AD A139404

DTIC FILE COPY

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

UNITED STATES AIR FORCE
SUMMER FACULTY
RESEARCH PROGRAM

1983

TECHNICAL REPORT
VOLUME 2 OF 2

SOUTHEASTERN CENTER FOR
ELECTRICAL ENGINEERING EDUCATION
(SCEEE)

WARREN D. BELL

PROGRAM DIRECTOR, SCEEE

ATHENS, GEORGIA

DTIC
ELECTE
MAR 29 1984
E

This document has been approved for public release and sale; its distribution is unlimited.

84 03 22 133

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AFOSR-TR- 84-0155	2. GOVT ACCESSION NO.: AD-A139404
3. RECIPIENT'S CATALOG NUMBER:	
4. TITLE (and Subtitle): UNITED STATES AIR FORCE SUMMER FACULTY RESEARCH PROGRAM(1983), TECHNICAL REPORT, VOL 2 of 2	5. TYPE OF REPORT & PERIOD COVERED: FINAL
	6. PERFORMING ORG. REPORT NUMBER:
7. AUTHOR(s): Warren D. Peele, Earl L. Steele, Maj. Amos L. Otis	8. CONTRACT OR GRANT NUMBER(s): F49620-82-C-0035
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Southeastern Center for Electrical Engineering Education, 11th & Massachusetts Ave; St. Cloud, FL 32769	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 61102F 2301/D5
11. CONTROLLING OFFICE NAME AND ADDRESS: AFOSR/XOT, Building 410, Bolling AFB, DC 20332	12. REPORT DATE: December 1983
	13. NUMBER OF PAGES:
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office):	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

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

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

Cont'd on back

#20.) ABSTRACT CONT'D:

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 the Southeastern Center for Electrical Engineering.

UNITED STATES AIR FORCE

SUMMER FACULTY RESEARCH PROGRAM

1983

PROGRAM TECHNICAL REPORT

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

Volume II of II

Program Director, SCEEE
Warren D. Peele
Earl L. Steele

Program Manager, AFOSR
Major Amos L. Otis

Submitted to

Air Force Office of Scientific Research

Bolling Air Force Base

Washington, DC

December 1983

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 the Southeastern Center for Electrical Engineering.

The specific objectives of the 1983 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; and

(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 1983, 101 faculty members participated. These researchers were assigned to 25 USAF laboratories/centers across the country. This two volume document is a compilation of the final reports written by the assigned faculty members about their summer research efforts.

LIST OF PARTICIPANTS

NAME/ADDRESS	DEGREE, SPECIALTY, LABORATORY ASSIGNED
Dr. Richard Anderson Professor University of Rolla Physics Department Rolla, MO 65401 (314) 341-4797	Degree: Ph.D., Physics, 1959 Specialty: Atomic and Molecular Physics and Optics Assigned: RADC/Griffiss
Dr. Richard Andrews Associate Professor University of Michigan Graduate School of Business Administration Ann Arbor, MI 48109 (313) 662-4106	Degree: Ph.D., Statistics, 1973 Specialty: Finite Population Sampling and Data Assigned: LMC
Dr. Francesco Bacchialoni Associate Professor University of Lowell Electrical Engineering Department Lowell, MA 01854 (617) 452-5000	Degree: Doctor in Engineering, 1946 Specialty: Control Systems, Digital Signal Processing, Microprocessors Assigned: GL
Dr. Pradip Bakshi Research Professor Boston College Physics Department Chestnut Hill, MA 02167 (617) 969-0100	Degree: Ph.D., Theoretical Physics, 1962 Specialty: Theoretical Plasma Physics, Quantum Theory, Mathematical Physics Assigned: GL
Dr. Daniel Barr Assistant Professor Virginia Military Institute Electrical Engineering Department Lexington, VA 24450 (703) 463-6236	Degree: Ph.D., Electrical Engineering, 1978 Specialty: Fuzing, Millimeter Waves, Superconductivity, Semiconductors Assigned: AD
Mr Ernesto Barreto Senior Research Associate State University New York Atomic Science Research Center Albany, NY 12309 (518) 457-4930	Degree: MS, Physics, 1960 Specialty: Electrostatics, Electrical Discharges, Combustion Assigned: FDL

List of Participants (continued: page 2)

Dr. Stanley Bashkin
Professor
Univeristy of Arizona
Physics Department
Tucson, AZ 85711
(602) 626-2322

Degree: Ph.D., Physics, 1950
Specialty: Accelerator-based Atomic Physics, Atomic Spectroscopy
Assigned: GL

Dr. Joseph Baumgarten
Professor
Iowa State University
Mechanical Engineering Department
Ames, Iowa 50010
(515) 294-1380

Degree: Ph.D., Mechanical Engineering, 1958
Specialty: Vibrations, Machinery Dynamics
Assigned: AEDC

Dr. Michael Becker
Associate Professor
University of Texas
Electrical Engineering Department
Austin, TX 78712
(512) 471-3628

Degree: Ph.D., Electrical Engineering, 1973
Specialty: Lasers, Laser Material Interactions, Nonlinear Optics
Assigned: WL

Dr. Henry Bertoni
Professor
Polytechnic Institute of NY
Electrophysics Department
Brooklyn, NY 11201
(212) 643-4832

Degree: Ph.D., Electrophysics, 1967
Specialty: Ultrasonic and Electromagnetic Propagation and Diffraction
Assigned: RADC

Dr. Willie Bragg
Program Manager
University of Cincinnati
College of Education
Cincinnati, OH 45221
(513) 475-4542

Degree: Ph.D., Special Education, 1979
Specialty: Mental Retardation, Early Childhood Education, Speech and Hearing
Assigned: SAM

Dr. M. Quinn Brewster
Assistant Professor
University of Utah
Mechanical Engineering Department
Salt Lake City, Utah 84106
(801) 581-7105

Degree: Ph.D., Mechanical Engineering, 1981
Specialty: Thermal Radiation in Particulate Media
Assigned: RPL

Mr Louis W. Buckalew
Associate Professor
Alabama A&M University
Psychology Department
Normal, AL 35762
(205) 322-2489

Degree: MS, Psychology, 1969
Specialty: Drug Response Factors, Human Engineering, Alcohol
Assigned: SAM

List of Participants (continued: page 3)

Dr. Chester Canada
Associate Professor
Oklahoma State University
Computer Science Department
Stillwater, OK 74078
(409) 624-5724

Degree: Ph.D., Physics, 1968
Specialty: Materials Response to Dynamic Loading
Assigned: ESC

Dr. Jack Chatelain
Professor
Utah State University
Physics Department
Logan, UT 84321
(801) 750-2936

Degree: Ph.D., Physics, 1957
Specialty: Theoretical Physics
Assigned: WL

Mr. John Cicero
Instructor
Illinois Institute of Technology
Electrical Engineering Department
Chicago, IL 60532
(312) 567-3400

Degree: MSEE, Electrical Engineering, 1978
Specialty: Microprocessor Hardware and Software Design. Multiple Access Satellite Network Protocols
Assigned: HRL/W

Dr. Frank Collins
Professor
University of Tennessee
Aerospace Engineering Department
Tullahoma, TN 37388
(615) 455-0631

Degree: Ph.D., Mechanical Engineering 1968
Specialty: Physical Fluid Dynamics
Assigned: AEDC

Dr. Richard Conte
Assistant Professor
Manhattan College
Mechanical Engineering Department
Bronx, NY 10471
(212) 920-0148

Degree: Ph.D., Mechanical Engineering,
Specialty: Heat Transfer, Numerical Computer Analysis, Solar Energy
Assigned: APL

Dr. Billy Covington
Assistant Professor
Sam Houston State University
Physics Department
Huntsville, TX 77340
(713) 294-1606

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

Dr. Peter Crane
Assistant Professor
University of Pittsburgh
Psychology Department
Johnstown, PA 15905
(814) 266-9661

Degree: Ph.D., Experimental Psychology, 1979
Specialty: Cognitive Psychology/ Human Factors
Assigned: AMRL

List of Participants (continued: page 4)

Dr. Carolyn Crouch
Associate Professor
University of Alabama
Computer Science Department
University, AL 35486
(205) 348-6363

Degree: Ph.D., Computer Science, 1971
Specialty: Information Storage and Retrieval, Operating Systems and Systems Programming
Assigned: AD

Dr. Donald Crouch
Associate Professor
University of Alabama
Computer Science Department
University, AL 35486
(205) 348-6363

Degree: Ph.D., Systems Engineering, 1972
Specialty: Design of Languages, DBMS, Information Systems
Assigned: AD

Dr. Carol Deakyne
Assistant Professor
College of The Holy Cross
Chemistry Department
Worcester, MA 01610
(617) 793-3367

Degree: Ph.D., Theoretical Chemistry, 1976
Specialty: Applications of Molecular Orbital Theory
Assigned: AD

Dr. Terry L. Dickinson
Visiting Professor
Old Dominion University
Psychology Department
Norfolk, VA 23508
(804) 440-4235

Degree: Ph.D., Psychology; 1968
Specialty: Industrial/Organizational Psychology and Psychometrics
Assigned: HRL/B

Dr Fred E. Domann
Associate Professor
University of Wisconsin
Physics Department
Platteville, WI 53818

Degree: Ph.D., Physics, 1975
Specialty: Surface Science
Assigned: WL

Dr. James W. Dooley
Assistant Professor
Wright State University
HPER Department
Dayton, OH 45435
(513) 873-3259

Degree: Ph.D., Physical Education, 1979
Specialty:
Assigned: SAM

Dr. George R. Doyle
Associate Professor
University of Dayton
Mech. Engineering Department
Dayton, OH 45469
(513) 229-2835

Degree: Ph.D., Mechanical Engineering, 1973
Specialty: Dynamics
Assigned: APL

List of Participants (continued: page 5)

Dr. Jonathan C. Dutton
Assistant Professor
Texas A&M University
Department of Mech. Engineering
College Station, TX 77834
(713) 845-5011

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

Dr. John Eoll
Assistant Professor
Lenoir-Rhyne College
Physics Department
Hickory, NC 28601
(704) 328-1741

Degree: Ph.D., Astrophysics, 1976
Specialty: Radiaton Transport, Fluid Dynamics, Nuclear Weapons Effects
Assigned: WL

Dr. Amir Faghri
Associate Professor
Wright State University
Engineering Department
Dayton, OH 45435
(513) 873-2501

Degree: Ph.D., Mechanical Engineering, 1976
Specialty: Heat Transfer, Fluid Mechanics, Engineering Analysis
Assigned: APL

Dr. Hans Fellner
Associate Professor
Slippery Rock State College
Physics Department
Slippery Rock, PA 16057
(412) 794-7781

Degree: Ph.D., Physics, 1973
Specialty: Light Scattering, Liquid Crystals, Phase Transitions
Assigned: AMRL

Dr. Robert Foley
Assistant Professor
Virginia Poly. Inst.
IEOR Department
Blacksburg, VA 24061
(703) 961-7142

Degree: Ph.D., Industrial and Operation Engineering, 1979
Specialty: Operations Res., Queueing Network Theory, Stochastic Processes
Assigned: HRL/WP

Dr. Eddie Fowler
Associate Professor
Kansas State University
Electrical Engineering Department
Manhattan, KS 66506
(913) 532-5600

Degree: Ph.D., Electrical Engineering, 1969
Specialty: Modeling and Simulation/ Speech Recognition and Vocal Tract Modeling
Assigned: WL

