there is congestion that can result in violent frontal headaches, nasal obstruction, runny nose, and occasional nosebleed. A dry cough usually accompanies throat inflammation.

Shortness of breath, hoarseness, cyanosis, spasms and anxiety--depending on severity--are associated with laryngeal inflammation.

In the mid-respiratory tract the acute reaction is characterized by bronchospasm--the extreme case being asthma (29). Pulmonary edema and pneumonitis develop from acute irritation of the deep respiratory tract. Primary pulmonary edema resulting in extravasation of fluid and cells from the capillary beds into the alveoli, is due to direct toxic action on the capillary walls. Exposure to ozone and oxides of nitrogen commonly precipitate this reaction. Pneumonitis, on the other hand, is an inflammation of the lung parenchyma in which cellular infiltration rather than fluid extravasation predominates. Beryllium and cadmium can cause acute pneumonitis. The most widespread causes of acute responses are irritant gases (29). Factors such as intensity and duration of exposure determine their site of action.

Pulmonary fibrosis is the most readily recognized work-related chronic pulmonary reaction. When the intensity and duration of exposure are too great to be defended against effectively, chronic intrapulmonary inflammatory and fibrotic changes cause distortions of pulmonary architecture leading to functional derangements (28). The two basic types of fibrosis occurring in the lung tissues are localized and nodular, usually peribronchial and diffuse interstitial fibrosis (29).

All of the fibrotic changes occuring in lung tissue, due to dusts are collectively termed pneumoconioses (27,29). The clinical features of all pnuemoconioses are similar exhibiting initial nonproductive cough, shortness of breath of increasing severity, and, in the later stages, productive cough, distant breath sounds, and signs of right heart failure.

The biological responses to different types of dusts vary with their physical properties, e.g., size, solubility, chemical and electrical properties, and whether they are inhaled singly or in combination with other dust particles and their concentrations and duration of exposure (27,28,29). Other factors which are important to the development and severity of the responses are the worker's age, past respiratory infections (especially tuberculosis), pulmonary reserve, the functional efficiency of the lungs and the presence of concurrent pulmonary disease (27).

The pneumoconioses may be associated with one of two etiologies--those arising from organic dust particles versus the induction from inorganic dust particles. Diseases of the respiratory tract which are attributable to exposure to organic dusts most often elicit immunologic reactions. The antigenic particles, are usually between 1 and 3 microns in diameter and thus are capable of penetrating into the alveoli (30). Inorganic dust deposition calls forth a simple foreign-body reaction which stimulates the recruitment and accumulation of macrophages in the area of deposition (27,28). Particles lodged on or just beneath the alveolar surface induce proliferative reactions and become overgrown by surface cells. These reactions can evolve to include cells such as lymphocytes, macrophages, and polymorphonuclear leucocytes. In addition, the fibroblasts present in the interstitium of the lungs may be stimulated to form excessive amounts of reticulum or collagen (28).

Many of the organic pneumoconioses have similar pathogenic mechanisms resulting in a characteristic clinical and pathologic syndrome. Common classification considers them hypersensitivity disorders, i.e., extrinsic allergic alveolitis (31). The disorders are generally recognized by names descriptive of their occupational or antigenic origin. Farmer's Lung and Byssinosis are the most prominent of these disorders and are discussed in some detail and serve as - prototypes for other organic pneumoconioses such as, Bagassosis, Suberosis, Malt Worker's Lung and Mushroom Workers' Lung (30,31).

The clinical manifestations of Farmer's Lung disease vary, depending on the extinct and frequency of exposure to clouds of dust containing fungal spores (found in hay due to heat of the decomposing vegetable compost), individual susceptibilities and coexistent pulmonary diseases. In most individuals affected or exposed to the mold spores, symptoms of dyspnea, chills, severe coughing, malaise, headache, and myalgia and fever occur 4 to 8 hours after exposure. The acute response possibly mimics pulmonary edema with respiratory failure. However, after repeated exposure, a chronic state develops which is characterized by permanent shortness of breath. Lung biopsy reveals extensive granulomatous interstitial pneumonitis. As the acute stage progresses, nodular densities creating a generalized granular pattern occurs. If the worker eliminates sources of exposure, a complete clearing occurs

gradually over a period of weeks or months. In more advanced and chronic stages, however, the lungs display a coarse, reticular pattern consistent with interstitial fibrosis.

Byssinosis is a respiratory disorder caused by exposure to the dust of cotton, flax and hemp. It is encountered primarily in factories where these fibers are combed and machines are cleaned during service operations. After several years of exposure, the disease is characterized by the gradual onset of symptoms typical of a syndrome called "Monday fever" (31). After a weekend off, the worker returns to a dusty atmosphere and experiences tightness in the chest, dyspnea and coughing which last the complete day. As time progresses, the individual experiences severe and irreversible ventilatory impairment, having permanent shortness of breath, cough, and phlegm. Byssinosis is asthmatic in nature, with bronchoconstriction possibly caused by the release of histamine.

Holness et al. (32) reported the effects of exposure to cotton dust, with attention to the mill fever syndrome on a healthy non smoking woman exposed for 6 hours to dust from a cotton felt producing facility. At 2 hours of exposure, the subject began coughing and sneezing, with myalgias and chills at 6 hours, fever, nausea, and headache at 12 hours. By 24 hours symptoms began to subside and by 60 hours they were gone. Noweir (33) in studying women in cotton gins noted the occurrence of Byssinosis in several workers. As with some other occupational illnesses, Noweir found that women were more affected by exposure than men.

Inorganic Pneumoconioses

The fibrotic changes occurring in the pneumoconioses caused by inorganic dust particles vary according to the particular kind of dust as well as the duration and intensity of exposure. Some of the more common particles encountered by women as well as men in the labor force are silica, asbestos, talc and beryllium.

Pneumoconioses in women have received little attention in the literature. However, a great number of women work in industries such as ceramics, textile mills, etc., that provide environments for exposure to toxic dusts. Silicosis, and asbestosis, are possibly the most prominent of lung diseases experienced by women in industry, especially in the

ceramic industry. Gerhardsson and Ahlmark (34) noted that among approximately 4,700 cases reported to the Swedish Pneumoconiosis Registry in the period 1931 through 1980, 53 cases were women with silicosis, 42 of whom had worked in the ceramic industry. In comparison to men, the prediagnosis duration of exposure to dust was significanlty shorter for the women (20.5±8.6 yr) than for men (28±10.1 yr). Roentgenographic evidence of progression of the lesions was also more pronounced in the women. These same reporters noted other investigators who also cited a greater incidence of silicosis occurring in women than men while working with pottery (e.g., finishing, polishing, decorating) and manufacturing scouring powder. In addition, the exposure time noted in these studies were considerably shorter in women as compared to the exposure time of men.

Silicosis is induced by chronic exposure and the deposition of crystalline silica--the only form which produces the characteristic nodular form of this pneumoconiosis--in the lungs (27,28,29). Epidemiologic studies suggest that crocidolite dust fibers cause more severe asbestosis than other forms of asbestos, e.g., chrysotile, amosite and anthophyllite, because dust particles containing crocidolite have a much larger number of very fine fibers per mg of dust (30,35-37).

Talcosis is induced by the inhalation of talc dust and develops gradually after 15 to 20 years exposure. As with silicosis and asbestosis, women tend to be more susceptible and vunerable to exposure of talc dust than men. Unlike silicosis and asbestosis, however, talcosis demonstrates massive fibrosis totally altering the normal architecture of the lungs (38).

Beryllium is one of the most extensively used metals today, because of its physical properties of low density and strength--properties which it confers on its compounds and alloys. The pure substance is used primarily in the aerospace industry as ideal structural material for space vehicles, space mirrors, satellite antennae, inertial guidance systems and airplane brakes (27,39,40)

During the mining, manufacturing and processing of beryllium, the substance becomes airborne as a dust or fume. The main routes of intake for men as well as animals are inhalation and ingestion (40,41). While the absorption of ingested beryllium is probably quite insignificant,

the chemical properties of the substance allow it to undergo transformation from soluble to insoluble forms in the lung thereby facilitating a greater retention time.

The exact biochemical mechanism(s) by which the metal produces its toxic effects, inspite of considerable investigative efforts, is still rather poorly understood. Extensive literature cites beryllium as a causative agent in producing systemic disorders and manifestations (39,42) that may involve almost any organ or tissue. It has an especially marked predilection for the lungs. As a result of the systemic disorders, it has been questioned whether beryllium should be included among the pneumoconioses. However, the fibrogenic properties, the chronic granulomatous changes, and marked interstitial infiltration of lung tissues, resulting from exposure, warrants its consideration as a type of pneumoconiosis (27,32,42). In addition, some forms of beryllium disease resemble clinically, roentgenographically, and pathologically other granulomatous diseases such as sarcoidosis and tuberculosis (42).

There are two forms of the disease, an acute pneumonitis and a chronic granulomatous and fibrotic lung reaction. The lung reaction is dependent upon the dose of beryllium inhaled. While beryllium may be found in lung tissues, its presence is not essential for diagnosis, particularly of the chronic disease. Lung clearance of the offending exposure may take place prior to the time of examination (30). The National Beryllium Case Registry has reported chronic cases where manifestations of the disease occurred up to 25 years after exposure (40).

In the acute disease, the appearance is essentially that of non-specific acute and subacute bronchitis and pneumonitis with irritated respiratory mucosa, coughing and moderate breathlessness which may occur within 72 hr of heavy exposure (30,42). The alveolar septa are sometimes widened by interstitial edema which may occassionally be massive. In addition, in the acute form, there is fever, cyanosis, tachycardia, and tachypnea with rales throughout the lungs. Most patients contracting acute beryllium disease recover (42). It should be noted that most exposures linked to the acute form of the disease generally are to less-soluble aerosols (39).

Processes linked with chronic pulmonary disease are those processes where beryllium oxide is either developed or utilized, e.g., chemical extraction of beryllium from ore, calcination of beryllium hydroxide, handling beryllium powder, as well as formulating and using beryllium for ceramics (40). In some instances (6.2% of all cases) as reported from Registry data (43), acute beryllium disease is followed by chronic illness with and without further exposure. The onset of chronic illness may be insidious with only a slight cough and fatigue. It may occur 1 or 2 years after exposure, or as stated earlier, after a longer latency period lasting 5, 10 or even up to 25 years. However, when treated with steroids, no further progress of symptoms may be noted (42,43,44). In other cases, workers experience progressive pulmonary insufficiency, dyspnea on exertion and at rest, weight loss, anorexia, weakness, burning chest pain and a constant, non productive hacking cough. Cyanosis and clubbing of the fingers will be seen in about 30-40% of all cases, and due to increasing pulmonary fibrosis and resistance of pulmonary circulation, cor pulmonale is another frequent feature.

Pulmonary pathology in the chronic disease is characterized by granulomatous changes and interstitial pneumonitis. Frequently calcium metabolism is altered and serum calcium levels are elevated as manifested by the occurrence of hypercalcemia, calcification in the necrotic centers of nodular foci found in the pulmonary parechyma and hilar nodes (40,42).

Even when exposed to excessively high levels of beryllium oxide or beryllium salts, only 1-4% of exposed individuals actually acquire the disease, with a greater incidence of the chronic disease occurring in women (45). With only a small percentage of individuals contracting chronic beryllim disease, it is suggested that individuals possibly have a marked predisposition or susceptibility for the disease. For example, Hardy (46) cited the case of a 29 year old woman dying after an emergency appendectomy in a tuberculosis sanatarium where she had been a patient for some weeks. After the woman's death, her medical history was searched. She had worked in a fluorescent lamp factory. Three other women from the same flourescent-lamp manufacturing plant were under medical care for the abrupt onset of weight loss, cough and dsypnea. This was later followed by 17 cases with similar clinical

histories, chest X-ray findings and prognosis. Within 3-4 years, six of the 17 women died. After much investigation, the deaths and illnesses of the individuals were attributed to beryllium poisoning.

In a neighborhood case study of beryllium toxicity, pregnancy was associated in time with the onset of beryllium poisioning in 18 of 28 females. The high incidence of mortality in these persons suggested that some individuals represent cases where a special host response occurred, perhaps a genetically determined failure in defense (47). Incidence such as these generated considerable concern and efforts to elucidate the cause for the onset and fatality of the chronic disease.

Sterner (48) in 1951, postulated that beryllium disease is induced by two general types of reactions: (1) a regular, primary irritant chemical intoxication and (2) a modified immunologic response in which beryllium is the specific allergen. In some instances, Sterner suggested that both mechanisms may be involved. To substantiate his hypothesis, the investigator noted a tremendous difference in the reactions of individuals at given exposure levels--some individuals with exposures very small, both as to concentration and as to time, contract severe or fatal diseases, while others under similar or much greater exposures show no sign of injury. This pattern is not consistent with tnat of other chronic industrial intoxications. This resulted in the postulation by Sterner that beryllium disease could be explained on an "allergic" basis. The persistence of inflammatory reaction many days after initial application and the formation of granulomas along with this persistence, further substantiated the authors claim.

In subsequent decades, the pathogenesis of chronic beryllium lung disease has not been defined beyond the suggestions of Sterner and is still poorly understood. Deodhar et al. (49) continuing in the same frame of reference as Sterner (48), suggested that immunologic mechanisms play an important role in chronic beryllium disease of the lungs. Their invitro studies of blast transformation of lymphocytes in the presence of beryllium sulfate provided evidence for the existence of cellular immune reactivity to beryllium in patients with the chronic disease and further suggested the participation of these mechanisms in the pathogenesis of the disease.

Clary and Stokinger (50) proposed a mechanism for the induction of beryllium disease inwhich some forms of stress, e.g., surgery, infection, etc., caused beryllium translocation, lysosomal instability, and rupture and consequent cell death, resulting in the onset of chronic disease. The authors suggested that the long latent period observed in the onset of disease in the industrial worker may be based on the accumulation of beryllium in a latent form in the lysosome. This latent period could persist for years until the general decline in adrenal output, due to aging, combined with an adrenal stress, decreases hormone levels sufficiently to destabilize cell membranes leading to lysosomal rupture.

Although it is suggested that physiologic stress situations such as surgery or pregnancy increase the vunerability of an individual to beryllium exposure and thus the chronic condition (40), much investigative research is still necessary to establish the actual mechanism(s) involved in the development of the disease.

VI. Recommendations

1. Industries and businesses manufacturing and utilizing toxic substances should familiarize workers with possible health hazards associated with handling these agents.
2. Workers should avail themselves of information of their work environment and take the necessary precautions to protect themselves.
3. Efforts to define the effects of the various toxic chemicals on humans should be continued by the scientific community inorder to provide means to safeguard the workers' environment.
4. Since women seem more susceptible to most toxic substances, they should evaluate and assess the risk factors and the advantages of a particular occupation, regardless of laws providing the right and opportunity to work.
5. Definitive experiments must be designed to elucidate the mechanism(s) of action of beryllium on pulmonary architecture and function.
6. As a possible participant of the mini-grant program, I propose to design experimental protocol to investigate the degenerative effects of beryllium in the lungs as a result of physiological stress.

References

1. Stellman, Jeanne M., Women's Work, Women's Health: Myths and Realities, New York, New York, Pantheon Books-Random House, 1977, p. 3.
2. Hunt, Vilma R., Work and the Health of Women, Boca Rouge, Florida, Press, Inc., 1979, pp. 1-3.
3. Conibear, Shirley, "Women as a High-Risk", Proc.: Conference on Women and the Workplace, SOEH, Washington, D.C., 1977, p. 168.
4. Vena, John E. et al., Morality of a Municipal Worker Cohort: II Females, Amer. J. Indus. Med., 9, 1986, pp. 159-169.
5. Cox, S., T. Cox, and J. Steventon, Women at Work: Summary and Overview, Ergonomics, 27, 1984, pp. 597-605.
6. Jougla, E. et al., Health and Employment of a Female Population in an Urban Area, Inter. J. Epidem., 12, 1983, pp 67-76.
7. Colin, J.M. and C.M. Bishop, Occupational Health of Women at Work: Some Human-Factors Consideration, Ergonomics, 27, 1984, pp. 489-498.
8. Legator, M.S. and S.J. Rinkus, "The Chemical Environment and Mutagenesis", Proc.: Conference on Women and the Workplace, SOEH, Washington, D.C., 1977, p.5.
9. Browning, Ethel, Toxic Solvents, London, England, Edward Arnold & Co., 1953, pp. v-vi.
10. Ferguson, D.M., "Toxicological Problems Related to the Employment of Women", Health of Women at Work, Proceedings of Symposium of Society of Occupational Medicine Research Panel, London, England, 1976, pp. 41-52.
11. Frumkin, Howard, "Toxins and Their Effects, Occupational Health: Recognizing and Preventing Work-Related Disease, B.S. Levy and D.H. Wegman (eds.) , Boston, Massachusetts, 1983, pp. 131-134.
12. Kronenberg, M.H., Working Conditions for Female Employees, J. Amer. Med. Assoc. 124, 1944, pp. 677-683.
13. Messite, J. and M.B. Bond, "Occupational Health Considerations for Women at Work", Developments In Occupational Medicine, Carl Zenz (ed.), Chicago, Illinois, Year Book Medical Publishers, Inc., 1980, p. 43.
14. Hirakawa, T., Some Observations on Benzene Poisoning, Arch. Mal. Prof., 21, 1960, pp. 46-49.

15.Kaplan, M. and S. Knutson, Systems Approach to the Study of Women in the Workplace, Madison, Wisconsin,State of Wisconsin Department of Industry and Labor and Human Relations, 1977, pp. 677-683.

16.McCormick, W.E., "Legislation and Legislative Trends", Patty's Industrial Hygiene and Toxicology, G.D. Clayton and F.E. Clayton (eds.), New York, New York, John Wiley & Sons, 1, 1978, p. 24.

17.Alexander, Mary, "The Right of Workers to Refuse to Work Under Hazardous Conditions: Impact of the Whirlpool Supreme Court Case", Legal and Ethical Dilemmas in Occupational Health, J.S. Lee and W.N. Rom (eds.), Ann Arbor, Michigan, Ann Arbor Science Publishers, 1982, p. 24.

18.Trebilock, A.M., OSHA and Equal Employment Opportunity Laws for Women, Prev. Med., 7, 1978, pp. 372-384.

19."Equal Employment Opportunity Commission, Chapter XIV", Code of Federal Regulations, Office of the Federal Register, National Archives and Records Services and General Services Administration, 29, 1984, p. 177.

20.Brunt, M. and A. Hricko, "Problems Faced by Women Workers", Occupational Health: Recognizing and Preventing Work-Related Disease, B.S. Levy and D.H. Wegman (eds.), Boston, Massachusetts, 1983, pp. 403-416.

21."Equal Employment Opportunity Commission, Chapter XIV", Code of Federal Regulations, Office of the Federal Register, National Archives and Records Services and General Services Administration, 29, 1984, pp. 138-143.

22.Lucas-Wallace, K., "Legal Considerations Bearing on the Health and Employment of Women Workers", Work and the Health of Women, V.R. Hunt (ed.), Boca, Raton, Florida, CRC Press, Inc., 1979, pp. 181-199.

23.Barnard, R.C., Legal Issues: Reproductive Health Hazards, Ann. Am. Conf. Gov. Ind. Hyg., 3, 1982, pp. 157-163.

24.Whorton, D.M., "Considerations about Reproductive Hazards", Legal and Ethical Dilemmas in Occupational Health, J.S. Lee and W.N. Rom (eds.), Ann Arbor, Michigan, Ann Arbor Science Publishers, 1982, p. 399.

25.Taylor, R.A., A Growing Concern with Indoor Air, EPA Journal, 2, 1985, p.20.

26. Rylander, Ragnar, Organic dusts and Lung Reactions-Exposure Characteristics and Mechanisms for Disease, Scand. J. Work Environ. Health, 11, 1985, pp. 199-206.
27. Mayers, M.R., Occupational Health: Hazards of the Work Environment, Baltimore, Maryland, The Williams and Wilkins Company, 1969, p. 223.
28. Wright, G.W., "The Pulmonary Effects of Inhaled Inorganic Dust", Patty's Industrial Hygiene and Toxicology, G.D. Clayton and F.E. Clayton (eds.), Baltimore, Maryland, John Wiley & Sons, 1978, p. 165.
29. Wegman, D.H., "Respiratory Disorders", Occupational Health: Recognizing and Preventing Work-Related Disease, B.S. Levy and D.H. Wegman (eds.), Boston, Massachusetts, Little Brown and Company, 1983, p. 267.
30. Levy, S.A., "Occupational Pulmonary Disease", Occupational Medicine: Principles and Practical Applications, Carl Zenz (ed.), Chicago, Illinois, Year Book Medical Publishers, Inc., 1982, p. 139.
31. Smith, J.P. and R.W. Stone, "Occupational Pulmonary Disorders", Clinical Medicine for the Occupational Physician, M.H. Alderman and M.J. Hanley (eds.), New York, New York, Marcel Dekker, Inc., 1982, p. 422.
32. Holness, D.L., I.G., Taraschuk, and R.S. Goldstein, Acute Exposure to Cotton Dust, J. Amer. Med. Assoc., 247, 1982, pp. 1602-1603.
33. Noweir, M.H., Some Observations on Epidemiologic Studies in Egyptian Gins, Cotton Pressing Plants and Cottonseed Oil Extraction Plants, Chest, 79, 1981, pp. 15s-20s.
34. Gerhardsson, L. and A. Ahlmark, Silicosis in Women: Experience from the Swedish Pneumoconiosis Register, J. Occup. Med., 27, 1985, pp. 347-350.
35. Enterline, P.E. and V. Henderson, Type of Asbestos and Respiratory Cancer in the Asbestos Industry, Arch. Environ. Health, 27, 1973, pp. 312-317.
36. Gross, P. and D.C. Braun, "Asbestos", Toxic and Biomedical Effects of Fibers, Park Ridge, New Jersey, Noyes Publications, 1984, pp. 9-117.
37. Seilikoff, I.J., J. Churg, and E.C. Hammond, The Occurrence of Asbestosis among Insulation Workers in the United States, Ann. N.Y. Acad. Sci., 132, 1965, p. 139.

38.Gielec, L. and J. Osinska, Talcosis of the Lungs, Medycyna Pracy, 25, 1974, pp. 287-292.

39.Dickerson, O.B., "Metals and Metalloids: Antimony and Its Compounds", Occupational Medicine: Principles and Practical Applications, Carl Zenz (ed.), Chicago, Illinois, Year Book Medical Publishers, Inc., 1982, pp. 613 and 626.

40.Preuss, O.P., "Beryllium and Its Compounds", Occupational Medicine: Principles and Practical Applications, Carl Zenz (ed.), Chicago, Illinois, Year Book Medical Publishers, Inc., 1982, p. 619.

41.Health Assessment Document for Beryllium: Review Draft, Sprinfield, Virginia (U.S.) Evironmental Protection Agency, Research Triangle Park, N.C., U.S. Department of Commerce National Technical Information Service, 1984, p. iv.

42.Freiman, D.G. and H.L. Hardy, Beryllium Disease: The Relation of Pulmonary Pathology to Clinical Course and Prognosis Based on a Study of 130 Cases from the U.S. Beryllium Case Registry, Human Path., 1, 1970, pp. 25-44.

43.Hardy, H.L., E.M. Rabe and S. Lorch, United States Beryllium Case Registry (1956-1966): Review of its Methods and Utility, J. Occup. Med., 9, 1967, pp. 271-276.

44.Lurie, M.B., Immunological Aspects of Steroid Therapy, Arch. Environ. Health, 11, 1965, pp.235-241.

45.De Nardi, J.M., Long-Term Experience with Beryllium Disease, A.M.A. Arch. Ind. Health, 19, 1959, pp. 110-116.

46.Hardy, H.L., Beryllium Disease-Experience with Investigation Required to Establish Etiolotgy of Occupational Disease, Ann. N.Y. Acad. Sci.,107, 1963, pp. 525-538.

47.Hardy, H.L., Beryllium Poisoning-Lessons in Control of Man-Made Disease, New Eng. J. Med., 273, 1965, pp. 1188-1199.

48.Sterner, J.H. and M. Eisenbud, Epidemiology of Beryllium Intoxication, Ind. Hyg. and Occup. Med., 4, 1951, pp. 123-131.

49.Deodhar, S.D., B. Barna and H.S. Van Ordstrand, A Study of the Immunologic Aspects of Chronic Berylliosis, Chest, 63, 1973, p. 309-313.

50.Clary, J.J. and H.E. Stokinger, The Mechanism of Delayed Biologic Response Following Beryllium Exposure, J. Occup. Med., 15, 1973, pp. 255-259.

1986 USAF-UES SUMMER RESEARCH PROGRAM/

GRADUATE STUDENT SUMMER SUPPORT PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

UNIVERSAL ENERGY SYSTEMS, INC.

FINAL REPORT

MECHANISMS OF CHROMATIC CONTRAST

Prepared by:	B. R. Wooten
Academic Rank:	Professor
Department and	Psychology
University:	Brown University
Research Location:	Human Resources Laboratory
	Williams Air Force Base
USAF Researcher:	Dr. George Geri
Date:	October 9, 1986
Contract No.	F49620-85-C-0013

MECHANISMS OF CHROMATIC CONTRAST

by

Billy R. Wooten

ABSTRACT

The question of whether chromatic contrast is recurrent or non-recurrent was explored using psychophysical procedures. Stimuli were presented to four observers by a three-channel maxwellian view optical system. The target consisted of a four-element bulls-eye pattern. The central spot was varied in wavelength in order to determine each subject's unique yellow using a variety of color conditions for the several annuli. The basic experimental question was whether or not the contrast effect of a region of the visual field is influenced when its appearance is altered by contrast with another region of the visual field. Although more research is required, we tentatively conclude that contrast-induced color appearance has no effect on adjacent regions of the visual scene. Thus, chromatic contrast seems to behave like a non-recurrent network.

ACKNOWLEDGEMENTS

This report constitutes a summary of research conducted in the summer of 1986 at the Human Resources Laboratory, Williams AFB, Arizona. The author gratefully acknowledges support as a Summer Faculty Fellow sponsored by the U.S.A.F. Systems Command, Air Force Office of Scientific Research. Special thanks go to the staff at HRL for making my second visit as pleasant and productive as my first. I want to especially thank my Focal Point, Dr. Liz Martin, for providing the kind of warm and informal atmosphere that is so essential for good work. Dr. George Geri also deserves special gratitude for essentially turning his laboratory over to me for use during the 10 week visit.

Sections V and VI of this report describe experiments that would complete and extend the summer's work. It constitutes, in conjunction with the other Sections, a research proposal as part of the Mini Grant Program.

I. INTRODUCTION:

I received my graduate education in experimental psychology from Brown University with specialization in the sensory and perceptual aspects of color vision. Post-doctoral work was taken in the Biology Department at Harvard University where I studied the photochemical basis of color vision and color blindness. At Brandeis University, Freiburg University, and for the last 12 years at Brown University I have pursued numerous research projects in the area of color vision, color blindness, eye movements, and the optical properties of the human eye. My general area of interest could be described as the sensory, perceptual, and optical aspects of color appearance in simple and complex visual scenes.

The USAF is interested in all aspects of vision because of its obvious importance in pilot performance. At the Williams Air Force Base, where the primary interest is pilot training, a major effort is the development and refinement of flight simulators. A primary concern is accurate and realistic visual displays. Many issues remain unsolved concerning light levels, range of colors, stability of light sources, scattered light, etc. Several researchers at the HRL are working on various aspects, both applied and basic, of these questions. My background and interests overlap and supplement those of the HRL staff.

II. OBJECTIVES OF THE RESEARCH EFFORT:

Many factors in addition to the spectral composition of a light determine hue. Intensity, duration, image size, retinal location, and adaptation state all contribute. One of the most important parameters is the color of adjacent regions in the visual field. If, for example, a white circular test spot is surrounded by a red annulus it will appear tinged

with green. This effect is termed simultaneous chromatic contrast or chromatic induction. Since Chevreul's pioneering study, published in 1838(1), considerable work has been focussed on the phenomenon. In recent years neurophysiological techniques have supplemented the traditional psychophysical procedures.

On an empirical level, many basic aspects of the phenomenon have been explored. It is known that the induced hue is roughly complementary to the inducing hue, e.g., red induces green, yellow induces blue, red-blue induces green-yellow, etc. It is known that maximal chromatic induction occurs with minimal brightness differences. It is known that induction decreases sharply as the fields are separated; a 10 minute gap reduces the amount of induction by approximately 90 percent(2). It is known that the amount of induction increases as the inducing field is made larger.

Despite the long history of research on simultaneous chromatic contrast, the basic mechanisms are still poorly understood. It is clear that the perceptual effects are a manifestation of lateral interactions in the visual nervous system. The nature and level of these interactions, however, remain speculative. It is the goal of our research to explore certain perceptual aspects of chromatic indution that relate directly to the nature of the underlying neural processes. Specifically, we set out to see if chromatic contrast exhibits a phenomenon analogous to an effect termed "disinhibition," which is found in some neural networks. The existence of disinhibition, or the lack of it, has important implications concerning the basic organization and mechanisms that underly any lateral information processing system.

Disinhibition in the visual system was first demonstrated electrophysiologically by Hartline and Ratliff using the eye of the horseshoe crab, *Limulus*(3). They recorded from single optic nerve fibers while stimulating the retina with small spots of light. It was well

known that activity from a given retinal region, call it A, could be inhibited by stimulating a neighboring retinal region, call it B. Hartline and Ratliff showed that if a third retinal region (C), adjacent to B, was stimulated two effects could be observed: activity from B decreased, as expected, and activity from A *increased*. This latter effect is the phenomenon of disinhibition. The explanation is simply that the lateral inhibition of B onto A was decreased via C's lateral inhibition onto B. Similar effects can be observed in the considerably more complex vertebrate retina.

The importance of the demonstration of disinhibition lies not in the phenomenon itself, but in what it implies for the underlying neural substate. Figure 1, adapted from Ratliff, Hartline, and Miller(4), illustrates two fundamentally different kinds of lateral inhibitory systems. The recurrent type, shown on the left, is consistent with disinhibition. Notice that the lateral inhibition from each unit feeds onto its neighbors at or before the site X where impulses are generated. Thus, the inhibitory output of unit B depends upon its ultimate response, rather than just on the stimulus to it. The net effect is that if B is inhibited by a third unit, C, its inhibition onto A is decreased and A's activity is increased. Ratliff calls this recurrent or response-dependent inhibition and it clearly can account for the disinhibition effect. The non-recurrent type of inhibitory organization, shown on the right, is not consistent with disinhibition. Notice that the lateral inhibition from each unit feeds onto its neighbors *after* the site X where impulses are generated. In this scheme, the inhibitory output of unit B does not depend upon its ultimate response. Rather, it depends only upon the stimulus to it, i.e., the inhibition is non-recurrent or stimulus-dependent. Thus, if B is inhibited by a third unit, C, its inhibition onto A is unaffected and A's activity is unchanged. It is clear that this kind of non-recurrent inhibitory organization would not predict the disinhibition effect.

Ratliff summarizes why it is critical that a lateral inhibitory system be characterized as either recurrent or non-recurrent(5):

> The difference between these two forms of inhibition is important. ...in the recurrent system influences are not only exerted on neighbors, but also indirectly on others by way of those affected directly. The inhibition that one receptor exerts on another may thus be diminished by the activity of still other receptors that are in a position to inhibit the first. (from *Mach Bands*,, p. 131)

The non-recurrent system is quite different in that the inhibition that one receptor exerts on another is *not* diminished by other receptors.

The concept of recurrent versus non-recurrent systems, worked out so beautifully in the *Limulus* retina, can also be applied to the perceptual phenomenon of simultaneous chromatic contrast. The underlying physiology is, of course, immensely more complex involving many levels and huge numbers of neurones. Color appearance, however, is obviously in part determined by lateral interactions in the nervous system and the question of recurrent versus non-recurrent organization is thus a valid and important issue. Before any general model of color perception can hope to account for complex visual scenes, chromatic induction must be understood. And before chromatic induction can be understood, lateral effects must be characterized as either recurrent or non-recurrent. Our purpose was to explore how chromatic induction might be shown to be recurrent or non-recurrent.

III. APPROACH:

In terms of color perception, our basic experimental goal can be phrased as followed: Is the chromatic induction of a region B of the visual scent onto adjacent region A affected

when region B's hue is altered by induction from its neighboring region C? If the answer is "yes," chromatic induction can be considered to reflect an underlying recurrent organization. If the answer is "no," chromatic induction can be considered to reflect an underlying non-recurent organization.

Our approach can best be understood by considering a schematic of the actual stimulus arrangement as shown in Figure 2. It consists of a bullseye pattern where A is a circular test spot, B is an annulus surrounding A, and C is an annulus surrounding C. In principle, we first determine the wavelength that produces unique (or pure) yellow for A when seen alone. We then surround A with a B that induces a change in A, say a greenish B that causes A to be tinged with red. We then re-determine the wavelength that again produces a unique-yellow A, now with the B surround. This is our measure of the standard chromatic induction effect. Next we surround B with a C that induces a change in B, say a reddish C that makes B even greener due to chromatic induction. Then we again determine the wavelength for A that is required for unique yellow, now surrounded by B which is in turn surrounded by C. If induction is recurrent, surrounding B with C should have an effect upon A, i.e., B's change in appearance via induction should in tern affect A's appearance. If induction is non-recurrent, B's altered appearance should have no effect on A's appearance. Our response measure is the wavelength that is required for a unique-yellow A in the latter two conditions.

The apparatus used to generate the stimuli was a standard three-channel, maxwellian-view optical system that was constructed at the HRL during the summer of 1985. The details were given in last year's project report. Necessary modifications and calibrations required for the chromatic induction experiments were made in June and July of this year. Considerable effort was required to achieve photographic apertures that

allowed the precise registration of the bullseye's components. The outside diameters of fields A, B, and C were determined to be 1.04, 2.37, and 4.0 degrees of visual angle. The three fields were always flashed simultaneously for 0.5 sec every 15 sec. Retinal iluminance of the stimulus was 100 trolands.

The wavelength required to produce unique yellow for the various conditions was determined by a method of constant stimuli. Wavelengths were presented in quasi-random order and the subjects were required to respond either "red" or "green." For each session the wavelength spanned a region from clearly reddish yellow to clearly greenish yellow. With repeated presentations it was possible to generate enough responses at each wavelength so that a smooth curve could be drawn through the data to accurately estimate the 50 percent red-50 percent green point. That wavelength was defined as the one giving unique yellow.

IV. RESULTS:

The results for three subjects are shown in Figure 3. The number entered for each subject in each condition is the wavelength required for unique yellow. It is the mean of four determinations, each done on a different day. The number in parenthesis is the standard deviation. Means are based on eye-ball fits to the data. We are now in the process of re-determining these values based upon a computerized procedure. Preliminary results indicate that the informal fits are extremely accurate. We are also in the process of determining statistical significance based upon t-tests. Our conclusions at this point must be viewed as tentative, but it seems likely that the t-tests will confirm our judgment based upon the means & standard deviations. At the left of the figure the relevant stimulus arrangements are diagrammed. Although incompletely analyzed at this point and not presented, complete data from a fourth subject confirms our tentative conclusions.

A series of steps was necessary to build up to the full bullseye condition shown in figure 2. First, the unique-yellow wavelength for each subject was determined for A alone. The values for subjects SU, LI, and LA are shown in the row labelled "a," which refers to the stimulus condition. The values are fully consistent with those found in the literature for these conditions. The small standard deviation (0.8 to 1.8 nm) attest to the precision of the method. Condition "b" shows the result when unique-yellow was determined for the B field alone. As expected, they do not differ significantly from the "a" condition. The next condition, labelled "c," is the crucial one. We determined unique yellow for A with a unique-yellow B (from b condition) and a red C of 640 nm. The red C field did in fact induce green into the B field, a standard indiction effect. The unique-yellow wavelengths for A in the c condition are shown in the c row. Upon comparing them with the a condition, it is seen that the differences are small but consistent: in every case the c condition gives a *longer* wavelength for unique yellow. This would indicate that the red C field causes the A field to look slightly greener. Thus, to counter this the yellow locus is shifted slightly in the long-wave (red) direction. Note that this is in the wrong direction from what is expected if induction is recurrent. A recurrent system would require that the greenish (from induction) B field would induce redness and thus result in a shift in the short-wave (green) direction. A non-recurrent system would require, of course, no shift at all from the a result. (Control experiements d and e, not shown, demonstrated that the change in appearance of B in the c condition was sufficient to have caused a significant recurrent-type shift in the unique-yellow wavelength of A.)

The results from conditions a-e were totally unexpected and revealed that some aspect of our assumptions about the underlying mechansisms must be wrong. Ths most likely problem seemed to be our assumption that in the full bulleye stimulus, the c condition, C would have no direct effect upon A. This seemed reasonable in advance of our study since

it is well known that even a small gap between two fields wipes out almost all of an induction effect. The visual angel separating A and C in our stimulus is 46 minutes, extremely large in this context. In addition, it is filled with component B that we expected to also stop induction from C directly onto A. Walroven has shown, however, that some induction is transmitted across large gaps. This raises the possibility that these may be two contrast effects: a large one between contiguous fields and a small, but not trivial, one between even remote fields. For the sake of brevity, we refer to the former as the near-contrast-effect (NCE) and the latter as the far-contrast effect (FCE). Given this possibility our c results could be interpreted as follows: near contrast (NC) is non-recurrent, thus the green B field does not induce redness into A; far contrast (FC) from the red C field not only induces greeness into B, but also into A. Another view is, however, also equally plausible: NC is recurrent, but the induced redness from the green B is off-set by the slightly greater FC induction of greeness from the red C field.