Dr. Victor S. Frost
Assistant Professor
University of Kansas
Lawrence, KS 66045
(913) 864-4615

Degree: Ph.D., Electrical Engineering, 1982
Specialty: Communications and Signal Processing
Assigned: RADC

List of Participants (continued: page 6)

Dr. Patrick Garrett
Professor and Chairman
University of Cincinnati
Electrical Engineering Department
Cincinnati, OH 45221
(513) 465-4651

Degree: Ph.D., Electrical Engineering, 1970
Specialty: Computer Interface Design for Data Acquisition and Digital Control
Assigned: AL

Dr. Richard Gill
Assistant Professor
Wright State University
Engineering Department
Dayton, OH 45435
(513) 873-2701

Degree: Ph.D., Mechanical Engineering, 1981
Specialty: Human Factors Engineering
Assigned: AMRL

Dr. John Giolma
Assistant Professor
Trinity University
Science Department
San Antonio, TX 78209
(512) 736-7563

Degree: Ph.D., Electrical Engineering, 1975
Specialty: Signal Processing, Modeling and Simulation
Assigned: SAM

Dr. Samuel Green
Associate Professor
Auburn University
Psychology Department
Auburn, Alabama 36830
(205) 826-4412

Degree: Ph.D., Psychology, 1975
Specialty: Statistics, Measurement Theory, Industrial Psychology
Assigned: LMDC

Dr. Arthur Gutman
Associate Professor
Florida Inst. of Technology
Psychology Department
Melbourne, Fla 32901
(305) 723-3701

Degree: Ph.D., Psychology, 1975
Specialty: Experimental and Bio-psychology; Computer Applications
Assigned: HRL/L

Dr. Terry Herdman
Associate Professor
Virginia Tech
Mathematics Department
Blacksburg, VA 24060
(703) 961-5279

Degree: Ph.D., Mathematics, 1974
Specialty: Volterra Integral Equations Ordinary and Functional Differential Equations
Assigned: FDL

Dr. Stuart Hirschfield
Assistant Professor
Hamilton College
Mathematics Department
Clinton, NY 13323
(315) 859-4136

Degree: Ph.D., Computer and Information Sciences, 1978
Specialty: Artificial Intelligence, Software Engineering
Assigned: RADC

Dr. Craig Holt
Assistant Professor
Tufts University
Electrical Engineering Department
Medford, MA
(617) 628-5000

Degree: Ph.D., Electrical Engineering, 1981
Specialty: Distributed Digital Systems Fault Tolerance
Assigned: ESD

Dr. Kathleen Howell
Assistant Professor
Purdue University
Aeronautics & Astronautics Dept.
West Lafayette, IN 47906
(317) 494-5786

Degree: Ph.D., Aeronautics and Astronautics, 1983
Specialty: Orbit Mechanics, Spacecraft Dynamics
Assigned: RPL

Dr. Gwendolyn Howze
Associate Professor
Texas Southern University
Biology Department
Houston, TX 77021
(713) 527-7005

Degree: Ph.D., Molecular Biology, 1974,
Specialty: Electron Microscopy and and Biochemistry of Cell Nuclei
Assigned: AMRL

Dr. Medhat Ibrahim
Professor
California State Univ., Fresno
Electrical Engineering Department
Fresno, CA 93612
(209) 294-4020

Degree: Ph.D., Electrical Engineering, 1969
Specialty: Electromagnetic Systems and Control and Power Systems
Assigned: APL

Dr. Charles Ih
Professor
University of Delaware
Electrical Engineering Department
Newark, DE 19711
(302) 738-8173

Degree: Ph.D., Physics
Specialty: Optical Communication, Electro-Optics, Holography and Applications
Assigned: RADC

Dr. Gregory Jones
Associate Professor
Utah State University
Computer Science Department
Logan, UT 84322

Degree: Ph.D., Mathematics, 1972
Specialty: Computability, Software Engineering, Operating Systems
Assigned: HRL/L

Dr. Amir Karimi
Assistant Professor
University of Texas San Antonio
Engineering Department
San Antonio, TX 78230
(512) 691-4490

Degree: Ph.D., Mechanical Engineering, 1982
Specialty: Thermal Sciences - Condensation Heat Transfer, Metastable Thermodynamics
Assigned: SAM

List of Participants (continued: page 8)

Dr. Jerome Keating
Assistant Professor
University of Texas SA
Mathematics Department
San Antonio, TX 78285
(512) 691-4452

Degree: Ph.D., Mathematical Sciences, 1980
Specialty: Mathematical Statistics
Assigned: SAM

Dr George Kirby
Assistant Professor
West Virginia University
Mech. & Aerospace Eng. Department
Morgantown, West Virginia 26505
(304) 293-4111

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

Mr. James Kirkpatrick
Associate Professor
Alabama A&M University
Mathematics Department
Normal, Alabama 35762
(205) 859-7239

Degree: MS, Physics, 1981
Speciality: Applied Mathematics
Assigned: AEDC

Dr. Stephen Krause
Assistant Professor
Arizona State University
Mech. & Aerospace Eng. Department
Tempe, AZ 85287
(602) 965-2050

Degree: Ph.D., Engineering Materials, 1981
Specialty: Electron Microscopy, X-Ray Diffraction, & Properties of Polymers
Assigned: ML

Dr. Madakasira Krishna
Associate Professor
South Carolina State University
Computer Science Department
Orangeburg, SC 29117
(803) 536-7120

Degree: Ph.D., Fluid Mechanics, Numerical Analysis, Computer Science
Specialty: Computational Fluid Mechanics
Assigned: FDL

Dr. William Kyros
Associate Professor
University of Lowell
Mechanical Engineering Department
Lowell, Mass 01854
(617) 459-9357

Degree: Ph.D., Education, 1980
Specialty: Mechanical Behavior of Materials
Assigned: ML

Dr. David C. Lai
Professor
University of Vermont
Comp. Sci. & E.E. Department
Burlington, VT 05405
(802) 656-3330

Degree: Ph.D., Electrical Engineering, 1960
Specialty: Signal Processing and Pattern Recognition
Assigned: RADC

List of Participants (continued: page 9)

Dr. Stella Lawrence
Professor
Bronx Community College, CUNY
Electrical Engineering Department
Bronx, NY 10467
(212) 220-6044

Degree: MS, Mathematics, 1951
Specialty: Interference Tolerant Communications Systems, Coding and Data Compression, Fault Tolerant Computers
Assigned: AL

Dr. David Lee
Associate Professor
University of Dayton
Management Department
Dayton, OH 45409
(513) 229-4249

Degree: Ph.D., Industrial Engineering, 1972
Specialty: Operations Management
Assigned: BRMC

Dr. Stanley Lee
Professor
Kansas State University
Industrial Engineering Department
Manhattan, KS 66506
(913) 532-5606

Degree: Ph.D., Chemical Engineering, 1962
Specialty: Systems Engineering, Optimization and Control, Modeling
Assigned: LC

Dr. Mark Lewittes
Assistant Professor
University of Texas SA
Engineering Department
San Antonio, TX 78285
(512) 691-4490

Degree: Ph.D., Electrical Engineering 1982
Specialty: Laser Applications
Assigned: SAM

Dr. Sigurd Lillevik
Assistant Professor
Oregon State University
Elec. & Comp. Engineering Dept.
Corvallis, OR 97331
(503) 758-5318

Degree: Ph.D., Electrical Engineering, 1978
Specialty: Multiprocessor Computers
Assigned: AL

Dr. Stephen Lin
Associate Professor
North Carolina Central University
Chemistry Department
Durham, NC 27707
(919) 683-6463

Degree: Ph.D., Physical Chemisty, 1970
Specialty: Molecular Spectroscopy, Photoelectron Spectroscopy, Computer Simulation of Protein Folding
Assigned: ESC

Dr. Leonard Lion
Assistant Professor
Cornell University
Environment Engineering Dept.
Ithaca, NY 14853
(607) 256-7571

Degree: Ph.D., Environmental Engineering, 1980
Specialty: Chemistry of Water and Wastewater, Applied Aquatic Chemistry
Assigned: ESC

List of Participants (continued: page 10)

Dr. Daryl Logan
Associate Professor
Rose-Hulman Institute
Civil Engineering Department
Terre Haute, IN 47803
(812) 877-1151

Degree: Ph.D., Structural Engineering, 1976
Specialty: Structural Design/ Mechanics
Assigned: ESC

Dr. Charles Mastin
Professor
Mississippi State University
Mathematics & Statistics Dept.
Mississippi State, MS 39762
(601) 325-3414

Degree: Ph.D., Mathematics, 1969
Specialty: Computational Fluid Dynamics
Assigned: AD

Dr. Tapaz Mazumdar
Associate Professor
Wright State University
Mathematics & Statistic Dept.
Dayton, OH 45435
(513) 873-2785

Degree: Ph.D., Mathematics, 1971
Specialty: Partial Differential Equations and Related Functional Analysis
Assigned: FDL

Dr. Michael McKee
Assistant Professor
Auburn University
Chemistry Department
Auburn, Alabama 36849
(205) 862-4043

Degree: Ph.D., Chemistry, 1977
Specialty: Inorganic Chemistry, Theoretical Chemistry
Assigned: FJSRL

Dr. Robert McLauchlan
Assistant Professor
Texas Tech University
Mechanical Engineering Dept.
Lubbock, TX 79423
(806) 742-3563

Degree: Ph.D., Mechanical Engineering 1978
Specialty: Fluid-Structure (Acoustic) Interactions
Assigned: WL

Dr. Donald Michelsen
Assistant Professor
Virginia Tech
Chemical Engineering Department
Blacksburg, VA 24061
(703) 961-5157

Degree: Ph.D., Chemical Engineering, 1967
Specialty: Mass Transfer, Hazardous Waste Treatment
Assigned: ESC

Dr. George Miner
Associate Professor
University of Dayton
Physics Department
Dayton, OH 45469

Degree: Ph.D., Physics, 1965
Specialty: Magnetic Resonance, Transport Properties
Assigned: ML

List of Participants (continued: page 11)

Dr. Don Mittleman
Professor
Oberlin College
Mathematics Department
Oberlin, OH 44074
(216) 775-8385

Degree: Ph.D., Mathematics, 1951
Specialty: Applied Mathematics, ODE, Number Analysis, Geom.PDE,OR
Assigned: FDL

Dr. Charles Moseley
Associate Professor
Ohio State University Lima
Chemistry Department
Lima, OH 45804
(419) 228-2641

Degree: Ph.D., Organic Chemistry, 1967
Specialty: Organic Chemistry
Assigned: ML

Dr. Dale Moses
Associate Professor
San Diego State University
Aerospace & Engineering Mech. Dept.
San Diego, CA 92182
(619) 265-6074

Degree: Ph.D., Aerospace Engineering, 1981
Specialty: Subsonic and Transonic Wind Tunnel Testing
Assigned: FDL

Dr. Randolph Moses
Instructor
Virginia Poly. Inst.
Electrical Engineering Dept.
Blacksburg, VA 24061
(703) 961-5114

Degree: MS, Electrical Engineering,
Specialty: Digital Signal Processing
Assigned: RADC/Griffiss

Dr. James Mrotek
Associate Professor
Meharry Medical College
Physiology Department
Nashville, TN 37208
(615) 327-6288

Degree: Ph.D., Biology, 1973
Specialty: Physiological Response of Cultured Adrenal Cells to Stressors
Assigned: SAM

Dr. Frederick Nagle
Professor
University of Miami
Geology Department
Coral Gables, FL 33124
(305) 284-4254

Degree: Ph.D., Geology, 1966
Specialty: Field Mapping Geology Caribbean Geology, Minerology
Assigned: GL

Dr. Philip Olivier
Assistant Professor
Louisiana State University
Elec. & Comp. Engineering Dept.
Baton Rouge, LA 70803
(504) 388-5241

Degree: Ph.D., Electrical Engineering, 1980
Specialty: Mathematical Systems Theory
Assigned: RADC

Dr. Albert Payton
Associate Professor
Hampton Institute
Chemistry & Phys. Sci. Dept.
Hampton, VA 23668
(804) 727 5609

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

Dr David Pegg
Professor
University of Tennessee
Physics Department
Knoxville, TN 37916
(615) 974-5478

Degree: Ph.D., Physics, 1970
Specialty: Atomic Physics; Spectroscopy of Fast Moving Ion Beams. Lasers
Assigned: WL

Dr. Gerald Poje
Assistant Professor
Miami University
Zoology Department
Oxford, OH 45056
(513) 529-3624

Degree: Ph.D., Biology & Environmental Health Sciences, 1981
Specialty: Ecological Toxicology
Assigned: AMRL

Dr. L. Pujara
Associate Professor
Wilberforce University
Natural Science Department
Wilberforce, OH 45384
(513) 376-2911

Degree: Ph.D., Mathematics,1971
Specialty: Control Systems, Model Reduction, Mathematical Analysis
Assigned: AL

Dr. David L. Questad
Assistant Professor
Pennsylvania State University
Eng. Sci. & Mech. Department
University Park, PA 16802
(814) 863-2367

Degree: Ph.D., Mechanics & Material Sciences, 1981
Specialty: Polymer Science and Engineering
Assigned: RPL

Dr. Dallas Russell
Professor
Auburn University
Electrical Engineering Dept.
Auburn, AL 36849
(205) 826-4330

Degree: Ph.D., Electrical Engineering, 1975
Specialty: Multivariable Control Systems
Assigned: AD

Dr. Herman Senter
Associate Professor
Clemson University
Mathematical Sciences Dept.
Clemson, SC 29631
(803) 656-3434

Degree: Ph.D., Mathematics, 1973
Specialty: Mathematics, Statistics
Assigned: HRL/B

List of Participants (continued: page 13)

Dr. M. Paul Serve
Professor
Wright State University
Chemistry Department
Dayton, OH 45435
(513) 873-2835

Degree: Ph.D., Organic Chemistry, 1965
Specialty: Organic Chemistry Synthesis Toxicology, Drug Chemistry
Assigned: AMRL

Dr. Robert Sigman
Senior Research Engineer
Georgia Institute of Technology
Aerospace Engineering Department
Atlanta, Georgia 30332
(404) 894-3041

Degree: Ph.D., Aerospace Engineering, 1970
Specialty: Computational Fluid Mechanics, Aeroacoustics, Combustion
Assigned: AD

Dr. Bruce Simon
Associate Professor
University of Arizona
Aero & Mech. Engr. Department
Tucson, AZ 85712
(602) 626-3752

Degree: Ph.D., Mechanical Engineering, 1972
Specialty: Theoretical Mechanics
Assigned: AMRL

Dr. William Squires
Assistant Professor
Texas Lutheran College
Biology Department
Sequin, TX 78155
(512) 379-4161

Degree: Ph.D., Exercise Physiology, 1979
Specialty: Biology
Assigned: SAM

Dr. James Steelman
Associate Professor
New Mexico State University
Electrical Engineering Dept.
Las Cruces, NM 88003
(505) 646-4111

Degree: Ph.D., Electrical Engineering, 1968
Specialty: Tracking Filters, Pointing Systems
Assigned: WL

Dr. James Strickland
Professor
Texas Tech University
Mechanical Engineering Dept.
Lubbock, TX 79409
(806) 742-3563

Degree: Ph.D., Mechanical Engineering, 1973
Specialty: Thermal Fluid Sciences
Assigned: FJSRL

Dr. Timothy Su
Associate Professor
Southeastern Mass University
Chemistry Department
North Dartmouth, MA 02747
(617) 999-8235

Degree: Ph.D., Physical Chemistry, 1971
Specialty: Ion-Molecule Reactions, Gas Kinetics, Atmospheric Chemistry
Assigned: GL

List of Participants (continued: page 14)

Dr. Patrick Sweeney
Associate Professor
University of Dayton
Mechanical Engineering Dept.
Fairborn, OH 45324
(513) 229-2238

Degree: Ph.D., Mechanical Engineering, 1977
Specialty: OPS Research, Modeling
Assigned: ML

Dr. Paul Szydlik
Professor
State University of New York
Physics Department
Plattsburgh, NY 12901
(518) 564-2048

Degree: Ph.D., Physics, 1964
Specialty: Theoretical and Computational Physics; Solar Energy
Assigned: APL

Dr. Enoch Temple
Associate Professor
Alabama A&M University
Math Department
Huntsville, AL
(205) 859-7239

Degree: Ph.D., Statistics, 1980
Specialty: Applied Statistics
Assigned: LMDC

Dr. William Terry
Associate Professor
University of Toledo
Industrial Engineering Department
Toledo, OH 43551
(419) 537-2412

Degree: Ph.D., Industrial Engineering, 1977
Specialty: Time Series, Statistics, Stochastic Processes, Computer Integrated Manufacturing
Assigned: BRMC

Dr. John Tomchick
Assistant Professor
Pennsylvania State University
Eng. Sci. & Mechs. Department
State College, PA 16801
(814) 237-1683

Degree: Ph.D., Physics, 1974
Specialty: Solid State-Semiconductors: Transport, Defects, Electron Phonon Int.
Assigned: AL

Dr. Arthur Thorbjornsen
Associate Professor
University of Toledo
Electrical Engineering Dept.
Toledo, OH 43606

Degree: Ph.D., Electrical Engineering, 1972
Specialty: Computer-Aided Circuit Design, Semiconductor device Modeling
Assigned: AL

Dr. Jon Tolle
Professor
University of North Carolina
Mathematics & Operations Research
Chapel Hill, NC 27514
(919) 962-8401

Degree: Ph.D., Mathematics, 1966
Specialty: Optimization, Numerical Analysis
Assigned: AD

List of Participants (continued: page 15)

Dr. George Trevino
Associate Professor
Del Mar College
Physics Department
Corpus Christie, TX
(512) 881-6213

Degree: Ph.D., Applied Mathematics, 1969
Specialty: Turbulence and Stochastic Processes
Assigned: WL

Dr. Keith Walker
Professor, Head
Bethany Nazarene College
Physics Department
Bethany, OK 73008
(405) 789-6400

Degree: Ph.D., Physics, 1971
Specialty: Atomic & Molecular Physics
Assigned: APL

Dr. Shih-sung Wen
Professor
Jackson State University
Psychology Department
Clinton, MS 39056
(601) 968-2371

Degree: Ph.D., Educational Psychology, 1971
Specialty: Cognitive Psychology, Learning, Psychological Measurement
Assigned: HRL/W

Dr. Kenneth Williamson
Associate Professor
Oregon State University
Civil Engineering Department
Corvallis, OR
(503) 754-2751

Degree: Ph.D., Environmental Engineering, 1973
Specialty: Environmental Engineering
Assigned: AMRL

PARTICIPANT LABORATORY ASSIGNMENT (Page 1)

1983 USAF/SCEEE SUMMER FACULTY RESEARCH PROGRAM

AERO PROPULSION LABORATORY
(Wright-Patterson Air Force Base)
1. Dr. Richard Conte - Manhattan College
2. Dr. Gregory Doyle - University of Dayton
3. Dr. Jonathan Dutton - Texas A&M University
4. Dr. Amir Faghri - Wright State University
5. Dr. Medhat Ibrahim - California State Univeristy/Fresno
6. Dr. Paul Szydlik - State University of New York/Plattsburgh
7. Dr. Keith Walker - Bethany Nazarene College

AEROSPACE MEDICAL RESEARCH LABORATORY
(Wright-Patterson Air Force Base)
1. Dr. Peter Crane - University of Pittsburgh/Johnstown
2. Dr. Hans Fellner - Slippery Rock State College
3. Dr. Richard Gill - Wright State University
4. Dr. Gwendolyn Howze - Texas Southern University
5. Dr. Gerald Poje - Miami University
6. Dr. M. Paul Serve - Wright State University
7. Dr. Bruce Simon - University of Arizona
8. Dr. Kenneth Williamson - Oregon State University

ARMAMENT DIVISION
(Eglin Air Force Base)
1. Dr. Daniel Barr - Virginia Military Institute
2. Dr. Carolyn Crouch - University of Alabama
3. Dr. Donald Crouch - University of Alabama
4. Dr. Charles Mastin - Mississippi State University
5. Dr. Dallas Russell - Auburn University
6. Dr. Robert Sigman - Georgia Institute of Technology
7. Dr. Jon Tolle - University of North Carolina

ARNOLD ENGINEERING DEVELOPMENT CENTER
(Arnold Air Force Station)
1. Dr. Joseph Baumgarten - Iowa State University
2. Dr. Frank Collins - University of Tennessee Space Inst.
3. Dr. James Kirkpatrick - Alabama A&M University

AVIONICS LABORATORY
(Wright-Patterson Air Force Base)
1. Dr. Patrick Garrett - University of Cincinnati
2. Dr. Stella Lawrence - Bronx Community College
3. Dr. Sigurd Lillevik - Oregon State University
4. Dr. John Thomchick - Pennsylvania State University
5. Dr. Arthur Thorbjornsen - The University of Toledo

BUSINESS RESEARCH MANAGEMENT CENTER
(Wright-Patterson Air Force Base)
1. Dr. David Lee - University of Dayton
2. Dr. William Terry - University of Toledo

EASTERN SPACE & MISSILE CENTER
(Patrick Air Force Base)

ELECTRONICS SYSTEMS DIVISION
(Hanscom Air Force Base)
1. Dr. Craig Holt - Tufts University

ENGINEERING & SERVICES CENTER
(Tyndall Air Force Base)
1. Dr. Chester Canada - Oklahoma State University
2. Dr. Stephen Lin - North Carolina Central University
3. Dr. Leonard Lion - Cornell University
4. Dr. Daryl Logan - Rose-Hulman Institute of Technology
5. Dr. Donald Michelson - Virginia Tech.

FLIGHT DYNAMICS LABORATORY
(Wright-Patterson Air Force Base)
1. Dr. Ernesto Barreto - State University of New York/Albany
2. Dr. Terry Herdman - Virginia Polytechnic Inst. & State Univ.
3. Dr. George Kirby - West Virginia University
4. Dr. Madakasira Krishna - South Carolina State College
5. Dr. Tapas Mazumdar - Wright State University
6. Dr. Don Mittleman - Oberlin College
7. Dr. Dale Moses - San Diego State University
8. Dr. L. Rai Pujara - Wilberforce University

FRANK J. SEILER RESEARCH LABORATORY
(USAF Academy)
1. Dr. Michael McKee - Auburn University
2. Dr. Albert Payton - Hampton Institute
3. Dr. James Strickland - Texas Tech. University

GEOPHYSICS LABORATORY
(Hanscom Air Force Base)
1. Dr. Francesco Bacchialoni - University of Lowell
2. Dr. Pradip Bakshi - Boston College
3. Dr. Stanley Bashkin - University of Arizona
4. Dr. Carol Deakyne - College of the Holy Cross
5. Dr. Frederick Nagle - University of Miami
6. Dr. Timothy Su - Southeastern Massachusetts University

HUMAN RESOURCES LABORATORY/ADVANCED SYSTEMS DIVISION
(Wright-Patterson Air Force Base)
1. Dr. Robert Foley - Virginia Polytechnic Inst. & State Univ.