From the point-of-view of deciding whether NC is recurrent or non-recurrent, it must be admitted that experiment c is ambiguous. The confounding influence of a possible FCE must be either eliminated or controlled. As a step towards sorting these factors apart, we decided to first determine whether or not the FCE is in fact a significant influence in our conditions. To do this we simply set up a condition where the B field was blocked leaving A and C with a gap between them. The results for this condition are shown in Figure 3 and labelled "f." Comparing the f results with the c results, it is seen that the directions of the shift are the same, i.e., the remote red C field does in fact induce greeness into A across a 46 min gap. In fact for two of the subjects the shift is larger in f than in c. The magnitude of shifts in the two conditions is difficult to interpret, but it is clear that in our conditions FC is potent and cannot be ignored.

We conclude from experiment f that the FCE complicates the meaning of experiment c. Thus far, we have not separated the effects of FC and NC. Thus, we cannot decide whether NC is recurrent or non-recurrent. Further experiments are required to settle this issue.

V. RECOMMENDATIONS:

Our experiments of last summer we regard as fruitful, but incomplete. They certainly helped us to focus on the nature of recurrent versus non-recurrent systems and they revealed the complicating factor of FC. In the time available we feel that we accomplished a great deal. A number of important aspects of the research, however, was left undone at the summer's end. We propose to use the Mini Grant to tie up the loose ends and begin some new studies that are needed to resolve the basic issues. In addition, we propose to use the time and support provided by the Mini Grant to formulate a major research proposal to be submitted to the USAF Office of Scientific Research. The plan would be an extension of the experiments and concepts generated from the summer-program support and the Mini-Grant Award.

Improvements on old experiments

A number of things need to be either completed or extended from the work outlined above:

1.) The data from the fourth subject will be analyzed.

2.) The entire body of data will be analyzed using computer curve fitting and statistical analysis.

3.) The experiments were all done at one luminance level. It would be of great interest to vary the luminance of field C over as large a range as possible.

The point is that this might give larger induction effects and therefore be easier to establish significant differences.

New Experiments

The major problem not addressed by last summer's work was the apparent confounding of NC and FC effects. Specifically, the results of c could be interpreted as resulting from either a non-recurrent NCE in conjunction with a small FCE or a recurrent NCE slightly out-weighed by a sizeable FCE. It is critical that these two choices be evaluated by experiment. This seems possible if we could set up a condition that would retain the size of the effect of FC on A, but eliminate any possible recurrent induction of B onto A. If the small long-wave shift observed in c is entirely from FC, i.e., NC is non-recurrent, then the new results should be the same as c. If, on the other hand, the c results originate from a balance between a recurreent NC (short-wave shift) and a slightly stronger FC (long-wave shift), then the new results (with NC eliminated) should show a much larger long-wave shift compared with c. It may not be possible to perfectly achieve this new condition but it can be closely approximated by creating a new field D that surrounds C and blocking C itself. This would create a gap between D and B which would almost totally destroy any recurrent NC by eliminating the green formerly induced from C onto B. FC would, however, be unaffected since a Walroven conclusively showed FC is unaffected by even relatively large gaps. Thus, one could test the predictions from the two possible interpretations just outlined. In this way, one should be able to settle the original issue of recurrent versus non-recurrent organization.

Other experiments could be done to further explore the mechanisms of contrast. Of particular interest would be conditions designed to examine the properties of the FCE. Very little work has been done on this topic beyond Walroven's initial study. For example,

the effects of FC and NC need to be carefully compared in order to explore whether they are fundamentally different mechanisms or merely one mechanism expressed to different degrees. It is important to establish, for example, that the FCE is not merely an artifact of chromatic adaptation from the scattered light.

The experiments outlined so far and proposed refer only to the red-green apparent system. It would also be interesting to explore the yellow-blue system by measuring the unique-green locus. We know that the two systems do not always behave in the same fashion. Thus, any general model of color perception must include characteristics of both systems. It would also be possible and worthwhile to examine the black-white mechanism from the same perspective of recurrent versus non-recurrent organization.

All of these experiments could be done on an existing apparatus in my laboratory at Brown University. The apparatus does need a new power supply and certain other minor items are needed as outlined in the next section.

VI. PROPOSED BUDGET FOR A MINI-GRANT:

The proposed budget is meant to fund the conclusion at the old experiments and the new studies for a 1 year period.

Principal Investigator time $9,600.

(During the academic year the PI will devote 20% of his time to the research at no cost to the grant; he will devote 100% of 2 summer months at 2/9 of his academic year salary.

Fringe Benefits $2,800.

Travel $1,200.

(For PI to attend scientific meetings; includes air fare and per diem)

Equipment $4,200.

(Power supply for apparatus at Brown is old, broken, and out-dated; this item is to purchase a new xenon power supply by Photochemical Associates)

Suppplies $1,000.

(1 replacement xenon lamp; neutral density filters; metal for mounts; plastic, etc.)

Graduate Assistant $1,200.

(2 summer months)

TOTAL $20,000.

REFERENCES

1.) Chevruel, M. (1838). *The principles of harmony and contrast of colours and their applications to the arts.* (Originally published in 1838. Translation by C. Martel, 1899, 3rd edn.) Bell, London.

2.) Walraven, J. (1973). Spatial characteristics of chromatic induction. *Vision Res. 13*, 1737-1753.

3.) Hartline, K and F. Ratliff (1957). Inhibitory interaction of receptor units in the eye of *Limulus*. *J. Gen. Physiol. 40*, 357-376.

4.) Ratliff, F., K. Hartline, and W. Miller (1963). Spatial and temporal aspects of retinal inhibitory interaction. *J. Opt. Soc. Am. 53*, 110-120.

5.) Ratliff, F. (1965). *Mach bands*, Holden-Day, San Francisco.

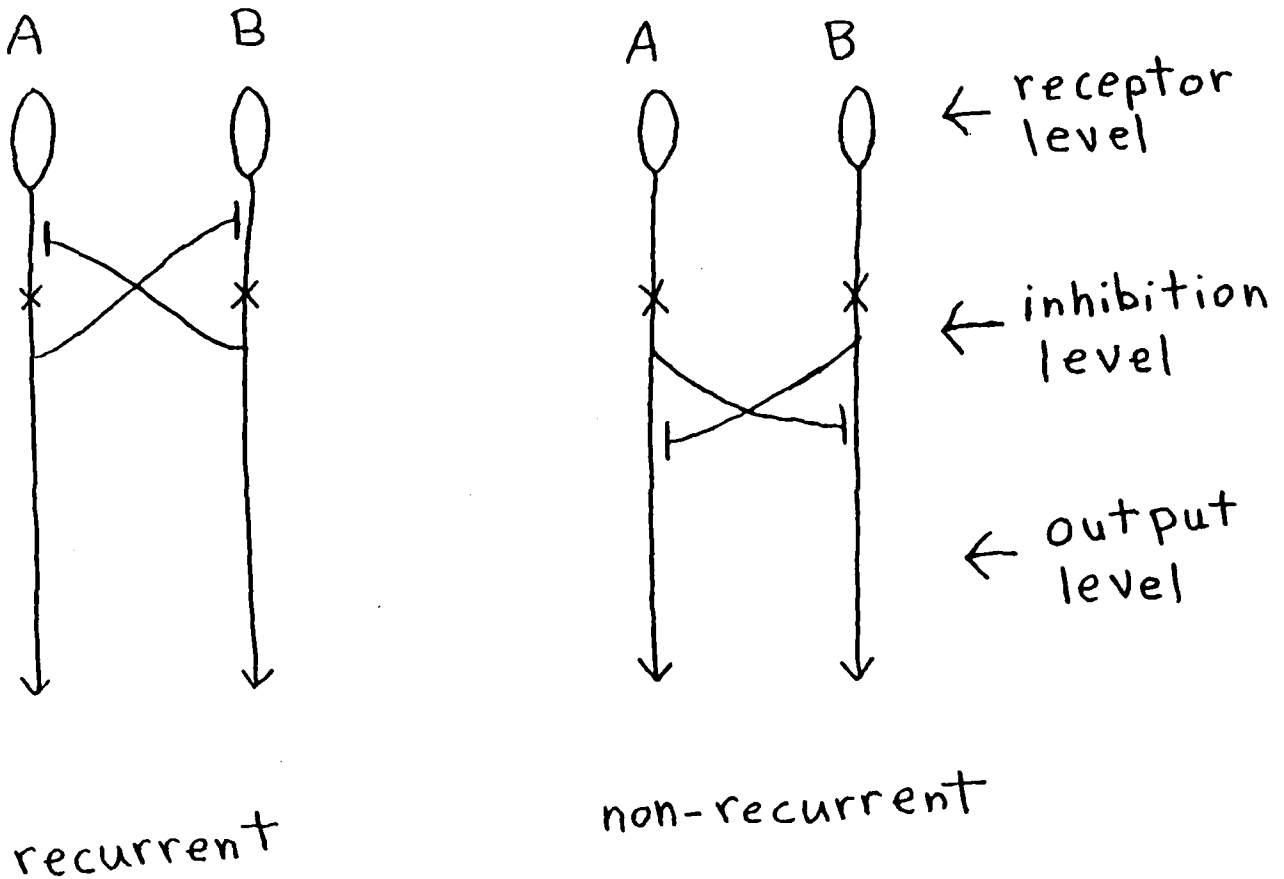


Fig. 1 see text for details

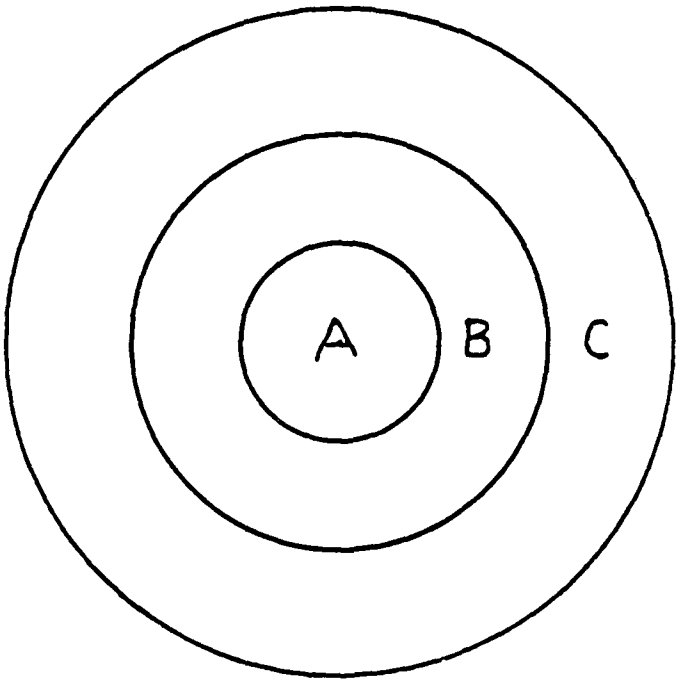


Fig. 2 see text for details

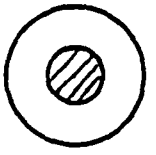
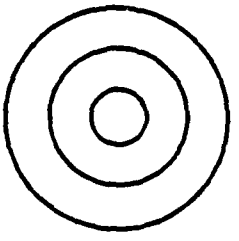
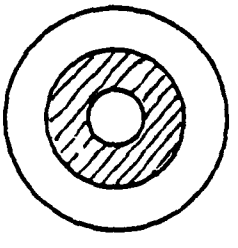
Stimulus		SU	LI	LA
a.		574.5 (1.8)	578.9 (1.27)	579.3 (.84)
b.		576.0 (2.04)	580.6 (1.49)	579.2 (.70)
c.		575.5 (1.69)	581.8 (1.02)	581.8 (2.27)
f.		582.0 (1.26)	581.1 (1.85)	584.9 (.92)

Fig. 3 see text for details

1986 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT SUMMER SUPPORT PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

Continuum Analysis of Low Pressure Tube and Nozzle Flows

Prepared by:	Daniel W. Yannitell
Academic Rank:	Associate Professor
Department and	Mechanical Engineering Department
University:	Louisiana State University
Research Location:	Air Force Rocket Propulsion Laboratory Edwards AFB, CA 93523
USAF Researcher:	David Weaver

Date:	September 15, 1986
Contract No:	F49620-85-C-0013

Continuum Analysis of Low Pressure Tube and Nozzle Flows

by

Daniel W. Yannitell

ABSTRACT

Fluid dynamics codes available on the Air Force Rocket Propulsion Laboratory computer system were evaluated as possible means to provide input data for Monte Carlo analysis of tube and nozzle lip area expansion flow. This is to be compared with experimental data to be taken at the Laboratory. The VNAP2 code was chosen for further study, and several numerical experiments were run to test the feasibility of its use for this project. It is concluded that careful use of the code should provide sufficiently accurate and detailed results for single specie gases.

ACKNOWLEDGEMENTS

I wish to thank the Air Force Systems Command and the Air Force Office of Scientific Research for making this project possible, and Universal Energy Systems for management of the program.

I also thank many people at the Air Force Rocket Propulsion Laboratory. These include David Weaver and David Campbell with whom I worked most closely, Yigal Kronzon who helped immesurably with the VNAP2 code and plotting routines, all the members of the Combustion Research branch, all the computer center personnel who made it easy to learn and work with a new system, the several members of the Tehachapi carpool who made the long commuting trip pleasant, and Wayne Roe, my effort focal point.

I also thank my family for its support and willingness to spend the summer away from home.

I. Introduction:

I have been teaching mechanics at Louisiana State University for several years. My research efforts have been concentrated primarily in the area of fluid dynamics. The principal investigators in charge of the project described herein have their primary expertise in analytical rarefied gas dynamics/free molecular flow and experimental diagnostics of combustion and gas flow. The addition of a continuum dynamicist to the team was deemed appropriate and I was fortunate to be chosen.

II. Objectives:

The long term goals of the Air Force Rocket Propulsion Laboratory project on which I worked include theoretical and experimental investigations of the rapidly expanding flow of gas at and near the lip of a nozzle which is exhausting into a vacuum (e.g. at very high altitudes or in space). The flow in this region must be accurately characterized in order to determine the potential of exhaust specie gas contamination of delicate optical or thermal surfaces on spacecraft.

It is well understood that there is a transition region, which extends outward from a nozzle lip, in which the exhaust gases expand rapidly from densities which permit the assumption of continuum flow to those which allow the assumption of free molecular flow. At present, the only theoretical tool which appears to be appropriate for analyzing this transition region is the Direct Simulation Monte Carlo method, and several studies using this method have been performed (see references 1 and 2). There remains, however, no direct comparison of

Monte Carlo results with detailed experimental data. The current project will attempt this comparison.

Experimental studies are to be performed at the Rocket Propulsion Laboratory in a medium size vacuum chamber. Because of size and pumping capacity limitations the nozzles tested must be small, and mass flow rates and temperatures are limited. In addition, it is desirable that the studies progress from the simple to the complex so that the effect of each complication can be assessed independently. The same argument applies to the theoretical and numerical studies.

My assignment as a USAF-UES Summer Faculty Research Fellow was to consult with the researchers in charge of this long term project and to determine how to provide the continuum flow start-line data necessary for input to the Monte Carlo analysis. This report describes the results of that assignment.

III. Project design:

It was decided that, in the interest of simplicity, the experimental investigations should begin with tests on a straight tube through which a monatomic gas at near room temperature is exhausted to a vacuum. In this way both the experiment and the analysis with which it will be compared are kept as simple as possible. Subsequent experiments could then add one additional complexity at a time allowing an informed assessment of the relative importance of each effect. These studies could include variations in nozzle shape, exit lip geometry, gas species, gas temperatures and, finally, realistic nozzle/exhaust gas combinations.

For the theoretical analysis, the Direct Simulation Monte Carlo Method developed by Graham Bird will be used, and the research team has already consulted with Dr. Bird and will continue to do so. This technique actually simulates the statistical nature of rarefied gas flow by modeling the random collisions between molecules. For this procedure to work, the computer code must be provided with inlet conditions or a "start line". At the inlet certain flow parameters such as velocities, pressures, temperatures must be known, and these are generally provided by continuum flow analysis which is valid where densities are appropriately high. As a gas expands and rarefies, the continuum hypothesis breaks down and the Monte Carlo simulation must take over. The remainder of this report deals with my attempts to determine the best available methods of conducting the continuum flow analysis to provide start-line conditions or parameters for the Monte Carlo code.

IV. Computer code evaluation:

The primary requirement for the continuum codes is the ability to provide the necessary accuracy and resolution of flow parameters. Three computer codes, readily available to Rocket Lab personnel, and structured to analyze nozzle flow, are CONTAM, SPP, and VNAP2 (references 3, 4, 5 and 6). Of these, the first two are extensive rocket codes of which the nozzle flow analysis is only a portion of the whole. That portion can, however, be used independently of the remaining parts, with some effort and experience. The nozzle flow portions are quite similar, consisting of finite difference

computation in the converging, subsonic inlet, switching to a method of characteristics scheme for the inviscid, supersonic core flow in the divergent outlet, followed by a separate analysis of the boundary layer along the nozzle wall. The codes' accuracy is suspect for small nozzles with thick boundary layers (for which an iteration with core flow analysis should be done) and for nozzles with small expansion ratios. In addition these codes cannot handle straight tube flow at all, because no supersonic region exists. For these reasons CONTAM and SPP were eliminated for the present studies.

The third computer code, VNAP2, is a finite difference code well structured for low expansion ratio nozzles and tube flow problems. Thus, although it has limitations such as being restricted to single specie gases with constant specific heat ratio (no chemistry) it was selected as the appropriate code for further study. This code, developed by Michael Cline, uses a straightforward finite difference scheme for interior grid points, and a careful analysis of the various possible types of boundary grid points. It is well suited for the various geometries and boundary conditions encountered in this project, the main disadvantages being those mentioned above. This should not hinder its use for the early stages of the project where single specie gases at relatively low temperatures will be used. It provides several useful options for boundary types, turbulence models, etc. Details of the mathematical formulation, options, and input/output formats are presented in the manual (ref. 6).

V. Numerical studies with VNAP2:

Several numerical experiments were performed using VNAP2 as installed on the main-frame computer at the AFRPL. Since the exact dimensions and conditions of the experiments had not yet been determined, representative situations were chosen for these numerical studies. First, small bore (~ 4 mm) straight tubes were considered, with room temperature argon as the flowing gas.

Studies were performed to assess the effect on the numerical analysis of such things as the grid geometry, inlet/outlet conditions, wall boundary conditions, choice of turbulence model, pressure range, etc. As the code solves steady state problems by iterating the time dependent equations until a steady solution is reached, an important factor in evaluating the results of the code is the manner in which the flow field approaches this state. Thus the variation of flow parameters with time step was carefully observed. The most important influence on this approach to steady state was found to be the grid geometry. Figures 1 and 2 illustrate the time variation of one parameter (the pressure at about the center of the tube) with different grid cell aspect ratios. For figure 1 the axial spacing of grid points was about 4.5 times the radial spacing, whereas for figure 2 the factor was about two. Other variables behaved similarly. It is obvious that this aspect ratio should be kept small.

Another criterion for acceptance of the results of the code can be mass conservation. As the code uses the primative variables rather than the conservation form of the equations, this is not automatically guaranteed. In all the tube cases run the mass flux varied by no more than two percent in the axial direction. This is believed to be

acceptable. For the cases examined, the flows should be laminar. The same problems were run, however, with input specifying a turbulence model. As hoped, the results differed very slightly, indicating that the code will produce laminar type profiles even if the investigator assumes turbulent flow.

Both adiabatic and isothermal wall boundary conditions were tried and both were successful, with only minor differences appearing when everything is initially at the same (room) temperature. However, if the wall temperature is maintained constant while the gas near the axis is cooling as it expands, a thermal boundary layer develops at the wall. Since conditions near the wall near the tube exit will probably be very important as input to the Monte Carlo analysis, it will be necessary to carefully design the experiments to know which boundary conditions are appropriate. Finally, the flow angle at the upstream inflow boundary was varied to see what effect experimental variations in tube inlet conditions might have. It was found that the influence of such variations seemed to die out by about eight to ten tube diameters along the axis.

The conclusion reached is that VNAP2 can be used, with care, to provide the detailed analysis of the continuum flow for the tube flow experiments. Figures 3 through 5 show some typical results.

The next set of numerical experiments was performed on small nozzles with low (< 10) expansion ratios. As no nozzle shape has been chosen, a simple circular arc inlet/conical convergent section/ circular arc throat/conical divergent section nozzle was chosen for these tests, and studies similar to those described for the tube flows were performed.

Again, the most important input affecting the dependability of the results appears to be the grid geometry. The aspect ratio of the computational cells must be kept small, prefereably below 2.5. Since the radial grid spacing varies automatically as the nozzle radius, care must be taken to vary the axial spacing appropriately. This is especially true if it is desired to refine the grid near the wall to provide enhanced resolution in the boundary layer. It may require more than one attempt to produce the desired resolution while maintaining reasonable cell aspect ratios. The "quick-solver" option of the VNAP2 code should help keep computer time down, but it was not investigated during these studies due to lack of time.

Some numerical results of nozzle tests are shown in figures 6 through 8. The gas was again argon with stagnation conditons set at room pressure and temperature, and the exhaust pressure set well below critical. It is apparent that these conditions must be changed for the experimental work as the predicted centerline exit temperatures are so low that condensation would probably occur. The profiles shown were produced using plotting routines developed by Dr. Yigal Kronzon at the AFRPL for use with the VNAP2 code. Although the nozzle problems done are only representative of what might be used in the project, the quality of the results indicates that reasonable solutions, from which Monte Carlo start-line data can be taken, can be obtained through use of the code.

The third and final type of problem attempted was a tube exit problem, using the "dual flow space" feature of VNAP2. This type of analysis might provide better resolution near the lip and additional comparison data for both the Monte Carlo and experimental investiga-

tions. Although at least one test seemed to be successful, others exhibited unexplained instabilities. This type problem needs further investigation.

V. Conclusions and recommendations:

In conclusion, it appears that VNAP2 can be used to analyze continuum gas flows in both tubes and low expansion ratio nozzles to provide start-line data for Monte Carlo investigations of flow at the exit plane, around the lip, and into the backflow region. The primary limitation for the long range project is the inability to analyze multispecie gases or variable specific heat ratio. The principal caveat is that care must be taken to keep the aspect ratio of the computation cells small (around 2 or smaller) to avoid very slow convergence with time and thus excessive computer costs.

REFERENCES

1. Heuser, J. E., L. T. Melfi, Jr., G. A. Bird and F. J. Brock, "Analysis of Large Solid Propellant Rocket Engine Exhaust Plumes Using the Direct Simulation Monte Carlo Method", AIAA 22nd Aerospace Sciences Meeting, Reno, Nevada, January 1984, AIAA-84-0496.

2. Heuser, J. E., L. T. Melfi, Jr., G. A. Bird and F. J. Brock, "Rocket Nozzle Lip Flow by Direct Simulation Monte Carlo Method", AIAA 20th Thermophysics Conference, Williamsburg, VA., June 1985, AIAA-85-0995.

3. Hoffman, R. J., M. A. Hetrick, Jr., G. R. Nickerson and F. J. Jarossy, "Plume Contamination Effects Prediction, CONTAM III", AFRPL TR-82-033, December 1982.

4. Hoffman, R. J. and A. H. Kawasaki, "Update Report: CONTAM 3.2 Thick Nozzle Boundary Layers and Exact Nozzle Shock", AFRPL TR-84-069, September 1984.

5. Nickerson, G. R., D. E. Coats, R. W. Hermsen and J. T. Lamberty, Jr., "A Computer Program for the Prediction of Solid Propellant Rocket Motor Performance (SPP)", AFRPL TR-83-036, September 1984.

6. Cline, Michael C., "VNAP2: A Computer Program for Computation of Two-Dimensional, Tune-Dependent, Compressible, Turbulent Flow", Los Alamos National Laboratory Report LA-8872, UC-32, August 1981.

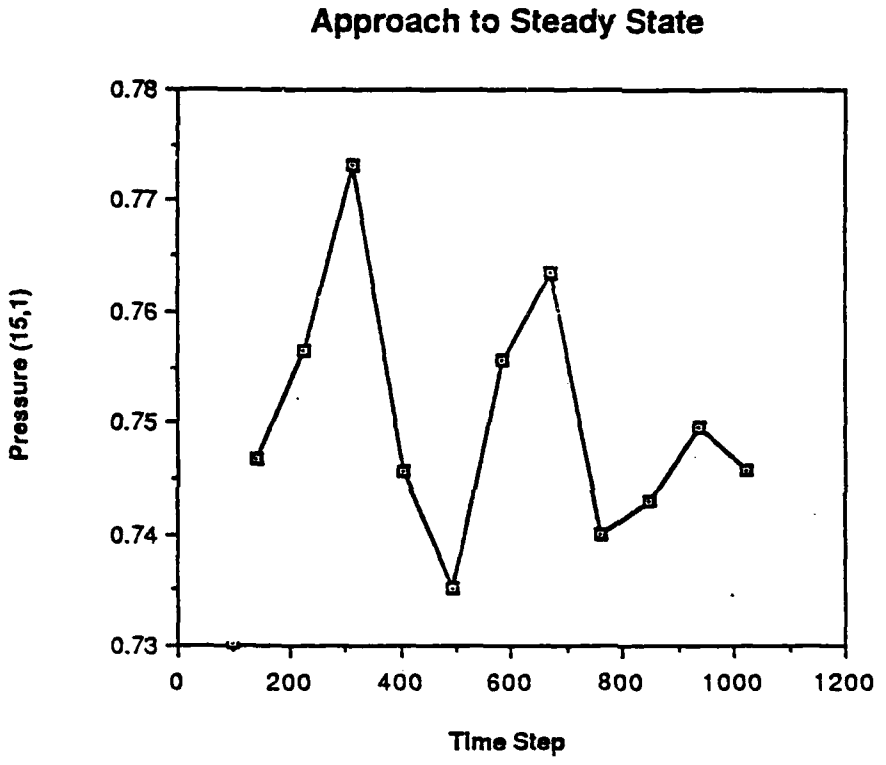


Figure 1. Convergence of pressure at tube center, AR=4.5

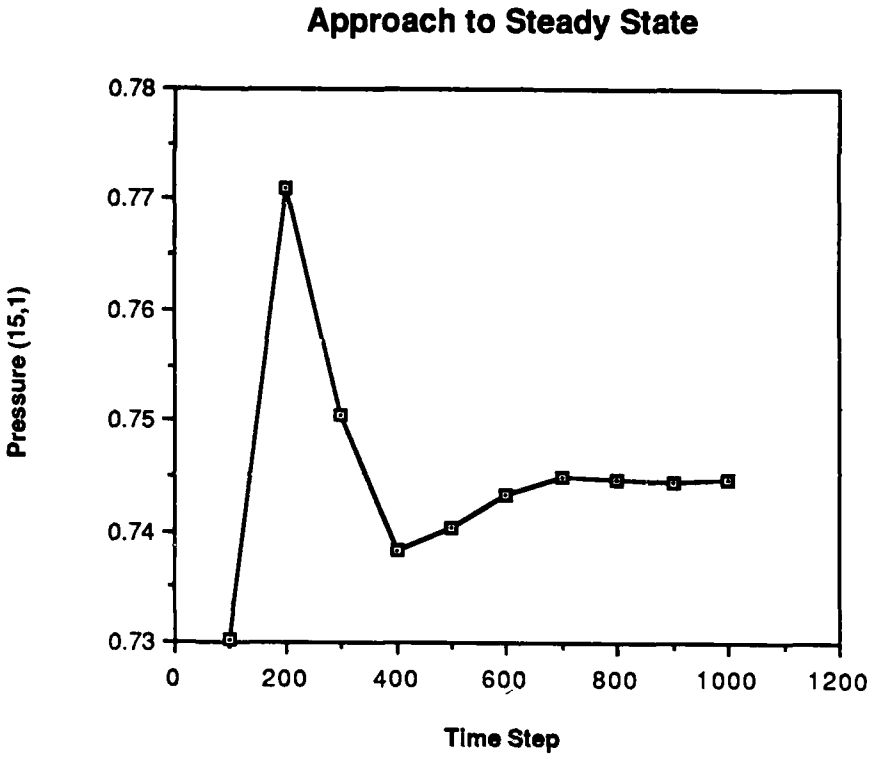


Figure 2. Convergence of pressure at tube center, Ar = 2

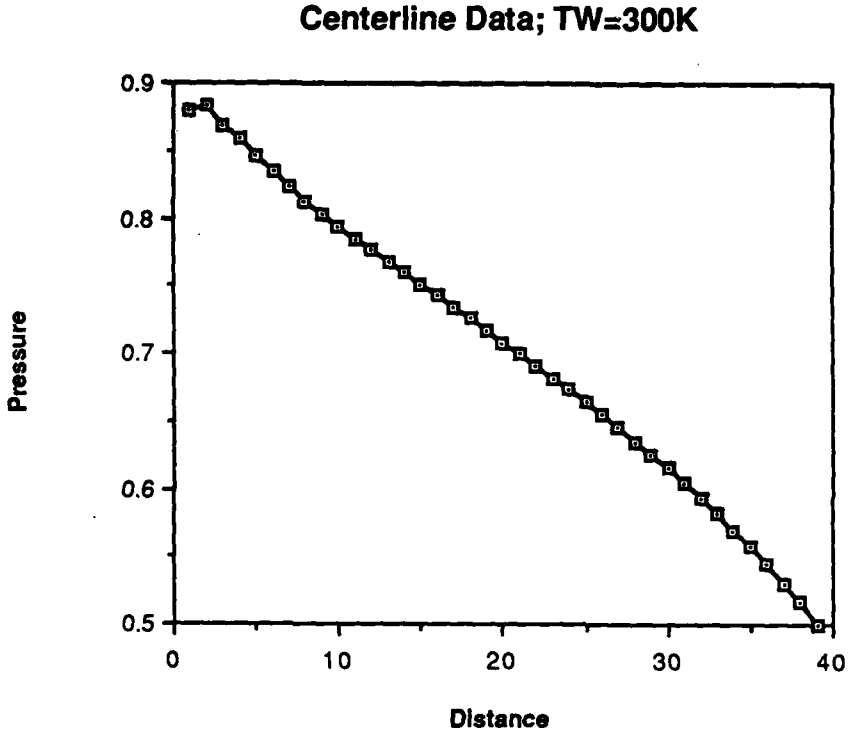


Figure 3. Pressure drop along the tube centerline

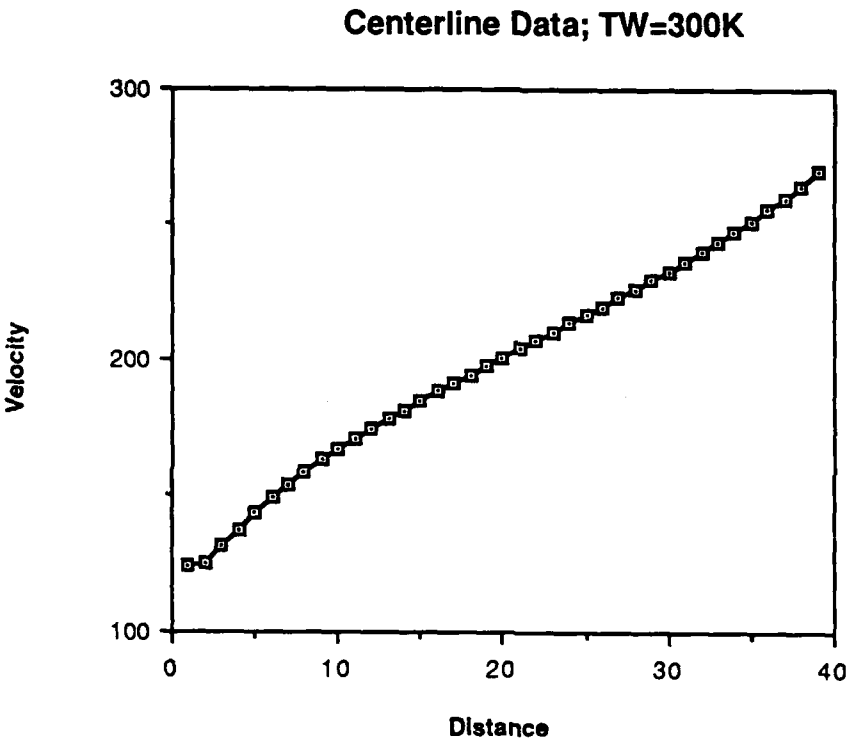


Figure 4. Fluid acceleration along the tube centerline

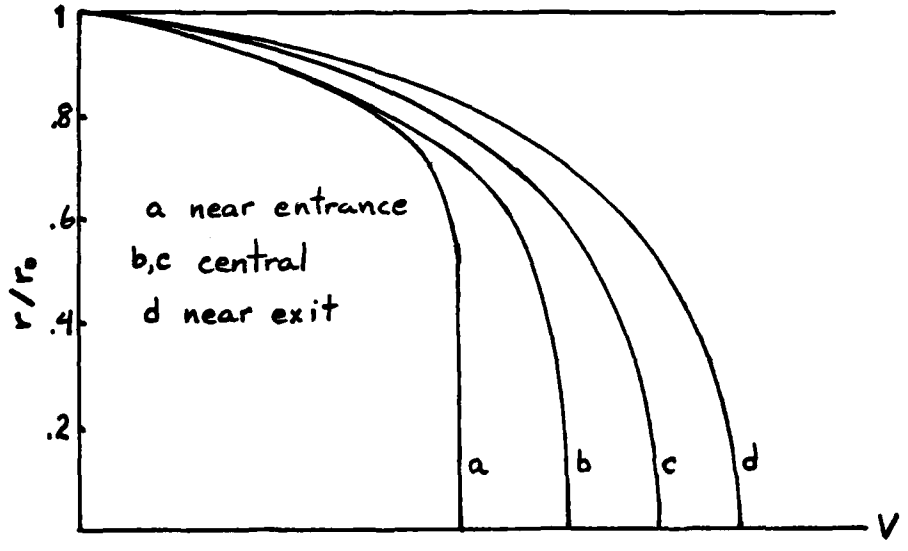


Figure 5. Velocity profiles in tube, showing acceleration and profile development