HUMAN RESOURCES LABORATORY/FLYING TRAINING DIVISION
(Williams Air Force Base)
1. Dr. John Cicero - Illinois Institute of Technology
2. Dr. Shih-sung Wen - Jackson State University

PARTICIPANT LABORATORY ASSIGNMENTS (Continued: page 3)

HUMAN RESOURCES LABORATORY/PERSONAL RESEARCH DIVISION
(Brooks Air Force Base)
1. Dr. Terry Dickinson - Old Dominion University
2. Dr. Herman Senter - Clemson University

HUMAN RESOURCES LABORATORY/TECHNICAL TRAINING DIVISION
(Lowry Air Force Base)
1. Dr. Arthur Gutman - Florida Institute of Technology
2. Dr. Gregory Jones - Utah State University

LEADERSHIP & MANAGEMENT DEVELOPMENT CENTER
(Maxwell Air Force Base)
1. Dr. Samuel Green - Auburn University
2. Dr. Enoch Temple - Alabama A&M University

LOGISTICS COMMAND
(Wright-Patterson Air Force Base)
1. Dr. E. Stanley Lee - Kansas State University

LOGISTICS MANAGEMENT CENTER
(Gunter Air Force Base)
1. Dr. Richard Andrews - The University of Michigan

MATERIALS LABORATORY
(Wright-Patterson Air Force Base)
1. Dr. Billy Covington - Sam Houston State University
2. Dr. Stephen Krause - Arizona State University
3. Dr. William Kyros - University of Lowell
4. Dr. George Miner - University of Dayton
5. Dr. Charles Moseley - Ohio State University/Lima
6. Dr. Patrick Sweeney - University of Dayton

ROCKET PROPULSION LABORATORY
(Edwards Air Force Base)
1. Dr. M. Quinn Brewster - University of Utah
2. Dr. Kathleen Howell - Purdue University
3. David Questad - Pennsylvania State University

ROME AIR DEVELOPMENT CENTER
(Griffiss Air Force Base)
1. Dr. Richard Anderson - University of Rolla
2. Dr. Victor Frost - University of Kansas
3. Dr. Stuart Hirshfield - Hamilton College
4. Dr. David Lai - University of Vermont
5. Dr. Randolph Moses - Virginia Polytechnic Inst. & State Univ.
6. Dr. Philip Olivier - Louisiana State University

ROME AIR DEVELOPMENT CENTER/ELECTRONICS TECHNOLOGY
(Hanscom Air Force Base)
1. Dr. Henry Bertoni - Polytechnic Institue of New York
2. Dr. Charles Ih - University of Delaware

SCHOOL OF AEROSPACE MEDICINE
(Brooks Air Force Base)

1. Dr. Willie Bragg - University of Cincinnati
2. Dr. Louis Buckalew - Alabama A&M University
3. Dr. James Dooley - Wright State University
4. Dr. John Giolma - Trinity University
5. Dr. Amir Karimi - University of Texas/San Antonio
6. Dr. Jerome Keating - University of Texas/San Antonio
7. Dr. Mark Lewitter - University of Texas/San Antonio
8. Dr. James Mrotek - Meharry Medical College
9. Dr. William Squires - Texas Lutheran College

WEAPONS LABORATORY
(Kirtland Air Force Base)

1. Dr. Michael Becker - University of Texas/Austin
2. Dr. Jack Chatelain - Utah State University
3. Dr. Fred Domann - University of Wisconsin
4. Dr. John Eoll - Lenoir-Rhyne College
5. Dr. Eddie Fowler - Kansas State University
6. Dr. Robert McLauchlan - Texas Tech. University
7. Dr. David Pegg - University of Tennessee
8. Dr. James Steelman - New Mexico State University
9. Dr. George Trevino - Del Mar College

RESEARCH REPORTS
1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Volume I Report Number	Title	Research Associate
1	The Feasibility of Recording Atmospheric Interferograms	Dr. Richard Anderson
2	Testing the Representativeness of the Supply Data Bank	Dr. Richard W. Andrews
3	Intelligent Controller for Space Experiments	Dr. Francesco L. Bacchialoni
4	Infrared Earthlimb Emission Lineshapes as Signatures of Atmospheric Parameters	Dr. Pradip M. Bakshi
5	An Evaluation of a Cantilever Beam Solid State Accelerometer	Dr. Daniel W. Barr
6	The Gas Heating Phase in Electrical Breakdown	Dr. Ernesto Barreto
7	Laser-Induced Helium Plasma	Dr. Stanley Bashkin
8	Investigation of Liquid Sloshing in Spin-Stabilized Satellites	Dr. Joseph R. Baumgarten
9	Searching for Precursors to Laser-Induced Damage	Dr. Michael F. Becker
10	Propagation Loss in Electrostatically Variable Saw Delay Lines	Dr. Henry L. Bertoni
11	Dilemmas of Combat Psychiatry: World War II and Vietnam	Dr. Willie A. Bragg
12	Effects of Radiative Heat Feedback on Solid Rocket Propellant Combustion	Dr. M. Quinn Brewster
13	Aircrew-Relevant Man-Monkey Analogs for Evaluation of CD Agents: Pitch and Alcohol	Dr. L. W. Buckalew

Report Number	Title	Research Associate
14	Combined Blast and Fragment Loading on Reinforced Concrete	Dr. Chester E. Canada
15	Finite Element Preliminaries in EMP Environments	Dr. Jack E. Chatelain
16	Multiple Cockpit Combat Mission Trainer Network	Dr. John A. Cicero
17	The Use of a Unique Heat Tranfer Probe to Measure Spacecraft Rocket Plume Contamination	Dr. Frank G. Collins
18	A Computer Program for the Automatic Generation of a Two-Dimensional Finite Difference Mesh to Investigate the Heat Transfer Characteristics on Arbitrary High-Temperature Turbine Blades	Dr. Richard V. Conte
19	Raman and Infrared Spectroscopy of Extrinsic P-Type Silicon	Dr. B.C. Covington
20	Human Factors Comparison of Touch Screen and Voice Command Data Entry on a Command, Control, and Communications System	Dr. Peter M. Crane
21	Performance Analysis and Evaulation in a Local Area Network	Dr. Carolyn J. Crouch
22	The Impact of ADA on USAF Computational Support Services	Dr. Donald B. Crouch
23	A Molecular Orbital Study of OH^-. H_2O, $H^+(CH_3CN)_k(H_2O)_m$, And $H+(HCN)_n$ Cluster Ions	Dr. Carol A. Deakyne
24	Models for Evaluating the Validity and Accuracy of Performance Rating	Dr. Terry L. Dickinson
25	Searching for Precursors to Laser-Induced Damage	Dr. Fred E. Domann
26	Hyperbaric Oxygenation (HBO) Alteration of Metabolism and Cardiovascular Function During and Following Exercise Conditioning	Dr. James W. Dooley

Report Number	Title	Research Associate
27	Natural Frequencies and Mode Shapes of Uniform Beams	Dr. George R. Doyle, Jr.
28	Time-Dependent Calculations of Swirling Nozzle Flow	Dr. J. Craig Dutton
29	An Evaluation of Two Nuclear Weapons Effects Computer Programs	Dr. John G. Eoll
30	Axial Variation of Local Heat Flux Along the Condenser Section of a Double-Wall Artery High Capacity Heat Pipe	Dr. Amir Faghri
31	Windscreen Haze Characteristics Studies	Dr. Hans G. Fellner
32	Mirem and Mission Phasing	Dr. Robert D. Foley
33	Hemp Vulnerability/Survivability of Computer Networks	Dr. Eddie R. Fowler
34	An Approach to the Design of an Adaptive Spread Spectrum Modem	Dr. Victor Frost
35	Quantization Error Analysis for the DeAnza Image Processor	Dr. Patrick Garrett
36	Pilot Workload and G-Stress	Dr. Richard T. Gill
37	Identification of Rapid Eye Movement by Computer During Discrete Tracking Tasks	Dr. John P. Giolma
38	An Evaluation of the Measurement System Used by the Leadership and Management Development Center for the Assessment of Its Consulting Efforts	Dr. Samuel B. Green
39	Reinforcement Induced Stereotypy of Sequential Behavior	Dr. Arthur Gutman
40	A Two-Dimensional Aeroelastic System	Dr. Terry L. Herdman
41	An Integrated Approach to Interface Design	Dr. Stuart H. Hirshfield

Report Number	Title	Research Associate
42	Message Routing Methods for a Tactical Air Control System Communication Network	Dr. Craig S. Holt
43	Attitude Control Issues for Large Flexible Space Systems	Dr. Kathleen Howell
44	A Scanning Electron Microsopical Study of Periosteum from Rat and Monkey	Dr. Gwendolyn B. Howze
45	Cycloconverter Modeling for Variable Speed Drives	Dr. Medhat A.H. Ibrahim
46	A Novel Modulation Technique for FDM for Optical Fiber Communication	Dr. Charles S. Ih
47	Software Fault-Tolerance/ Diagnostics for Single-User Systems	Dr. Gregory W. Jones
48	A Thermal Evaluation of the "LSSI" Liquid-Cooled System: An Engineering Perspective	Dr. Amir Karimi
49	A Statistical Method for the Serial Comparison of Vector Cardiograms	Dr. Jerome P. Keating
50	Short Crack Behavior for Flaws Emanating from Fastener Holes	Dr. George C. Kirby
51	An Evaluation of the Mathematical Process and Formulation for Case Mounted Displacement Sensors	Dr. James Kirkpatrick
52	Tem Morphology Study of Molecular Composites of Polymers	Dr. Stephen J. Krause
53	Analytical Representation of Afterbody Surface of X-24C-10D Reentry Vehicle	Dr. Madakasira Krishna
54	Mechanical Characterization of Carbon-Carbon Composites for Turbine Engine Design: A State-of-the-Art Review	Dr. William Kyros

Report Number	Title	Research Associate
Volume II		
55	Performance of Image Restoration Filters in Machine Recognition	Dr. David C. Lai
56	Evaluation of Proposed Integrated Communication Navigation Identification Avionics (ICNIA) Architectures for Their Fault Tolerance Characteristics and Potentials	Dr. Stella Lawrence
57	Decision Aids for Selecting Air Force Manufacturing Technology Projects	Dr. David R. Lee
58	An Assessment of Wartime Availability of Recoverable Items	Dr. E. Stanley Lee
59	Laser Densitometer Design	Dr. Mark E. Lewittes
60	Real-Time Data Quality Assessment of Distributed Data Acquisition Systems	Dr. Sigurd L. Lillevik
61	Photochemical Reactions in a Small Indoor Smog Chamber	Dr. Stephen F. Lin
62	Partitioning Equilibria of Volatile Pollutants in Three Phases Systems	Dr. Leonard W. Lion
63	Evaluation of Projectile Impact on Earth Covered Structures	Dr. Daryl L. Logan
64	Numerical Solution of the Euler Equations on Dynamic Grids	Dr. C. Wayne Mastin
65	Approximate Evaluation of Optimal Control Minimizing Noncoercive Cost-Functionals Over Unbounded Sets; Hyperbolic Systems	Dr. Tapas Mazumdar
66	A Study of the CH_2NO_2 Radical Using a Multiconfigurational Approach	Dr. Michael L. McKee
67	Investigation of Vibration Problems with Heterodyne Holographic Interferometer	Dr. Robert McLauchlan

Report Number	Title	Research Associate
68	Use of Colloidal Gas Aphrons (CGA's) for Treating Hazardous Wastes	Dr. Donald L. Michelsen
69	Transport and Electron Paramagnetic Resonance Studies of Infrared Detector Materials	Dr. George K. Miner
70	Synergetic Maneuvers	Dr. Don Mittleman
71	Approaches to Synthesis of Some Novel Polybenzimidazole Monomers	Dr. Charles G. Moseley
72	Boundary Corrections for Low Speed, Solid Wall Wind Tunnels	Dr. Dale F. Moses
73	Combined Time Space Filtering for HF Antenna Array Systems	Dr. Randolph L. Moses
74	Raman Spectroscopy of Inhibited and Stimulated, Normal and Neoplastic Cultured Human and Mammalian Cells	Dr. James J. Mrotek
75	Combined Magnetic and Gravity Anomalies: A Guide to Crustal Type and Tectonics in the Southeastern Indian Ocean and Caribbean Regions	Dr. Frederick Nagle
76	Using Artificial Intelligence in Avionic Fault Isolation	Dr. Philip D. Olivier
77	Organic Reactions in Room Temperature Chloroaluminate Molten Salts	Dr. Albert L. Payton
78	A Method of Sensing Small Changes in the Angular Separation of Crossed Laser and Neutral Particle Beams	Dr. David J. Pegg
79	Evaluation of Naphthalene Toxicity Using Short Term Exposure and the Amphipod, Gammarus Tigrinus	Dr. Gerald V. Poje
80	Simplification of Nonlinear Systems	Dr. L. Rai Pujara

Report Number	Title	Research Associate
81	The Effect of Large Deformation on the Fracture Mechanics of Solid Propellants	Dr. David L. Questad
82	The Application of an Extended Kalman Filter to the Design of a Bank-to-Turn Missile Autopilot	Dr. Dallas W. Russell
83	Analysis of Pilot Selection Data	Dr. Herman F. Senter
84	Effect of Jet Fuel JP-4 Fractions on Fischer 344 Male Rats	Dr. M. Paul Serve
85	Secondary Muzzle Flash in Rapid Fire Cannons	Dr. Robert K. Sigman
86	Poroelastic Models of the Intervertebral Disk	Dr. Bruce R. Simon
87	Effects of Fluid Shifts and Hypovolemia in Individuals with Different Working Capacities While Resting at a Five Degree Declination	Dr. William G. Squires
88	Pinhole Beam Sensors II	Dr. James Eldon Steelman
89	Dynamic Stall: A Study of the Constant Pitching Rate Case	Dr. James H. Strickland
90	Experimental and Theoretical Investigations of Negative Ion-Polar Molecule Reactions	Dr. Timothy C.K. Su
91	A Dynamic Mini-Model for Space Technology Resource Allocation	Dr. Patrick J. Sweeney
92	Early Performance of the Gallium Arsenide Photovoltaic Array on the Living Plume Shield (Lips) Satellite	Dr. Paul P. Szydlik

Report Number	Title	Research Associate
93	Building a Multiple Regression Equation When Many Variables Are Available	Dr. Enoch C. Temple
94	A Preliminary Investigation of of the Utility of Linear Digital Filters for Analyzing Economic System Performance Data	Dr. William R. Terry
95	Shallow Donor Impurity Binding Energies in Asymmetric Quantum Wells	Dr. John Thomchick
96	Statistical Simulation and Correlation Studies of GAAS Mesfets	Dr. Arthur R. Thorbjornsen
97	Delivery Accuracy	Dr. Jon W. Tolle
98	An Introductory Dynamical Theory for Fully Compressible Trubulence	Dr. George Trevino
99	Electronic Excitation of Atomic Xenon By Electron Impact	Dr. Keith G. Walker
100	Effects of Graphic Information on Reading Comprehension: Eye Movements in Reading Text with Graph	Dr. Shih-sung Wen
101	Intragastric Administration of Dibromomethane to Rats	Dr. Kenneth J. Williamson

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

PERFORMANCES OF IMAGE RESTORATION FILTERS IN MACHINE RECOGNITION

Prepared by:	Dr. David C. Lai
Academic Rank:	Professor
Department and University:	Department of Electrical Engineering and Computer Science, The University of Vermont
Research Location:	Rome Air Development Center, Intelligence & Reconnaissance Division Image Systems Branch, Image Exploitation Section
USAF Research	John Potenza
Date:	29 July 1983
Contract No:	F49620-82-C-0035

PERFORMANCE OF IMAGE RESTORATION FILTERS IN MACHINE RECOGNITION

by

David C. Lai

ABSTRACT

Various image restoration filters are evaluated and compared based on their performances in machine classification under a variety of blur and noise conditions. These filters are analyzed to provide in-sight that may lead to ideas for designing new filters and the development of guidelines for applications of these filters in practice. Experimentation is designed to evaluate and compare performance of selected filters under various conditions. Performance measures appropriate to our application are derived. Results are reported and suggestions for further research are offered.

ACKNOWLEDGEMENT

The author wishes to thank the Air Force Systems Command, the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for giving him the opportunity to spend an interesting and exciting summer at the Rome Air Development Center, Griffiss AFB, N.Y. He wishes also to acknowledge the hospitality and good working conditions provided by the Image Exploitation Section at RADC.

Of course, this work would not have been possible without the support of many individuals at RADC. The author wishes to thank particularly Mr. John Potenza for his collaboration and many fruitful discussions, and Mr. Frederick W. Rahrig for his helping hand on the computer facility and valuable suggestions.

Finally, he would like to thank Mr. Kevin Verfaille of the University of Vermont for carrying out the experimental work.

I. INTRODUCTION:

The image restoration problem is mathematically ill conditioned, hence there is no unique solution. A variety of restoration filters have been proposed [1, 2, 3, 4, 5, 6, 7, 8], notably the inverse, weiner, parametric wiener, geometrical mean, power spectrum equalization (PSE), pseudo-inverse, homomorphic, Kalman, MAP, ML, MEM, median etc filters. These filters were designed according to different sets of specific mathematical criteria based on various assumptions of the image models. In practice filter performance varies with the type of image, the blur and the noise conditions. It is nacural to ask the following question: How well do the various filters enhance human interpretation of the image or facilitate machine recognition? Or, which filter is better than the other and under what conditions? Systematic evaluation and assessment of relative performance between filters is badly needed, since it will facilitate judicious selection of filters for use in Air Force automatic target classification and identification systems using high resolution reconnaissance imagery. Furthermore the resultant information will provide insight and a sound basis in designing new filters for similar Air Force image exploitation missions.

This project concentrates on the investigation and evaluation of performance of six image resotration filters; viz, inverse, wiener, parametric wiener ($\gamma = \frac{1}{4}$), geometrical mean ($\alpha = \frac{1}{2}, \gamma = \frac{1}{2}$), PSE, and median filters. The basic model for a permanently recorded or observed image g (x,y) is usually given as:

$$g(x,y) = s\{H[f(x,y)]\} + \sigma \times n(x,y) \qquad (1)$$

where s is a function representing the image sensor response; H, an operator representing the image formation process; f(x,y), the original undistorted image or original object images; n(x,y), a random noise (formulated as either an additive or multiplicative) process due to the record medium and/or electronic circuit used in image recording. The goal of image restoration is to retrieve

f(x,y) from g(x,y). To be more precise, the restoration problem may be stated as: given the recorded and distorted image g(x,y), knowledge about the type of the noise n(x,y), possibly some knowledge of the original image f(x,y), and some knowledge about the image formation in terms of the point-spread-function (PSF), estimate the original undistorted image or original object intensity distribution f(x,y). The model given in Eq. (1) is too general to be useful in designing restoration filters. In practice, assumptions and approximations are made to render the design and implementation of restoration filters to be reasonably manageable. In many instances, the non-linear function or operator is approximated by a linear one; the image formation is approximated by convolution with a space-invariant point-spread function (SIPSF); and the noise is assumed to be additive white Gaussian. It is apparent that the ill conditioned nature of the restoration problem together with the variety of image model, assumptions and approximations, and mathematical criteria for optimazation results in the many different restoration filters available. Moreover, it would be impossible to test and evaluate all of the filters. With the desire to encompass a wide range of restoration filters and to keep computation within the support of "IRAMS" (Image Restoration and Manipulation Software) implemented by PAR Corp for RADC, the above six filters were selected for evaluation.

Since the Air Force is emphasizing machine classification and identification and the restoration filters are intended to be used as preprocessor for a pattern recognition system, it seemed more appropriate to evaluate the restoration filters based on how well they perform for machine classifiers rather than for photointerpreters. Most of image restoration work was aimed at producing "nice looking" pictures to please human observers.[9] However, lack of knowledge about the psychophysical processes of human vision and a universal criterion for "beauty" hindered efforts in a systematic evaluation and

comparison of restoration filters. Cannon, et al [10] did work on the evaluation and comparison of restoration filters based on a mathematical criterion and the judgment of a panel of photointerpreters. They compared three filters; viz, Wiener, PSE, and MAP. They concluded that based on the subjective quality of photointerpreters, all filters perform equally well on Gaussian blurred images regardless of signal-to-noise ratio (SNR) level, blur severity, or image type; however, MAP filter seems to work best on defocused image in high signal-to-noise environment. Based on the particular mathematical criterion (minimum mean square error), the Wiener filter, as expected performed best. Another closely related work[11] aimed at evaluating the effect of degradation of images on photo interpretability and subjective quality without any restoration performed on the images. The criteria which was used in measuring the degradation effect on human observers are:

(1) Performance by trained photointerpretaters (PI's) in the extraction of a set of essential elements of information from degraded photos, and

(2) Subjective quality as scored by the PI's based on the 10-point NATO standardized scale.

Their conclusion was, indeed, that the interpretabiity has been reduced in proportion to the degree of degradation by both blur and noise; however, the noise had less effect in reducing interpretability. Somehow, the PI's can see through noise but not blur. The subjective quality as measured by the NATO Scale was worsened in either case when noise was increased but blur was decreased or when noise was decreased but blur was incresed. For individual PI's, the correlation between subjective quality and interpretability was low; however, for PI's as a group, the mean correlation was high. To the best knowledge of the author, there has been no work reported on for evaluating image restoration filters for machine recognition. However, the effects of median filtering on machine analysis of images, for edge detection , shape analysis,

and texture analysis were reported.[12]

II. OBJECTIVES:

The main objective of this project is to systematically evaluate and compare the performances of different image restoration filters in machine classification so as to ascertain their usefulness in facilitating machine recognition. More specifically, the following items were planned:

(1) To analyze image restoration filters to gain insights that may lead to the development of guidelines for practical applications and ideas for designing new filters.

(2) To test and evaluate six restoration filters: inverse, Wiener, parametric wiener, geometrical mean, PSE, and median filter on images of various degrees of degradation based on their performances in machine classification of the scene in the image.

(3) To derive an appropriate performance measure in order to quantify performances of the filters.