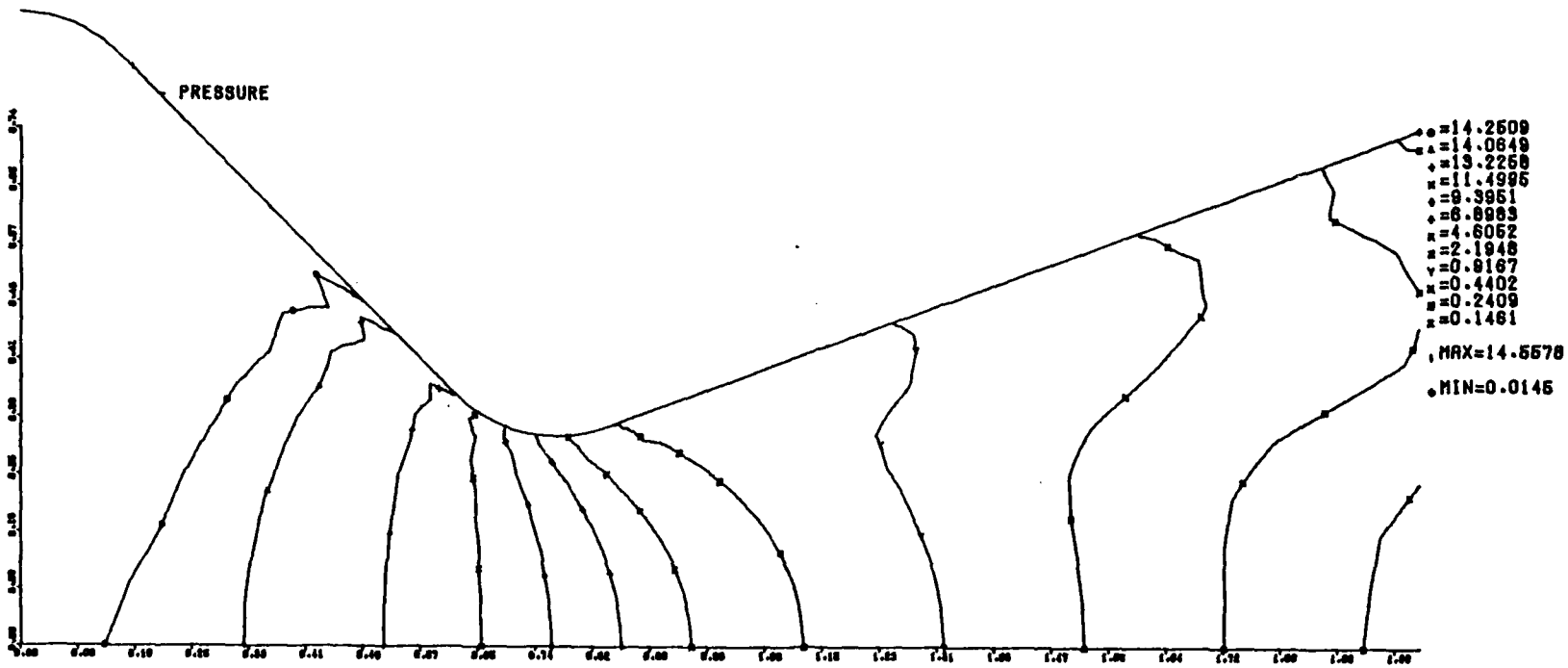


Figure 6. Pressure distribution in nozzle

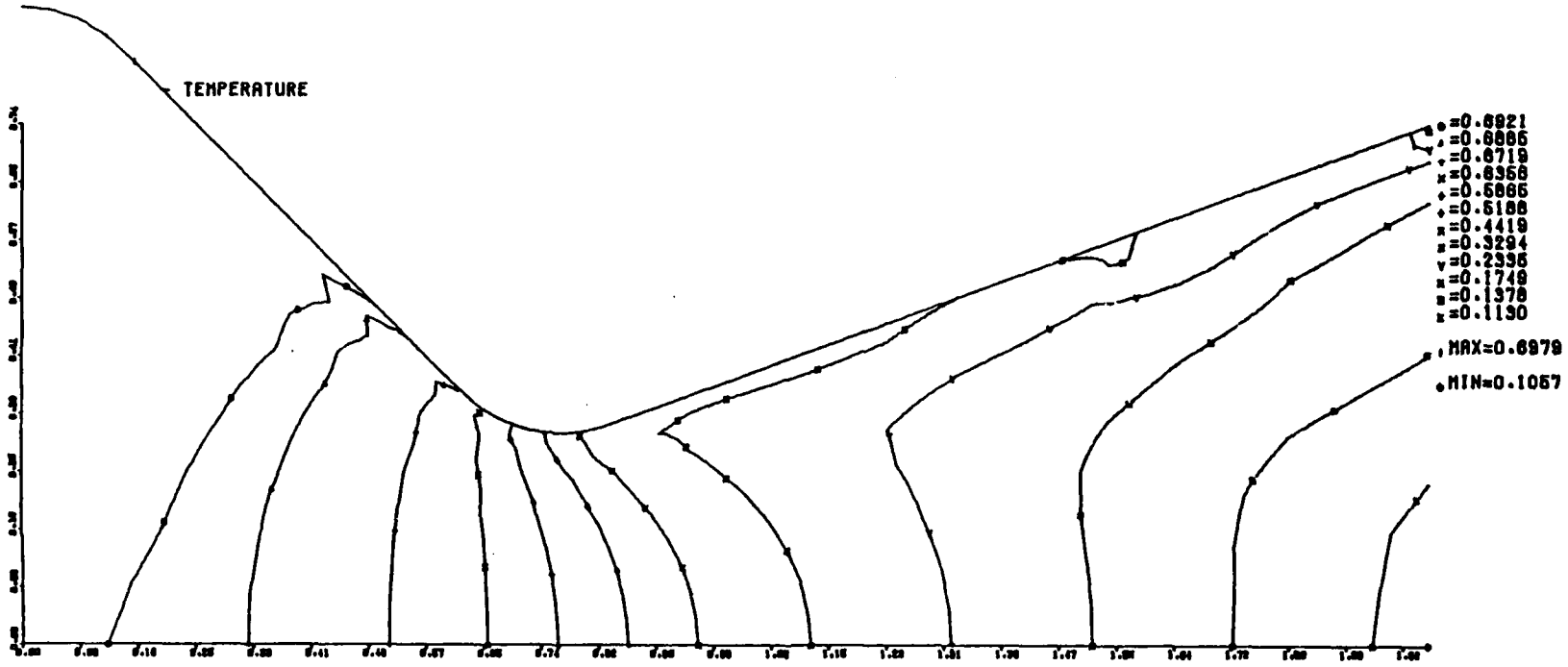


Figure 7. Temperature distribution in nozzle

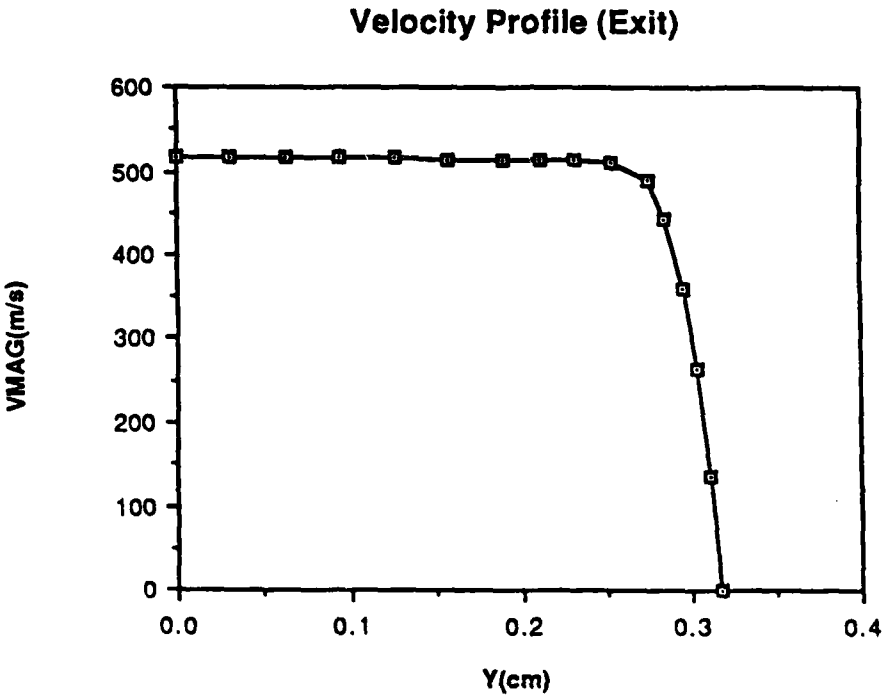


Figure 8. Velocity profile at nozzle exit

1986 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT SUMMER SUPPORT PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

CHEMICAL KINETICS OF HIGH TEMPERATURE AIR FOR MACH 5 - 14 FLIGHT

Presented by:	T. Gary Yip
Academic Rank:	Assistant Professor
Department and University:	Department of Aeronautical & Astronautical Engineering The Ohio State University
Research Location:	Flight Dynamics Laboratory Wright-Patterson Air Force Base Aeromechanics Division High Speed Aero Performance Branch Aerothermo Group
USAF Researcher:	Arthur Lewis
Date:	September 8, 1986
Contract No.:	F49620-85-C-0013

CHEMICAL KINETICS OF HIGH TEMPERATURE AIR FOR MACH 5 - 14 FLIGHT

by

T. Gary Yip

ABSTRACT

A chemical kinetics model for Mach 5 - 14 hypersonic flow fields has been formulated. The model consists of 20 reactions. Previously published models all have a wrong NO removal mechanism. A correct reaction path is adopted in the present model. For Mach 5- 14, there is only one dominant ionization process identified by earlier experimental and theoretical studies. It is the ionizing atomic collision of N and O atoms to form NO^+ and e^-. This reaction is also in the model.

The rate constant expressions for the 20 reactions have been selected through a review of old and more recent rate data obtained in kinetics studies using more precise experimental technique.

A one-dimensional, steady, inviscid, chemically non-equilibrium flow model has also been formulated for calculating the species and flow properties profiles based on the chemical kinetics model of this study. The calculation could not be finished because of the limited time available in the research program. Continuing efforts are recommended and will be included in the research program to be proposed for a Mini Grant support.

ACKNOWLEDGMENTS

I would like to thank the Air Force Systems Command and the Air Force of Scientific Research for Sponsorship of my research. My gratitude is also extended to the Aerothermo Group in the Flight Dynamics Laboratory, Wright-Patterson Air Force Base. I would like especially to thank Mr. Arthur Lewis for his support and many helpful discussions about the research work. I would also like to thank Mr. Richard Newman for his assistance during the ten weeks that I spent in the group.

I. INTRODUCTION

I received my Ph.D. in Aeronautical & Astronautical Engineering from the University of Illinois, Urbana. My major emphases were gas dynamics and combustion. In the last two years, as an assistant professor in the Aeronautical & Astronautical Engineering Department at the Ohio State University, I was teaching aerothermochemistry, physical and chemical gas dynamics, propulsion and supersonic wind tunnel experiments. In the summer of 1985, I entered the area of hypersonics. Since then I have reviewed a large volume of literature on hypersonic flow research performed in the past. Of particular interest to me are the real gas effects and the similitude of hypersonic flows. I have also received a Seed Grant from the OSU to support preliminary experiments in the OSU Arc-Driven Shock Tube.

My research effort during the ten weeks of the summer faculty research program was focused on the chemical reactions in the shocked air behind the leading shock wave of hypersonic vehicles. This effort is aimed at assisting my contract monitor at the Flight Dynamics Laboratory in the development of a computer code for predicting low density real gas flows about hypersonic vehicles.

II. OBJECTIVES OF THE RESEARCH EFFORT

The objectives of the research effort are:

1. to identify the chemical reactions which are important in Mach 5 through 14 non-equilibrium flow fields,
2. to identify the free-electron production reactions that are dominant in this Mach number ranges, and
3. to review the existing rate data of these reactions.

III. APPROACH

The approach chosen consists of two phases. In the first phase a chemical kinetics model for high temperature air is to be formulated. This includes an in-depth literature review to identify the reactions and to determine the reliability of their rate data. The second phase is to use the reactions and rate data obtained in the first phase in a one-dimensional, inviscid, chemically non-equilibrium flow model to calculate the species profiles and flow properties behind the vibrational relaxation zone of a normal shock front. It is intended that the results will be compared with existing experimental data to establish the validity of the chemical kinetics model.

The first phase has been completed as planned. The second phase could not be carried out to its completion because of the limited time available in the summer faculty research program. Near the end of the research program, the

flow model was formulated. An attempt was made to write a computer program to solve the governing equations. No numerical result has been obtained yet. This effort will be continued as a part of the research program in a Mini Grant proposal.

IV. CHARACTERISTICS OF MACH 5 THROUGH 14 SHOCK WAVES

For flights at hypersonic speeds in the earth's atmosphere, the air passing through the leading shock wave is heated to high temperatures at which real gas processes play an important role in altering the gas properties. Such alterations often give rise to significant deviations from the surface heating rate and aerodynamic performance predicted by computer codes which use an ideal gas model. Examples of these processes are the internal excitations of the gas molecules, dissociation, formation of new molecular species, ionization and radiation. They are often non-equilibrium in nature and occur simultaneously in a hypersonic flow field. Furthermore, some of these processes are coupled.

However, it is known that the significance of the different non-equilibrium phenomenon in a hypersonic flow field varies with the flight Mach number and the composition of the atmosphere. The flight Mach number and altitude of interest in this study are Mach 5 through 14 anJ from

100,000 feet (30.5 km) to 200,000 feet (61 km), respectively. For this altitude range, it is appropriate to approximate the composition of air as 79% N_2 and 21% O_2 by mole.

The structure of a normal shock and the energy available in the shocked gas flow change with the Mach number. At Mach numbers up to about 10, translational and rotational excitations of N_2 and O_2 occur first and are followed by vibrational excitation. Around Mach 14 all three modes have been observed to take place simultaneously within the shock front (Wray & Freeman, and Ref. 2). This is because the collisions are so energetic that the direct transfer of energy from the translational to vibrational mode is possible. Around Mach 16 dissociation has been observed to occur in the same zone. These observations are very important to the numerical modeling of hypersonic flow fields because they suggest that for Mach number up to 14 the upstream boundary, immediately behind the shock front, air composition is the same as the ambient and the temperature is the equilibrium temperature of N_2 and O_2.

In this study, a computer program has been written to estimate the equilibrium temperature of air behind the front of Mach 5 through 14 normal shocks. Based on an ambient temperature of 300 K, the highest shocked air temperature is about 10,000 K. At this temperature, the radiative heat

transfer is known to be only a small fraction of the total energy balance (Stupochenko, 1967). Therefore, radiative heat transfer is not considered in the chemical kinetics model to be presented here.

V. A CHEMICAL KINETICS MODEL FOR MACH 5 THROUGH 14 FLOW FIELDS

The reactions being considered in this study include the followings:

a. the dissociation of O_2 and recombination of O

$$O_2 + M == O + O + M \ , \qquad (1)$$

b. the dissociation of N_2 and recombination of N

$$N_2 + M == N + N + M \ , \qquad (2)$$

c. the dissociation of NO and recombination of N and O

$$NO + M == N + O + M \ , \qquad (3)$$

d. the production of NO

$$O + N_2 == NO + N \qquad (4)$$

$$N + O_2 == NO + O \ , \qquad (5)$$

e. the removal of NO

$$NO + NO == N_2O + O \qquad (6)$$

$$O + N_2O == N_2 + O_2 \ , \qquad (7)$$

f. the formation of NO^+ and free-electrons

$$N + O == NO^+ + e^- . \qquad (8)$$

In reactions (1), (2) and (3), M is the collision partner which can be O_2, N_2, O, N or NO. Therefore, there are 15 reactions in (1) through (3) and a total of 20 reactions in the model.

VI. A REVIEW OF THE REACTION RATE DATA

The kinetics data for the twenty reactions are reviewed in this section. For most of the reactions, only the expression for the forward or reverse rate constant is presented. The other rate constant can be obtained from the definition of the equilibrium constant.

a. $O_2 + M == O + O + M$

The dissociation of O_2 and the reverse three-body atom recombination are probably the most studied reactions of interest to hypersonic flow. Atomic oxygen is the most effective collision partner followed by O_2 and N_2 in order. No data is available on the effective of N and NO. Therefore, they are often assumed to have the same effectiveness as Argon. From here on the symbol $k_d(A,B)$ will be used to represent the dissociation rate of species A with B as the collision partner. The unit of k_d used in this report is cc/mole-sec.

There is a general consensus about the relative magnitude of the rate constants for the dissociation of O_2 with the various collision partners. The consensus is that

$k_d(O_2,O_2)$ is one order of magnitude larger than $k_d(O_2,Ar)$,

$k_d(O_2,O) = 3 \times k_d(O_2,O_2)$, and

$k_d(O_2,N_2) = 1/4 \times k_d(O_2,O_2)$.

Wray (1962) obtained the expression $k_d(O_2,Ar) = 2.5 \times 10^{16} \times T^{-0.5} \exp(-59400/T)$ by curve fitting data up to 18000 K. This expression has been found to be inaccurate in its temperature dependence (Johnston, 1968). Bortner (1969) curve fitted the data from nine different studies and obtained an expression which yields a rate about a factor of 3 smaller at 10,000 K. The difference decreases with temperature and at 4000 K the two rates are about the same. The most reliable data was obtained by Breshares et al (1972) using the laser schlieren technique. Their data has only a factor of two uncertainty. They also found that O_2 is 9.6 times more effective than Argon in causing the dissociation of oxygen. Therefore, the recommended rate expressions for the dissociation of O_2 are:

$$k_d(O_2,Ar) = 7.87 \times 10^{13} \exp(-52760/T), \text{ (Breshares et al)}$$

$$k_d(O_2,O_2) = 7.60 \times 10^{14} \exp(-52760/T),$$

$$k_d(O_2, O) = 2.28 \times 10^{15} \exp(-52760/T),$$

$k_d(O_2,N_2) = 1.90 \times 10^{13} \exp(-52760/T)$, and

$k_d(O_2, N) = k_d(O_2,NO) = k_d(O_2,Ar)$.

b. $N_2 + M = N + N + M$

Unlike the dissociation of O_2, the efficiency of different collision partners in the dissociation of N_2 are not related simply by a constant. For example, the relative efficiency between N_2 and N is temperature dependent.

The dissociation of N_2 has not been studied as much as that of O_2 and resulted in the rate data being sparse. A gap of 5400 K is found between the highest temperature recombination data (600 K) and the lowest temperature shock tube dissociation data (6000 K). The most extensive data are for $M = N_2$. Inspite of the large intervening temperature range, the two set of data extrapolate into one another very well (Byod & Burns, 1982). The $k_d(N_2,N_2)$ expressions recommanded by Bortner (1969) and Wray et al (1962) are in very good agreement. Baulch et al (1973) recommanded the expression obtained by Appleton et al (1968), but also suggested error limits of a factor of 3 and cautioned users about the unresolved disparity between their experimental

data and computed results. Therefore, the rate constant expression recommanded by Bortner is adopted here:

$$k_d(N_2,N_2) = 5.0 \times 10^{18} T^{-.75} \exp(-113{,}200/T).$$

For M = N, there are even fewer data. The rate constant expressions presented by Wray, Bortner and Hansen & Baganoff (1972) agree to within a factor of 2 at temperatures below 10,000 K. The expression Bortner obtained by curve fitting data from four studies is recommended here:

$$k_d(N_2, N) = 3.0125 \times 10^{22} T^{-1.5} \exp(-113{,}200/T).$$

No data on the dissociation of N_2 with NO, O and O_2 as collision partners is available. Their effectiveness in dissociating N_2 is often assumed to be the same as Argon. For M = Ar, Cary (1965) and Bryon (1966) obtained high temperature data, 6000 - 10,000 K, which differ by less than a factor of 2. These data were curve fitted by Bortner resulting in a rate constant expression which gives rates closer to those of Bryon. Again Appleton's results differ from the above by a factor of 3 - 4. The expression recommended here is that of Bryon:

$$k_d(N_2,Ar) = 1.9 \times 10^{17} T^{-0.5} \exp(-113{,}200/T), \text{ and}$$

$$k_d(N_2,O) = k_d(N_2,NO) = k_d(N_2,O_2) = k_d(N_2,Ar)$$

c. NO + M == N + O + M

There is no reliable NO dissociation rate data available. Freedman & Daiber (1961) and Wray and Teare (1962) studied NO dissociation at 3000 - 4300 K and 3000 - 8000 K, respectively. Unfortunately , other reactions, such as NO + NO -- N_2O + O and NO + N -- N_2 + O, also occured in the test gas, thus complicating the data analysis. By assuming O_2 and N_2 to be equally effective as Ar as a collision parnter and O and N as effective as NO, Wray and Teare obtained the rate constant for k_d(NO, NO) and found k_d(NO, Ar) to be 20 times smaller. These results are in fair agreement with those obtained by Freedman and Daiber. When Wray's expression for $k_d(NO,N_2)$ was extrapolated to low temperatures, Bortner discovered that the rates were an order of magnitude higher than the low temperature data reported by other investigators. Therefore, the following rate constant expressions are adopted here:

$$k_d(NO,NO) = 8.0 \times 10^{21} T^{-1.5} \exp(-75{,}400/T) \quad \text{(Wray \& Teare)}$$

$$k_d(NO, N) = k_d(NO, O) = k_d(NO, NO)$$

$$k_d(NO,N_2) = 9.0 \times 10^{14} \exp(-75{,}400/T) \quad \text{(Bortner), and}$$

$$k_d(NO,O_2) = k_d(NO, N_2).$$

d. $O + NO = N + O_2$ and $N + NO = N_2 + O$

The formation of NO in high temperature air is primarily controlled by these two reactions. Their rate determining nature has prompted many studies of their rate constants, especially by researchers in the field of combustion. By now the rate constants are reasonably well characterised.

For $N + O_2 == NO + O$, there are adequate low temperature data. Wray & Teare reported the only high temperature data at 5000 K. Hanson et al (1974) reported data at 2500 - 4100 K and an rate constant expression that fits all the data from 300 - 5000 K with extremely small scatter. Therefore, expression due to Hanson et al is adopted here:

$$k_r = 2.36 \times 10^9 \, T \exp(-19430/T),$$

with the subscript r representing the backward reaction.

For $N_2 + O == NO + N$, Glick et al (1957) and Wray & Teare determined the forward rate. Their results were latter substantiated by a number of combustion studies. Baulch et al curve fitted all the data and the resulting rate constant expression is recommended here:

$$k_f = 1.7 \times 10^{13} \exp(-38{,}000/T)$$

with the subscript f representing the forward reaction. An error factor of 2 in the temperature range 2000 - 5000 K is expected.

e. $NO + NO == N_2O + O$ and $N_2O + O == N_2 + O_2$

The reaction $NO + NO == N_2 + O_2$ was mistaken as the primarily NO removal mechanism until Camac and Feinberg (1966) found in a shock tube study of NO decomposition that N_2O and O were produced at a rate compatible with the removal of NO by $NO + NO == N_2O + O$. Other researchers re-evaluated their data and found agreement with the results of Camac and Feinberg. The rate constant expression derived by Kaufman and Kelso, also recommended by Baulch, is adopted here:

$$k_f = 1.3 \times 10^{12} \exp(-32{,}100/T), \text{ and}$$

$$k_r = 1.0 \times 10^{14} \exp(-14{,}100/T).$$

It must be noted that all published computer simulations of NO decomposition, including the different kinetics model being used in hypersonic flow field calculations, have incorporated the wrong NO removal mechanism. Determining the effect of the correct mechanism on the predicted hypersonic flow fields is not possible until it is incorporated.

The removal of N_2O is known to occur via the reactions

$$N_2O + O \text{ -- } N_2 + O_2 \qquad (9)$$

$$N_2O + O \text{ -- } NO + NO \qquad (10)$$

The relative efficiency of the two reactions is often represented by the ratio k_9/k_{10} and have been found to be

around unity at temperatures above 1000 K. The rate constant expression for $N_2O + O == N_2 + O_2$ recommended by Baulch is adopted here:

$$k_f = 1.0 \times 10^{14} \exp(-14{,}100/T), \text{ and}$$

$$k_r = 6.3 \times 10^{13} \exp(-55{,}200/T).$$

f. $N + N == NO^+ + e^-$

There are different types of ionization processes which can take place in high temperature air. The dominancy of the different types depends largely on the available energy in the system. For Mach 5 - 14 flight, the available energy is rather small for most of the ionization processes to occur except for the ionizing atomic collision

$$N + O == NO^+ + e^-$$

This has been confirmed by Lin and Teare (1963). This reaction requires less initial energy than the ionization potential of NO by an amount released when NO is formed from N and O atoms. The rate constant expressions obtained by Bortner are adopted here:

$$k_f = 5.181 \times 10^{11} \exp(-31900/T), \text{ and}$$

$$k_r = 1.5 \times 10^{21} T^{-1.5}.$$

Other ionic species are not included in the present chemical kinetics model because their concentrations are typically at least two orders of magnitude smaller (Lin & Teare).

VII. GOVERNING EQUATIONS FOR STEADY, ONE-DIMENSIONAL, INVISCID, CHEMICALLY NON-EQUILIBRIUM FLOW

In order to evaluate the chemical kinetics of high temperature air proposed in the last section, the reactions can be incorporated into a flow model. The calculated species and flow properties profiles can be compared to available experimental data to determine the appropriateness of the kinetics. One of the simplest flow model that can be used for this purpose is the steady, one-dimensional, inviscid flow model. The governing equations are as following:

1. Global Continuity Equation $$\frac{du}{dx} + u \frac{d\rho}{dx} = 0 ,$$

2. Momentum Equation $$\frac{dp}{dx} = - \rho u \frac{du}{dx} ,$$

3. Energy Equation $$\frac{dh}{dx} - \frac{1}{\rho} \frac{dp}{dx} = 0 ,$$

by definition $$\frac{dh}{dx} = c_p \frac{dT}{dx} + \Sigma h_i \frac{dy_i}{dx} ,$$

4. Species Continuity Equations $\frac{dy_i}{dx} = \frac{M_i}{\rho u}\frac{dc_i}{dt}$,

where i = 1,2,3,...,n-1,

n = number of species in the system,

y_i = mass fraction of the i^{th} species,

c_i = concentration in moles/unit volume, and

M_i = molecular mass of the i^{th} species,

5. Equation of State $p = \rho RT$,

6. Species Conservation Equation $\Sigma y_i = 1$

There are n + 4 equations and the same number of unknowns, namely, $y_1, \ldots\ldots, y_n$, u, ρ, p, and T. The enthalpy h_i is a thermodynamic function depends only on T for thermally perfect gas. The term $\frac{dc_i}{dt}$ is determined from the master rate equation which consists of the rate constant expressions and the species concentrations. The above equations are to be solved simultaneously to obtain the species and flow properties prolifes as a function of x.

VIII. RECOMMENDATIONS

1. A chemical kinetics model has been formulated as planned. At the end of the summer research program, an attempt was

made to solve the one-dimensional, chemically non-equilibrium flow equations, but the effort came to an end because of the limited time available in the research program. Therefore, the chemical kinetics model has not been validated yet. I will include this task as a part of the research program in a Mini Grant proposal. This task is very important for a few reasons. All previous high temperature air kinetics models have used the wrong NO removal mechanism, which is NO + NO == N_2 + O_2. The present model has the correct mechanism. How this mechanism will affect the heat transfer and flow field characteristics remains to be found out. Secondly, some of the rate constant expressions in the present model are better because they are obtained with the consideration of the kinetics studies using more reliable experimental technique. Thirdly, this model is specifically formulated for Mach 5 - 14 hypersonic flow fields and it has a free-electrons producing reaction which is the only dominant ionization process in this flight Mach number range. The eliminations of the much less significant ionization processes can help to reduce the computer time needed to obtain a solution.

2. The rate constant expressions all have a factor of uncertainty. The more reliable data have factors of 2 or less while the less reliable data have much larger factors. Therefore, I recommend that the rate constants

be systematically varied to test the sensitivity of the predicted results to the uncertainty of the rate data.

3. Heat transfer is of high priority in hypersonic flow. The energy transfer between the flow and the surface of a vehicle depends largely on the chemistry and energy available in the flow. I recommend the present chemical kinetics model be used in a simple two-dimensional or axisymmetric flow model to test its effects on heat transfer.

These recommendations are directed at the continuation of my research effort during the ten weeks that I have spent at the Flight Dynamics Laboratory. Therefore, these recommended future efforts will be included in the research plan to be proposed for a Mini Grant support.

References

1. Wray, K.L. & Freeman, T.S., "Shock Front Structure of O2 at High Mach Numbers," J. Chem. Phy., 40:2785-2789 (1964)
2. A Private Communication between P.V. Marrone and K.L. Wray.
3. Stupochenko, Y.V., Relaxation in Shock Waves, p.311, Springer-Verlag, New York, 1967.
4. Johnston, H.S., "Gas Phase Reaction Kinetics of Neutral Oxygen Species," NSRDS-NBS 20, National Bureau of Standards, 1968.

5. Boyd and Burns, Shock Waves in Chemistry, Ed. Assa Lifshitz, p. 161, Marcel Dekker Inc., New York, 1981.
6. Bortner, M.H., "Review of Rate Constants of Selected Reactions of Interest in Re-entry Flow Fields in the Atmosphere," NBS Tech Note 484, 1969.
7. Wray, K.L., "Chemical Kinetics of High Temperature Air," Hypersonic Research, Ed. F.R. Ridell, Academic Press, 1962.
8. Baulch, D.L., Drysdale, D.D. and Horne, D.G.,Evaluated Kinetic Data for Hight Temperature Reactions, Vol. 2, Butterworth, London, 1973.
9. Appleton, J.P., Steinberg, M. and Liquornik, D.J., "Shock Tube Study of Nitrogen Dissociation using U.V. Light Absorption," J. Chem. Phy., 48:599 (1968)
10. Hanson and Baganoff, AIAA J., 10:211 (1972)
11. Cary, B., Phy. of Fluid, 8:26 (1965)
12. Bryon, S., "Shock-Tube Measurement of the Rate of Dissociation of Nitrogen," J. Chem. Phy,, 4:1378 (1966)
13. Freedman E. and Daiber, J.W., J. Chem. Phy. 34:1271 (1961)
14. Wray, K.L. and Teare, J.D., "Shock-Tube Study of the Kinetics ofNitric Oxides at Hight Temperatures," J. Chem. Phy., 16:2582-2596 (1962)
15. Glick, H.S., Klein, J.J., and Squire, W., J. Chem. Phy., 27:850 (1957)

16. Hanson, R.K., Flower, W.L., and Krager, C.H., "Determination of the Rate Constant for the Reaction O + NO == N + O2," Comb. Sci. and Tech., 9:79-86 (1974)

17. Kaufman, F. and Kelso, J.R., J. Chem. Phy. 23:1702 (1955)

18. Camac, M. and Feinberg, R.M., "Formation of NO in Shock-Heated Air," 11th Symposium (International) on Combustion, p.137, 1966.

19. Lin, S., Neal, R.A., adn Fyfe, W.I., "Rate of Ionization Behind Shock Waves in Air. I. Experimental Results," J. Chem. Phy. 5:1633-1648 (1962)

20. Lin, S. and Teare, J.D., "Rate of Ionization Behind Shock Waves in Air. II. Theoretical Interpretations," J. Chem. Phy. 6:355-375 (1963)

1986 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/
GRADUATE STUDENT SUMMER SUPPORT PROGRAM

Sponsored by the
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
Conducted by the
Universal Energy Systems, Inc.
Final Report

Physiological Correlates of Behavioral Performance on the Mathematical Processing Subtest of the CTS Battery

Prepared by:	Robert L. Yolton
Academic Rank:	Associate Professor
Department and University:	College of Optometry, Pacific University
Research Location:	Armstrong Aerospace Medical Research Laboratory, HEG, Wright-Patterson AFB, OH
USAF Researcher:	Glenn Wilson, PhD
Date:	September 1986
Contract No:	F49620-85-C-0013

Physiological Correlates of Behavioral Performance on the Mathematical Processing Subtest of the CTS Battery

by
Robert L. Yolton

ABSTRACT

The Criterion Task Set (CTS) is a battery of behavioral tests designed to assess mental abilities and information processing. One of the sub-tests within the CTS evaluates mathematical processing ability by asking subjects to solve equations with 1, 2 or 3 mathematical operators (only + or - operators were used). The number of operators establishes the difficulty level of the equations, thus 3 levels were investigated.

In this project, I trained 10 subjects to a criterion level of performance at each difficulty level and then measured the mean time required for them to solve the equations. The following physiological correlates of their performance were also recorded: electroencephalographic (EEG) signals from 3 scalp locations, heart rhythms and peripheral temperatures. Additionally, subjects gave a verbal rating of their perceived workload for the three difficulty levels.

Results indicate that there are complex changes in the late components of the EEG signals which correlate with the time required to solve the equations and with the subject's reports of workload. Variations in peripheral temperature and heart rhythm were small and did not correlate systematically with changes in difficulty levels. Changes in certain physiological systems such as the EEG can, therefore, provide an indication of the degree of workload a subject is experiencing but more refined analyses than were available for use in this study will be needed to detect changes in other systems (eg., heart rhythm).

ACKNOWLEDGMENTS

I am pleased to acknowledge the support of the Air Force Systems Command, Air Force Office of Scientific Research, the Armstrong Aerospace Medical Research Laboratory at Wright-Patterson AFB, OH and Universal Energy Systems, Inc., that together made this summer project possible. I also wish to thank the employees of SRL, Inc., including Penny Fullenkamp, Brian Porter, Phil Rang, Kathy McCloskey, Claud Orr, Barbara Culbertson, Walter Rosenthal and especially Iris Davis for their help, encouragement and understanding.

Gary Reid, Bill Perez, Tom Eggemeier, Paul DeRego and Andy Junker provided me with papers and gave me insights into their work for which I am most grateful.

Finally, and most importantly, I thank Dr. Glenn Wilson for directing my project and for making my research in his laboratory so valuable and enjoyable.

I. INTRODUCTION: I received BS, MA and PhD degrees in psychology from the University of Wisconsin, California State University at Sacramento and the University of Texas, respectively. My masters program provided me with a background in testing and assessment while my doctoral work dealt with visual physiology and function in lower species. Following my PhD, I obtained an applied understanding of human visual function while completing my Doctor of Optometry degree at the New England College of Optometry. For the past 11 years, I have been at the Pacific University College of Optometry, first as Director of Research and now as Associate Professor. At Pacific, my research program has concentrated on electrophysiological measurement of the neurological events related to the receipt, encoding and processing of information delivered via the visual system.

II. OBJECTIVES OF RESEARCH EFFORT: The goal of my project was to determine if physiological measures, including electroencephalographic (EEG) signals, heart rhythm and peripheral temperature would correlate with behavioral performance and subjective perception of workload on a mathematical processing task in which subjects mentally solved equations presented to them on a video display screen. This task had three levels of difficulty obtained by varying the number of mathematical operators in the equations the subjects solved.

My specific objectives were as follows:

1. Select and train ten normal, adult humans to criterion performance at the three difficulty levels of the math processing task.
2. Assess the learning curves and final performance levels of the subjects with respect to response accuracy, time required to solve the equations, and variability of these times.
3. Use the trained subjects to determine if EEG, heart rhythm and/or peripheral temperature data correlate with behavioral performance (reaction times, etc.) and perception of workload associated with the different task difficulty levels.
4. Use any resulting correlations to further demonstrate that changing the number of operators in the math processing task equations does indeed change the difficulty/workload level of the task.
5. Use any correlations to add to the the knowledge base regarding how the body reacts physiologically to different levels of mental workload.

III. BACKGROUND: The Criterion Task Set (CTS) is a battery of performance tests

developed to provide a standardized metric for assessing a wide range of mental performance abilities (Shingledecker, 1984; Acton and Crabtree, 1985; Shingledecker and Schlegel, 1985). The battery consists of nine sub-tests which measure skills including spatial processing, eye-hand coordination, mental time keeping, grammatical reasoning and mathematical processing. All but one of the sub-tests have multiple levels of difficulty incorporated into them.

Development of the CTS was based on the multiple resource (Wickins, 1981) and stage (Sternberg, 1969 a, b) models of mental processing and the CTS has been used to test various implications of these theories. The multiple difficulty levels built into the sub-tests have also allowed the CTS to be used to generate tasks with known workload requirements and to assess the effects on the body of these different degrees of workload.

In the experiment described in this report, the mathematical processing sub-test of the CTS was used to demonstrate that the three difficulty levels built into the sub-test produce different workload requirements which, in turn, produce different levels of physiological response in human subjects.

IV. SUBJECTS: Ten subjects (6 males and 4 females) were drawn from a pool of paid volunteers. Their mean age was 23.2 (SD = 3.7). All were free from visual or systemic pathology and had at least equivalent 6/6 (20/20) near visual acuity with 40 arc sec of stereo acuity. Dominant hand was determined by subject report and dominant eye was established by using an aperture sighting task with a distant target. Of the 9 right-handed subjects, 5 had mixed eye-hand dominance; the left-handed subject had corresponding dominance.

Prior to participation, subjects completed an Informed Consent Form and each cooperated fully and enthusiastically during all phases of the study.

V. TRAINING: During a pre-training briefing, subjects were told that they would solve arithmetic equations presented on a video display screen. They were instructed to always work from left to right through the equation, and indicate by pushing a button whether the solution was less than 5 (1,2,3 or 4) or greater than 5 (6,7,8 or 9). It was emphasized that the solution could never be 5 nor outside of the ranges specified. It was also explained that the equations could have 1, 2 or 3 arithmetic operators (+ or -) but that information on how many operators to expect for each block of equations would be provided. All of the subjects had previous experience with experiments of this type and readily understood what was expected of them.

For training purposes, a subject was comfortably seated in a dimly lit, sound attenuating chamber, 1.0 M from a Commodore Model 1526 video monitor. Equations were presented in the center of the display as white 6.0 mm high by 3.0 mm wide (20.6 arc min by 10.3 arc min) characters on a black background. When an equation was presented, the subject had 4.0 sec to press one of two push button switches indicating that the solution was greater or less than 5. The 1.0 cm square switches were mounted in a 9.0 cm wide by 12 cm long by 3 cm deep box which the subjects were told to rest on their thigh and hold with the thumb and fourth fingers of their dominant hand. If they were right-handed, their first and second fingers were used to press the less-than and greater-than switches, respectively. The left-handed subject pressed the less-than switch with her second finger and the greater-than switch with her first finger.

If a subject did not press either switch during the 4.0 sec equation presentation, the screen was blanked and a miss was recorded by the Commodore 64 computer controlling the experiment. Approximately 3.0 sec later a new equation was presented. If the subject did press a switch during the equation presentation, the computer recorded the correctness of the response and its latency (reaction time) as measured from the time the computer began to draw the raster line which would form the upper-most part of the equation. After a response, the screen was blanked for approximately 3.0 sec and then a new equation was presented. This process continued until 50 equations had been presented.

Each equation presentation constituted a trial and each set of 50 equations made up a block. All trials within a block had the same number of operators and subjects were told before each block how many operators to expect. After a block, which took about 3.0 min to complete, a rest period of about 4.0 min was provided. Training was given 3 times per week with each session lasting 1.0 H. The proportion of 1, 2 and 3 operator blocks per session was approximately 1:2:4, with training continued for six sessions until each subject completed 7, 14, and 30 blocks at the 1, 2 and 3 operator levels, respectively. The values were based on a report by Shingledecker (1984) which indicated that stable performance should be reached at these levels of training.