(4) To compare these filters under different distortion conditions. This project will not be concerned with the interrelationship between restoration filtering, feature extraction and classifying. In other words, the mutual effets between filters, feature extractors, and classifiers will not be considered. Moreover, the many image enhancement techniques which , undoubtedly, facilitate machine recognition will also not be considered here.

III. ANALYSIS OF IMAGE RESTORATION FILTERS

The image model given in Eq. (1) provides an accurate characterization of image formation but it is too general to be useful for filter design. For most linear filters proposed, the sensor function s is approximated by a linear function; the image or blur operator H is assumed to be linear and represented by spatial-invariant point-spread function; and the noise is assumed to be

additive. Then the observed image becomes:

$$g(x,y) = h(x,y) \circledast f(x,y) + n(x,y) \qquad (2)$$

where $\circledast$ signifies two-dimensional convolution and h(x,y), the PSF. Six filters were chosen for testing and evaluation as mentioned prevously. The first five; viz., inverse, Wiener, parametric Wiener, geometrical mean, and PSE filters, represent essentially the entire range of linear filters. They were derived based on the linear model given in Eq. (2). Wiener filtering is perhaps the most common restoration method used. Parametric Wiener filtering seems to give visibly good results. [13] Geometrical mean filtering allows the inverse filter to boost the high-frequency components in a controlled fashion.[14] PSE filtering seems to give a pleasing clear picture.[15] The sixth, median filter, is a nonlinear filter which is useful for noise reduction and edge preservation.[8] Thus, the filters chosen for evaluation represent a wide range of available filters. Derivation of these filters can be found in References (1) and (2) and will not be repeated here. A brief analysis of these filters are given below:

(1) Inverse filter - Based on the model given in Eq. (2) and the criterion of minimizing the norm (i.e., power) of the noise, the inverse filter expressed as transfer function is derived as:

$$H_I(\omega_x, \omega_y) = \frac{1}{H(\omega_x, \omega_y)} \qquad (3)$$

where $H(\omega_x, \omega_y)$ is the optical transfer function (OTF); i.e. the 2-D Fourier transform of the PSF h(x,y) in Eq. (2). The restored image is then

$$\hat{f}(x,y) = \mathcal{F}^{-1}\left[\frac{G(\omega_x, \omega_y)}{H(\omega_x, \omega_y)}\right] \qquad (4)$$

Using the operator matrix and vector notation, it is seen that

$$\hat{f} = \underline{H}^{-1}\underline{H}f + \underline{H}^{-1}n = f + \underline{H}^{-1}n \qquad (5)$$

which gives the original image f plus the transformed noise. In the absence of noise and singularities of the filter, one could restore the distorted image to

the original form perfectly. It is seen that the inverse filter concentrates on regaining the original image but disregards (even amplifies) the noise. It is reasonable to expect that the inverse filter will perform well if the noise level is low and there are no singular values of $H(\omega_x, \omega_y)$ for any ω_x and ω_y. In general, OTF's exhibit a low frequency "hump", and thus $H^{-1}(\omega_x, \omega_y)$ tends to boost the high-frequency portion of an image.

(2) Wiener filter - Using the linear model in Eq. (2) and minimizing the mean squared error between the original and the restored image one obtains the Wiener filter.

$$H_w(\omega_x, \omega_y) = \frac{H^*(\omega_x, \omega_y)}{|H(\omega_x, \omega_y)|^2 + [S_n(\omega_x, \omega_y)/S_f(\omega_x, \omega_y)]} \qquad (6)$$

where $H(\omega_x, \omega_y)$ is again the OTF; $S_n(\omega_x, \omega_y)$, the noise power spectrum density (PSD); and $S_f(\omega_x, \omega_y)$, the PSD of original (or desired) image. In practice, d.c. components in S_n and S_f are subtracted, if there is any. It is obvious that Wiener filtering reduces to inverse filtering in the absence of noise. For this reason, the noise-to-signal ratio term $[S_n/S_f]$ in the expression of Wiener filter may be viewed as a modification function that smoothes the inverse filter and even eases the problem created by the singularities of OTF to provide optimum restoration in the presence of noise. The restored image in the frequency domain is:

$$\hat{F}(\omega_x, \omega_y) = \frac{H^*(\omega_x, \omega_y) F(\omega_x, \omega_y)}{|H(\omega_x, \omega_y)|^2 + [S_n(\omega_x, \omega_y)/S_f(\omega_x, \omega_y)]} + \frac{H^*(\omega_x, \omega_y) N(\omega_x, \omega_y)}{|H(\omega_x, \omega_y)|^2 + [S_n(\omega_x, \omega_y)/S_f(\omega_x, \omega_y)]} \qquad (7)$$

where $F(\omega_x, \omega_y)$ and $N(\omega_x, \omega_y)$ are the Fourier transforms of sample functions of the original image process and the noise process, respectively. This shows that the restored image consists of somewhat distorted original and reduced noise. In other words, Wiener filter does not concentrate all its effort in regaining (deblurring) the original image but spares some ammunition for combating noise as represented by the $[S_n/S_f]$ term in the denominator. Since $|H|^2$ and $[S_n/S_f]$ in the denominator are equally weighted, the effort of Wiener filter in deblurring the image and reducing the noise

is equally distributed. If it is desired that more emphasis should be put on deblurring and less emphasis on reducing noise, then the influence of the $[S_n/S_f]$ term in Eq. (7) should be reduced so that the Wiener filter behaves more like inverse filter. This leads to the idea of adjusting the effect of $[S_n/S_f]$ term by multiplying it by a factor γ in Eq. (6) which gives rise to the so-called parametric Wiener filter.

The presence of $H^*(\omega_x, \omega_y)$ in the numerator of Wiener filter tends to make the filtering biased towards the low frequency end. To combat this low-frequency bias the so called geometrical mean filter was developed and will be discussed later.

(3) Parametric Wiener Filters - Although parametric Wiener filter could be derived via the heruistic approach discussed previously, it does not provide a sense of optimallity nor insight for obtaining the best γ value to multiply to the $[S_n/S_f]$ term. Mathematically, using the model in Eq. (2) and minimizing the effective noise-to-signal ratio of the restored image subject to the constraint that the residual norm between the observed image and the reblurred restored image equal to the norm of the noise results in

$$H_{pw}(\omega_x, \omega_y) = \frac{H^*(\omega_x, \omega_y)}{|H(\omega_x, \omega_y)|^2 + \gamma [S_n(\omega_x, \omega_y)/S_f(\omega_x, \omega_y)]} \qquad (8)$$

where, in order to obtain optimum filtering in the constrained least-squares sense, the factor γ must be chosen such that the constraint is satisfied. An algorithm for determining this γ value exists.[13]

It is seen from Eq. (8) that when $\gamma = 1$, parametric Wiener filter becomes Wiener filter and when $\gamma = 0$, it becomes inverse filter. Choosing arbitrarily a value ($\gamma \geq 0$) will cause the Wiener filter either to behave more like inverse filter ($\gamma < 1$) or to combat noise more vigorously ($\gamma > 1$). The γ value cannot be arbitrarily assigned to obtain optimum filtering. It must be chosen to satisfy the above mentioned constraint.

(4) Geometrical mean filters - From the previous discussion, it is understood that the inverse filtering regains the original image completely in the absence of noise provided the filter is non-singular. However, in the presence of severe noise and/or filter singularities, the inverse filtering will perform badly; it not only will amplify noise but also accentuate both noise and image signal at the singularities. Wiener filter, on the other hand, will restore the image (not perfectly) as well as reduce noise and it can never be singular in the presence of noise; however, it does tend to favor the low frequency portions of the image and thus loses sharp edges. One thought is to combine inverse filter and parametric Wiener filters in such a way as to parameterize the ratio of their effects on restoration. By judicious choice of the parameters, it is hoped that the low frequency dominance of Wiener filter may be lessened while some singularities of the inverse filter may be avoided. Therefore, geometrical mean filters are defined as:

$$H_{GM}(\omega_x,\omega_y) = \left[H_I(\omega_x,\omega_y)\right]^{\alpha}\left[H_{pw}(\omega_x,\omega_y)\right]^{1-\alpha} = \left[\frac{H^*}{|H|^2}\right]^{\alpha}\left[\frac{H^*}{|H|^2+\gamma[S_n/S_f]}\right]^{1-\alpha} \tag{9}$$

with $0 \le \alpha \le 1$ and $0 \le \gamma < \infty$

For $\gamma = 1$, H_{GM} changes from a complete Wiener filter to inverse filter when α changes continuously from 0 to 1. At the geometric mean ($\alpha = \frac{1}{2}$) with $\gamma = 1$, and for symmetrical PSF; i.e. H (ω_x,ω_y) is a real function and $H_{GM}(\omega_x,\omega_y)$ in Eq. (9) becomes

$$\left[\frac{S_f(\omega_x,\omega_y)}{|H(\omega_x,\omega_y)|^2 S_f(\omega_x,\omega_y)+S_n(\omega_x,\omega_y)}\right]^{\frac{1}{2}} = \left[\frac{S_f(\omega_x,\omega_y)}{S_g(\omega_x,\omega_y)}\right]^{\frac{1}{2}} \tag{10}$$

which has the same form as the magnitude of PSE filter.

For the range of $\gamma < 1$ and any α or $\alpha > \frac{1}{2}$ and any γ, the inverse filtering effect dominates. On the other hand, for $\alpha < \frac{1}{2}$ and $\gamma > 1$, the Wiener filtering effect dominates the scene. Hunt[14] showed that, for moderate blur and low SNR,

the geometrical mean filter with $\alpha = \frac{1}{2}$ and $\gamma = \frac{1}{4}$ produced good result.

(5) PSE (Power Spectrum Equalization) filter - As stated earlier, the inverse filter restores perfectly in the absence of noise, provided there is no singularity in the blur function; however, in the presence of noise, it attempts also to regain the image information in the noise bands and thus accentuates the noise. One way to remedy this is to limit the gain of the restoration filter for the low SNR bands of the image spectrum. Based on the image model in Eq. (2) and with the constraint that the PSD of the restored image is equal to that of the original image, the PSE filter was derived as:

$$\left| H_{PSE}(\omega_x, \omega_y) \right| = \left[\frac{S_f(\omega_x, \omega_y)}{S_g(\omega_x, \omega_y)} \right]^{\frac{1}{2}} \qquad (11)$$

and the phase is set equal to the negative of the phase of the blur spectrum.[15] It is seen that the design of this filter does not require knowledge about the noise (spectrum).

The PSE filter can also be obtained by setting $\alpha = \frac{1}{2}$ and $\gamma = 1$ in the design of geometrical mean filters. Thus PSE filtering has the combined effect of inverse filtering and Wiener filtering in equal strength.

(6) Median filter - Median filters are non-linear filters. This family of filters is not derived on the basis of an image model and a set of mathematical criteria. Median filtering was first suggested by Tukey[16] in 1971 for smoothing time series and noticed its property in prserving large sudden changes of level (edges) in time series. It has later been adapted for use in image processing. Median filtering is performed by moving a window over the pixels of an image and replacing the pixel at the center of the window by a pixel whose value is the median of the original pixel values within the window. It was shown that median filtering preserves sharp edges and is very efficient in smoothing "salt-and-pepper" noise but not very effective in reducing Gaussian noise.[8]

The mathematical definition of median filtering is given as follows:

$$m(x_i, y_i) = \underset{W}{median}[g(x_i, y_i)] \triangleq median[g(x_{i+r}, y_{j+s}); (x_r, y_s) \in W]$$

where W is the filter window; (x_i, y_i) denotes a pixel of the image usually the center pixel of the window and (x_r, y_s), a pixel inside the window. r and s are varied to cover all the pixels in the window and varying i and j to move the window to cover another portion of the image. The filter window may take many forms such as line segments, rectangles, squares, circles, crosses, square frames, etc. For border points, the median is usually computed on those points covered by the window (without padding zeros). In general, the number of pixels inside the window is odd; otherwise, the mean of two middle points is taken as the median value. Because of the difficulty involved in the theoretical analysis of median filters, there has been practically no published results. However, efficient median filter algorithms have been developed.[17,18] Furthermore many attempts were made to implement median filters in hardware which will perform in real time.[19, 20]

A median filter was chosen for evaluation because of its simlicity in design and potential for VLSI implementation. It should be noted that this family of filters were designed mainly for suppresion of noise while preserving sharp edges with no attempt or regard to restore a blurred image to its original form. It must be used with caution.

IV. EXPERIMENT FOR FILTER EVALUATION AND COMPARISON:

Because of time limitation for this project, only one picture was chosen for the image restoration filter evaluation experiment. It is a section of an aerial photo of a test site. The image selected consists of a B-52 on a pedestal, shadows, and soil. It was chosen mainly for the reason that the scene can be easily classified by machine. This image is shown in its original undistorted form and blurred forms in Fig. 1. The image size is 256 x 256

pixels. In order to create a diverse group of distorted images to allow different filters to work under various conditions, the image was subjected to different degrees of degradation.

Due to, again limitations in time, only one kind of blur was chosen; viz; the blur with a Gaussian PSF. Three levels of Gaussian blurs with standard deviations σ_h = 3.5, 4.5 and 6.5 were used. The original and the three blurred images were then subjected to further degradation by additive zero-mean white Gaussian noise. The original is not completely noise free. It has a small amount of noise. Two levels of additive noise corresponding to SNR of 12dB and 5 dB were used. Thus, it produced 12 images as our data base for testing. They include the original, three images degraded by blur only (shown in Fig. 1) two images degraded by additional noise only, and the rest degraded by both blur and noise in combinations of degrees of severity. The degraded images were restored by six different filters which are described in the previous section. Thus 66 restored images were produced along with the 11 test images and the original. They were submitted for machine classification. The scene of each of those images was classified into three categories; viz., metal (airplane), shadow, and soil. The number of pixels classified into each class was then counted. The result of classification performed on the original image was taken as the reference (truth) based on the assumption that the machine can do no better on the degraded images than on the original. It was observed that noise alone had less effect on machine classification. For the ease to compare the effect due to blur, the 77 images (11 degraded but unrestored and 66 restored) were divided into three groups according to their noise levels. Each group contains 28 images except the one with the original SNR which has only 21 images. They were then evaluated and compared numerically based on the performance measures which will be described in next section.

For the parametric Wiener filter used here, γ was arbitrarily set equal to ¼. No attempt was made to determine the best γ value. Similarily, for the geometrical mean filter, α and γ value were arbitrarily chosen to be ½ and ¼, respectively. Therefore, the results reported here for the parametric Wiener and the geometrical mean filters are pertinent only to these special cases. For the median filter, a 3 x 3 filter window was used. No attempt was made to select the best window shape and size.

In the machine recognition scheme, two features (average intensity over a 3 x 3 window and minimum intensity over a 3 x 3 window) were measured and the condensed nearest neighbor classifier[21] was used. The features and the classifier were selected by cut-and-try method. No attempt was made to find the best features and the best classifier for use on each of these images. Similarly, a set of training regions was determined by cut-and-try on the original images and this set was used throughout the experiment so as to minimize the effect of training regions used. Undoubtedly, one could better the classification performance on each image by choosing the best features, classifier, and training regions for the particular image, however, that would defeat our purpose of evaluating and comparing the filters for facilitating machine recognition. For our purpose, the effect on classification due to factors such as training regions and features selected, classifier used, etc must be minimized. Therefore, in the strictest sense, our results pertain only to the specific set of experimental circumstances. However, it is felt that the results obtained may be generalized to other types of images, classifiers, etc. as far as that they are applicable.

There was also no attempt to select features which were improved or enhanced by particular restoration filters. In other words, the problem of feature selection in relation to the use of restoration filters as preprocessor to machine recognition was not addressed here. The understanding of the interrelation between the preprocessor (filters and enhancements) and the

feature extractor is important for any practical application of machine recognition.

Two software packages were employed to carry out the experiment. They are "IRAMS" (Image Restoration And Manipulation Software) and "AFES" (Automatic Feature Extraction System") on DEC 11/70 and 11/34; both were installed by PAR Technology Corp for RADC.

V. PERFORMANCE MEASURE

In the experiment, after performing classification on the images, pixels classified into each class were counted. Comparing directly these pixel counts of images restored by various filters is not only cumbersome but also leads to difficulty in ranking the filters. The conventional percentage of correct classification or confusion matrix cannot be used here since it is not feasible to identify each individual pixel as being wrongly or correctly clasified; only the number of pixels classified into each class can be counted with little difficulty.

The deviation of the pixel counts per class obtained on the restored images from those counted on the original image gives a measure of the size of the area being wrongly classified. It was thought that this deviation for the particular class would serve as a figure of merit to measure the filter's capability in facilitating the clasification of pixels. However, with a little more reflection, the severity of each deviation should be considered since it is more severe to wrongly classify 2 pixels out of 10 pixels than to misclassify 2 pixels out of 100. Therefore, the misclassification index for the i^{th} class was formulated as:

$$\varepsilon_i \triangleq \frac{N_{ai} - N_{ei}}{N_{ei}} \qquad (12)$$

where N_{ei} = no. of pixels truly belonging to the i^{th} class.

N_{ai} = no. of pixels actually classified into the i^{th} class by machine.

In our case, N_{c_i} is determined by performing machine classification on the original undegraded image as discussed before. It is noted that ε_i has a range of

$$-1 \leq \varepsilon_i \leq \frac{N_t - N_{c_i}}{N_{c_i}}$$

where N_t = total no. of classified pixels in the image. The extreme points in the range indicate that either no pixel is classified into the i^{th} class or all pixels are classified into it. It is seen then that a negative value for ε_i signifies that the i^{th} class is underclassified and a positive value, overclassified. When $\varepsilon_i = 0$, it indicates the best performance for classifying pixels into the i^{th} class. However, it does not provide a single-number as an overall performance measure.

Often, it is desireable to have a single-number index to gauge the overall performance in classifying pixels into all classes under consideration. Summing over ε_i would not do since that allows underclassification and overclassification to cancel each other but both kinds of misclassification are equally harmful. A meaningful overall performance measure is thus defined as

$$\varepsilon \triangleq \sum_{i=1}^{n} \frac{|N_{t_i} - N_{c_i}|}{N_{c_i}} = \sum_{i=1}^{n} |\varepsilon_i| \qquad (13)$$

where n designates the number of clases.

This measure has a theoretical range of

$$0 \leq \varepsilon \leq \varepsilon_{max}$$

where ε_{max} is derived under the worst situation when every pixel is classified into the class of the smallest size. Let N_{c_s} be the number of pixels in the smallest class, then

$$\varepsilon_{max} = (n-1) + \frac{|N_t - N_{c_s}|}{N_{c_s}} = (n-2) + \frac{N_t}{N_{c_s}} \qquad (14)$$

However, in practice, it is more likely that every pixel is swamped into the

class of the largest size. In that case, the largest value of $\mathcal{E}$ that will ever result will be less than the $\mathcal{E}_{max}$ in Eq. (14). A practical range of $\mathcal{E}$ values is

$$0 \leq \mathcal{E} \leq (n-2) + \frac{N_t}{N_{cl}}$$

where N_{cl} is the number of pixels in the largest class. Therefore, in order to apprehend the significance of the values for $\mathcal{E}$, one should always analyze the range of $\mathcal{E}$ values that may arise in any particular application. This realistic maximum value may be used to normalize $\mathcal{E}$ so that the index varies between 0 and 1. The best performance occurs when $\mathcal{E} = 0$. This misclassification index gives a good indication of the performance of a filter pertinent to the machine classification for all n classes under consideration.

Both types of the misclassification indices were used herein. The class-by-class index was employed to evaluate and compare the image restoration filters for their capability in aiding the classification of pixels into a particular class; the overall measure was used to evaluate and compare the filters for their overall performance in classification for all the classes under consideration.

VI. RESULTS AND DISCUSSIONS - The numerical evaluation and comparison is based on the misclassification indices (classification errors) described in the previous section. These results are presented in Tables I, II, II, and IV. The values of Table I represent the classification errors for the metal (airplane); those of Table I!, the errors for the soil; those of Table III, the errors for the shadow; and the values of Table IV, the classification errors for all three classes. The consistency and patterns exhibited by these values show the utility of the performance measures derived. The range for the values of each table is listed below:

Table	Range
I	$-1 \sim 6.0115$
II	$-1 \sim 0.3304$
III	$-1 \sim 8.4608$
IV	$0 \sim 10.4608$

The range for the overall misclassification index (Table IV) was computed based on the assumption that every pixel would be classified as shadow when the image is hopelessly degraded. In all cases, a zero value indicates the best performance. Hence, the closer to zero the index value, the better the performance.

The metal is much brighter than soil or shadow. It is easier to classify. However, shadow and soil are quite similar to begin with and become more alike after degradation of the image. They are much harder to separate. It is seen from Table I that practically all of the filters facilitated the classification of pixels into metal under any distortion condition. From Tables II and III, it is seen that the performances of filters are not clear cut; some of them facilitated the classification for soil at the expense of shadow and vice versa. For example, Wiener filter improved the classification of pixels into soil and shadow at all noise levels and blur severities of $\sigma_h = 4.5$ and 6.5; however, at SNR = 12 db and $\sigma_h = 6.5$, it did not improve the soil classification but helped greatly for the shadow classification; on the other hand, at SNR = 5 dB and $\sigma_h = 4.5$ it performed well for the soil but badly for the shadow. For the overall performance; i.e., facilitating the classification for all three classes, it is seen from Table IV that at high noise level and severe blur, all filters except the inverse and parametric Wiener filters improved overall classification. The behavior of inverse filter at high noise level and severe blur is expected. The erratic performances of parametric Wiener (PW) and

geometrical mean (GM) filters proved again the importance of selecting the γ value (for the PW filter) and the α and γ values (for the GM filter). The value for γ used for PW filter here does not seem to work well for high level of noise and severe blur condition. The PSE filter has the worst record (3.5845) among all the filters; this occurred at the original SNR and blur of σ_h = 6.5. It did improve metal classification but misclassified many soil pixels into the shadow class. As mentioned earlier, the PSE filter limits the gain for the low SNR bands of the image spectrum. After severe blur, at the original SNR, the SNR of the soil band became relatively lower. Hence, many soil pixels became darker and thus were classified as shadow.

The median filter (3 x 3 window) improved the classification of metal for images of high noise and severe blur; for classification of shadows and soil, its performance was mixed. This is in line with its character in preserving contrast while reducing noise. It also showed good overall performance in facilitating classification of pixels into all three classes.

Examining Table IV, filters did not help under noise alone situation though supposedly that filters had filtered out noise. This shows that noise alone has little effect on classification. The filters improved classification in the situations when high noise and severe blur were combined. These results agree somewhat with the results reported in Ref. 11 for human observers as stated in Sec. I of this report. Of course, the results obtained could be due to the characteristics of our particular features (measurements) used in the pattern recognition system. Any generalization of our results must be taken with caution.