During the first 4 training sessions, emphasis was placed on the speed and accuracy of the responses to the equations, and feedback on performance was provided at the end of each block. For the last two training sessions, subjects were

asked to maintain or improve their speed and accuracy, but to establish a rhythm as they solved the equations. This was done to reduce the standard deviation of the reaction times for each block of trials; most subjects were able to reduce their standard deviations with only a transient sacrifice in response latency and/or accuracy.

VI. RESULTS OF TRAINING: Block by block performance means for all 10 subjects are shown on Figure 1. Rapid learning is shown on the reaction time and percent correct curves, but reaction times within blocks remained quite variable, especially at the three operator level, even after stable performance was reached. The curves show that there was little performance change after the third, seventh and fifteenth blocks for the 1, 2 and 3 operator levels, so the mean values from the last 4, 7 and 15 blocks representing subject capabilities at these levels are shown in Table 1.

TABLE 1 - SUMMARY OF MEAN TRAINING DATA FOR LAST 4, 7 AND 15 BLOCKS

	1 OPERATOR (4 BLOCKS)	2 OPERATORS (7 BLOCKS)	3 OPERATORS (15 BLOCKS)
MEAN REACTION TIME FOR CORRECT TRIALS (MSEC)	681	1243	1850
MEAN BLOCK BY BLOCK STANDARD DEVIATIONS OF REACTION TIMES	132	307	415
MEAN OF MEDIAN REACTION TIMES FOR ALL CORRECT BLOCKS (MSEC)	657	1149	1792
MEAN OF MEDIAN REACTION TIMES FOR LESS-THAN 5 BLOCKS (MSEC)	666	1174	1768
MEAN OF MEDIAN REACTION TIMES FOR GREATER-THAN 5 BLOCKS (MSEC)	645	1202	1794
MEAN PERCENTAGE OF TRIALS CORRECT PER BLOCK	98	96	95

Subjects responded more quickly and accurately, and with less response time variability, with fewer operators in the equations. Taking into account differences in experimental procedures, the mean response times in Table 1 are in reasonable accord with values reported by Shingledecker (1984) for a similar study.

Based on the median data, subjects responded more quickly for the trials with a less-than 5 correct answer at the 2 and 3 operator levels, but this difference was not statistically significant at the 0.05 level (paired comparisons t-tests). The

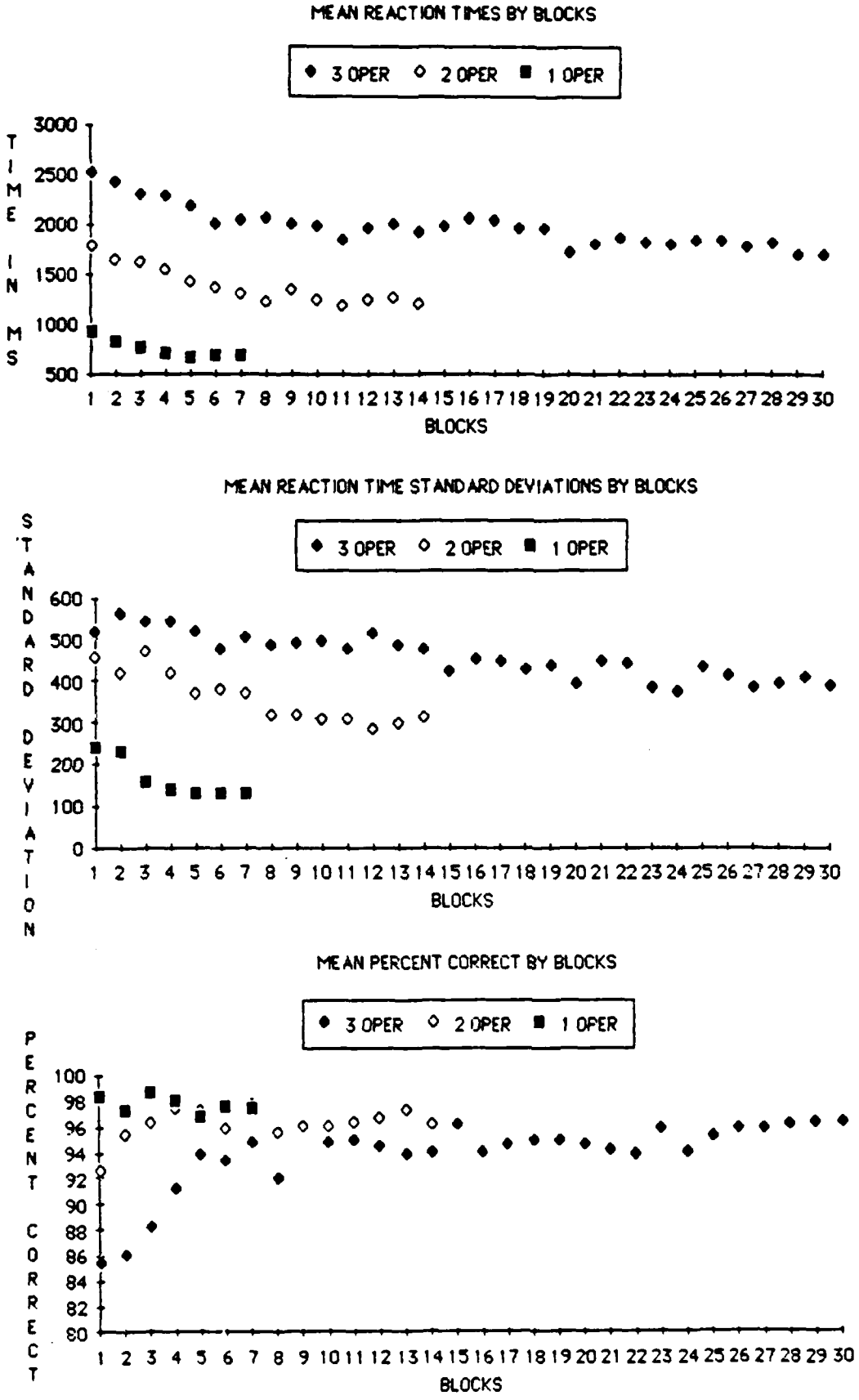


FIGURE 1. REACTION TIME, STANDARD DEVIATION AND PERCENT CORRECT CURVES FOR TRAINING BLOCKS.

statistical significance of the differences between reaction times, standard deviations and percents correct at the three different operator levels is shown in Table 2. For each statistical comparison, a repeated-measures ANOVA was used with Newman-Keuls follow-up testing if the F-ratio from the ANOVA reached the 0.05 level. In the Table, "SIG" indicates significance at the 0.05 level and "NS" means not significant.

TABLE 2 - STATISTICAL ANALYSES OF DATA FROM LAST 4, 7 AND 15 TRAINING BLOCKS

	ANOVA	NEWMAN-KEULS RESULTS	
	F-RATIO	2 OPERATORS	3 OPERATORS
MEAN REACTION TIMES	SIG		
1 OPERATOR		SIG	SIG
2 OPERATORS		–	SIG
MEAN STANDARD DEVIATIONS	SIG		
1 OPERATOR		SIG	SIG
2 OPERATORS		–	SIG
MEAN PERCENT CORRECT	SIG		
1 OPERATOR		NS	SIG
2 OPERATORS		–	NS

The statistical tests demonstrate that changing the number of operators in the equations does have a significant effect on the time required for solution even for highly trained subjects. If reaction time is an indication of effort, the subjects workload increased as the number of operators increased. The significant difference in accuracy for the 1 versus 3 operator levels also indicates a difference in workload for these two operator levels (Mulder, in press).

VII. TESTING SESSIONS: Following completion of training, subjects participated in individual testing sessions during which EEG, heart rhythm, eye blink and peripheral temperature data were obtained while the math processing tasks were being performed. Subjective ratings of perceived workload associated with each operator level were also obtained during this session by use of the Subjective Workload Assessment Technique (SWAT) (Reid, Shingledecker, Nygren and Eggemeier, 1981; Reid, Shingledecker and Eggemeier, 1981; Reid, 1985; Eggemeier, 1985). The procedures used during the testing sessions will be discussed following a description of the physiological recording equipment.

VIII. PHYSIOLOGICAL RECORDING EQUIPMENT AND DATA ACQUISITION:
EEG recordings were made using conventional skin preparation and electrode placement procedures. Silver/silver-chloride electrodes were placed on the scalp at locations Cz, Pz and Oz; one mastoid served as reference and the other was ground. All inter-electrode impedances were 5.0 Kohm or less. EEG signals were amplified by a factor of 50,000 and filtered with half-amplitude values of 0.1 and 30.0 Hz.

Eye blinks were recorded with a silver/silver-chloride electrode placed 0.5 cm above the brow of the dominant eye. Ground and reference electrodes were the same as for EEG recording. Amplifier filters were set at 0.1 and 100.0 Hz. Blink data were used in later analyses of EEG signals to reject trials contaminated by blinks.

To record heart rhythms, electrodes were placed on the suprasternal superior aspect of the manubrium and between the lower two ribs on the left side. Amplifier cut-off values of 0.1 and 100.0 Hz were used.

Peripheral temperature was recorded with a digital thermometer. The thermocouple probe was taped over the fingerprint area of the first finger of the non-dominant hand and the hand was loosely covered with a tissue to prevent excessive air flow over the thermistor.

EEG and heart signals were stored on analog tape for off-line analysis and digitized on-line by a Neurological Workload Test Battery (NWTB) system which consists of a customized DEC 11-73 computer (Wilson, 1985). Digitalization rates were 5.0 msec for the EEG and 1.0 msec for the heart signals. Data from the digital thermometer were recorded manually at the beginning and end of each block.

IX. TEST PROCEDURES: Following electrode placement and a complete briefing on the testing sequence, the subject was seated in the same sound attenuating chamber used for training and given a 1.0 min eyes-closed rest period. This was followed by a 3.0 min period during which the subject viewed a stationary display on the video monitor and the computer recorded resting base-line heart data. Following a 5.0 min period (required to insure that all data acquisition and analysis equipment was functioning properly) testing began.

An operator level (1, 2 or 3) was selected from a pre-prepared list which counterbalanced the order of presentations and the subject was informed

regarding how many operators to expect. The subject was also told that the first block of 50 equations would have the letter "G" substituted for each of the numbers in the equations and that she or he was to view the letters and press either of the response buttons in about the time it would have taken to solve a numeric equation with the specified number of operators. Thus the letter trials simulated the real equations in terms of visual input and motor response, but they required little mental effort.

After the block of letter equations, the subject was given a warm-up block of equations with numbers. No data were recorded for this block. The warm-up block was followed by two 50 equation blocks (DATA 1 and DATA 2) during which data were recorded. Each block required about 3.0 min to complete and 4.0 min rest periods were provided between blocks. After the second data block, the subject was asked to give a SWAT rating for the workload associated with data blocks 1 and 2 combined. The testing sequence was then repeated for the second and the third operator levels with 5.0 min rest periods provided between sequences.

X. RESULTS OF TESTING - PERFORMANCE DATA: Reaction times and related values were combined for testing session data blocks 1 and 2 and the across subjects means are shown in Table 3. To determine if performance

TABLE 3 - SUMMARY OF MEAN TESTING SESSION DATA

	1 OPERATOR	2 OPERATORS	3 OPERATORS
MEAN REACTION TIME FOR CORRECT TRIALS (MSEC)	650	1142	1730
MEAN BLOCK BY BLOCK STANDARD DEVIATIONS OF REACTION TIMES	113	294	365
MEAN OF MEDIAN REACTION TIMES FOR ALL CORRECT BLOCKS (MSEC)	630	1079	1749
MEAN OF MEDIAN REACTION TIMES FOR LESS-THAN 5 BLOCKS (MSEC)	649	1087	1657
MEAN OF MEDIAN REACTION TIMES FOR GREATER-THAN 5 BLOCKS (MSEC)	618	1071	1724
MEAN PERCENTAGE OF TRIALS CORRECT PER BLOCK	98	95	93

during the testing sessions differed from the last 4, 7 and 15 training blocks, matched-pairs t-tests were used to compare the respective reaction time, standard deviation and accuracy data. No significant differences were found.

To assess the significance of differences in performance between operator levels, repeated measures ANOVA's with Newman-Keuls follow-up tests were used (Ferguson, 1981). Results of these tests show a pattern exactly like that in Table 2, i.e., all differences were significant except those between percents correct for 1 versus 2 and 2 versus 3 operator levels. Based on these results and the matched-pairs t-tests, performance during the testing sessions is representative of the subjects' performance during the stable training blocks.

XI. RESULTS OF TESTING - SWAT DATA: SWAT scores are scaled composites based on subject's three point ratings of time, effort and stress associated with a task. Mean SWAT scores for the 1, 2 and 3 operator equation blocks used in the testing sessions were 10.3 (SD=14.3), 31.1 (SD=22.8) and 44.4 (SD=18.7), respectively. A repeated measures ANOVA and Newman-Keuls tests indicate that these values are significantly different from each other at the 0.05 level. The subjects' perception of workload has thus been shown to increase with the number of operators in the equations.

XII. RESULTS OF TESTING - PERIPHERAL TEMPERATURE: It is well known that relaxation produces an autonomic nervous system mediated peripheral vasodilation and a corresponding increase in peripheral temperature. Since stress reduction increases temperature, it was hypothesized that peripheral temperature should be inversely related to workload. To test this hypothesis, finger temperature was measured at the beginning and end of each block of trials. Grand mean temperature change data (data from blocks 1 and 2 for equations containing numbers were combined) are shown in Table 4. A repeated measures ANOVA and Newman-Keuls

TABLE 4 - MEAN TEMPERATURE CHANGES IN DEGREES FAHRENHEIT

	1 OPERATOR	2 OPERATORS	3 OPERATORS
EQUATIONS WITH LETTERS	0.35 (SD=.6)	0.42 (SD=1.5)	0.15 (SD=1.1)
EQUATIONS WITH NUMBERS	-0.44 (SD=.8)	-1.15 (SD=1.2)	-0.35 (SD=.7)

tests show that the temperature decrease for the 2 operator equations with numbers is significantly greater than the decrease for the 1 and 3 operator equations. There

is no significant difference between the 1 and 3 operator temperature changes.

Matched-pairs t-tests between the number and letter equation changes indicate that only the 2 operator level shows a significant difference. It is difficult to understand why a significant difference should be seen only for the 2 operator level. A possible explanation is that two subjects showed unusually large changes (-3.2 and - 2.9 degrees) at this level and these two values may have had an inordinate effect on the data. No artifactual causes for these large changes could be found therefore they are considered to be valid data points.

XIII. RESULTS OF TESTING - EEG SIGNALS: As indicated by the many overlapping components shown in EEG recordings, the brain's response to a sensory stimulus is complex and multi-faceted. In general, response components with short latencies (up to 250 msec for the visual system) are associated with processing physical parameters (eg., color, contour, etc) of the stimulus while mid-latency components, such as the P300, may reflect analysis of the information content of the stimulus. Long latency components may be associated with programming and output of the motor activity resulting from the stimulus.

If an EEG response is elicited by a visual stimulus, the primary sensory components are largest in amplitude, hence most easily recorded, over the visual cortex near location Oz. Indicating its primary hippocampal generator, the mid-latency P300 is usually recorded over the parietal area (Pz) and motor signals are clearly seen at location Cz because of the proximity of motor cortex to this site.

The question of how stress/workload and attention interact with the various components of the EEG is still open (Hillyard and Kutas, 1983). It would seem that the earliest responses in the sequence would be affected least by workload because of the relatively "mechanical" nature of the analyses carried out by the structures generating these components, but there is controversy about this. Several researchers have suggested that the LGN might modulate sensory flow to the visual cortex, and since the LGN has inputs from the attentional system, early components could be sensitive to workload or attention changes.

The P300 has been the focus of most of the research on the effects of workload, but here again there is controversy. Heffley and Donchin (1978) found no difference in P300 amplitude or latency for arithmetic tasks at two different difficulty levels. Other researchers (eg., Mulder, in press: Poon, Thompson and Marsh,1976: Brookhuis, Mulder, Mulder and Gloerich, 1983) have found differences in P300

amplitude and/or latency caused by shifts in workload.

Based on evaluation of the available literature, our hypothesis was that the amplitudes of the early EEG components should increase with operator level because there would be more characters on the screen with more operators, hence total screen luminance would increase. The hypothesis regarding the P300 was that it would not show a significant change with operator level because the P300 indicates the brain's recognition that an equation had been presented and does not signal the solution of the equation. Subjects were expected to work longer to solve the 3 operator problems, but it did not follow that they would work harder or devote more attention to the 3 operator equations, hence no P300 effect was predicted.

To investigate these hypotheses, EEG data from the testing sessions were analyzed using 3.0 sec and 1.0 sec duration epochs. To obtain 3.0 sec data, blocks 1 and 2 for each subject were retrieved from analog tape and a Nicolet CA-1000 was used to ensemble average the data. Grand average curves for 7 subjects (for technical reasons, data for subjects RK, PR and BR were not available) recorded from scalp sites Cz, Pz and Oz are shown on Figure 2. The short vertical lines on the curves represent mean reaction times for the subjects at each operator level. The long vertical line on the left side of the curves represents stimulus onset (160 msec after the beginning of the 3.0 sec epoch).

Qualitative inspection of the CZ curves shows a rapidly rising and falling P160, the amplitude and latency of which do not appear to vary with operator level. A probable P300 occurs next and is best seen in isolation at the 3 operator level. The 3 operator curve returns to baseline at about 1100 msec and then shows a large positive component peaking at about 1800 msec. This component is also present in the 2 and 1 operator curves but peaks earlier. The relationship of these long latency positive components to reaction times and their frontal location suggest a motor origin.

Data from the Pz electrode shows the P160 that was seen at Cz, but the reduced amplitude indicates a frontal generator site for this component. Presumed P300's with latencies of about 500 msec are seen following the P160's. The P300's are themselves followed by large positive components similar to, but reduced in amplitude with respect to the comparable components at Cz. This again indicates a frontal generator site for these components of the waves.

Oz data show the expected early constellation of positive and negative

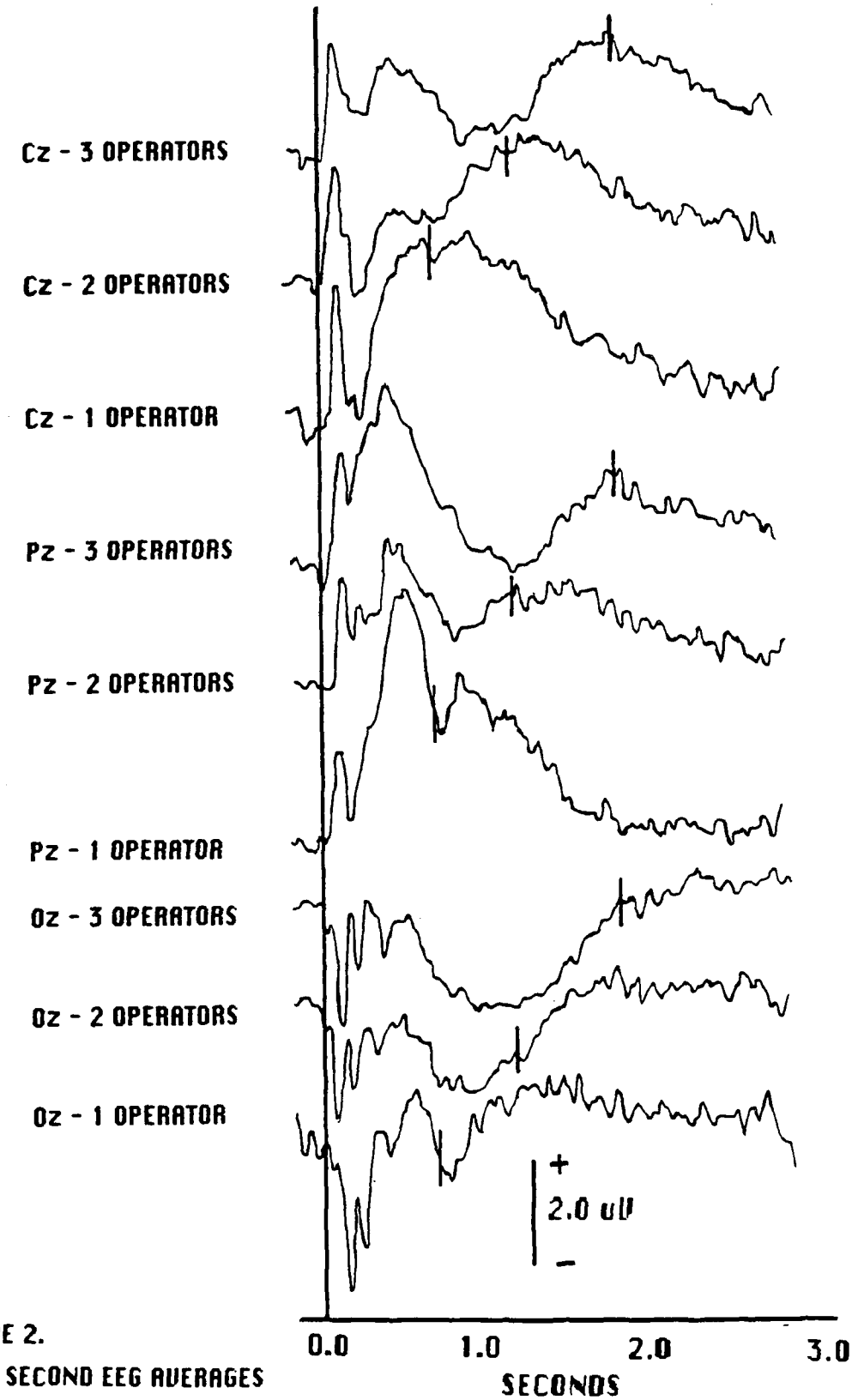


FIGURE 2.
THREE SECOND EEG AVERAGES

components probably associated with sensory processing of the luminance and contour components of the stimulus. The remainder of the Oz waves are similar in shape to those seen at Pz but the frontal origin of many of the later components is indicated by the reduced Oz amplitudes.

Figure 3 shows grand mean 3.0 sec average curves for equations with letters and for combined data blocks 1 and 2 with numbers in the equations. At Cz, the P160 component appears to be the same for numbers and letters, but the late positive component is reduced in amplitude for the letters. At Pz, the P300's are better defined and somewhat larger for the number equations, and at Oz there is a difference in the waveform configuration between 200 to 300 msec across all three operator levels.

One second analyses of the early portions of the EEG curves (150 msec and 850 msec after stimulus presentation) were accomplished by using the NWTB computer to digitize data for an IBM 370-3031 which averaged data across subjects. Figure 4 shows the resultant curves. At location Cz, the P160 component is well defined and does not seem to be related in amplitude or latency to operator level. For the 1 operator trials, a rapidly rising positive wave starting at about 250 msec is easily seen. This wave probably contains parts of the P300 and the motor components of the response.

At Pz, the P300 responses can be clearly seen with the 1 operator response larger than the responses for the 2 and 3 operator levels. At Oz, the primary visual evoked response components do not appear to differ significantly for the 2 and 3 operator levels but the first negative component is considerably larger for the 1 operator level.

These results indicate a difference in EEG waveforms, but the differences, except for those associated with the motor aspects of the task, are subtle and do not change in a regular fashion with operator level and reaction time.

XIV. RESULTS OF TESTING - HEART RHYTHMS: Researchers have suggested that changes in the activity level of the autonomic nervous system associated with increases in mental stress or workload can cause the heart rate to rise and the interbeat interval to decrease in variability (Mulder, in press; Ginsberg, Heslegrave, Scher, Wong and Furedy, 1980; Mulder and Mulder, 1981). To determine if this effect could be demonstrated for the three levels of the math processing task, heart rhythm data from the initial rest period and from the letter and

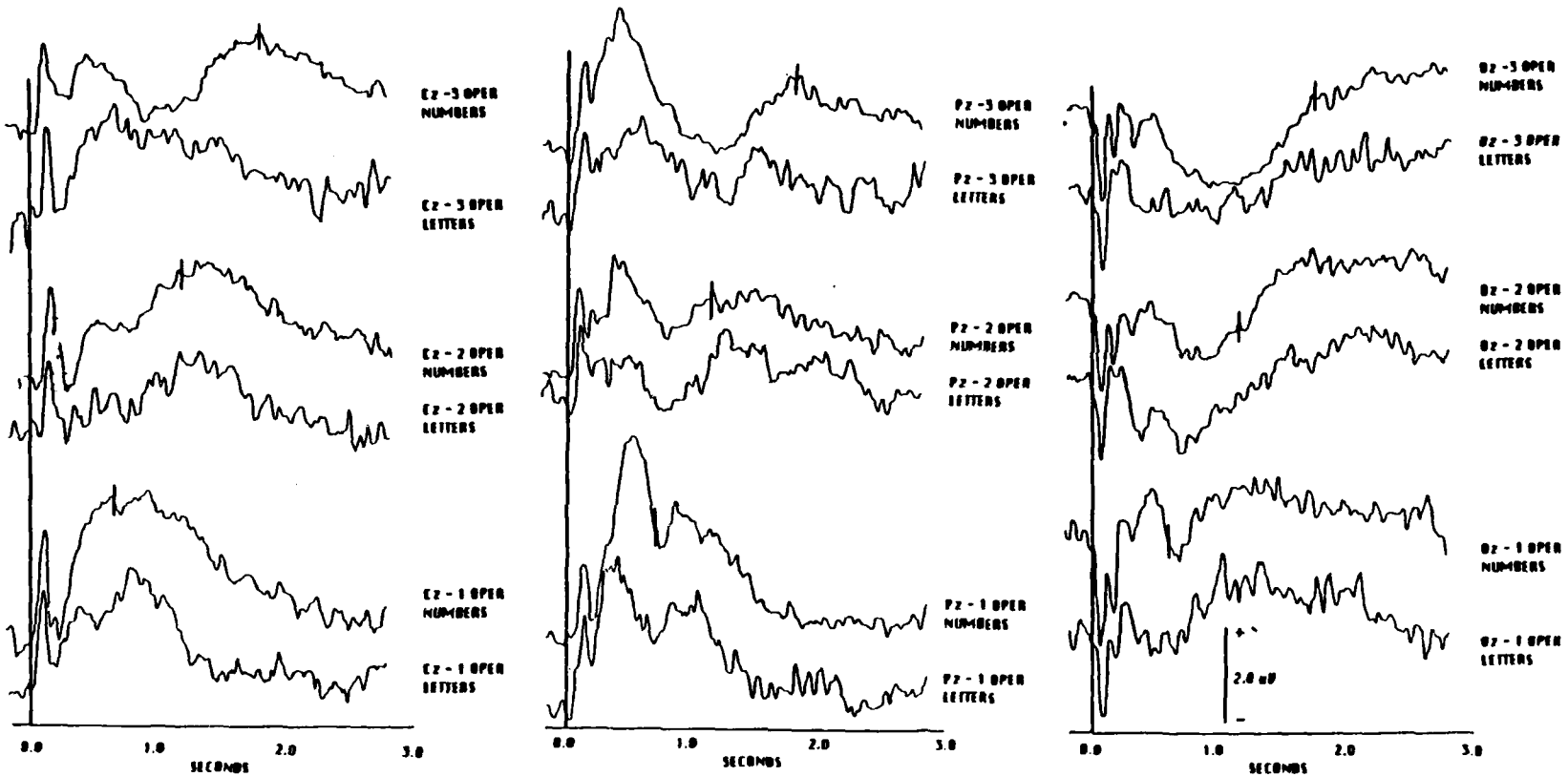


FIGURE 3. EEG AVERAGES FOR LETTER AND NUMBER EQUATIONS

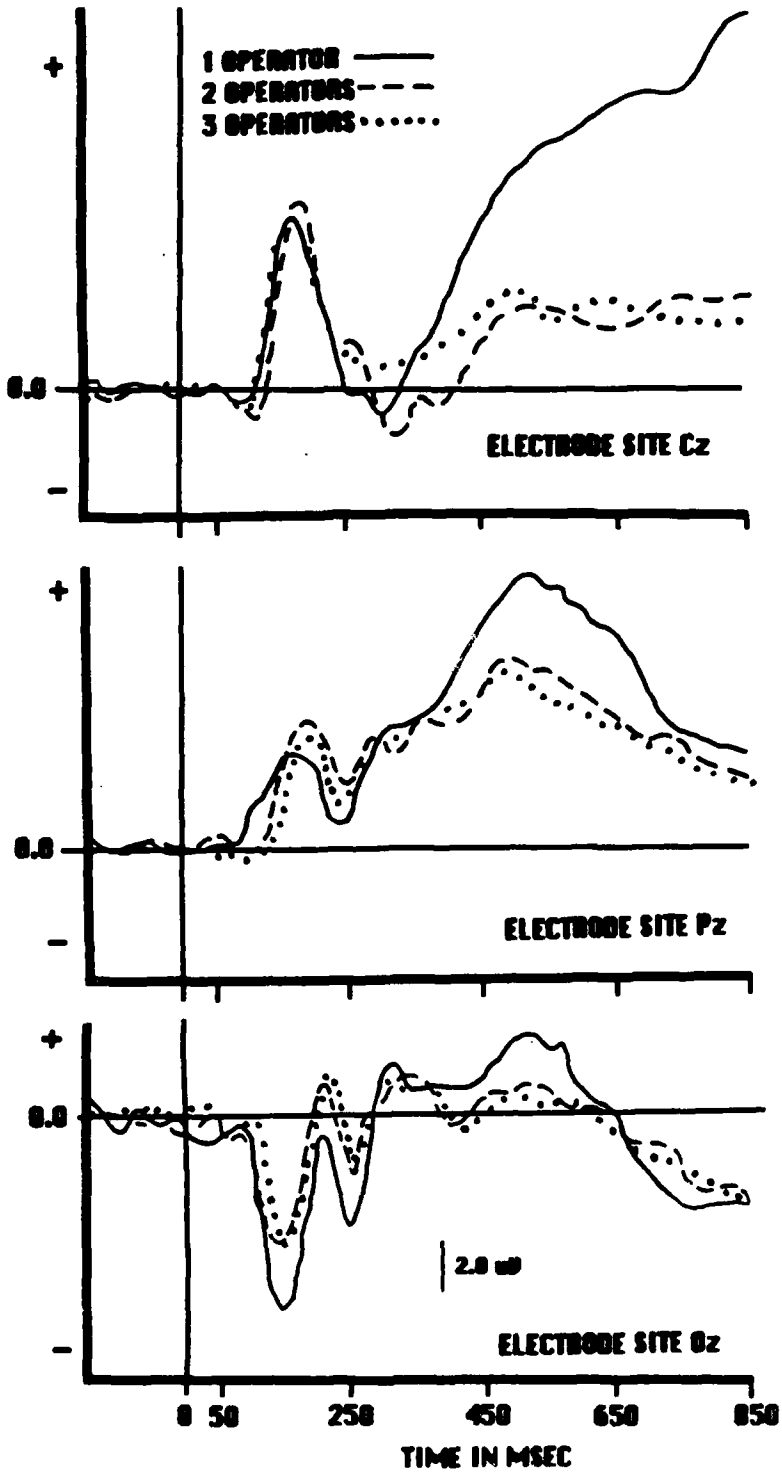


FIGURE 4. EVOKED POTENTIALS AVERAGED WITH 1.0 SEC EPOCHS

number equation blocks were averaged across subjects. Table 5 shows mean beats per minute, interbeat intervals (IBI's) and standard deviations of the interbeat intervals for the various experimental conditions. Also shown are the differences between the mean values for the letter and number equations (a negative value indicates that the magnitude of the value for the number condition is greater). Numbers in parentheses are standard deviations of the across subject means.

The heart rates, IBI's and standard deviations of the IBI's were compared across the resting condition and the three operator levels by use of repeated measures ANOVA's. None of the measures showed significant differences, i.e., none of the operator levels could be shown to change the heart rate, IBI or standard deviation of IBI's from the resting values.

TABLE 5 - HEART RHYTHM DATA

	MEAN BEATS PER MIN	MEAN INTER-BEAT INTERVAL	MEAN STANDARD DEVIATION OF INTER-BEAT INTERVALS
RESTING LEVEL	72.9 (13.5)	853.5 (182.7)	62.5 29.9)
ONE OPERATOR			
LETTERS-	71.1 (10.7)	862.9 (143.1)	63.1 (32.9)
NUMBERS-	71.1 (9.5)	858.7 (119.9)	63.8 (45.4)
DIFFERENCE-	0.0 (4.3)	4.2 (54.9)	-0.8 (27.0)
TWO OPERATOR			
LETTERS-	73.9 (11.4)	831.0 (135.1)	57.9 (34.3)
NUMBERS-	71.8 (10.6)	855.7 (134.9)	59.0 (33.4)
DIFFERENCE-	2.1 (4.7)	-24.7 (46.5)	-1.2 (18.4)
THREE OPERATORS			
LETTERS-	70.5 (11.4)	871.8 (139.4)	82.7 (53.0)
NUMBERS-	72.0 (10.0)	857.4 (133.4)	60.4 (37.2)
DIFFERENCE-	-1.6 (8.2)	14.4 (108.7)	22.3 (25.1)

XV. CONCLUSIONS: The goal of this project was to determine if physiological measures including EEG, heart rhythm and peripheral temperature would correlate with behavioral measures of performance and subjective ratings of workload on a 3 level mathematical task. We found that the levels of the task did indeed produce different reaction times and subjective ratings of workload, but neither heart rhythm or peripheral temperature was related systematically to task level. Except for the late components which were tied to motor reactions, EEG components also did not seem to show a systematic relationship to task level. There are, however, suggestive differences between early components and P300 amplitudes for the 1 versus 2 and 3

operator levels, but the 2 and 3 operator data appear almost identical. It is possible that a ceiling was reached at the transition from 1 to 2 operators, but it is more likely that either workload was not really changing as operator level changed for the highly practiced subjects used in this experiment, or that the physiological parameters measured were just not very good indicators of workload. More research will be needed to resolve these questions.

XVI. RECOMMENDATIONS:

1. The concept of integrating changes in physiological systems with behavioral and psychological measures to provide a better metric for workload is a very good one and should be continued. This approach will yield both basic and immediately applicable data that will be of value in the assessment of human performance.
2. To facilitate future projects in this area, computer programs should be acquired to allow the Fourier analysis of heart interbeat interval data. There is evidence that Fourier spectra could be useful in evaluating the degree of workload/stress being experienced by subjects.
3. The concept of the NWTB computer with built-in stimulus generation and data analysis capabilities is good, but the current system seems somewhat cumbersome and inflexible. For example, grand mean EEG data for 1.0 sec periods require data to be transferred from the NWTB to an IBM main frame computer and then back to the NWTB in the form of paper plots. Generation of 3.0 sec grand means could not be done with the NWTB at all.
4. Consideration should be given to developing an IBM PC-AT (or compatible) based system which would replace or supplement the NWTB. Such a system would take advantage of the off-the-shelf accessibility of current hardware and might allow the NWTB and CTS systems to be made available to many users in the form of software.
5. None of the sub-tests in the current CTS battery were specifically designed to elicit EEG or other physiological reactions. Consideration should be given to developing an additional sub-test (perhaps an odd-ball task) which would produce large amplitude P-300 responses from individual subjects. The addition of probe procedures to the CTS may, however, make a new sub-test unnecessary.
6. The validity of using physiological responses to assess workload rests on the assumption that a valid, independent measure of workload is available. Reaction times do not not necessarily measure workload, but the SWAT is designed

specifically to do this. Although data are available indicating that SWAT ratings are reliable, the seeming lack of certainty and arbitrariness of subjects when giving SWAT ratings does not inspire confidence in their validity. Since a valid measure of workload is the cornerstone of attempts to correlate changes in physiology and mental workload, emphasis should be placed on the continued development and validation of SWAT and/or SWAT-like procedures.

7. Preliminary tests conducted during my research program suggested that a significant decrement in performance and a corresponding increase in workload can be produced by making subjects monocular during the math processing task. Since delivery of monocular visual information is possible with certain types of aircraft displays, and temporary monocular vision could result from sun flash, laser exposure or exotropia caused by reduced oxygen tension at high altitudes, this effect should be studied in future experiments.

8. My final recommendation is to keep up the good work of the laboratory. I found the lab to be well managed and productive with a clear, if diverse, set of goals and objectives. It has been a pleasure to participate in the work of this laboratory.

REFERENCES

Acton, W.H. and Crabtree, M.S., Users Guide For The Criterion Task Set (U), AAMRL-TR-85-034, Armstrong Aerospace Medical Research Laboratory, Wright-Patterson AFB, OH, 1985.

Brookhuis, K.A., Mulder, G., Mulder, L.J.M. and Gloerich, A.B.M., "The P3 Complex as an Index of Information Processing: The Effects of Response Probability," Biological Psychology, 1983, pp. 277-296.

Eggemeier, F.T., "Considerations in the Application of Subjective Measures of Workload," In Brown, I.D., Goldsmith, R., Coombes, K. and Sinclair, M.A., (eds), Ergonomics International 85, London, Taylor and Francis, 1985, pp. 115-117.

Ferguson, G.A., Statistical Analysis in Psychology and Education, New York, McGraw-Hill, 1981.

Ginsberg, S., Heslegrave, R.J., Scher, H., Wong, K.K. and Furedy, J.J., "R-Wave Changes as a Function of Differing Degrees of Task Difficulty," Psychophysiology, 1980, p. 294 (Abstract).

Heffley, E. and Donchin, E., "Event-related Potentials Associated with Performance of Simple Mental Arithmetic," Society for Psychophysiological Research Abstracts, 1979, p. 173.

Hillyard, S.A. and Kutas, M., "Electrophysiology of Cognitive Processing," Annual Review of Psychology, 1983, pp. 33-61.

Mulder, G., "Mental Effort and its Measurement," In Hockey, G.R.J., Gaillard, A.W.K. and Coles, M., (eds), Energetics in Information Processing, Dordrecht, Reidel, In Press, 1986.

Mulder, G. and Mulder, L.J.M., "Information Processing and Cardiovascular Control", Psychophysiology, 1985, pp.392-401.

Poon, L.W., Thompson, L.W. and Marsh, G.R., "Average Evoked Potential Changes as a Function of Task Complexity," Psychophysiology, 1976, pp. 43-49.

Reid, G.B., "The Systematic Development of a Subjective Measure of Workload," In Brown, I.D., Goldsmith, R., Coombes, K. and Sinclair, M.A., (eds), Ergonomics International 85, London, Taylor and Francis, 1985, pp. 109-111.

Reid, G.B., Shingledecker, C.A., and Eggemeier, F.T., "Application of Conjoint Measurement to Workload Scale Development," Proc. Human Factors Society, 1981, pp. 522-526.

Reid, G.B., Shingledecker, C.A., Nygren, T.E. and Eggemeier, F.T., "The Development of Multidimensional Subjective Measures of Workload," Proc. 1981 International Conference on Cybernetics and Society, 1981, pp. 403-406.

Shingledecker, C.A., A Test Battery for Applied Human Performance Assessment, AFAMRL-TR-84-071, United States Air Force Aerospace Medical Research Laboratory, Wright-Patterson AFB, OH, 1984.

Shingledecker, C.A. and Schlegel, R.E., "A Resource Theoretical Test Battery for Workload and Human Performance," In Brown, I.D., Goldsmith, R., Coombes, K. and Sinclair, M.A., (eds), Ergonomics International 85, London, Taylor and Francis, 1985, pp. 112-114.

Sternberg, S., "Memory Scanning: Mental Processes Revealed by Reaction time Experiments," American Scientist, 1969a, pp. 421-457.

Sternberg, S., "The Discovery of Processing Stages: Extension of Donder's Method," Acta Psychologica, 1969b, pp. 276-315.

Wickens, C.D., Processing Resources and Attention, Dual Task Performance and Workload Assessment, Technical Memorandum EPL-81-3/ONR-81-3, Office of Naval Research, Arlington, VA, 1981.

Wilson, G.F., "A Neurological Test Battery for Workload Assessment," Proc. Human Factors Society, 1985, pp. 224-225.

1986 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT SUMMER SUPPORT PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

Application of Finite Element Analysis to Two Disparate Structural Problems: Thermomechanical Coupling and Optimal Sizing of Truss Members

Prepared by:	Richard W. Young
Academic Rank:	Associate Professor
Department and University:	Aerospace Engineering & Engineering Mechanics University of Cincinnati
Research Location:	AFWAL/Flight Dynamics Laboratory Structures and Dynamics Division Analysis and Optimization Branch Design and Analysis Methods Group (FIBRA)
USAF Researcher:	Dr. Vipperla B. Venkayya
Date:	29 August 1986
Contract No:	F49620-85-C-0013

Application of Finite Element Analysis to Two Disparate Structural Problems: Thermomechanical Coupling and Optimal Sizing of Truss Members

by

Richard W. Young

ABSTRACT

1. Thermomechanical coupling in viscoelastic solids.

The equations of thermal and dynamic equilibrium in polar coordinates of a viscoelastic solid are presented for the case of steady-state vibration. The equations are nonlinearly coupled due to temperature dependent material properties. Finite element techniques are used to reduce the system of seven second-order, nonlinear, partial differential equations to a partially linearized system of algebraic equations in 21 field and auxiliary variables. The equations are partially decoupled and an iterative solution procedure is developed. It is shown that traditional Gauss quadrature fails for this problem and an alternative Gauss quadrature rule is proposed.

2. Optimal sizing of truss members.

Minimum weight design of trusses is simplified by the fact that the elemental stiffness matrices are linear in the design variables (bar areas). This fact has been well exploited in optimality criteria approaches to iterative redesign algorithms. In the enforcement of the optimality criteria, however, dependence of the flexibility coefficients on the bar areas is neglected. This leads to a simple formulation but could result in a retarding of convergence. A reformulation to account for the missing dependence is presented. It can be used for an investigation of the computational trade-offs involved in the inclusion of the dependence.

ACKNOWLEDGMENTS

I would like to thank the Air Force Systems Command, Air Force Office of Scientific Research, for their sponsorship of the Summer Faculty Research/Graduate Student Summer Support programs.

I would also like to thank Dr. V. B. Venkayya of the AFWAL Flight Dynamics Laboratory for his diligent efforts to make optimal use of this program and to ensure that a productive work environment was is place for me and my colleagues. I thank the many individuals in the Design and Analysis Methods Group, especially Ms. Dedee Frantz, in ASIAC and in the computer center for their assistance.

Most importantly, I thank my wife, Laura, for her encouragement and extra hard work during this period of long hours and commuting.

HF

I. INTRODUCTION

My doctoral studies at Stanford University were on the thermomechanical coupling behavior of viscoelastic solids. My teaching responsibilities at the University of Cincinnati include graduate courses in finite element techniques. I have previously combined these topics in research conducted for AFRPL at Edwards AFB, California. In 1980 I attended a NATO ASI on structural optimization, an area in which I am just now developing a research effort.

These factors, the proximity of Wright-Patterson AFB to Cincinnati and a growing need for the consideration of thermal effects in structural design all led to my appointment as a Summer Fellow in the Design and Analysis Methods Group (FIBRA) in the Flight Dynamics Laboratory at Wright-Patterson AFB.

II. OBJECTIVES OF THE RESEARCH EFFORT

Preliminary work on the two-dimensional analysis of thermomechanical coupling has not led to meaningful numerical results. The primary objective of the research reported herein was to remedy that problem by reviewing the preliminary work, focusing upon:

(a) the appropriateness of the constitutive data

(b) the derivation and coding of the field equations

(c) the appropriateness of the finite element modeling.

A second goal of this research effort was to review work, done at the flight Dynamics Laboratory and elsewhere, in the field of combined structural/controls optimization. It is hoped that the previously discussed work will also find application to such structural problems as thermal source problems and modeling of structural damping.

III. THERMOMECHANICAL COUPLING: CONSTITUTIVE DATA

In cases of steady-state planar vibration of a solid without variation of the geometry or loading in the out-of-plane (z) direction, cylindrical components of displacement are expressed in terms of steady, in-phase and out-of-phase terms.

$$u(r,\theta,t) = \{u_1(r,\theta) + iu_2(r,\theta)\}e^{i\omega t} + u_3(r,\theta) \ . \tag{1}$$

$$v(r,\theta,t) = \{v_1(r,\theta) + iv_2(r,\theta)\}e^{i\omega t} + v_3(r,\theta) \ . \tag{2}$$

$$w(z,t) = \{(c_1 + ic_2)e^{i\omega t} + b\}z \ . \tag{3}$$

Equilibrium analysis shows that c_1 and c_2 must be zero.

Stress and strain components are likewise expressed in such terms and are further written in terms of volumetric and deviatoric components. The volumetric terms are

$$\sigma_V = \sigma_{rr} + \sigma_{\theta\theta} + \sigma_{zz} \tag{4}$$

and

$$\epsilon_V = \epsilon_{rr} + \epsilon_{\theta\theta} + \epsilon_{zz} \ . \tag{5}$$

The deviatoric terms are defined by

$$s_{ij} = \sigma_{ij} - \frac{1}{3}\sigma_V \delta_{ij} \tag{6}$$

and

$$e_{ij} = \epsilon_{ij} - \frac{1}{3}\epsilon_V \delta_{ij} \ , \tag{7}$$

where subscripts i and j stand for x, y and z in general and the Kronecker delta, δ_{ij}, is defined as unity if i=j and zero otherwise.

This investigator originally proposed constitutive relations of the form:

cyclic terms:

$$s_{ab1} + is_{ab2} = [G_1(\omega, T) + iG_2(\omega, T)](e_{ab1} + ie_{ab2}) \tag{8}$$

steady terms:

$$s_{ab3} = G_0 e_{ab3} \tag{9}$$

and volumetric term:

$$\sigma_V = E_K (\epsilon_V - 3\alpha(T - T_0)) \ . \tag{10}$$

During this research period it was decided that these formulations were not sufficiently general in that the steady terms were independent of temperature and that the the volumetric term did not allow for a phase shift between stress and strain. Therefore a new formulation was developed wherein the modulus G_o in equation (9) is treated as a function of temperature and equation (10) is replaced by

$$\sigma_{V1} + \sigma_{V2} = E_K(\omega, T)\{\cos\phi(\omega, T) + i\sin\phi(\omega, T)\}(\epsilon_{V1} + \epsilon_{V2}) \tag{11}$$

and

$$\sigma_{Vs} = E_{K0}\{\epsilon_{Vs} - 3\alpha(T - T_0)\} \ . \tag{12}$$

The effect of these changes is to roughly double the complexity of the equilibrium equations. It should be noted that the original formulation is retained as a special, probably physically unrealisitic, case of the general formulation.

With the help of ASIAC some original test data [1] refered to by previous investigators [2,3] and used in some detailed uniaxial analyses [4,5] was found. Unfortunately it does not look as if it will be a source of multidimensional constitutive data. The current work uses uniaxial data from [5], a reasonable guess for a variable Poisson's ratio and the cyclic elastic-viscoelastic analogue of Christensen [6].

IV. THERMOMECHANICAL COUPLING: REVIEW OF ANALYSIS AND CODING

The final form of the equations of dynamic equilibrium can be derived in the following basic stages:

(a) original statement in terms of stress field components

(b) use of constitutive equations to recast in terms of strain field components

(c) use of kinematics to rewrite in terms of displacement field components. The initial formulation (a),

$$\rho\ddot{u} = \sigma_{rr,r} + \frac{[\sigma_{r\theta,\theta} + \sigma_{rr} - \sigma_{\theta\theta}]}{r} , \tag{13}$$

$$\rho\ddot{v} = \sigma_{r\theta,r} + \frac{[\sigma_{\theta\theta,\theta} + 2\sigma_{r\theta}]}{r} \tag{14}$$

and

$$\rho\ddot{w} = \sigma_{zz,z} , \tag{15}$$

is simple enough and the strain-displacement relationships in (c),

$$\begin{pmatrix} \epsilon_{rr} \\ \epsilon_{\theta\theta} \\ \epsilon_{zz} \\ \epsilon_{r\theta} \\ \epsilon_{\theta z} \\ \epsilon_{zr} \end{pmatrix} = \begin{pmatrix} u_{,r} \\ \frac{(u+v_{,\theta})}{r} \\ w_{,z} \\ \frac{v_{,r}+(u_{,\theta}-v)/r}{2} \\ \frac{v_{,z}+(w_{,\theta})/r}{2} \\ \frac{(w_{,r}+u_{,z})}{2} \end{pmatrix} , \tag{16}$$

is not very complicated, but because the moduli used in equations (8,9,11 and 12) depend upon temperature, which is itself a function of r and θ, the differentiations called for in equations (13-15) give rise to quite a few terms.

For this reason the manual derivation of these equations was verified by comparison with the results of a symbolic manipulation program, MACSYMA [7], made

available by the computer center at the Flight Dynamics Laboratory. Except for the correction of one numerical coefficient the original derivations were verified. The final form of the equations remains somewhat cumbersome and space does not allow their inclusion in this report.

The energy balance equation is quite a bit simpler and was not affected by the modifications to the constitutive data, so it did not require verification by MACSYMA.

V. THERMOMECHANICAL COUPLING: FINITE ELEMENT FORMULATION

The fundamental concept in the least-squared-error finite element formulation is the approximation of field variables by local interpolation functions over subregions (the finite elements) of the domain being analysed. The interpolations are based upon point values of the field variables at a prechosen array of points (nodes) in the element. These nodal values become the primary unknowns (degrees of freedom) with respect to which measures of the error introduced by the interpolations are minimized.

This procedure requires the governing differential equations to be written in first-order form by introducing as "independent" quantities the first derivatives of the field variables and writing the second derivatives as first derivatives of the new (auxiliary) variables. This technique is discussed by Lynn and Arya [8].

The element chosen for this work is a hybrid iso- and sub-parametric triangular element which, in its regular geometry, is equilateral with nodes at its vertices (numbered 1,2,3) and at the middle of its sides (numbered 4,5,6 opposite vertices 1,2,3 respectively). The location of a point within the element can be given (redundantly)

in terms of three dimensionless coordinates, L_i, i=1,2,3, where

$$L_i = \frac{distance\ \ from\ \ point\ \ to\ \ side\ \ (i+3)}{distance\ \ from\ \ node\ \ i\ \ to\ \ side\ \ (i+3)} \ . \tag{17}$$

The auxiliary variables are linearly interpolated between the vertices by

$$q(\underset{\sim}{L}) = \begin{pmatrix} L_1 \\ L_2 \\ L_3 \end{pmatrix} \cdot \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} = \underset{\sim}{n}_1 \cdot \underset{\sim}{q} \tag{18}$$

where q_i is $q(\underset{\sim}{L})$ evaluated at node i. The original field variables are quadratically interpolated between all six nodes by

$$q(\underset{\sim}{L}) = \begin{pmatrix} 2L_1^2 - L_1 \\ 2L_2^2 - L_2 \\ 2L_3^2 - L_3 \\ 4L_2L_3 \\ 4L_3L_1 \\ 4L_1L_2 \end{pmatrix} \cdot \begin{pmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \\ q_5 \\ q_6 \end{pmatrix} = \underset{\sim}{n}_2 \cdot \underset{\sim}{q} \ . \tag{19}$$

These interpolation (or "shape") functions satisfy the necessary condition that each scalar function in the shape function vectors be unity at its "own" node and zero at all other nodes, that is

$$n_i(\underset{\sim}{L}_j) = \delta_{ij} \ . \tag{20}$$

Integration of the error measures over the domain of the problem must in general be carried out by quadrature. The quadrature rule suggested in Zienkiewicz [9] calls for equally weighted evaluations at the three midside nodes. During the debugging discussed in section IV the deficiencies of this quadrature finally became apparent. The first three scalar functions of $\underset{\sim}{n}_2$ must, by equation (20), necessarily be zero at all three of these points. There is, therefore, no way that the degrees of freedom associated with the field variables at the vertex nodes can affect the analysis. This in turn leads to singularities in the linear algebraic analysis.

For this reason an alternative Gauss quadrature suggested by Cowper [10] is now employed and the singularities are avoided. The new quadrature calls for equally weighted evaluations at points whose coordinates are

$$[\underset{\sim}{L}_1, \ \underset{\sim}{L}_2, \ \underset{\sim}{L}_3] = \frac{1}{6}\begin{pmatrix} 4 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 4 \end{pmatrix} \tag{21}$$

VI. THERMOMECHANICAL COUPLING: RECOMMENDATIONS

Despite the lack of good numerical output so far, this investigator remains convinced that the problem formulation is sound. The most likely sources of difficulty remain:

(a) further coding errors in the FORTRAN coding of the equations and iteration logic. These will have to be found by further debugging and perhaps by "desophistication" of the code, wherein attempts to save computations or storage are eliminated one-by-one to see if they are an error source.

(b) the constitutive data. Further search of the literature for appropriate test data is needed.

Once the code is debugged and running, it should be tested against several simple cases and then it should be ready for use in a predictive capacity: for specific problem analyses, for design work or for general study to see what kind of phenomena might be expected.

VII. TRUSS OPTIMIZATION: A MODIFIED ITERATION SCHEME

In seeking to minimize the weight of a truss, subject to constraints on displacements, stresses, compressive loads and the like, it is often advantageous to assume that the optimum, or at least a near optimal, design will occur when one or more

of the constraints are active. Such an optimality criterion approach is used by Khot, Berke and Venkayya [11]. A Lagrangian formulation introduces the active constraints and a Newton-Raphson iteration is used to enforce their activity.

The authors' optimality conditions are given by

$$\rho_i l_i A_i^2 = \sum_{j=1}^{p} \lambda_j E_{ij} \ , \qquad i = 1, m \ , \tag{22}$$

where i is the index number of a bar, j the index number of a constraint condition and ρ, 1 and A are the mass density, length and cross-sectional area (design variable) respectively of a bar. The Lagrange multiplier of the jth constraint is λ_j. E_{ij} is the flexibility coefficient for bar i and virtual load j (associated with the jth displacement constraint.) It is given by

$$E_{ij} = \underset{\sim}{r}_i^T [k_i] \underset{\sim}{s}_i^j A_i \ , \tag{23}$$

where $\underset{\sim}{r}_i$ and $\underset{\sim}{s}_i^j$ are the displacement degrees of freedom associated with bar i due to the physical and jth virtual load vectors respectively. The symmetric matrix $[k_i]$ is the elemental stiffness matrix associated with bar i. Importantly, it is linearly proportional to A_i and independent of all of the other bar areas.

Expanded to global (structural) size by zero fill, $[k_i]$ is written $[K_i]$. The global stiffness matrix of the structure is given by

$$[K] = \sum_{i=1}^{m} [K_i] \ . \tag{24}$$

The local displacement vectors are subsets of the global displacement vectors $\underset{\sim}{r}$ and $\underset{\sim}{s}^j$, which are found from the loads $\underset{\sim}{R}$ and $\underset{\sim}{S}^j$ by

$$[K]\{\underset{\sim}{r}, \ \underset{\sim}{s}^j\} = \{\underset{\sim}{R}, \ \underset{\sim}{S}^j\} \ . \tag{25}$$

It should also be noted that

$$\underset{\sim}{r}_i^T [k_i] \underset{\sim}{s}_i^j = \underset{\sim}{r}^T [K_i] \underset{\sim}{s}^j \ . \tag{26}$$

Finally, the constraint conditions are written

$$f_j = \sum_{i=1}^{m} \frac{E_{ij}}{A_i} - \overline{C}_j \ , \qquad j = 1, p \ , \tag{27}$$

where the $\overline{C}_j's$ are constants.

Application of the Newton-Raphson iteration requires calculation of partial derivatives of f_j with respect to the $\lambda's$. This is done by using the chain rule on equation (27),

$$\frac{\partial f_j}{\partial \lambda_q} = \sum_{i=1}^{m} \frac{\partial f_j}{\partial A_i} \frac{\partial A_i}{\partial \lambda_q} \ , \tag{28}$$

and using equation (22) to define the A's as functions of λ.

This procedure leads directly to the authors' result that

$$\frac{\partial f_j}{\partial \lambda_q} = -\frac{1}{2} \sum_{i=1}^{m} \frac{E_{ij} E_{iq}}{\rho_i l_i A_i^3} \tag{29}$$

if the dependence of E_{ij} upon A_i and, hence, upon λ is neglected in equation (22). This approximation leads to a fairly simple iterative algorithm, but it is not known what the cost might be in terms of retarded convergence. The following derivation sets forth the algebra required to account for the functional dependence $E_{ij}(A(\lambda))$.

Using equations (23, 24 and 26), equation (27) can be written as

$$f_j = \underline{r}^T [K] \underline{s}^j - \overline{C}_j \ . \tag{30}$$

We differentiate equation (30) with respect to λ_q and, noting from equation (25) that

$$\frac{\partial}{\partial \lambda_q} [\underline{r}, \ \underline{s}^j] = -[K]^{-1} \frac{\partial [K]}{\partial \lambda_q} [\underline{r}, \ \underline{s}^j] \ , \tag{31}$$

find that

$$\frac{\partial f_j}{\partial \lambda_q} = -\underline{r}^T \frac{\partial [K]}{\partial \lambda_q} \underline{s}^j \ . \tag{32}$$

We can write

$$\frac{\partial[K]}{\partial\lambda_q} = \sum_{n=1}^{m} \frac{\partial[K]}{\partial A_n}\frac{\partial A_n}{\partial\lambda_q} = \sum_{n=1}^{m} \frac{[K_n]}{A_n}\frac{\partial A_n}{\partial\lambda_q} \ . \tag{33}$$

We now use equation (22) to find the partial derivative of A_m with respect to λ_q. We first use equation (23) top rewrite equation (22) as

$$\rho_i l_i A_i = \sum_{j=1}^{p} \lambda_j \underline{r}^T [K_i] \underline{s}^j \ . \tag{34}$$

Differentiating with respect to λ_q

$$\rho_i l_i \frac{\partial A_i}{\partial\lambda_q} = \underline{r}^T [K_i] \underline{s}^q + \sum_{j=1}^{p} \lambda_j \frac{\partial}{\partial\lambda_q} \{\underline{r}^T [K_i] \underline{s}^j\} \ . \tag{35}$$

From equation (31) and the dependence of $[K_i]$ on A_i we have

$$\rho_i l_i \frac{\partial A_i}{\partial\lambda_q} = \underline{r}^T [K_i] \underline{s}^q + \sum_{j=1}^{p} \lambda_j \underline{r}^T \left(\frac{[K_i]}{A_i}\frac{\partial A_i}{\partial\lambda_q} - \frac{\partial[K]}{\partial\lambda_q}[K]^{-1}[K_i] - [K_i][K]^{-1}\frac{\partial[K]}{\partial\lambda_q} \right) \underline{s}^j \ . \tag{36}$$

From equation (34) it is seen that the coefficients of $\partial A_i/\partial\lambda_q$ on both sides of the equation are equal. Thus we are left with

$$\underline{r}^T [K_i] \underline{s}^q = \underline{r}^T \left(\frac{\partial[K]}{\partial\lambda_q}[K]^{-1}[K_i] + [K_i][K]^{-1}\frac{\partial[K]}{\partial\lambda_q} \right) \underline{s}^* \tag{37}$$

where

$$\underline{s}^* = \sum_{j=1}^{p} \lambda_j \underline{s}^j \ . \tag{38}$$

Finally we use equation (33) and are able to write

$$\underline{r}^T [K_i] \underline{s}^q = \underline{r}^T \left(\sum_{n=1}^{m} \frac{1}{A_n} \{[K_n][K]^{-1}[K_i] + [K_i][K]^{-1}[K_n]\} \frac{\partial A_n}{\partial\lambda_q} \right) \underline{s}^* \ . \tag{39}$$

Equation (39) provides a set of linear algebraic equations of the partial derivatives of the areas with respect to the Lagrange multipliers,

$$[B] \left(\frac{\partial \underline{A}}{\partial\lambda_q} \right) = \underline{C}^q \tag{40}$$

where

$$\underset{\sim}{A} = (A_1,\ A_2,\ A_3, \ldots,\ A_m)^T\ , \tag{41}$$

the i'th element of $\underset{\sim}{c}^q$ is

$$C_i^q = \underset{\sim}{r}_i^T [k_i] \underset{\sim}{s}_i^q \tag{42}$$

and the matrix $[B]$ has entries

$$b_{ij} = \frac{1}{A_j} \underset{\sim}{r}^T \{[K_i][K]^{-1}[K_j] + [K_j][K]^{-1}[K_i]\} \underset{\sim}{s}^* \ . \tag{43}$$

Numerical inversion of this system provides the sensitivities of the areas with respect to the Lagrange multipliers for use in equation (33) and thus in equation (32).

There is no question that use of equations (43), (33) and (32) will be far more computationally expensive than use of equation (29), although several efficiencies can be used to hold down the cost:

(a) equation (42) need only be evaluated on the elemental scale.

(b) the second matrix product in equation (43) is the transpose of the first.

(c) once $[K]^{-1}$ has been calculated, subsequent matrix products need be performed only on the elemental scale. Use of elemental dimensions affords a great savings. For a space truss having m bars and n nodes, the global stiffness matrix is (3n-6) square, but each of the elemental matrices will be 6x6. Further computation can be saved by writing

$$[K] = [L][D][L]^T\ , \tag{44}$$

$$[K]^{-1} = [L]^{-T}[D]^{-1}[L]^{-1}\ , \tag{45}$$

$$[K]^{-1} = \{[L]^{-T}[D]^{-\frac{1}{2}}\}\{[D]^{-\frac{1}{2}}[L]^{-1}\} \tag{46}$$

and finally

$$[K]^{-1} = \{[K]^{-\frac{1}{2}}\}^T [K]^{-\frac{1}{2}} \tag{47}$$

Then m products of the form

$$\{[k]^{-\frac{1}{2}}\}_{6x6}[k_i] = [\beta_i] \tag{48}$$

can be performed and equation (43) can be written as

$$b_{ij} = \frac{1}{A_j} \underset{\sim}{r}^T \{[\beta_i]^T [\beta_j] + [\beta_j]^T [\beta_i]\} \underset{\sim}{s}^* \ . \tag{49}$$

Finally it should be noted that once the changes in the Lagrange multipliers are calculated we can use them to recalculate areas as suggested in equation (40). Equation (22) should not be used here because A_i^2 is actually hidden in E_{ij} on the right-hand side and should drop out of the equation.

VIII. TRUSS OPTIMIZATION: RECOMMENDATIONS

It would be worthwhile to apply the logic developed in Section VII to the problem solved in [11]. In so doing it might be found that further efficiencies are possible with this approach and, even if not, comparison with the cost of the presently used Newton-Raphson scheme can be made.

REFERENCES

1. Jones, J., J. E. Fitzgerald and E. Francis, "Thermal Stress Investigation of Solid Propellant Grains," Report No. 578-F, Lockheed Propulsion Company, May 1963 (Contract AF04(611)-8013, Edwards AFB, California).

2. Schapery, R. A., "Effect of Cyclic Loading on the Temperature in Viscoelastic Media with Variable Properties," AIAA Journal 2, 1964, p. 827.

3. Huang, N. C. and E. H. Lee, "Thermomechanical Coupling Behavior of Viscoelastic Rods Subjected to Cyclic Loading," J. APPL. MECH. 34, (Trans. ASME 89 E), 1967, p. 127.

4. Mukherjee, S., "Thermal Response of a Viscoelastic Rod Under Cyclic Loading," J. Appl. Mech. 41, (Trans. ASME 96 E), 1974, p. 229.

5. Young, R. W., "Thermomechanical Response of a Viscoelastic Rod Driven by a Sinusoidal Displacement," IJSS 13, 1977, p. 925.

6. Christensen, R. M., Theory of Viscoelasticity, an Introduction, Academic Press, New York, 1971.

7. VAX UNIX MACSYMA™ Reference Manual, Release 309.1, Document SMI0501030.011, Symbolics, Inc., 1985.

8. Lynn, P. P. and S. K. Arya, "Finite Elements Formulated by the Weighted Discrete Least Squares Method," IJNME 8, 1974, p. 71.

9. Zienkiewicz, O. C. The Finite Element Method, Ed. 3., McGraw-Hill, London, 1977.

10. Cowper, G. R., "Gaussian Quadrature Rules for Triangles," IJNME 7, 1973, p. 405.

11. Khot, N. S., L. Berke and V. B. Venkayya, "Comparison of Optimality Criteria Algorithms for Minimum Weight Design of Structures," AIAA Journal 17 #2, 1979, p. 182.

1986 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/
GRADUATE STUDENT SUMMER SUPPORT PROGRAM

Sponsored by the
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
Conducted by the
Universal Energy Systems, Inc.

FINAL REPORT

(M,N)-APPROXIMATION: A SYSTEM SIMPLIFICATION METHOD

Prepared by:	Ajmal Yousuff
Academic Rank:	Assistant Professor
Department and	Department of Mechanical Engineering and Mechanics
University:	Drexel University
Research Location:	AFWAL/FIGC Wright-Patterson AFB, Dayton, OH
USAF Researcher:	Lt. Timothy E. McQuade
Date:	September 12, 1986
Contract No.:	F 49620-85-C-0013

(M,N)-APPROXIMATION: A SYSTEM SIMPLIFICATION METHOD

by

Ajmal Yousuff

ABSTRACT

This report presents a method for the simplification of systems made up of N interacting subsystems, by approximating only M of the N subsystems. It is proposed to reduce each of the M subsystems while all interactions are active, and to preserve the identities of all subsystems in the overall approximation, as demanded in many situations. Problems of conventional model reduction, controller or compensator reduction, reduction of finite element models, reduction of decentralized controllers, etc., can be treated as specific applications of this method. In this preliminary report, conventional model reduction methods, but constrained to meet the above objectives, are employed. As an application to controller reduction, a new version of balanced controllers is shown to be generated by this method. A numerical example is included which compares the new method with other existing balancing methods, and the results are shown to be in favor of the new method.

ACKNOWLEDGMENTS

I would like to thank the Air Force Systems Command and the Air Force Office of Scientific Research for sponsoring my research. I am grateful to the Flight Dynamics Laboratory for providing an excellent research environment and would like to express my appreciation to all personnels in the group AFWAL/FIGCA for their hospitality and assistance during my summer research period.

I would also like to thank the research group of Professor Skelton for bringing to my attention the work of Liu and Anderson.

In view of my marriage next week, I would like to thank my fiancée Ruth for making all the wedding arrangements.

I. INTRODUCTION:

Systems, such as large space structures, consist of several interacting subsystems, often requiring decentralized control. These decentralized controllers also can be viewed as subsystems of the overall closed loop system. Moreover, as demanded by the mission objectives of recent vehicles, such as high speed and supermaneuverable aircrafts, and also due to hardware limitations, controllers are required to be simple. Consequently, a need for simplifying systems consisting of interacting subsystems arises.

I have been involved in the control of flexible structures, including large space structures, with particular attention to model and controller reduction, during and after my Ph.D. at Purdue University. The Control Analysis Group (AFWAL/FIGCA) of Flight Dynamics Laboratory at Wright Patterson Air Force Base has been exploring the applicability of H^∞ compensator synthesis technique[1] to real world problems. The complexity of the controllers resulting from this technique has made the Control Analysis Group investigate means of simplifying these controllers. As a consequence, the Control Analysis Group's interest in controller reduction and my desire to research reduction of interconnected systems (of which controller reduction is a special case), has led to my being assigned to the Control Analysis Group of Flight Dynamics Laboratory.

In this report, the objectives (Section II) of the research during the summer period, the approaches taken and the results obtained(Sections III and IV) are presented. Section V concludes this report with some recommendations for future research.

II. OBJECTIVES OF THE RESEARCH EFFORT:

Since the reduction of H^∞ compensators is a special case of approximation of interconnected systems, and since both H^∞ compensator synthesis technique and H^∞-norm minimizing approximation require internal balancing, the following objectives were set forth for the summer period:

A. Development of a procedure to approximate interconnected systems;

B. Application to LQG-controller reduction by balancing;

C. Application to reduction of H^∞ compensators.

The rest of the report presents the approaches taken to realize Objectives A and B and the results obtained. Objective C could not be attempted during the summer period, and

is suggested for future work. The organization of the report is as follows: Section III formulates the approximation of interconnected systems as an (M,N)-approximation problem, and presents a procedure to solve it. Section IV presents the controller reduction problem as a (1,2)-approximation problem, and develops a new reduction scheme based on internal balancing[2]. A numerical example is included to compare the new scheme with other existing balancing methods.

III. (M,N)-APPROXIMATION:

Definition: An (M,N)-approximation is defined as an approximation of a system consisting of several interacting subsystems by reducing the models of only M ($M < N$) subsystems, with due regard to their simultaneous interactions.

Note that the conventional model reduction is a (1,1)-approximation, since, herein, the system itself is its subsystem (i.e. N=1), and this entire system is reduced (hence M=1) without regard to its internal structure. A controller reduction is a (1,2)-approximation since the overall system consists of N-2 subsystems - the plant and the controller - and only the controller (hence M=1) needs to be simplified.

In order to present our approach* to (M,N)-approximation, we begin with (1,1)-approximation, the standard model reduction. Consider the linear system S(n) given below

$$\dot{x} = Ax + Bw \; ; x \in \mathbb{R}^n$$
$$y = Cx + Dw \; ; y \in \mathbb{R}^k$$

Now, depending upon the nature of inputs w (such as Gaussian, energy bounded, etc.), one could use an appropriate reduction method (such as q-COVER[4], internal balancing[2], Hankel-norm approximation[5], etc.) to obtain an approximation $\underline{S}(r<n)$ that preserves certain properties of S(n) (such as root-mean-square values, covariance sequences, grammians, Hankel norm, etc.). Let such an $\underline{S}(r)$ be

$$\dot{\underline{x}}_r = \underline{A}_r \underline{x}_r + \underline{B}_r w \; ; \; \underline{x}_r \in \mathbb{R}^r$$
$$\underline{y}_r = \underline{C}_r \underline{x}_r + \underline{D}_r w \; ; \; \underline{y}_r \in \mathbb{R}^k$$

---

* The method proposed herein for (M,N)-approximation is a generalization of that used in [3].

Many of the approximation schemes that use state space representations proceed as follows. Transform S(n) to a suitable set of coordinates/basis as demanded by the properties to be preserved in $\underline{S}(r)$. Let $x=\underline{T}\underline{x}$ be the transformation. Such a $\underline{T}$ is obtained, for example, via internal balancing[2], cost-decoupling[6], etc.. Let

$$[\underline{T}_r\ \underline{T}_t] = \underline{T} = \underline{L}^{-1} = [\underline{L}_r^T\ \underline{L}_t^T]^T\ ;\ \underline{T}_r \in \mathbb{R}^{(n,r)} \quad \& \quad \underline{x} = [\underline{x}_r^T\ \underline{x}_t^T]^T$$

Then $\underline{S}(r)$ is produced from the representation of S(n) in the basis $\underline{x}$. It is essential to point out that the basic need for computing $\underline{T}$ is to identify the states $\underline{x}_r$ and $\underline{x}_t$ as $\underline{x}_r \in \mathrm{Im}(\underline{T}_r)$ and $\underline{x}_t \in \mathrm{Ker}(\underline{L}_r)$. Then $\underline{S}(r)$ is obtained either (a) by deleting $\underline{x}_t$, as in balancing[2], q-COVER[4], etc.; or (b) by absorbing some information about $\underline{x}_t$ in $\underline{x}_r$ and then deleting $\underline{x}_t$, as in singular perturbation[7], Hankel norm approximation[5], etc.. In this report we will focus only on the former case(a), in which case the parameters $(\underline{A}_r, \underline{B}_r, \underline{C}_r, \underline{D}_r)$ are obtained as $\underline{A}_r = \underline{L}_r A \underline{T}_r$, $\underline{B}_r = \underline{L}_r B$, $\underline{C}_r = C\underline{T}_r$, $\underline{D}_r = D$. With some abuse of language, they also can be viewed as $\underline{A}_r = \underline{P}A\underline{P}$, $\underline{B}_r = \underline{P}B$, $\underline{C}_r = C\underline{P}$, where $\underline{P} = \underline{T}_r\underline{L}_r$ is an oblique projection (since $\underline{L}_r\underline{T}_r = \underline{I}_r$) on $\mathrm{Im}(\underline{T}_r)$ along $\mathrm{Ker}(\underline{L}_r)$. Clearly, $\mathrm{Im}(\underline{P}) = \mathrm{Im}(\underline{T}_r)$ and $\mathrm{Ker}(\underline{P}) = \mathrm{Ker}(\underline{L}_r)$. (Strictly speaking, the above equations are not defined, since for example, the maps $\underline{B}_r$ and $\underline{P}B$ are defined on different but isomorphic spaces). From this point of view many of the existing approximation schemes basically identify the subspaces $\mathrm{Im}(\underline{P})$ and $\mathrm{Ker}(\underline{P})$, since a projection is uniquely defined by its image and kernel.

We will now formulate the (M,N)-approximation problem in this context. Suppose that S(n) consists of N interacting subsystems $S_i(n_i)$, each described by its associated states x_i, i=1,2,...N, and that there are N_w input vectors w_i and N_y output vectors y_i, such that

$$x^T = [x_1^T, x_2^T, \ldots, x_N^T]\ ;\ w^T = [w_1^T, w_2^T, \ldots, w_{N_w}^T]\ ;\ y^T = [y_1^T, y_2^T, \ldots, y_{N_y}^T]\ ;$$

$$x_i \in \mathbb{R}^{n_i}\ ; \qquad w_i \in \mathbb{R}^{m_i}\ ; \qquad y_i \in \mathbb{R}^{k_i}\ ;$$

$$n = \sum_{i=1}^{N} n_i\ ; \qquad m = \sum_{i=1}^{N_w} m_i\ ; \qquad k = \sum_{i=1}^{N_y} k_i\ .$$

The S(n) is then written in its subsystem form as

$$\dot{x}_i = \sum_{j=1}^{N} A_{ij}x_j + \sum_{j=1}^{N_w} B_{ij}w_j \ , \ i=1,2, \ldots ,N$$

$$y_i = \sum_{j=1}^{N} C_{ij}x_j + \sum_{j=1}^{N_w} D_{ij}w_j \ , \ i=1,2, \ldots ,N_y \ .$$

In order to obtain an approximation S(r) which preserves the identities of all subsystems $S_i(.)$, i=1,2,....N, we seek S(r) in the form

$$\begin{aligned} \dot{x}_{ir} &= \sum_{j=1}^{N} A_{ijr}x_{jr} + \sum_{j=1}^{N_w} B_{ijr}w_j \ ; \ x_{ir} \in \mathbb{R}^{r_i} \ , \ 0 \le r_i \le n_i \ , \ i=1,2, \ldots ,N \\ y_{ir} &= \sum_{j=1}^{N} C_{ijr}x_{jr} + \sum_{j=1}^{N_w} D_{ijr}w_j \ ; \ y_{ir} \in \mathbb{R}^{n_i} \ , \qquad i=1,2, \ldots ,N_y \end{aligned} \tag{1}$$

so that each state vector x_{ir} is associated with $S_i(r_i)$ which is an approximation of $S_i(n_i)$ and $r = r_1+r_2+\ldots+r_N$. Note that (a) $r_i = 0$ implies that $S_i(n_i)$ is approximated by a static system, and if in addition $D_{ij} = 0$ for all j, then $S_i(n_i)$ is completely ignored in the overall approximation S(r); and (b) $r_i = n_i$ implies that $S_i(n_i)$ is retained as it is (without any reduction) in S(r). In an (M,N)-approximation (N-M) r_is should satisfy $r_i = n_i$.

Note that in the development of the oblique projection $\underline{P}$ (through the transformation $\underline{T}$), eventhough the interactions among all subsystems is taken into account, the requirement that the approximation $\underline{S}(r)$ should be of the form (1) is not enforced. As a consequence, $\underline{S}(r)$ does not preserve the identities of the subsystems. One way of enforcing this requirement is to constrain the oblique projection to be of the form

$$P = \text{block diag}(P_1, P_2, \ldots , P_N) \tag{2}$$

where (N-M) P_is are identity maps, and the remaining M arbitrary projections P_is satisfy $P_i = T_{ir}L_{ir}$ for some T_{ir} and L_{ir} satisfying $L_{ir}T_{ir} = I_{ri}$. This would yield $A_{ijr} = L_{ir}A_{ij}T_{jr}$, $B_{ijr} = L_{ir}B_{ij}$, $C_{ijr} = C_{ij}T_{jr}$, thus clearly indicating the approximation of $S_i(n_i)$ by $S_i(r_i)$.

Thus far two oblique projections have been presented: (1) the (1,1)-approximation projection $\underline{P}$ which focuses on the properties to be preserved in S(r), but does not satisfy the sufficient condition (2) required of an (M,N)-approximation; (2) the (M,N)-approximation projection P that satisfies (2) but does not consider the properties. therefore, since there are M arbitrary P_is in (2), we choose these P_is in P such that P approximates $\underline{P}$ in a suitable norm sense. In order to obtain such a P we propose the following two steps:

1. Determine X = block diag($X_1, X_2, \ldots X_N$); X_i:(n_i,n_i) to minimize

$$\delta = \| \underline{P} - X \|_2 \qquad (3)$$

2. Determine $P_i = T_{ir}L_{ir}$ such that $L_{ir}T_{ir} = I_{ri}$. and $\mathrm{Im}(P_i) = \mathrm{Im}(X_i)$ and $\mathrm{Ker}(P_i) = \mathrm{Ker}(X_i)$.

The solution X to minimize (3) is not unique[8]: there exists a central solution, which is $X_i = \underline{P}_{ii}$, i = 1,2,...N, where $\underline{P}_{ii}$ is the (i,i) block entry of $\underline{P}$; and several non-central solutions. The central solution is also a 1-norm and ∞-norm minimizing solution. Obtaining the non-central solution for (1,2)-approximation becomes a constant matrix dilation problem, for which solutions are known. However, the non-central solution to general (M,N)-approximation seems to be an open problem.

For any solution (central or non-central) X-block diag(X_i) in step1. above, the explicit construction of P_i, i=1,2,....N in step2. can be obtained as given below.

Theorem 1. Given X_i, the projection $P_i = T_{ir}L_{ir}$ such that $L_{ir}T_{ir} = I_{ri}$. and $\mathrm{Im}(P_i) = \mathrm{Im}(X_i)$ and $\mathrm{Ker}(P_i) = \mathrm{Ker}(X_i)$ is constructed as

$$L_{ir} = V_{1i}^T ; \quad T_{ir} = U_{1i}S_i^{-1}$$

where U_{1i}, V_{1i}, and S_i are obtained from the singular value decomposition of X_i:

$$X_i = [U_{1i}\, U_{2i}] \begin{bmatrix} \Sigma_i & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_{1i}^T \\ V_{2i}^T \end{bmatrix} \quad ; S_i = V_{1i}^T U_{1i} \; .$$

Corollary 1. If X_i is idempotent for some i, then the P_i obtained by Theorem1 satisfies $P_i = X_i$.

We summarize our approach to solving the (M,N)-approximation problem in the following procedure:

(M,N)-approximation Procedure:

Step1: Construct the parameters (A,B,C,D).

Step2: Determine $\underline{P}$.(This step is problem dependent; namely, it depends on the properties to be preserved in S(r).

Step3: Determine the structure of P. (This depends on the internal structure of the system).

Step4: Obtain P_i to minimize $\| \underline{P} - P \|$.

(a1) For central solution, set $X_i = \underline{P}_{ii}$, i = 1,2, ... ,M.

(a2) For (1,2)-approximation only: for non-central solution , compute X_i by using the expressions given in[1].

(b) Compute L_{ir} and T_{ir} according to Theorem1.

Step5: Obtain the parameters $(A_{ijr}, B_{ijr}, C_{ijr})$ of $S_i(r_i)$ as $A_{ijr} = L_{ir}A_{ij}T_{jr}$, $B_{ijr} = L_{ir}B_{ij}$, $C_{ijr} = C_{ij}T_{jr}$.

IV. APPLICATION TO LOG-CONTROLLER REDUCTION BY BALANCING:

Let $S_1(n_1)$ be a model of a plant:

$$\dot{x}_1 = A_{11}x_1 + B_1w_1 + Ey_2$$
$$y_1 = C_{11}x_1 \qquad w_i \in \mathbb{R}^{m_i};\ y_i \in \mathbb{R}^{k_i},\ i = 1,2$$
$$z = Mx_1 + w_2$$

where w_1 is a zero-mean white Gaussian disturbance with intensity $W_1 > 0$, y_2 are the control variables, y_1 are the controlled variables, and z are the measurements corrupted by the zero-mean Gaussian noise w_2 with intensity $W_2 > 0$. Also let the LQG-controller $S_2(n_2)$

$$\dot{x}_2 = A_{22}x_2 + Fz\ ; \qquad x_2 \in \mathbb{R}^{n_2}$$
$$y_2 = Gx_2$$

be minimal with $n_2 = n_1$ and minimize

$$\lim_{t\to\infty} E\int_0^{\infty} \{y_1^T(\sigma)Q_1y_1(\sigma) + y_2^T(\sigma)Q_2y_2(\sigma)\}d\sigma$$

We wish to obtain a reduced controller $S_2(r_2)$ from $S_2(n_2)$ so that $S_2(r_2)$ yields $y_{1r}(t) = y_1(t)$ and $y_{2r}(t) = y_2(t)$. This is a (1,2)-approximation. We now illustrate Steps 1,2 and 3 of the (M,N)-approximation procedure.

Step1: In order to include the interactions between $S_1(n_1)$ and $S_2(n_2)$ in the reduction process, we first need to 'close the loop' as below. In closed loop form the overall system S(n) has dimension $n = n_1+n_2 = 2n_1$ with two subsystems. The parameters in the subsytem representation of S(n) are now given by $A_{11}=A_{11}$. $A_{12}=EG$. $B_{11}=B_1$. $B_{12}=0$. $A_{21}=FM$. $A_{22}=A_{22}$. $B_{21}=0$. $B_{22}=F$. $C_{11}=C_{11}$. $C_{12}=0$. $C_{21}=0$. $C_{22}=G$. $D_{ij}=0$:i,j=1,2.

Step2: Since we wish to use internal balancing[1], we elaborate this step as shown below.

a. Solve the following equations for Y and Z:

$$A^TY + YA + C^TQC = 0 \; ; \; Q = \text{block diag}\{Q_1, Q_2\}$$

$$AZ + ZA^T + BWB^T = 0 \; ; \; W = \text{block diag}\{W_1, W_2\}$$

b. Obtain a $\underline{T}$ such that

$$\underline{T}^{-1}Z\underline{T}^{-T} = \underline{T}^TY\underline{T} = \Sigma = \text{diag}\{\sigma_1, \sigma_2, \ldots, \sigma_{2n_1} ; \sigma_i \geq \sigma_{i+1}\}$$

c. Partition $\underline{T}$ and its inverse as

$$\underline{T} = [\underline{T}_r \; \underline{T}_t] \; ; \; \underline{T}^{-1} = \begin{bmatrix} \underline{L}_r \\ \underline{L}_t \end{bmatrix} \; ; \; \underline{T}_r \in \mathbb{R}^{(2n_1, n_1+r_2)} \; ; \; \underline{L}_r \in \mathbb{R}^{(n_1+r_2, 2n_1)}$$

d. Compute $\underline{P} = \underline{T}_r\underline{L}_r$.

Step3: Since we do not wish to reduce the model of the plant, we set $P_1 = I_{n1}$. Hence, the required structure for P is $\text{diag}(I_{n1}, P_2)$.

Proceeding with the (M,N)-approximation procedure, we obtain the reduced controller as

$$\dot{x}_{2r} = L_{2r}A_{22}T_{2r}x_{2r} + L_{2r}Fz$$

$$y_{2r} = GT_{2r}x_{2r}.$$

Note that for both central and non-central solution to minimizing $\| \underline{P} - P \|$, we have

$$\min \| \underline{P} - P \| \geq \max \{ \| \underline{P}_{11} - I \|, \left\| \begin{matrix} \underline{P}_{11} - I \\ \underline{P}_{12} \end{matrix} \right\| \}$$

Hence, the natural question that arises is 'does there exist a block diagonal $\underline{T}$ (hence, block diagonal $\underline{P}$) which balances the closed loop system?'. We do not believe that such a $\underline{T}$ exists. However, there exists a

$$\underline{T} = \begin{bmatrix} T & T \\ 0 & T \end{bmatrix}$$

such that the controllability and observability grammians of the closed loop system are "nearly" balanced in the sense that the grammians have the following structures respectively:

$$\begin{bmatrix} * & 0 \\ 0 & \Sigma_2 \end{bmatrix}, \quad \begin{bmatrix} * & -\Sigma_2 \\ -\Sigma_2 & \Sigma_2 \end{bmatrix},$$

where $\Sigma_2 = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_{n2})$. One can then look at the magnitudes of σ_is to determine the order of the reduced controller and which states to be deleted. This $\underline{T}$ yields

$$\underline{P} = \begin{bmatrix} I & \underline{P}_2 - I \\ 0 & \underline{P}_2 \end{bmatrix}, \text{ where } \underline{P}_2 \text{ is obtained from the partition of } T \text{ and } T^{-1}.$$

Then the central solution in Step4 of (M,N)-approximation procedure yields $P_2 = \underline{P}_2$, and $\min \| \underline{P} - P \| \geq \| I - \underline{P}_2 \|$. It turns out that the resulting reduced controller can be obtained from balancing the following representation of $S_2(n_2)$.

$$\dot{x}_2 = (A_{11} - EG)x_2 + F\tilde{z}\,; \quad \tilde{z}: \text{zero-mean, with intensity } W_2$$

$$y_{e2} = \begin{bmatrix} \sqrt{Q_1}\, C_{11} \\ \sqrt{Q_2}\, G \end{bmatrix} x_2$$

We present this methodology in the form of the following algorithm:

Balanced Controller Reduction 2:

Step1: Solve for Y_{22} and Z_{22} from

$$(A_{11} - EG)^T Y_{22} + Y_{22}(A_{11} - EG) + C_{11}^T Q_1 C_{11} + G^T Q_2 G = 0$$

$$(A_{11} - EG) Z_{22} + Z_{22}(A_{11} - EG)^T + F W_2 F^T = 0$$

Remark: Y_{22} is the regulator solution; Z_{22} is the covariance of the controller states.

Step2: Compute T such that

$$T^{-1}Z_{22}T^{-T} = T^{T}Y_{22}T = \Sigma_2 = \text{diag}\{\sigma_1,\sigma_2,\ldots,\sigma_{n_1} ; \sigma_i \geq \sigma_{i+1}\}$$

Step3: Partition

$$T = [T_{2r}\ T_{2t}]\ ;\ T^{-1} = \begin{bmatrix} L_{2r} \\ L_{2t} \end{bmatrix}.$$

Step4: Construct the reduced controller as

$$\dot{x}_{2r} = L_{2r}A_{22}T_{2r}x_{2r} + L_{2r}Fz$$
$$y_{2r} = GT_{2r}x_{2r}.$$

There exist several balancing schemes for reducing LQG-controllers [9-13]. We wish to compare the new method developed in this report, with a few of these balancing schemes.

Example: The example presented herein is taken from [14], and converted to an LQG setting.The system contains four states: $x_1(t) = u_x(t)$ - incremental speed along x-axis; $x_2(t) = u_z(t)$ - speed along z-axis; $x_3(t) = \theta(t)$ - pitch angle; $x_4(t) = q(t)$ - pitch rate. The control input to the system is $y_2(t) = \delta_e(t)$ - elevator deflection. The control objective is to maintain the pitch close to commanded pitch without excessive control deflection. The measurement available for control law implementation is assumed to be $z = \theta(t)$. To cast this problem in the LQG setting, it is assumed that a zero-mean Gaussian noise $w_1(t)$ enters into the system through the control inputs, and that the measurements are corrupted with the zero-mean Gaussian noise $w_2(t)$. The numerical values for the system parameter can be obtained from [14]. Additional parameters pertaining to the LQG setting are: $n_1=n_2=4$, $W_1=1.$, $W_2=1.0E-04$, $Q_1=1.$, and $Q_2=\rho$ a variable design parameter.

The following methods have been compared:

BRRA- Balanced Riccati Reduction Algorithm. [9,10]

BCRA- Balanced Controller Reduction Algorithm. [11]

BCR2- Balanced Controller Reduction 2 (The new method)

BAND- Balanced Anderson's Method [13].

Results and Observations:

The results are presented in Figures 1-6. Figure.1 displays the effect of reducing the order of the controller by BRRA on the trade-off between regulation (RMS value of

pitch) and the control (RMS value of elevator deflection). The effect of reduction is noticeable only in the case of r_2=1; for r_2=3,2 the reduced controllers yield essentially the same performance as the full order controller. This same observation was made with all the above controller reduction methods. Hence, a comparison of different methods was made only with respect to the first order controllers. These results are presented in the remaining figures. Figure 2 shows the performance achievable from first order controllers derived using the three methods: BRRA, BCRA, and BCR2; the solid line corresponds to the full order LQG controller, hence, the optimal performance. Figures 3-6 show the pitch response and control deflection for a command input of one degree. Figures 3 and 4 correspond to reduced controllers produced from the LQG controller for ρ = 10 and 1, respectively. For ρ = 1, BRRA yielded unstable system, and is not included in the figure. Notice, therefore, from Figures 2-6, that the new method consistently yields better controllers. Figure 5 shows the effect of using central and non-central solution to the norm minimization problem embedded in the (M,N)-approximation procedure; the (1,1)-approximation projection $\underline{P}$ was produced from balancing the full order(n=8) closed loop system. The results from BCR2 are also included for comparison. All the three methods produce fairly close performances. The last figure compares the methods BCR2 and BAND - the results are in favor of BCR2; in particular, with respect to control deflection.

V. RECOMMENDATIONS:

a. The controller reduction problem has been viewed as a special case of the general (M,N)-approximation problem, and a new version of balanced controller reduction has been developed. The results from a numerical example illustrate that this method is better than many existing methods. The results and the methodology have been derived only for LQG controller reduction. Extension of this work to reduction of H^∞ compensators should be investigated.

b. The development of the (M,N)-approximation procedure needs to be justified more rigorously; for instance, the structure of the (M,N)-approximation projection P in (2) is only sufficient but not necessary.

c. There exist several schemes for controller reduction; only balancing method has been reported. The effect of employing other schemes is a fruitful area to explore.

d. Application of (M,N)-approximation to other problems, such as finite element model reduction, reduction of decentralized control, warrants further investigation.

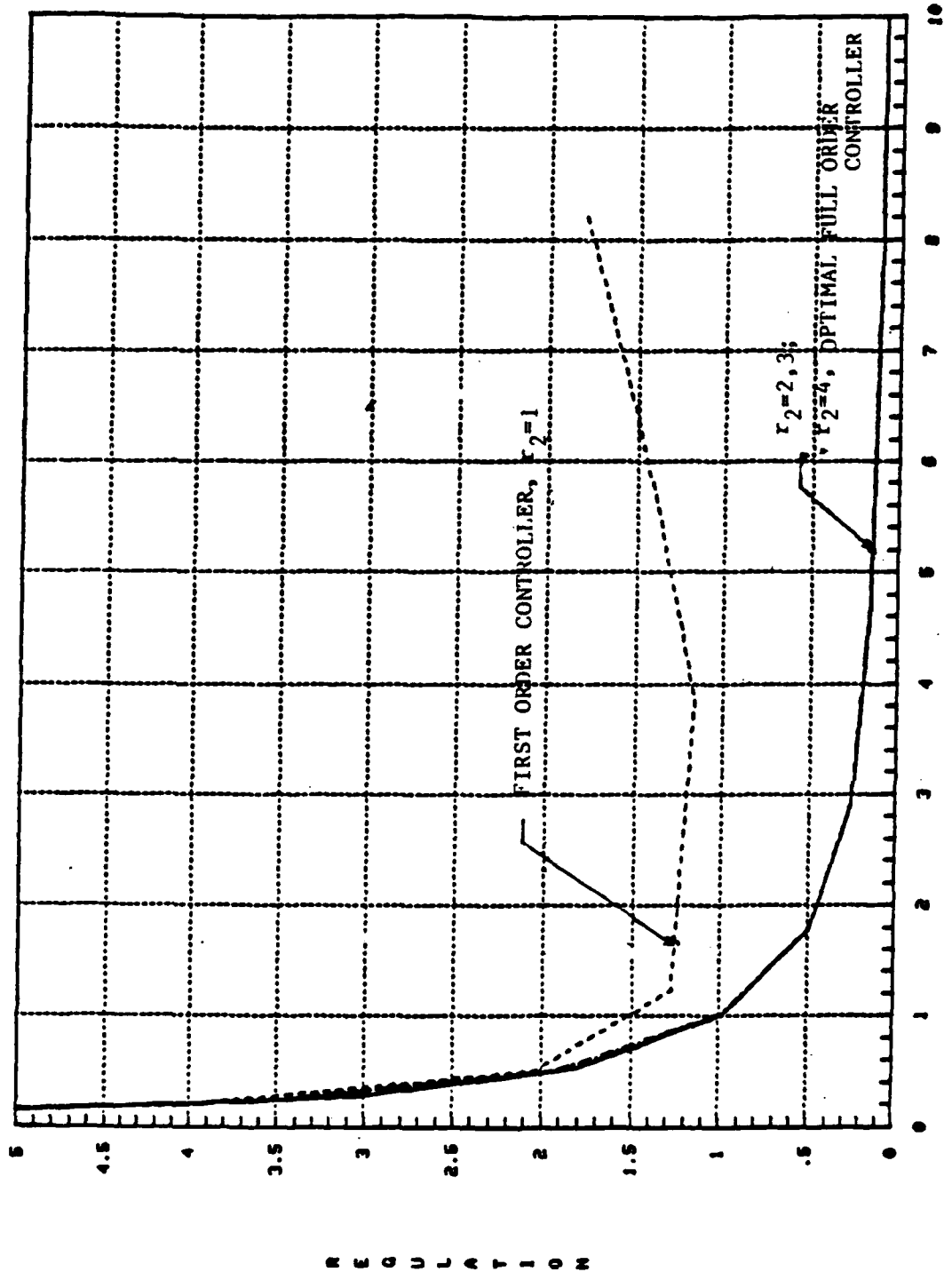


FIG.1 PERFORMANCE PLOT (BODE)

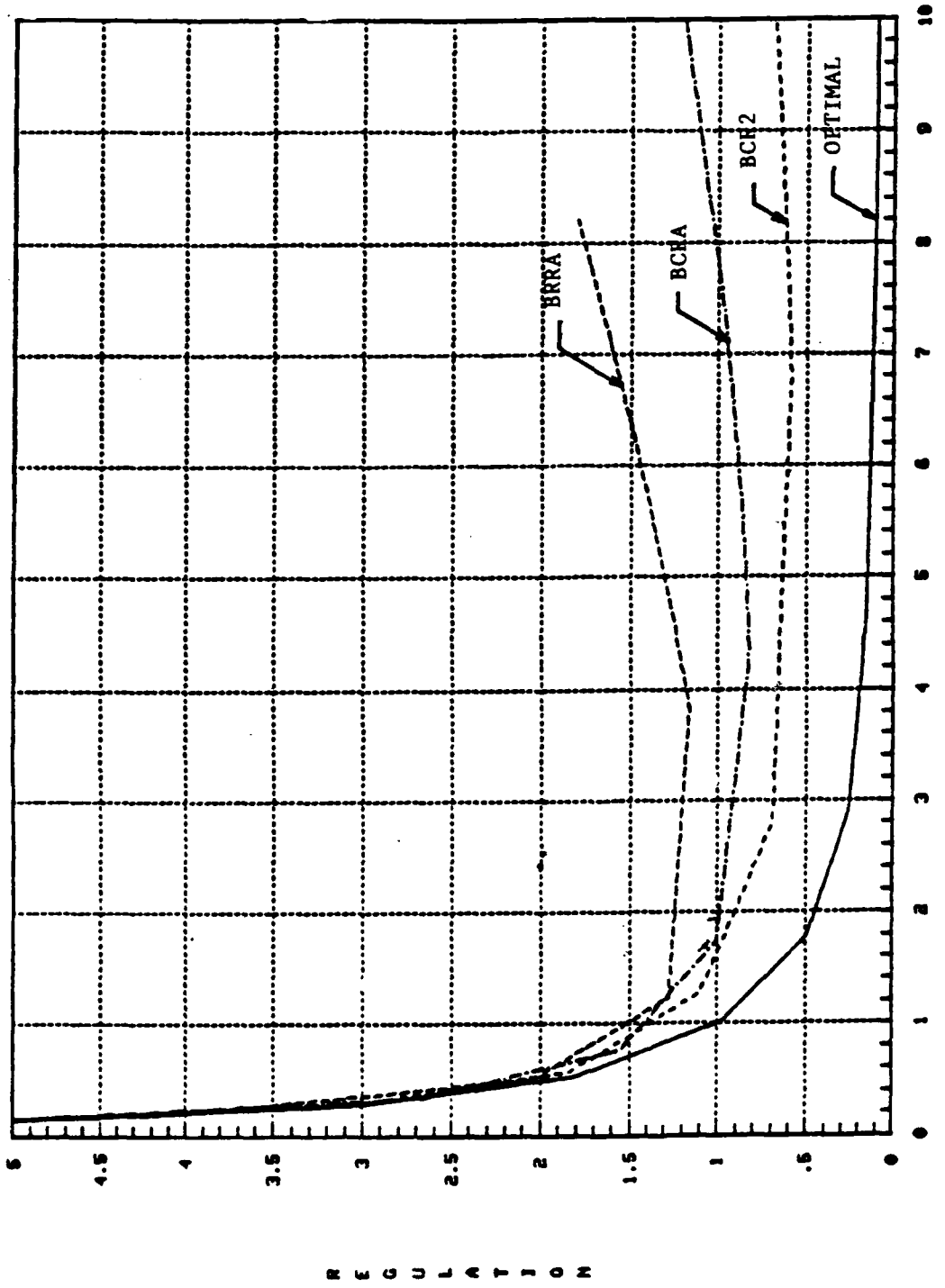


FIG.2 PERFORMANCE PLOT: FIRST ORDER CONTROLLERS

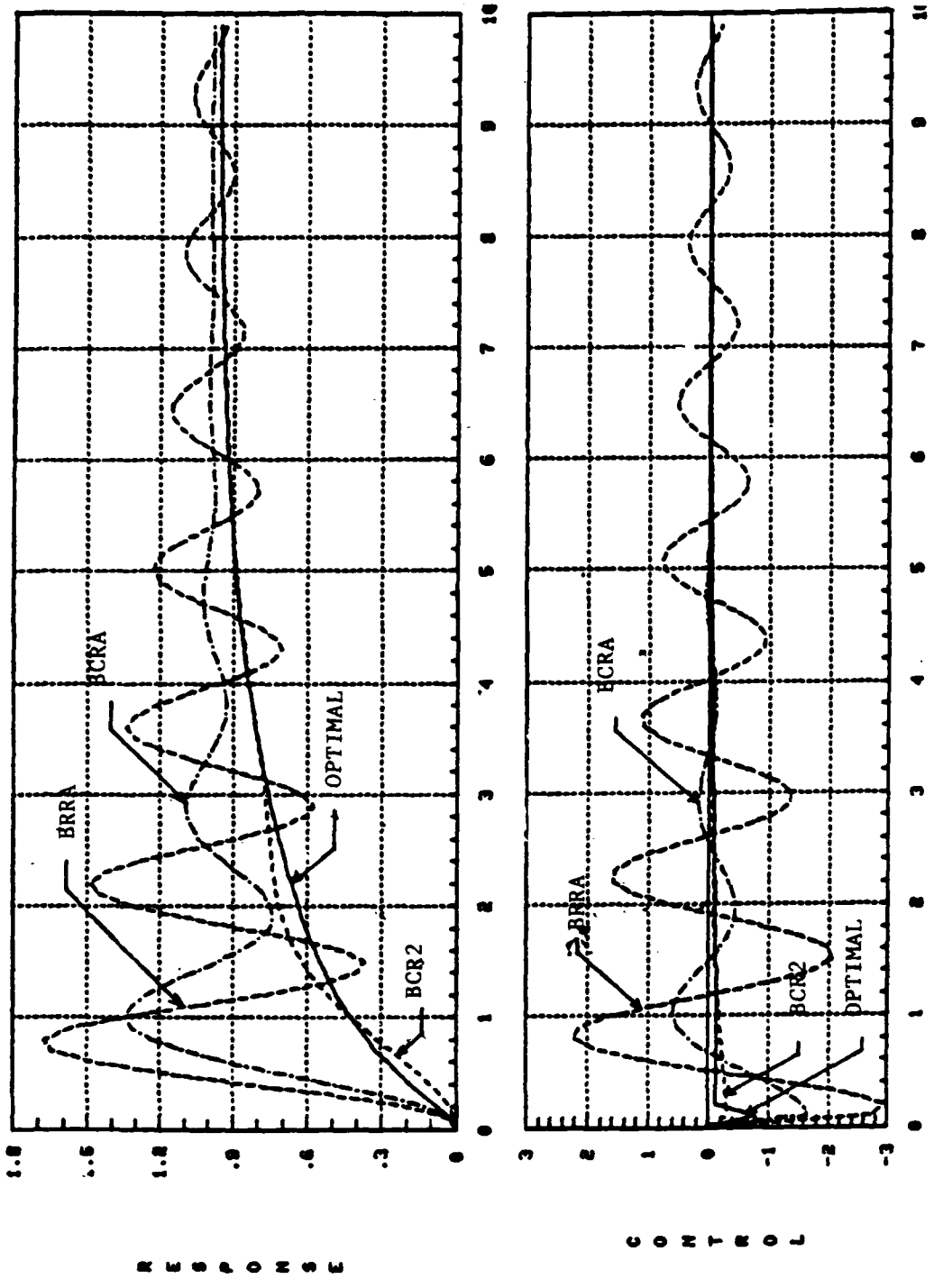


FIG.3 STEP RESPONSE: FIRST ORDER CONTROLLERS (RHO=10)

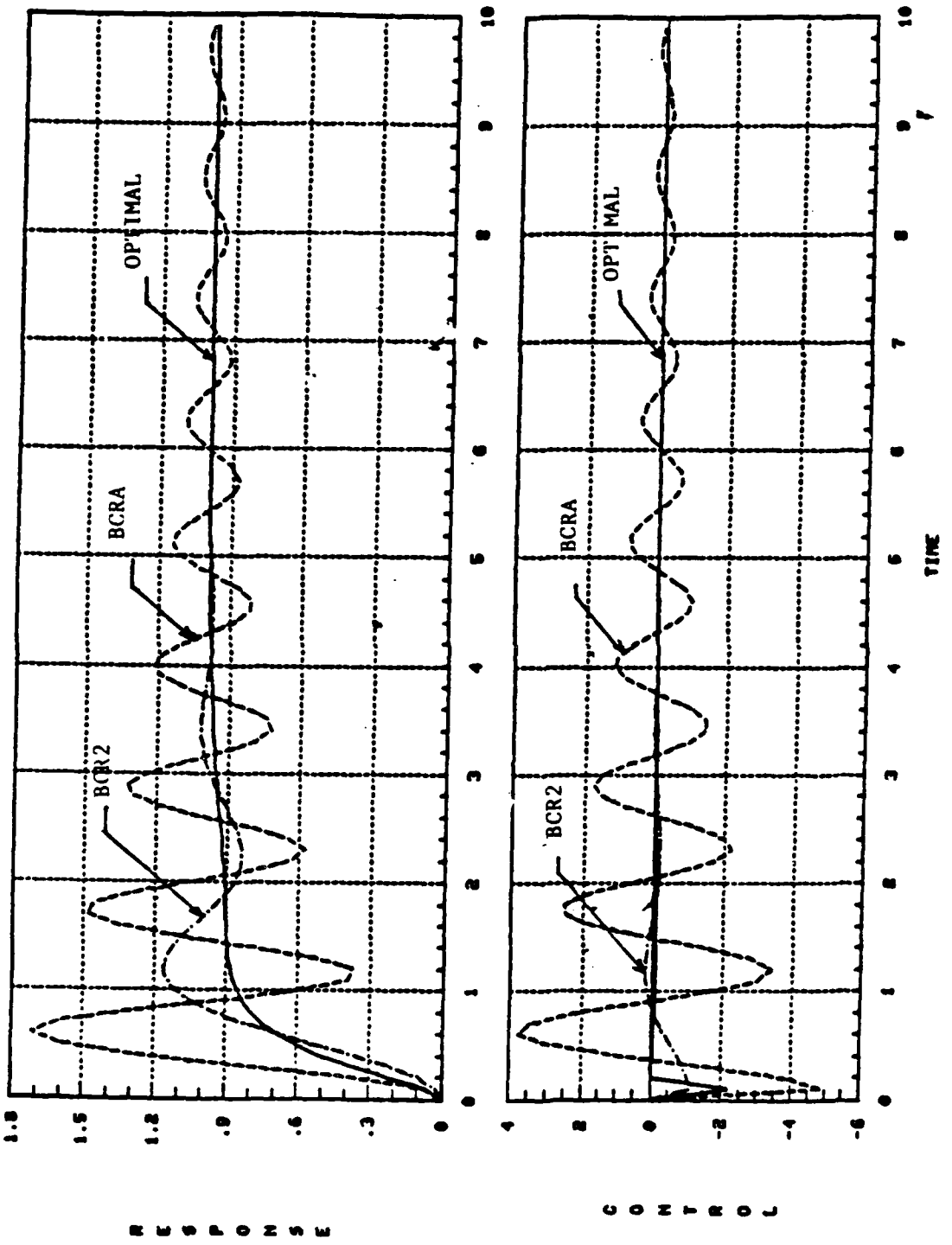


FIG.4 STEP RESPONSE: FIRST ORDER CONTROLLERS (RHO=1)

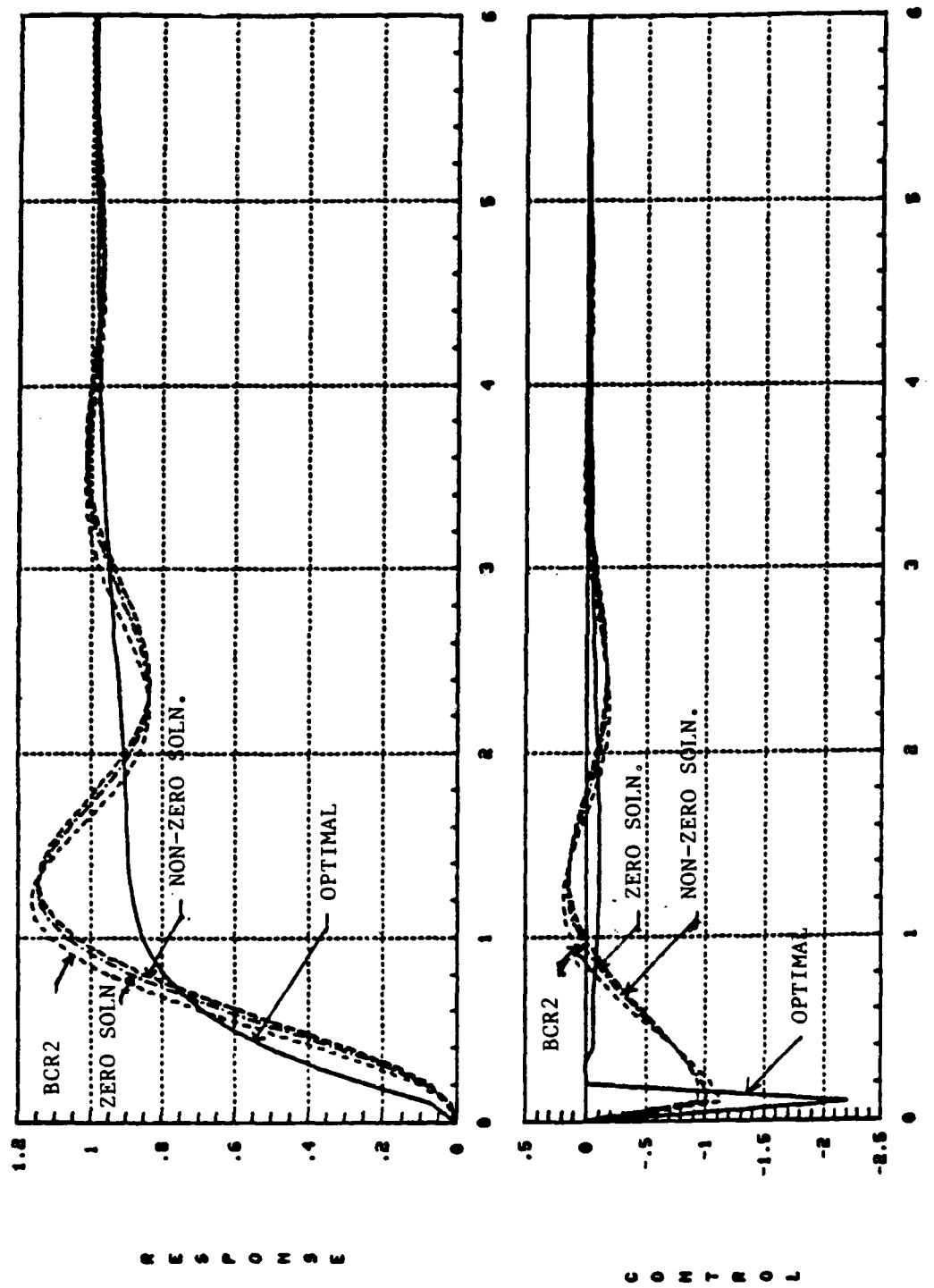


FIG.5 STEP RESPONSE: EFFECT OF NON-ZERO SOLN.

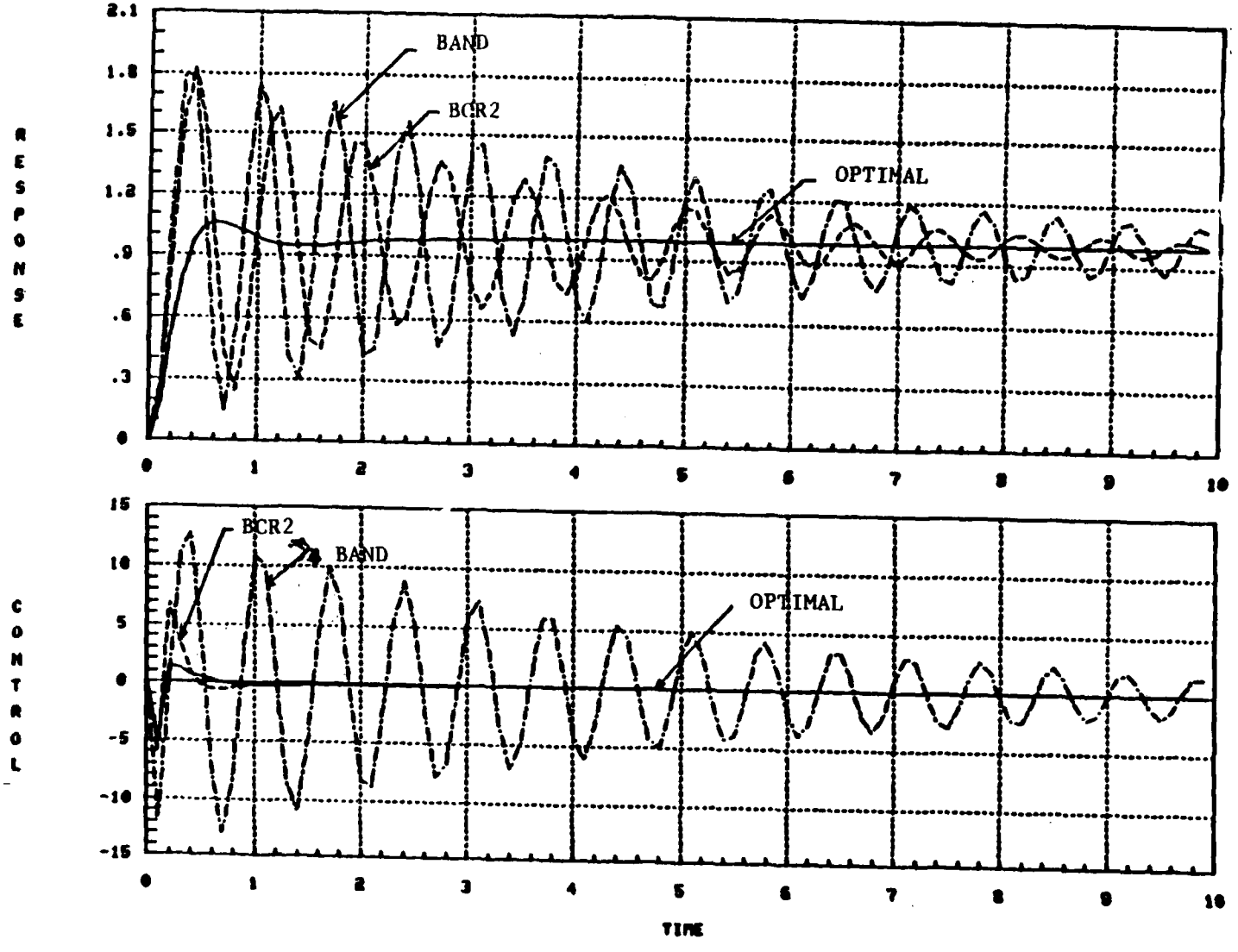


FIG.6 STEP RESPONSE: NEW METHOD .VS. ANDERSONS

REFERENCES

[1] C.C.Chu, "H^{∞} Optimization and Robust Multivariable Control," Ph.D. Dissertation, University of Minnesota, MN, Sep.1985.

[2] B.C.Moore, "Principal Component Analysis in Linear Systems: Controllability, Observability, and Model Reduction," IEEE Trans. Auto. Control, Vol.AC-26, pp.17-32, 1981.

[3] A.Yousuff, R.E.Skelton, "An Optimal Controller Reduction by Covariance Equivalent Realizations," IEEE Trans. Auto. Control, Vol.AC-31, pp.56-60, 1986.

[4] A.Yousuff, D.A.Wagie, R.E.Skelton, "Linear System Approximation via Covariance Equivalent Realizations," J. Math. Anal. Appl., Vol.106, pp.91-115, 1985.

[5] K.Glover, "All Optimal Hankel-norm Approximations of Linear Multivariable Systems and their L^{∞}-error bounds," Int. J. Control, Vol.39, pp.1115-1193, 1984.

[6] R.E.Skelton, A.Yousuff, "Component Cost Analysis of Large Scale Systems," Int. J. Control, Vol.37, pp.285-304, 1983.

[7] P.V.Kokotovic, R.E.O'Malley, P.Sannuti, "Singular Perturbation and Order Reduction in Control Theory - An Overview," Automatica, Vol.12, pp.123-132, 1976.

[8] C.Davis, W.M.Kahan, H.F.Weinberger, "Norm-preserving Dilations and their Applications to Optimal Error Bounds," SIAM J. Numer. Anal., Vol.19, pp.445-469, 1982.

[9] E.A.Jonckheere, L.M.Silverman, "A New Set of Invariants for Linear Systems - Applications to Approximations," IEEE Trans. Auto. Control, Vol.AC-28, pp.953-964, 1983.

[10] E.I.Verriest, "Suboptimal LQG-design and Balanced Realization," in Proceedings of 20th IEEE Conf. on Decision and Control, SanDiego, CA, pp.686-687, 1981.

[11] A.Yousuff, R.E.Skelton, "A Note on Balanced Controller Reduction," IEEE Trans. Auto. Control, Vol.AC-29, pp.254-257, 1984.

[12] J.A.Davis, R.E.Skelton, "Another Balnced Controller Reduction Algorithm," Systems & Letters, Vol.4, pp.79-83, 1984.

[13] Y.Liu, B.D.O.Anderson, "Controller Reduction via Stable Factorization and Balancing," Int. J. Control, Vol.44, pp.507-531, 1986.

[14] T.E.McQuade, S.S.Banda, "A Preliminary Investigation of H^{∞} Optimization," in Proceedings of AIAA Guid., Navig., Control Conf., paper #86-2197, pp.671-680, 1986.

1986 USAF-UES Summer Faculty Research Program/

Graduate Student Summer Support Program

Sponsored by the

Air Force Office of Scientific Research

Conducted by the

Universal Energy Systems, Inc.

Final Report

THE COMPOUND EYE: AN INTRODUCTION TO THE VARIETY OF VISUAL CAPABILITIES, GOALS, AND APPROACHES FOUND IN THE CLASS INSECTA.

Prepared by:	David D. Zeigler
Academic Rank:	Assistant Professor
Department and University:	Department of Biological and Allied Health Sciences, Bloomsburg University, Bloomsburg, PA 17815
Research Location:	The Air Force Armament Laboratory Advanced Seeker Division Electro Optical Terminal Guidance Branch Eglin Air Force Base, Fl 32542
USAF Researcher:	Dennis Goldstein
Date:	August 8, 1986
Contract No.:	F49620-85-0013

The Compound Eye: An Introduction to the Variety of Visual Capabilities, Goals, and Approaches Found in the Class Insecta.

by

David D. Zeigler

ABSTRACT

The following report contains one slightly abbreviated major chapter from a 73-page review paper on insect visual capacity. It was agreed that the variety and depth of topics discussed could not be effectively condensed to meet the 20-page requirements, so the chapter on distance estimation and an abbreviated introduction and bibliography were prepared for this report. The full-length review is on file with Dennis Goldstein at the Air Force Armament Laboratory/Advanced Seeker Division, Eglin A.F.B., FL 32542. Distance information can be gained from the relative "retinal" size and/or speed of visual targets on the surface of the compound eye. Such cues would work most effectively if the true size of the target was in some way "known". Such knowledge is most likely instinctive and "hardwired" for important biological targets which have been consistently encountered over evolutionary time. Speed or movement cues can also result from movement of the observer (motion parallax), and some insects have devised special behaviors (peering, flight oscillations, etc.) to utilize such cues. Binocular or stereoscopic vision may also be utilized by some insects, and there exists a range of mechanisms (in terms of complexity) which can use information from both eyes to estimate target distance.

ACKNOWLEDGEMENTS

I would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research, and the Air Force Armament Laboratory for sponsoring this effort. I also thank the Advanced Seeker Division and the Electro-Optical Terminal Guidance Branch for providing space, assistance, and interest during my work here.

I must thank Dr. Sam Lambert for his interest and encouragement in terms of the summer research associate program. Thanks to Dennis Goldstein and Rick Wehling for their interest and collaboration, to David Lawson and Kurt Perry for their assistance in obtaining reference sources, to David Crane for his assistance with some of my technical questions, and to Cub Sullivan and Judy Kenney for their help and assistance with the word processor. Lastly I must thank my wife, Christina, and my daughter, Catherine, for their sacrifice(s) in allowing me to participate in this program.

THE COMPOUND EYE: AN INTRODUCTION TO THE VARIETY OF VISUAL CAPABILITIES, GOALS, AND APPROACHES FOUND IN THE CLASS INSECTA

I. INTRODUCTION

II. OBJECTIVES

III. EFFICIENCY IN NATURE AND THE VISUAL SYSTEMS OF INSECTS

IV. VISUAL CUES IN INSECTS

- a. Simplicity of Functional Visual Cues
- b. Cue Priority and/or Preference
- c. Cue/Mode Relationships

V. MECHANISMS, ADVANTAGES, AND LIMITATIONS OF INSECT VISION

- a. Panoramic View
- b. Effective Distance
- c. Flicker-Fusion Frequency
- d. Visual Acuity
  - (1) Resolution
  - (2) Sensitivity
  - (3) Acute zones
  - (4) Movement detection
  - (5) Imaging
- e. Color Hue Discrimination
- f. Ultraviolet Light Detection
- g. Polarized Light Detection

VI. DISTANCE ESTIMATION

- a. Relative "Retinal" Size and Speed
- b. Motion Parallax
- c. Binocular Interactions

VII. VISUAL SPATIAL MEMORY

VIII. CONCLUSIONS AND RECOMMENDATIONS

I. INTRODUCTION—My research background is entirely as an animal behaviorist working specifically with insect communication. I have a strong general background in entomology, and I realize, perhaps more than the non-entomologist, the immense variety of insect behaviors, goals, and environments which have shaped the various insect visual systems such that they can solve a number of particular biological problems. My basic role is to fill in missing information pertinent to the functioning of multiaperture systems by using the workings of the insect compound eye as my information source. Artificial multiaperture optical systems may have a variety of possible strategic applications, one of which is in target recognition and homing control by target seeking missles.

II. OBJECTIVES—The original goal of this report was to find and correlate information concerning insect eye structure (morphology) with information concerning insect visual capabilities and visually mediated behavior across a wide range of insect groups. Precise correlations were not always forthcoming, but a number of interesting relationships surfaced which are herein discussed. The major value of this work is that it brings together in a unique way a number of previously known facts and relationships from close to 100 research and review sources. I have also tried to present a number of optical situations and problems unique to multiaperture visual systems using the insect compound eye as the working model.

VI. DISTANCE ESTIMATION—Like most higher animals, insects often find themselves in a variety of situations where they need to know their true or relative distance from certain biologically important "targets" such

as mates, rivals, prey, food, predators, landing surfaces, etc.. This section will discuss some of the ways in which insects visually determine these biologically important distances.

VI(a). Relative "Retinal" Size & Speed--Due to the basic architecture of the compound eye, objects will fill or stimulate different numbers of ommatidia depending on their distance from the eye surface (Fig. 7). This distance-relative variation in stimulation could be a cue to distance, but only if the size of the object is known. Two targets at different distances could stimulate equal numbers of ommatidia if they were of different sizes (Fig. 7). Thus for "retinal size" to be a useful cue in distance determination, it would seem that the actual size of the target must somehow be known. People who have lost their sight in one eye (and thus their stereoscopic 3-D vision) constantly judge distances to familar objects in part by knowing the true size of chairs, cars, houses, people, etc.. In the absence of other cues such as shading, texture, or neighboring objects of known size, the distance of unfamiliar objects could not be easily determined by the monocular person.

It is very likely that insects have come to "know" instinctively the size of very important biological targets over a range of important distances which their species has encountered consistently over evolutionary time. This "hardwired" information would be crucial for some insects in distance estimations. For important targets which move in predictable ways over evolutionary time, the apparent "retinal speed" may be neurally correlated with apparent retinal size in a hardwired fashion over a range of biologically important distances. This

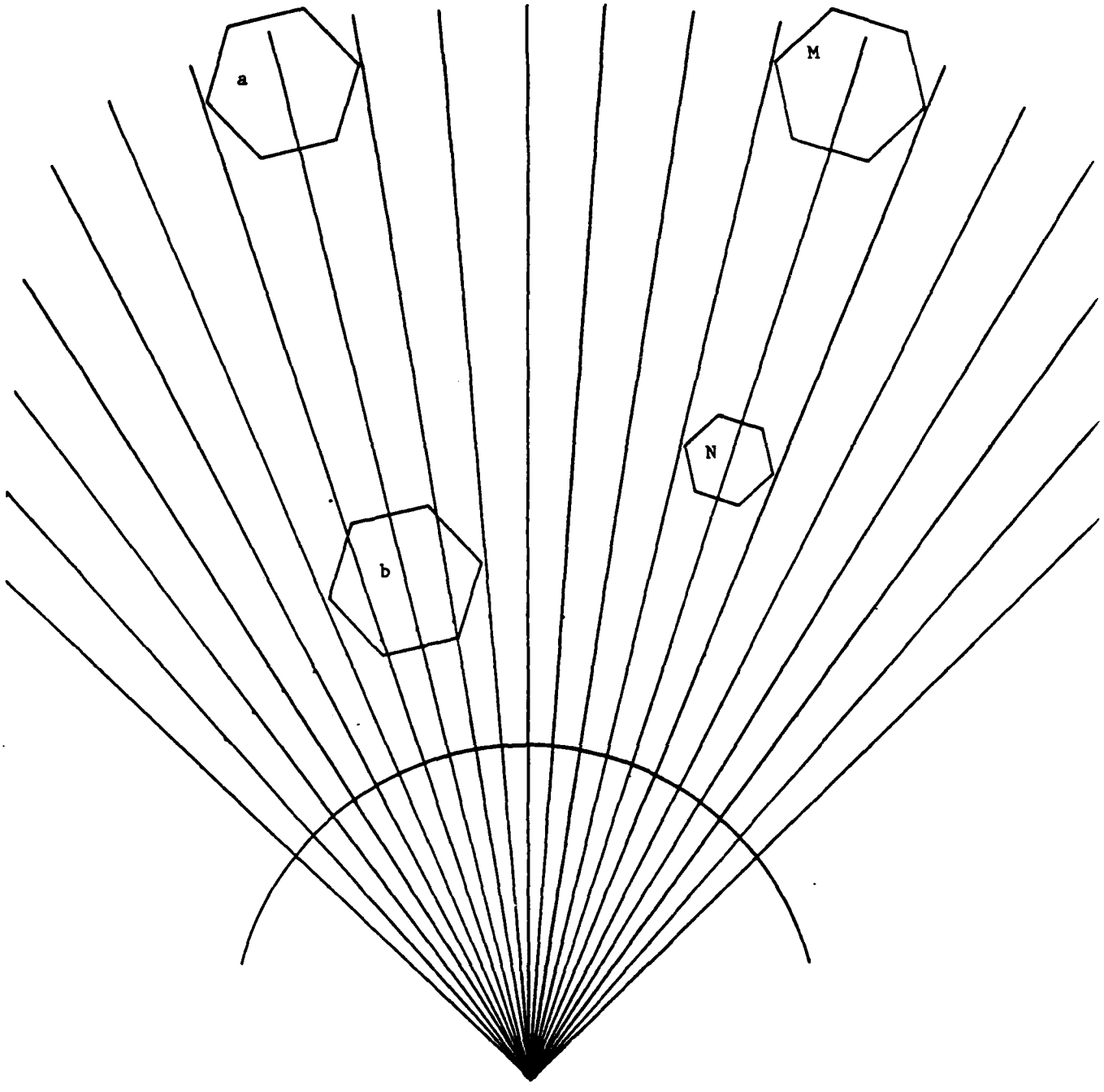


Figure 7. The same object can stimulate different numbers of ommatidia at different distances: position a = 2 ommatidia stimulated; position b = 4 ommatidia stimulated. Also, two differently sized objects (M & N) can stimulate equal numbers of ommatidia (2) if they are at different distances from the eye surface.

instinctive knowledge would allow the insect to reliably and quickly initiate the appropriate behavior patterns for dealing with these high priority images and situations (22). Many fast-flying insects may control their flight speed to accomodate their innate sensitivity by using the relative retinal velocity of background patterns as their cue (27).

The apparent retinal size of objects seems to be a important visual cue used by hoverflies to maintain their distance from one another in aerial chases (22), by bulldog ants to estimate appropriate snapping distances (92), and by some bees and wasps to judge their distance from familiar landmarks (15). As has already been stated, the relative apparent size increases as the distance between observer and target declines. This effect is known as looming (24). If the actual closing speed of observer and target is constant, it should be obvious that the apparent retinal speed at which the edge of the looming image travels across the surface of the compound eye will increase as the observer/target distance decreases. The relative speed at which looming occurs, together with the apparent retinal size, seems to be a cue in the landing response of some insects (18, 29,).

In the hoverfly, males seem to have a file of hardwired information which correlates the size and speed of typical hoverfly females in flight. Once a moving target of appropriate size at initial distance is sighted, the male plots an interception course with the target on the basis of his instinctive knowledge of female flight parameters (24). If the target is a female, the male's interception course is usually effective. If the female is intercepted successfully, mating will ensue. If the

target is not a conspecific female, but another type of insect or an artificial target launched by a curious researcher, then the speed and flight path is unlikely to match that of a female hoverfly and the male fails to manage a successful interception (24). The male probably trys to intercept inappropriate targets because at the initial sighting he is too far off to accurately image the target. Nor can he waste time in trying to make independent retinal size and speed correlations to insure that the image is a female hoverfly. Instead he must react almost instantaneously and begin his interception maneuver based on his hardwired knowledge and "hope" the target is a female. Though mistakes are made, the number of successes have apparently been adequate to maintain the species.

VI(b). Motion Parallax—Since compound eyes cannot be moved independently of the head, movements of the head and/or body typically alter the position of the retinal image produced by a stationary target (59). For two targets at different distances but of the same angular or retinal size (i.e. size cues absent), a head movement would displace the nearer object's image to a greater degree than that of the distant object. This effect is termed "motion parallax", and the degree of displacement should be a valuable cue in distance estimations (25). This cue would, however, be much less useful if the observer and target were both in independent motion (22).

It should be equally obvious that motion parallax is a cue to target size (25). A large retinal image represents a large actual object if motion parallax cues reveal only a small image displacement, but that

same retinal image represents a smaller object if motion parallax reveals a large retinal displacement of the image. Many types of flying insects perform zig-zag vertical or horizontal movements in flight while keeping the head/body junction fixed and without introducing roll, pitch, or yaw into the movements. These movements often increase in frequency as the insect nears objects of interest (15), and they may play a key role in correlating relative distances with known sizes and positions of landmarks, foodsources, or other organisms to bring about correct behavioral responses.

A few insects such as grasshoppers, locusts, and mantids have developed a unique behavior called "peering" which applies the motion parallax effect to specific distance estimation problems (21, 43). In peering, the insect typically rocks the upper body and head laterally with the abdomen tip relatively stationary (86). The head adjusts for the radial body movement so that at all times the head midline axis is parallel to its initial allignment (21). Moving the body laterally by a "known" distance allows calculation not only of relative distance, it also allows more precise estimation of the actual distances involved (22). For a walking insect this might not be so important since it will simply pick the nearest attractive target and walk to it (86), but for a mantid peering at his intended prey or a grasshopper about to jump to a distant perch, the precise distance would be valuable information. Some experiments have shown the importance of this lateral peering distance by moving the locust perch in sync with his peering movements, thus effectively extending his lateral distance and

causing greater image displacement. The locust consistently jumps short of the target in this situation since he interprets the greater image displacement to mean a closer target (21).

In conditions of motion parallax where the observer "knows" and controls his speed of movement in a precise manner, the relative speed at which the image shifts across his eye surface could also be a distance cue. Faster movements would correspond to closer targets and vice versa (38).

VI(c). Binocular Vision--It is obvious from Table 1. that some insects have a wide overlap in the visual fields of the two compound eyes. This does not, however, mean that any insect with binocular overlap of visual fields has true stereoscopic vision, and thus a means of depth perception, similar to that of humans (25). Binocular vision only implies that both eyes can "see" the same object (92). Stereoscopic vision, which humans have, implies that the two images are "fused" somewhere in the nervous system such that 3-D information emerges from "disparity cues" (differences in spacing, shading, and intensity between the two images)(39).

A controversy has continued for some years as to whether or not insects have any stereoscopic abilities. Arguments against such a possibility have included the following points.

* The two eyes of most insects are too close together to allow much disparity of images except at very close ranges (9).

* True stereoscopic vision implies a complex interconnecting system of nerves to communicate and compare the two incoming images. Some feel

TABLE 1. (Abbreviated) A COMPARISON OF INSECT EYE PARAMETERS

Insect	Ommatidia/ eye	$\Delta\phi$	Binocular Overlap	Source
worker ant *Ponera punctatissima*	1			17
Drosophila melanogaster	700	5°		25, 39
cockroach *Periplaneta americana*	2000	10°		79, 12
wood cricket *Nemobius sylvestris*	1204		42°	50
stalk-eyed flies *Cyrtodiopsis whitei*	2600		135°	10
caddisfly *Banksiola selina*	3710			40
Musca domestica	2900			6
honey bee worker *Apis mellifera*	5000	1-2°	42°	78, 65
blowfly *Callophora erythrocephala*	5342	1.2-2.5°		6, 57
Notiophilus biguttatus	1000		74°	5
hoverfly *Syritta pipens*	6310	.6-1.4°	10°	6, 22
dragonflies	up to 30,000			79

$\Delta\phi$ = Interommatidial angle

that the small and limited space within an insect's head precludes this possibility (43).

* A number of behavioral studies have precluded stereoscopy in the landing response of houseflies (84) and blowflies (29), the chasing behavior of hoverflies (22), jump distance estimation in the locust (87) and grasshopper (31), and others.

True stereoscopic vision has seemingly been revealed in the praying mantis by the use of visual prisms which alter the position of the image on the two eye surfaces without affecting the apparent size or the effects of motion parallax (72). Other workers have claimed, but with less convincing results, stereoscopic abilities in water scorpions Ranatra sp. (19) and ground beetles Notiophilus biguttatus (5). Other workers have shown binocular effects involved in certain behaviors of insects such as the bulldog ant (92), hoverflies (22), and mantids (71), but without claiming stereoscopic abilities.

There are in fact a number of possible ways (hypothetical) the two eyes can work together in the area of binocular overlap other than to produce stereoscopic vision. One possible type of interaction which will yield distance information is called binocular triangulation, which in its most fully cross-correlated state, approximates stereoscopic vision. Comparison of "sets" of stimulated ommatidia in the binocular areas of the two eyes is accomplished neurally and is the key to distance determination by binocular triangulation (Fig. 12)(2).

This model need not include depth information resulting from shading effects since this would also require intensity matching and a variety of more complex neural interrelationships. Complete cross-correlation of

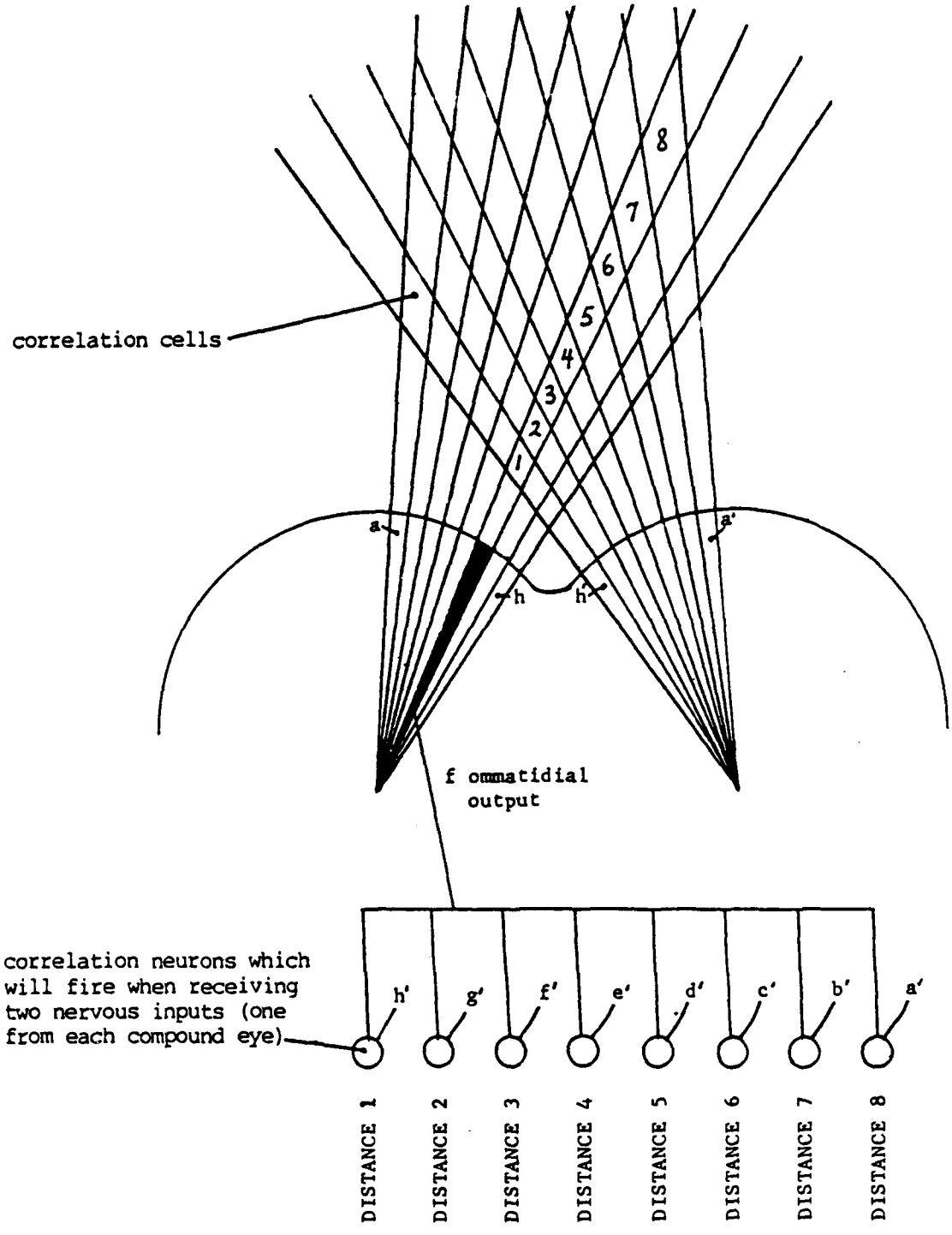


Figure 12. Binocular triangulation. Eight ommatidia are shown for each of the two compound eyes. Collectively, these ommatidia share a common visual area. Nerve connections could correlate each ommatidia with all those of the other eye which share sections of its visual field. Such nervous correlations are shown for ommatidia f of the left eye. Full cross-correlation of all ommatidia shown would require 64 (8 x 8) correlation nerurons. Or, a simplified system could use cross correlations for only one ommatidia/eye to all those of the opposite eye (as shown for ommatidia f) and vise versa, in which case correlations would equal only 15 (not 16 since two of these would be the same).

nerve nets for every biommatidial field crossing in the binocular zones would indeed involve many neurons and nervous circuits but this system could compare and estimate the distances of a number of objects within the binocular field. But the effective binocular field of insects is certainly distance limited since "correlation cells" (Fig. 12) get larger and thus give less precise position cues as distance is increased. It could be considered improbable that more than one object of interest would be found within this limited area at any one time. If binocular distance information is only required for one target at a time, then cross connections could be reduced such that the insect "fixates" the target or target edge with one eye, centering the image on only one or a few ommatidia and cross-correlating only those few receptors with those in the opposing eye which have visual field overlap with the field of the fixating ommatidia (Fig. 12)(71). Even if such a system were duplicated for the other eye, the number of nervous correlations and circuitry would be drastically reduced from that required for the fully correlated model Mantids may in fact use such a system since they often fixate and attack prey which are off-center as could be expected from this model, and this system would produce identical results to those found in the prisms studies mentioned earlier. These models are hypothetical, but their simplicity suggests a possible correspondence to the neural circuitry of real insects.

Some authors have suggested that the only possible function of binocular vision is for depth (distance) perception, but it has been suggested that in mantid fixation, hoverfly chases, and other behaviors that binocular vision serves as part of a centering mechanism to align

the head and body axis of an organism with its intended target (prey, mate, or rival). In the case of hoverfly chases, as the leading fly moves laterally from the visual midline of the chasing fly, he/she will partially leave the binocular visual field and thus be imbalanced on the two eye surfaces. It has been suggested that these insects can sense this imbalance and maneuver to rebalance and thus recenter the image (22).

VIII. CONCLUSIONS AND RECOMMENDATIONS--Insect compound eyes have been very heavily studied, but suprisingly few breakthroughs have occurred in our understanding of them. New theories abound on how certain insect groups may vary from others in their vision mechanisms and capabilities. Although these new ideas such as the diffraction theory of insect vision (11) and the neural superposition eye (33, 93) should be pursued, they are largely based on data taken from "parts" of insects which have been experimentally manipulated in ways which may have affected the results.

The problems of actually taking data from the eyes and nervous systems of such small animals present a major barrier to our understanding. Also, experiments with live insects are often performed in very unnatural environments which may not have counterparts in the real world.

Time was a limiting factor in the preparation of this review. Many other sources were located but were not used due to the available time being too short to read and synthesize their contents. Also, my background guided my concept of the paper and perhaps resulted in a work too generalized for its intended audience. Nevertheless, several

concepts and principles have been brought together which hopefully will guide other workers in their understanding of how multiaperture systems could work, along with their inherent problems and advantages.

The basic theory of the mosiac image produced by compound eyes is still a strong and overriding concept in our understanding of how multiaperture systems could work in solving the various problems which insects encounter daily. It is the underlying neural circuitry which we need to better understand before the idea of creating an effective artifical multiaperture visual system can bear fruit. Since visual processing in insects is a difficult and little understood area, perhaps we should look to what is known of vertebrate visual processing. Although the compound eye and the vertebrate eye have their differences, they also have some similarities. They probably vary most in their visual collection (sensor) systems, but are more similar at and below the level of nervous system input. Since more is known in this area (more workers, more applied knowledge, more funding), it might be wise to make a connection into this field and see how compatible the two areas could be.

Also, Dr. David Lawson and I had some very productive discussions and work sessions on nerve nets and visual processing. Although new to the field, we seemed to make progress because we used a functional approach in trying to model our nerve nets. In contrast, most other workers in this area are more concerned with accurately modeling and/or describing real nervous system processing, such that their theories will fit the empirical observations and data. I recommend our functional approach in situations where results are more important than fidelity to nature.

REFERENCES (Abbreviated)

2. Atkins, M.D., Introduction to Insect Behavior, New York, Macmillian Publishing Co., Inc. 1980.

5. Bauer, T., "Prey Capture and Structure of the Visual Space of an Insect That Hunts by Sight on the Litter Layer (Notiophilus biguttatus F., Carabidae, Coleoptera)," Behav. Ecol. Sociobiol., 8 (1981) 91-97.

9. Buchner, E., "Behavioural Analysis of Spatial Vision in Insects," In Photoreception and Vision in Invertebrates, M.A. Ali (ed.), New York, Plenum Press, 1984.

11. Burtt, E.T., and W.T. Catton, "A Diffraction Theory of Insect Vision I. An Experimental Investigation of Visual Acuity and Image Formation in the Compound Eyes of Three Species of Insects," Proc. R. Soc. B, 157 (1962) 53-82.

15. Cartwright, B.A., and T.S. Collett, "Landmark Learning in Bees," J. Comp. Physiol., 151 (1983) 521-543.

18. Chillemi, S., and C. Taddei-Ferretti, "Landing Reaction of Musca domestica," J. Exp. Biol., 94 (1981) 105-118.

19. Cloarec, A., "Distance and Size Discrimination in a Water Stick Insect, Ranatra linearis (Heteroptera)," J. Exp. Biol., 120 (1986) 59-77.

21. Collett, T.S., "Peering—A Locust Behaviour Pattern for Obtaining Motion Parallax Information," J. Exp. Biol., 76 (1978) 237-241.

22. Collett, T.S., and M.F. Land, "Visual Control of Flight Behaviour in the Hoverfly Syritta pipiens L.," J. Comp. Physiol. 99 (1975) 1-66.

24. Collett, T.S., and M.F. Land, "How Hoverflies Compute Interception Courses," J. Comp. Physiol., 125 (1978) 191-204.

25. Craig, C.L., "Orb-Web Visibility: the Influence of Insect Flight Behaviour and Visual Physiology on the Evolution of Web Designs Within the Araneoidea," Anim. Behav., 34 (1986) 54-68.

27. Downes, J.A., "The Swarming and Mating Flight of Diptera," Ann. Rev. Entomol., 14 (1969) 271-298.

29. Eckert, H., and K. Hamdorf, "Excitatory and Inhibitory Response Components in the Landing Response of the Blowfly, Calliphora erythrocephala," J. Comp. Physiol., 138 (1980) 253-264.

31. Eriksson, E.S., "Movement Parallax and Distance Perception in the Grasshopper (Phaulacridium vittatum (Sjostedt))," J. Exp. Biol., 86 (1980) 337-340.

33. Franceschini, N., "Retinal Mosaic of the Fly Compound Eye," In Photoreception and Vision in Invertebrates, M.A. Ali (ed.) New York, Plenum Press, 1984.

38. Goulet, M., R. Campan, and M. Lambin, "The Visual Perception of Relative Distances in the Wood-Cricket, Nemobius sylvestris," Physiol. Entomol., 6 (1981) 357-367.

39. Heisenburg, M. and E. Buchner, "The Role of Retinula Cell Types in Visual Behavior of Drosophila melanogaster," J. Comp. Physiol., 117 (1977) 127-162.

43. Horridge, G.A., "Insects Which Turn and Look," Endeavor N.S., 1 (1977) 7-17.

59. Mazokhin-Porshnzakov, G.A., Insect Vision, New York, Plenum Press, 1969.

71. Rossel, S., "Foveal Fixation and Tracking in the Praying Mantis," J. Comp. Physiol., 139 (1980) 307-331.

72. Rossel, S., "Binocular Stereopsis in an Insect," Nature, 302 (1983) 821-822.

84. Taddei-Ferretti, C., "Landing Reaction of Musca domestica, III: Dependence on the Luminous Characteristics of the Stimulus," Z. Naturforsh., 28c (1973) 568-578.

86. Wallace, G.K., "Visual Scanning in the Desert Locust Schistocerca gregaria Forskal," J. Exp. Biol., 36 (1959) 512-525.

87. Wallace, G.K., "Experiments on Visually Controlled Orientation in the Desert Locust, Schistocerca gregaria (Forskal)," Anim. Behav., 10 (1962) 361-369.

92. Via, S.E., "Visually Mediated Snapping in the Bulldog Ant: A Perceptual Ambiguity Between Size and Distance," J. Comp. Physiol., 121 (1977) 33-51.

93. Zeil, J., "Sexual Dimorphism in the Visual System of Flies: The Compound Eyes and Neural Superposition in Bibionidae (Diptera)," J. Comp. Physiol., 150 (1983) 379-393.

1986 USAF-UES Summer Faculty Research Program/

Graduate Student Summer Support Program

Sponsored by the

Air Force Office of Scientific Research

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

Microwave Impedance Matching for Optical Devices

Prepared by:	Henry Zmuda
Academic Rank:	Assistant Professor
Department and	Department of Electrical Engineering
University:	Stevens Institute of Technology
Research Location:	Rome Air Development Center
	Griffiss Air Force Base, NY
USAF Research:	Dr. D. Nicholson
Date:	September 1, 1986
Contract No.:	F49620-86-C-0013

MICROWAVE IMPEDANCE MATCHING FOR OPTICAL DEVICES

by

HENRY ZMUDA

ABSTRACT

One of the major thrusts in fiber optic communication is to replace waveguide. This requires matching of the microwave signal source to the modulating device. Addressed here are the matching requirements for systems using direct modulation of a laser diode, with particular attention to the question of broadband equalization.

ACKNOWLEDGMENTS

The author would like to thank the Air Force Systems Command and the Air Force office of Scientific Research for sponsoring this research. Gratitude is extended to Dr. D. J. Nicholson of RADC, not only for acting as the Effort Focal Point for this research, but especially for his encouragement and guidance.

The success of this project hinged upon the stimulating environment provided by the members of the RADC Wideband Communications division, and the author would like to express his thanks to all those with whom he had the pleasure of interacting. Special thanks are extended to Brian Hendrickson and Paul Sierak of RADC for their valuable comments and suggestions.

Last, but by no means least, the author wishes to thank Prof. H. J. Carlin of Cornell University. Prof. Carlin has been a source of inspiration and guidance over the years, and without him this research would not have been possible.

I. INTRODUCTION

I received my Ph.D. from Cornell University studying microwave modeling of optical waveguides. Though fundamentally a microwave engineer, my research required specific knowledge of optical devices giving me an interesting blend of both diciplines. Additionally my research group at Cornell was involved with developing innovative techniques for broadband matching of microwave devices. Thus, although it was not my primary area of study, I still had extensive exposure to broadband matching techniques and am very familiar with the area as a whole.

The research in which I was involved at RADC required knowledge in all of the above mentioned areas, an my background made me well suited to work on the problems associated with microwave matching to optical devices.

II. OBJECTIVES OF THE RESEARCH EFFORT

Fiber optic links have many applications for microwave systems. Though there are many factors which influence the performance of these fiber optic systems, this paper addresses the important question of microwave impedance matching by making use of recent results in the theory of broadband equalization [1].

Presented here are circuit designs for lossless matching networks which optimize the transducer power gain (i.e. the ratio of the average power delivered to the load to the maximum power available from the generator) of the system over a prescribed frequency band. This is done for the case of a resistive generator which directly modulates the laser diode, as well as for a photodiode loaded by a resistor.

The basis of the technique is a CAD procedure developed recently by Carlin and Yarman at Cornell University [2]. There the method has been successfully applied to the design of broadband multi-stage microwave FET amplifiers.

Two attractive features of Carlin's technique is that the design can proceed directly from measured (real frequency) data for the source and load, bypassing the need to obtain an approximate analytic (linear circuit) model. This is an especially desireable feature for the front end of the fiber link since the laser diode operates in a very nonlinear region. Secondly, the technique does not require an a priori choice of an equilizer topology. The optimum topology of the matching structure comes about as a consequence of the design process.

III. BACKGROUND

An excellent introduction to the theory of broadband matching can be found in [1-3], hence only the pertinent aspects of the theory are highlighted here. The most general form of the matching problem is illustrated in figure 1a.

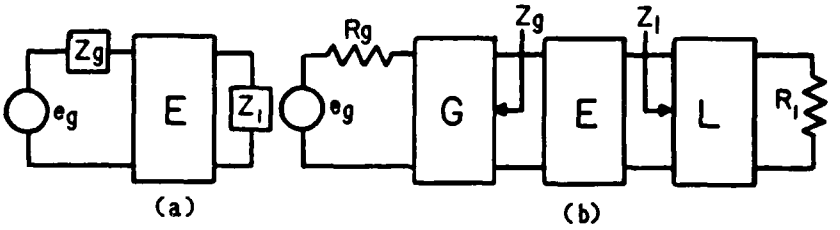


Figure 1
THE MATCHING PROBLEM

In words, given a generator with complex internal impedance Z_g, and a complex load impedance Z_1, determine a lossless matching network E which optimizes the power transfer from generator to load over a prescribed frequency band. An equivalent representation of the network, shown in figure 1b, is obtained from Darlington's theorem which states that a positive real impedance function can be realized as the input impedance of lossless two-port terminated in a resistance. Thus, for a given generator and load, we wish to find an equalizer such that maximum power is delivered to R_1, i.e. such that the transducer power gain $T(\omega)$ is maximized, where

$$T(\omega) = \frac{\text{Average power delivered to } R_1}{\text{Maximum power available from } e_g} \qquad (1)$$

For the network of figure 1, $T(\omega)$ is found to be

$$T(\omega) = \frac{|e_{21}|^2(1-|S_g|^2)(1-|S_1|^2)}{|1-e_{11}S_g|^2|1-S_2S_1|^2} \qquad (2)$$

where we have assumed the unknown reciprocal equilizer E has a scattering matrix S_e of the form

$$S_e = \begin{bmatrix} e_{11} & e_{21} \\ e_{21} & e_{22} \end{bmatrix} \qquad (3)$$

S_g and S_1 are the real normalized reflection factors for the generator and load respectively,

$$S_g=(Z_g-1)/(Z_g+1)$$
$$S_1=(Z_1-1)/(Z_1+1) \qquad (4)$$

Z_g and Z_1 are the real normalized generator and load impedances, and

$$S_2 = e_{22} + \frac{e_{21}^2 S_g}{1- e_{11}S_g} \qquad (5)$$

As for the structure of the scattering matrix S_e, employ a Belevitch representation, namely that the real canonic form for a scattering matrix of a real lossless two-port is [4]

$$S(s) = \frac{1}{g(s)} \begin{bmatrix} h(s) & f(s) \\ \pm f(-s) & \mp h(-s) \end{bmatrix} \qquad (6)$$

where:

$s=\sigma+j\omega$

f,g and h are polynomials
g is a Hurwitz polynomial
the degree of g is not smaller than the degrees of f and h

Since E is reciprocal, $f(s)=\pm f(-s)$, and together with the losslessness of E requires that S be a para-unitary matrix yielding

$$h(s)h(-s)\pm f^2(s)=g(s)g(-s) \qquad (7)$$

as the relationship to be satisfied between f,g and h.

As a matter of convenience it is assumed that E is a minimum phase structure with zeroes of transmission only at zero and/or infinity. This gives

$$f(s)=s^k \qquad (8)$$

It should be noted that the degree of the polynomial g (or h) is equal to the number of lossless elements in E. Furthermore, k, the order of the zero of transmission, is the number of high pass elements (series capacitors, shunt inductors) in E. Thus k=0 represents a low pass network whereas k>0 is band pass.

Hence, the scattering matrix of E is of the form

$$S_e = \frac{1}{g(s)} \begin{bmatrix} h(s) & s^k \\ s^k & -(-1)^k h(-s) \end{bmatrix} \qquad (9)$$

with

$$g(s)g(-s)=h(s)h(-s)\pm s^{2k} \qquad (10)$$

The computational procedure is as follows:

1. Fix upon the complexity of E by choosing the degree of h and k.
2. Take an initial guess for $h(s)=h_0+h_1 s+\ldots+h_n s^n$. Choosing $hi=\pm 1$ works well in most instances.
3. Construct the even polynomial $g(s)g(-s)=h(s)h(-s)+s^{2k}$
4. Perform spectral factorization of g(s)s(-s) and isolate the Hurwitz (left hand plane) roots. Use these to construct g(s). Note that this step guarantees realizability.
5. Compute $T(\omega)$ for some number of frequency points in the passband. Since a realization is guaranteed, any unconstrained optimization routine can be used to optimize $T(\omega)$. For the results to follow a modified Levenberg-Marquardt algorithm was used to minimize

$$\varepsilon = \sum_{\substack{\text{over m} \\ \text{frequency} \\ \text{points}}} \left[T(\omega_m) - T_i \right]^2 \qquad (11)$$

   in a least squares sense, where T_i is the constant idealized gain.
6. With the final optimized coefficients for h(s), construct S_e, and realize the equalizer E by synthesizing

$$Z(s)=[1+e_{11}(s)]/[1-e_{11}(s)] \qquad (12)$$

   as a lossless network terminated in a resistance via Darlington synthesis. The lossless 2-port with the resistor removed is the matching network.

This completes the summary of Carlin's technique. Let us now turn to some examples of matching microwave signals to optical devices.

IV. RESULTS

This section presents examples of matching microwave signal sources to directly modulated injection laser diodes as well as for matching photodiodes to a microwave load. The latter includes avalanche photodiodes (APD) and Shottkey Barrier PIN diodes.

The following comments should be noted. When matching a microwave source to a laser diode, it is assumed that the modulating source appears to have a purely resistive internal impedance. Similarly, it is assumed that a photodiode is matched to a purely resistive load. The reason for this is that when designing a fiber link the characteristics of the source or load are generally unknown except for their nominal impedances. The basic theory of broadband matching as well as the programs developed are general enough to include parasitic effects of the microwave terminations when they are known.[1]

For numerical efficiency, all circuit element values are normalized to both frequency and impedance level. Thus in all circuit designs, if R', L' and C' are the actual element values, the normalized values are

$$R=R'/Z_0,\ L=\omega_0 L'/Z_0,\ C=\omega_0 C' Z_0$$

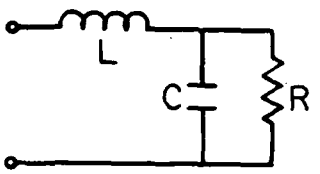
corresponding to a normalized radian frequency ω/ω_0, $\omega=2\pi f$. Further, Z_0 is assumed to be equal to 50 ohms throughout.

For illustrative purposes, a circuit model for the

---

[1]Results were obtained using a FORTRAN program developed at RADC on the Honeywell MULTICS computer system.

diode was used to generate real and imaginary parts for the impedance values. A typical diode model is shown in figure 2 along with typical element values [5,6,7].



TYPICAL VALUES			
	BH Laser Diode	PIN Photo Diode	
L	0.1 - 1.0	0.1	nH
C	1.0	10.0	pF
R	1.0 - 5.0	100.0	ohms

Figure 2
DIODE EQUIVALENT CIRCUIT

Example 1: Laser diode matched to a 50 ohm generator.
Input:

n=3, 3 elements in equilizer
k=0, low pass design
$h(s)=1-s+s-s^2$, initial guess
Frequency range of 0 to 10 GHz in 10 steps
Case (a): $T_i=0.95$
Case (b): $T_i=0.8$

Output:

Case (a): $h(s)=1.38-4.17s+3.81s^2-4.15s^3$
Case (b): $h(s)=1.38-2.86s+2.56s^2-1.46s^3$
Case (a): $g(s)=1.71+4.64s+4.28s^2+4.15s^3$
Case (b): $g(s)=1.71+3.26s+2.77s^2+1.46s^3$
Case (a): $T(\omega)=0.836\pm0.048$
Case (b): $T(\omega)=0.801\pm0.005$

Figure 3 shows the performance of the equalizers. The corresponding realization is shown in figure 4.

Example 2: Laser diode matched to a 50 ohm generator with:

n=3

k=2, 2 of the 3 elements are high pass

T_i=0.9

Frequency range of 5 to 10 GHz in 10 steps

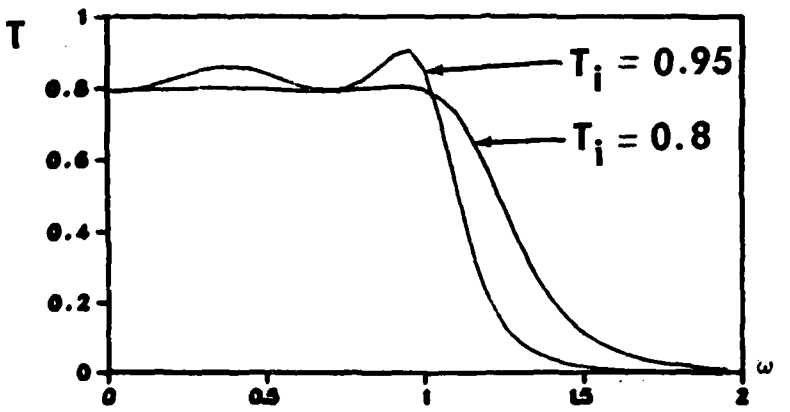
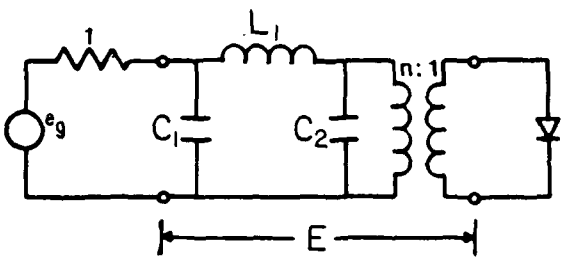


Figure 3
Example 1: PERFORMANCE



ELEMENT VALUES

	Normalized	Actual	Normalized	Actual
	T_i = 0.95		T_i = 0.8	
L_1	1.44	1.15 nH	1.20	0.96 nH
C_1	1.03	0.33 pF	0.55	0.18 pF
C_2	1.83	0.58 pF	1.44	0.46 pF
n	3.04		3.08	

Figure 4
Example 1: REALIZATION

This result in:

$$T(\omega)=0.9\pm0.009$$

The equalizer performance and its realization are shown in figures 5 and 6.

Example 3: Laser diode matched to a 50 ohm generator, to be designed without a transformer.

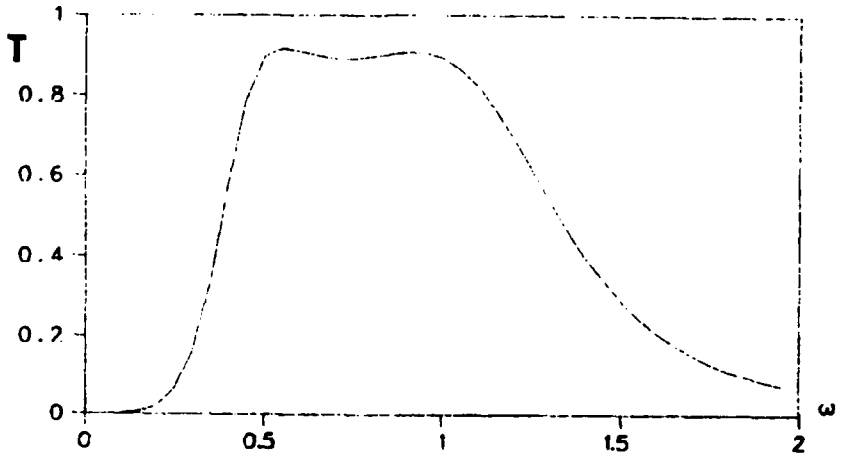
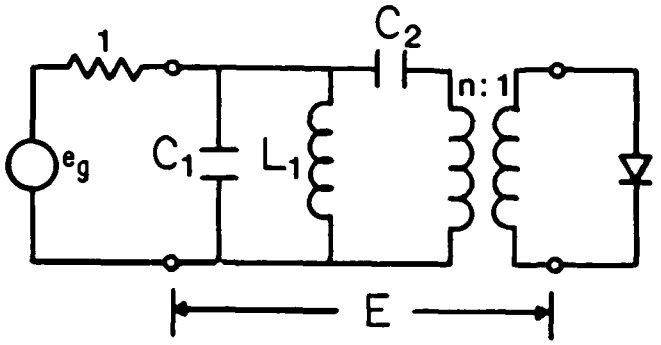


Figure 5
Example 2: PERFORMANCE



ELEMENT VALUES

	Normalized	Actual
L_1	1.69	1.35 nH
C_1	1.14	0.36 pF
C_2	1.28	0.41 pF
n	3.57	

Figure 6
Example 2: REALIZATION

Input:

n=5

k=0

$h(s)=0-s+s^2 -s^3 +s^4 -s^5$ (fixing h_0 =0 eliminates the transformer)

T_i=0.8

Frequency range from 5 to 10 GHz in 10 steps.

The performance of the network is

$$T(\omega)=0.803\pm0.012$$

as shown in figures 7 and 8.

Example 4: PIN diode(R=100, C=10pF) matched to a 50 ohm load, using a 2 element low pass design.

Result of optimization:

$$T(\omega)=0.501\pm0.012$$

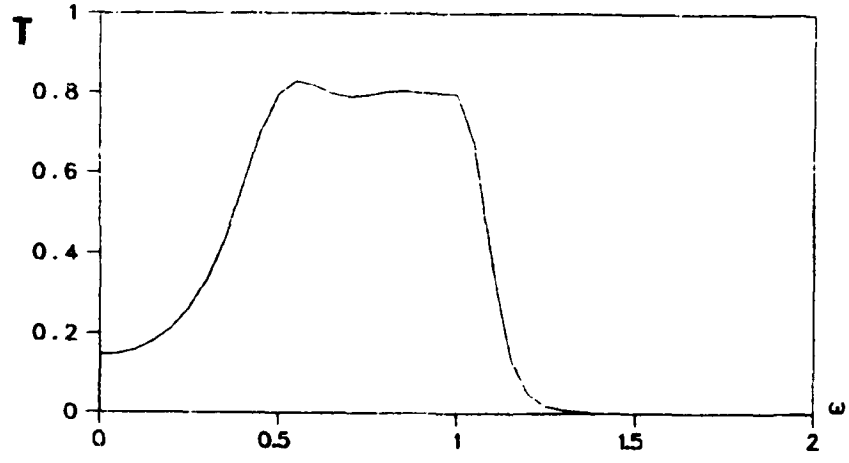
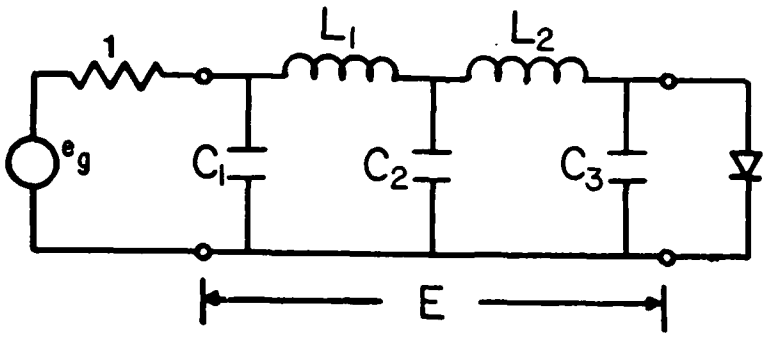


Figure 7
Example 3: PERFORMANCE



ELEMENT VALUES

	Normalized	Actual	
L_1	1.03	0.82	nH
L_2	0.52	0.41	nH
C_1	1.39	0.44	pF
C_2	4.69	1.49	pF
C_3	10.4	3.30	pF

Figure 8
Example 3: REALIZATION

The performance and realization are shown in figures 9 and 10.

In all the preceeding examples, if the idealized gain objective, T_i, is chosen too large for a passive realization the optimization routine simply fails to converge. This often manifests itself by the algorithm not being able to achieve a Hurwitz factorization for g(s). Thus active equalization would be necessary.

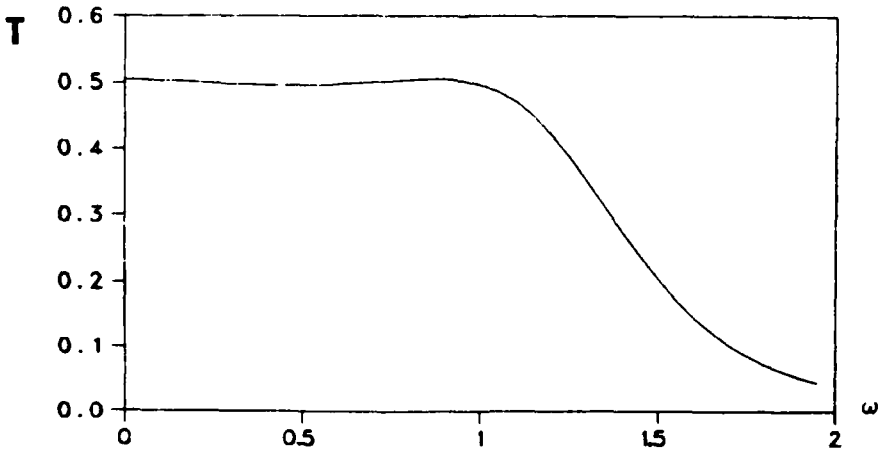
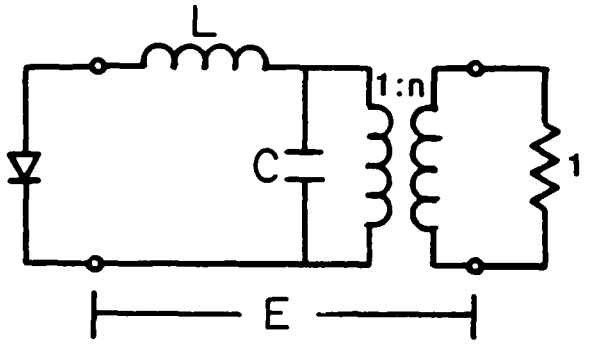


Figure 9
Example 4: PERFORMANCE



ELEMENT VALUES

	Normalized	Actual
L	0.26	0.20 nH
C	1.52	0.48 pF
n	1.7	

Figure 10
Example 4: REALIZATION

It should be noted however that the maximum possible pass band gain is not necessarily the most desirable solution. This is illustrated in example 1 where two ideal gain levels were specified. Clearly case (b) gave a flatter response, which is more desireable in most communication applications. By designing with the higher T_i, the optimization tried to meet this specification at the frequency sampling points but at the expense of lower gain between samples. This is consistent with the analytic theory of broadband matching. Although it is possible obtain to a match at any desired number of frequencies (provided the real part of the load impedance remains finite at those frequencies) it is incorrect to assume that one may obtain a reasonable match over a frequency band by correctly matching at a sufficiently large number of frequencies within the desired band [8].

Of course T_i need not always be chosen as a constant, but may include a taper if desired.

Example 2 illustrates the feature that an equalizer topology is not chosen in advance. Even though it is known that a three element matching network is desired,

the types of elements to be used and how they are to be placed is not clear, making it difficult to initiate commercially available CAD packages. Here the equalizer topology comes about directly from the synthesis of the input impedance function.

In example 3, we cannot allow $h_0=0$ and $k>0$ simultaneously since this violates the losslessness of E [2], namely

$$|e_{11}(j\omega)|^2 + |e_{21}(j\omega)|^2 = 1$$

The low frequency performance of the network is thus limited by the power transfer condition imposed by R_g and R_1.

V. RECOMMENDATIONS

It has been shown that Carlin's real frequency technique is appropriate for matching microwave signals to optical diodes. Distinct advantages of the technique are summarized as:

1. Readily implemented on most computers.
2. Design proceeds directly from measured real frequency data, bypassing the need for modeling the devices to be matched. Device nonlinearities are thus taken into account in a direct manner.
3. Realizability is guaranteed.
4. An equalizer topology is not chosen in advance.
5. Performance is superior to analytic techniques as well as other CAD techniques such as COMPACT.

Since the techniques employed here originated in the design of multi-stage microwave FET amplifiers, the question of active equalization is relatively straightforward and is currently being investigated.

Other topics currently being pursued include time domain performance and phase optimization of equilizer to yield constant group delay in the passband; front end matching to compensate for waveguide to microstrip discontinuities; realization using ultra wide-band distributed elements; and finally for matching microwave signals to systems employing indirect modulation techniques.

REFERENCES

1. H.J. Carlin, "A New Approach to Gain-Bandwidth Problems," IEEE Transactions on Circuits and Systems, Vol. CAS-24, No. 4, April 1977, pp. 170-175.

2. H.J. Carlin and B.S. Yarman, "The Double Matching Problem: Analytic and Real Frequency Solutions," IEEE Transactions on Circuits and Systems, Vol. CAS-30, No. 1, Jan. 1983, pp. 15-18.

3. B.S. Yarman and H.J. Carlin, "A Simplified 'Real Frequency' Technique Applied to Broad-Band Multistage Microwave Amplifiers," IEEE Transactions on Microwave Theory and Technique, Vol. MTT-30, No. 12, Dec. 1982, pp. 2216-2222.

4. V. Belevitch, Classical Network Theory, San Francisco, Holden Day, 1968, p. 278.

5. K.Y. Lau and A. Yariv, "Ultra-High Speed Semiconductor Lasers," IEEE Journal of Quantum Electronics, Vol. QE-21, No. 2, Feb. 1985, pp. 121-138.

6. R.S. Tucker and D.J. Pope, "Microwave Circuit Models of Semiconductor Injection Lasers," IEEE Transactions on Microwave Theroy and Techniques, Vol. MTT-31, No. 3, March 1983, pp. 289-294.

7. L. Figueroa, C.W. Slayman, and H.W. Yen, "High-frequency Characteristics of GaAlAs Injection Lasers," IEEE Journal of Quantum Electronics, Vol. QE-18, No. 10, Oct. 1982, pp. 1718-1727.

8. R.M. Fano, "Theoretical Limitations on the Broad-Band Matching of Arbitrary Impedance Functions," Journal of the Franklin Institute, Vol. 249, No. 1, Jan. 1950, pp.57-83.

1986 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT SUMMER SUPPORT PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

EVALUATION AND ANALYSIS OF VHSIC SOFTWARE TOOLS

Prepared by:	Dr. George W. Zobrist
Academic Rank:	Professor of Computer Science
Department and	Department of Computer Science
University:	University of Missouri - Rolla
Research Location:	Air Force Wright Aeronautical Laboratories
	Avionics Laboratory
	Microelectronics Branch/ AADE-3
USAF Researcher:	Mr. Tom Herbert
Date:	July 31, 1986
Contract No:	F49620-85-c-0013

Evaluation and Analysis of VHSIC Software Tools

by

Dr. George W. Zobrist

ABSTRACT

The evaluation and description of various VHSIC software tools is presented. The tools investigated were, the Intermetrics VHSIC hardware design language, the GE system designers workbench, and the RTI architecture design and assessment system. These tools have the capabilities of representing hardware in a design language that can be used for specifying the hardware and generating silicon, designing systems through a Buhr graphical representation which can be used to generate the VHSIC hardware design language and ADA code, simulating finite state machines, and the codesign of software modules and hardware processing elements. Recommendations for future work and enhancements in the VHSIC software tool support environment are presented.

Acknowledgements

I would like to thank the Air Force Wright Aeronautical Laboratories and the Air Force Office of Scientific Research for sponsorship of my research. The Avionics Laboratory provided an intellectually stimulating environment for the performance of my research. In particular I would like to thank several members of the Microelectronics Branch/AADE-3, Mr. Tom Herbert, Mr. Richard Wallace, Captain Robb Bellaccico, 1st Lieutenant Jack Tomlinson, and 2nd Lieutenant Rick Miller for giving me the opportunity, support, and guidance in this research endeavaor.

I. Introduction

I received my PhD from the University of Missouri - Columbia studying the problem of designing a software tool to implement Kirchhoff's third and fourth law on an IBM 1620. Since then my research experience and efforts have been in the area of computer aided design and analysis. In understanding, developing and doing research in this area, some of the programs I have worked on were, NASAP - a steady-state analysis program based upon signal flow graph theory, presenting workshops on ECAP, NET-1, SCEPTRE, and a project to benchmark USA and European CADA programs, and others.

The research problem at the Avionics Laboratory was similar in nature to previous efforts I have undertaken and because of this similarity I was assigned to this effort of evaluating VHSIC software design tools.

II. Objectives of the Research Effort

The overall objective of this effort was the analysis and evaluation of various VHSIC design and analysis software tools. The software tools investigated were, 1) The Intermetrics VHSIC hardware description language (VHDL) support environment, 2) The GE VHDL system designers workbench, and the 3) The RTI architecture design and assessment system (ADAS).

This involved the usage of these tools in a design environment and the understanding of the applicable theory used in the

software tools design, where applicable.

Specific objectives were,

- Usage of Intermetrics VHDL in a design capture.
- Usage of GE VHDL system designers workbench in the finite state machine editor and BRAT editor mode to generate both VHDL and ADA.
- Usage of RTI ADAS to implement software and hardware designs and simulate the same using the discrete event simulator and Petri net analyzer.
- Understanding of the theory applicable to the ADAS software tool.
- Presentation of a seminar on the relation of the ADAS model to Petri nets and partition/allocation approaches.

III. Description and Evaluation of the Software Tools

a) Intermetrics VHDL[1] - VHDL is a language that can be used to describe and specify hardware from simple logic gates to complex digital systems. This can be accomplished by specifying the inputs and outputs of the system and then either describing the system behaviorally or structurally. The behavioral description gives the outputs as a function of the inputs and the structural description gives the interconnection of various components which gives the desired output signals when the input signal is applied. These descriptions of a piece of hardware are called design entities. The design entity abstraction may be at the gate, register, component, or board level. The design entity

consists of an interface description and a set of one, or more, body descriptions.

The interface description is a specification of the systems input and output ports and their characteristics. In a sense this is a "black box" view of the design entity. The port declarations specify both their mode and type. The mode of a port specifies the direction in which data flows through the port. There are several types available; input, output, input/output (bidirectional), buffer, and linkage(direction unknown). The ports also must be typed, i.e., whether the data is bit, real, integer, character, etc.

An entity can be described by either an architectural or configuration body. An architectural body may be used to describe the behavior, or structure, of a design entity, while the configuration body is used to describe particular configurations of corresponding architectural bodies. The configuration body specifies which architectural body to use for various components within an architectural body being configured.

The architectural body contains the name of the architectural body, the design entity it is body of, and a block statement. The block statement contains a declarative and a statement part. The declarative part includes signal, types, components, and other declarations. The statement part contains component

instantiations, signal assignments and other concurrent statements.

A component is a physical or logical device that describes a particular operation characterized by an interface and a function. A component instantiation is the creation of a particular instance of a component.

The behavioral primitive within VHDL is the process statement and its associated sequential statements. There are three basic parts to a process statement, the input list which contains the signals that the process is initially sensitive to, the declarative part which contains the local variable declarations, and the statement part which contains the statements defining the algorithm which computes new outputs based upon changes in the inputs.

The process statements form a basis for the event driven simulation of the VHDL descriptions by executing a sequence of statements and possibly scheduling new events to occur.

The VHDL support environment consists of six components; the analyzer, the reverse analyzer, the simplifier, the simulator, the design library, and a design library manager.

The analyzer checks hardware descriptions for static errors, i.e., errors occuring without regard to simulation and

translates VHDL to an intermediate language. The reverse analyzer constructs VHDL source text from the intermediate language. The simplifier reorganizes the description of a hardware design. This may result from a top down design where a subcomponent is designed after the design of the higher level components. Prior to simulation the designer must bind the abstract components to actual design entities. The simplifier takes as input a configuration body describing a hierarchical description and produces a single design unit. The simulator computes the successive signal value which occur as a result of the time-dependent behavior of a particular hardware design. The design library contains the design data resources which can be used by the designer and the design library manager integrates the elements of the VHDL support environment.

b) GE VHDL system designers workbench[2] - The system designers workbench (SDW) is a collection of integrated tools for system design. The tools described are the; the BUHR representation editor (BRAT), and the state machine editor (SME).

The BRAT editor is a tool for designing systems using a graphic paradigm developed by R.J.A. Buhr. The user designs with BRAT by creating icons and connecting them, the units represented by icons can be used to model systems under investigation.[3]

Buhr diagrams are graphic representations of system components, originally they were used to represent ADA code objects. The objects contain sockets which represent entry points and communication between objects is accomplished by connections.

There are three types of program units from which programs can be constructed. They are; packages, tasks, and clouds. Packages allow specification of groups of logically related entities, tasks are code entities whose executions proceed in parallel, and clouds are conceptual features used by the designer before decisions are made on the particular usage of the program unit. Any subcomponent or connection valid for a package, or task, is valid for a cloud.

One feature of the BRAT editor is to associate a truth table with the program units input/output variables. Once this association is achieved then ADA code and/or a VHDL description can be generated for the proposed design.

The state machine editor allows the user to create and simulate state machines. These machines communicate through channels. State machines can be used to describe systems that move between states. These state changes are called transitions. Transitions have conditions associated with them and when the condition is met the transition will occur moving the machine between states.

After the state machine has been described graphically and the channel defined the machine can be simulated for one unit of simulation time whenever the step command is entered.

c) RTI Architecture Design Assessment System[4,5,6] - The ADAS package is an integrated set of tools used in the codesign of software and hardware. The tool set consists of an editor, a discrete event simulator, a Petri net analyzer and a software to hardware mapping tool. A VHDL generator is planned in a future version.

ADAS models the software and/or hardware processes by a directed graph methodology based upon a modified computational graph. A computational graph is a subclass of a Petri net.

A Petri net[7] is composed of several parts; a set of places, a set of transitions, an input function, and an output function. The input and output functions relate transitions and places. A marking is an assignment of tokens to the places of a Petri net. Tokens are used to define the execution of a Petri net. A Petri net executes by firing transitions. A transition can be fired if it is enabled. It will be enabled if each of its input places has as many tokens as arcs emanating from the place. A transition fires by removing all the enabling tokens from the input places incident upon the transition and places tokens in all places radiating from the transition being fired.

A computational graph is defined as a directed graph composed of vertices and arcs. Associated with each arc is a quadruple (I_{jk}, V_{jk}, W_{jk}, T_{jk}) and each arc represents a queue of tokens produced by node j and consumed by node k. The quadruple items are defined as; I_{jk} is the number of tokens initially in the queue, V_{jk} is the the number of tokens produced when node j fires, W_{jk} is the number tokens consumed by node k when it fires, T_{jk} is the number of tokens required before node k can be fired.[8]

A computation graph can be modeled by a Petri net, where each arc is represented by a place and each node is a transition. A computation graph is an extension of a marked graph.[9] The marked graph has $T_{jk} = W_{jk}$. *The distinguishing feature of* computation and marked graphs in their Petri net model is that all places have only one input transition and only one output transition.

The directed graph methodology is obtained by extending the computation graph by the process of attaching program procedures to each node. This then allows one to represent nodes as processing elements. The input queues to the node, and its associated procedure, may be regarded as the formal input parameters to the procedure and the output queues associated with the node, and its associated procedure, may be regarded as the formal output parameters of the associated

procedure.[10,11]

The graph queues can be used to transport both data items and synchronization and/or control tokens between nodes of the graph.

A further extension is to assign a firing delay[12] to each node and an execution priority.

After the software/hardware process has been built using the directed graph editor the designer has the option of simulating the execution of the software and/or hardware design and generating a performance analysis. The performance analysis report includes node utilization and arc queue size changes.

There is also a Petri net analyzer which generates steady-state statistics by analyzing the characteristics of the model. The performance analysis report includes node utilization statistics and arc data flow rates.

This information is obtained by utilization of the following theorem,

> Any finite place, finite transition, marked stochastic Petri net is isomorphic to a one-dimensional discrete space Markov process[13]

The steps are to find all the reachable states of the associated Petri net, solve for the steady-state probabilities of each marking (state), determine the probability of a token in each place (arc), and then determine the probability a transition (node) will fire. This will determine the node utilization and by dividing the node utilization by the node service time one can obtain the arc flow rate.[14,15]

This algorithm presumes that the firing rates (reciprocal of service time) are the averages for exponentially distributed transition (nodes) firing times. This algorithm has been deduced from the various papers presented by the RTI ADAS group.

The software to hardware mapping is a process whereby software graph modules are mapped onto hardware processor elements. The mapping tool optimizes system throughput by minimizing interprocessor communication costs and balancing processor load.

This is the partition/allocation problem[16] whereby first functions are partitioned into common software processing modules and then these partitions are allocated to hardware processing elements to optimize various items. In our case the software functional modules and the hardware processing elements are built using the directed graph editor and then through the utilization of the software to hardware mapping

the software modules are mapped onto the hardware processing elements. This mapping is combinatoric in nature, but the problem can be reduced by the addition of constraints and usage of various branch-bound algorithms.

IV. Recommendations

The VHSIC design software will provide powerful support for the designer. The tools evaluated cover the spectrum from chip to system design, and through discrete event simulation to analytical tools.

The following recommendations are a result of this study,

- A theoretical section should be included for each software tool, or portion of a software tool, where applicable. It is difficult for a designer to rely on a result if little is known about assumptions made in the analytics used on the programmed mathematical model. This applies, in particular, to the ADAS software tool.

- A study should be made on developing optimized VHDL code from the state machine and BRAT editor descriptions. This would include developing algorithms for going from state machine descriptions, to state minimization, to state transition tables, to next state maps for registers, to register maps, to optimized logic equations, to optimized VHDL code. This could be done for both synchronous and asynchronous machines.

- Expansion of types of finite state machines the state

machine editor is capable of describing, i.e., controllers, states modified internal to the state machine, event driven states, etc.

- Exploration of a Petri net simulator in addition to the extended computation graph model used by ADAS. This would allow incorporation of finite state machines. It would also remove the computation graph modeling restriction of places (arcs) with only one input/output arc.

- Investigation of formal verification techniques for verification of logic networks. At present one can use software tools, such as, the BRAT editor to generate packages and/or tasks and assign truth tables to the input/output sockets without regard to compatability. This could be implemented through incorporation of an expert system such as OPS-5.

REFERENCES

1. VHDL User's Manual, Volume I - Tutorial, Intermetrics Inc., August 1985.
2. VHDL Workstation - Draft Report, General Electric, 1985.
3. R. J. A. Buhr, System Design with ADA, Prentice Hall, Englewood Cliffs, N.J., 1984.
4. Architecture Design and Assessment System (ADAS) - User Manual, Version 2.0/CV, RTI, 1985.
5. Frank, C., C. Smith, and J. Cuadrado, "An Architecture Design and Assessment System for Software/Hardware Codesign", Proc. 22nd Design Automation Conference, Las Vegas, Nev., June 1985, pp. 460 - 463.
6. Frank, G., D. Franke, and W. F. Ingogly, "An Architecture Design and Assessment System", VLSI Design, August 1985, pp. 30 - 38.
7. Peterson, J. L., Petri Net Theory and the Modeling of Systems, Prentice-Hall, Englewood Cliffs, N.J., 1981.
8. Karp, R. and R. Miller, "Properties of a Model for Parallel Computation: Determinacy, Termination, and Queueing, "SIAM Journal of Applied Math, Vol. 14, No. 6, November 1966, pp. 1390 - 1411.
9. Commoner, A., A. Holt, S. Even, and A. Pnueli, "Marked Directed Graphs", Journal of Computer and System Sciences, Vol. 5, No. 5, October 1971, pp. 511 - 523.
10. Cuadrado, J., W. Honey, and A. Wenk, "Systems Programming Aids (SPA) Methodology", Proc of 1982 NAECON, 1982, pp. 1305 - 1311.

11. Cuadrado, J., and G. Linsenmayer, "Efficient High Speed Implementation of Directed Graph Signal Processing in a Distributed Processing System", Proc. of IEEE Compcon, Spring 1983, pp. 460 - 463.

12. Zuberek, W. M., "Timed Petri Nets and Preliminary Performance", Proc. of IEEE 7th Annual Symp on Computer Architecture, 1980, pp. 89 - 96.

13. Molloy, M., On the Integration of Delay and Throughput Measures in Distributed Processing Models, PhD Dissertation, Univ. California, Los Angeles, CA, 1981.

14. Molloy, M., "Performance Analysis Using Stochastic Petri Nets", IEEE Trans on Computers, Vol. C-31, No. 9, September 1982, pp. 913 - 917.

15. Florin, G. and S. Natkin, "Evaluation Based upon Stochastic Petri Nets of the Maximum Throughput in a Full Duplex Protocol", Application and Theory of Petri Nets: Lecture Notes in Computer Science, Springer-Verlag, N.Y., 1982, 52: 280 - 288.

16. Mariani, M. P. and D. I. Palmer, Distributed System Design, IEEE Tutorial, New York, New York, 1979.