For ease of comparison, graphs may be constructed from the tables. A typical graph is shown in Fig. 2. The graph represents the group of images at

12 db SNR in Table IV which portrays the overall classification performance. The abscissa of the graph shows the blur severity mesured in the standard deviation of the Gaussian blur and the ordinate represents the relative overall classification error within the group. The graph shows that the Wiener filter performed well for all blur severities at 12 dB SNR; the geometrical mean filter ($\alpha = \frac{1}{2}$, $\gamma = \frac{1}{2}$) is best for the blur of $\sigma_h = 4.5$ at 12 dB SNR; the PSE and Wiener filters performed at about the same level under the most severe blur; the inverse filter did not work well for the most severe blur at this noise level as expected; the parametric Wiener filter did not perform well because the γ value used here is not suitable for this combination of blur and noise; the median filter did not help at all since there were no sharp boundaries between shadows and soil. It also shows that even at this noise level filtering did not improve classification for less severe blur. It seems that the line connecting all the UR (unrestored) points in the graph and the line through the middle of the filters intersects at $\sigma_h = 3.5$. In other words, at this noise level for a blur less severe than a Gaussian blur of $\sigma_h = 3.5$, one should not use image restoration filter to facilitate machine classification. Other graphs can be constructed for groups of other noise levels. They show similar patterns of behavior except that most of the filters improved classification for less severe blur at high noise level. The Wiener filter had the most steady predictable performances.

In summary, the major objective of this project was to evaluate and compare various image restoration filters in facilitating machine classification. The experimental method used applied a variety of filters to images containing various degrees of degradation with the aim to determine under what conditions a given filter performed the best. The performance was determined by how well the machine classified the filter restored image. Results were obtained for a very

specific pattern recognition system. Hence, caution must be exercised in generalizing our results. It may be categorically stated that the image restoration filtering should be employed under a combination of high noise level and moderate or severe blur, and the combination of low noise and severe blur, furthermore care must be taken in selecting the filter for use.

VII. RECOMMENDATIONS:

Since the image restoration filtering problem is mathematically ill-conditioned, there is no unique solution. Traditionally, the selection of a specific solution from the infinite family of solutions was guided by some mathematical or ad hoc criterion or criteria [22] such as minimum mean square error, edge preservation, etc. The results of filtering were judged by human observers based on subjective quality or interpretability of the filtered images. However, in most machine pattern recognition problems, the features (measurements) and classifiers were selected based on the criterion of minimizing misclassification or maximizing correct recognition if feasible, compatibility between classifier and features, or convenience. Therefore, in order to design restoration filters to facilitate machine recognition, a recommendation for future research is made as follows:

(1) To solve the filtering problem based on a selected set of features (measurements) to be used in the machine recognition system.

Since there exists a variety of image restoration filters as attested in previous sections, it is natural to ask what features do these filters enhance so that the appropriate features may be extracted for use in the machine recognition system. The second recommendation is:

(2) To determine the set of features enhanced by a specific filter or a set of filters.

As the cliche goes "The proof of the pudding is in the eating"; the performance

of a machine recognition system should be examined in total. It is important to understand the interrelation between filters (as preprocessors), feature extractors, and classifiers in order to improve upon the performance of the system as a whole. A more fundamental problem for follow-on research is:

(3) To study how image restoration filters or enhancement techniques interrelate with feature extractors and classifiers.

Recommendations for other follow-on research include:

(4) To evaluate and compare filters realized in different forms such as various FIR and IIR implementations since in practice any image restoration filters must be implemented in one form or another.

(5) To design adaptive filters to adapt to the changing conditions of the feature extractor.

(6) To study the effect of median filter window shape and/or size on the filter's performance in machine recognition.

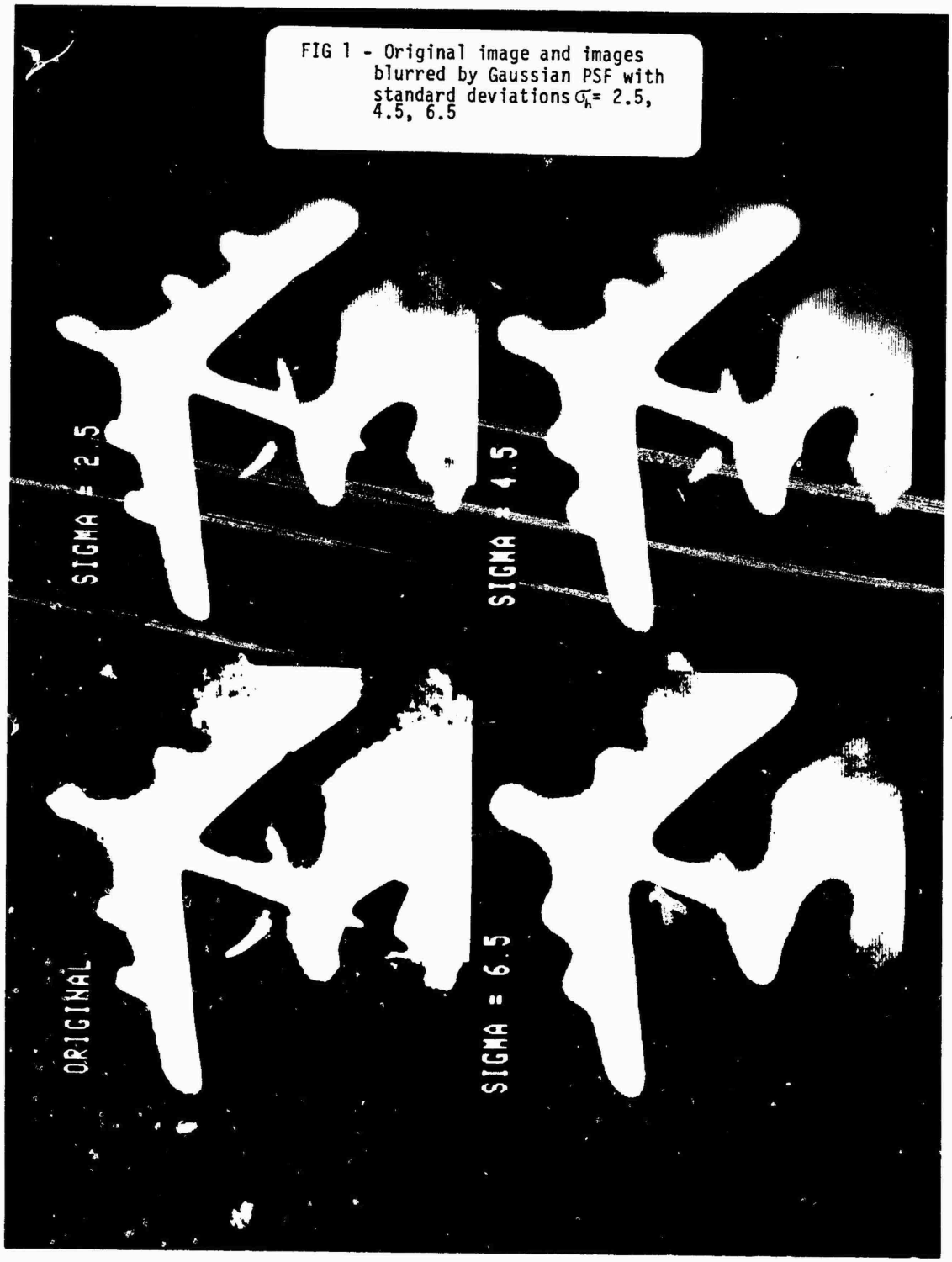


FIG 1 - Original image and images blurred by Gaussian PSF with standard deviations σ_h = 2.5, 4.5, 6.5

TABLE I - RELATIVE CLASSIFICATION ERORS FOR METAL (AIRPLANE)

SNR	σ_h	UR	I	W	PW	GM	PSE	M
Original	2.5	-.0462	-.0414	-.0678	-.0724	-.0625	-.0371	-.0722
	4.5	-.1500	-.1373	-.0623	-.1349	-.0884	-.1399	-.1514
	6.5	-.3528	-.1442	.0593	.1499	-.0645	-.0434	-.1544
12dB	0	.0051	.0063	-.0024	-.0007	.0001	.0077	.0255
	2.5	-.0536	-.0619	-.0422	-.0416	-.0388	-.0413	-.0287
	4.5	-.1470	-.0926	-.1148	-.1446	-.1299	-.1438	-.0750
	6.5	-.2032	.0221	-.2209	-.2237	.0623	-.1970	.1560
5 dB	0	.0136	.0077	.0221	.0111	.0020	.0392	.0506
	2.5	-.0298	-.0065	-.0092	-.0169	-.0102	-.0021	.0225
	4.5	-.1314	-.0635	-.0996	-.0758	-.1936	-.0810	-.0290
	6.5	-.1920	-.2238	-.1467	.0315	-.0980	-.2447	-.1256

Range: -1.000 ~ 6.0115

σ_h: standard deviation of Gaussian blur

UR: Unrestored

I: Inverse filtering

W: Wiener filtering

PW: Parametric Wiener filtering (= ½)

GM: Geometrical Mean filtering (= ½, - ½)

PSE: Power spectrum equalization filtering

M: Median filtering

Neg. No.: Underclassified

Pos. No.: Overclassified

TABLE II RELATIVE CLASSIFICATION ERRORS FOR SOIL

SNR	σ_h	UR	I	W	PW	GM	PSE	M
Original	2.5	.0125	.0029	.0150	.0115	.0168	.0112	.0141
	4.5	.0348	.0322	.0162	.0383	-.0121	.0303	.0373
	6.5	-.0345	.0172	.0026	-.1601	-.1312	-.4293	.0357
12dB	0	-.0020	.0182	.0036	.0053	.0105	.0025	-.0089
	2.5	.0141	.0212	.0175	.0207	.0209	.0250	.0131
	4.5	.0497	.0330	.0367	.0389	.0245	.0382	-.0380
	6.5	-.0233	-.1106	.0642	-.1082	-.0611	.0064	-2215
5 dB	0	-.0047	-.0039	.0157	.0095	.0081	-.0035	-.0042
	2.5	.0185	.0114	-.0116	-.0086	.0029	.0129	.0127
	4.5	.0300	.0216	.0125	-.0093	.0476	.0043	.0184
	6.5	-.0683	-.0668	-.0207	-.1479	-.0262	-.0475	-.0106

Range: $-1.000 \sim 0.3304$

Symbols explained in Table I.

TABLE III RELATIVE CLASSIFICATION ERRORS FOR SHADOWS

SNR	σ_h	UR	I	W	PW	GM	PSE	M
Original	2.5	-.0263	.0355	-.0150	.0157	-.0350	-.0296	-.0027
	4.5	-.0534	-.0440	-.0315	-.1009	.2055	-.0264	-.0608
	6.5	-.1959	.0709	-.0985	.9365	1.0204	3.1118	-.0456
20dB	0	.0074	.1390	-.0226	-.0367	-.0751	-.0285	.0291
	2.5	-.0283	-.0675	-.0674	-.0915	-.0962	-.1218	-.0541
	4.5	-.1834	-.1098	-.1061	-.0817	.0088	-.0773	.3714
	6.5	.4398	1.4954	-.1583	1.0712	-.3507	.2203	1.3[illegible]5
5 dB	0	.0153	.0175	-.1411	-.0825	-.0606	-.0281	-.0381
	2.5	-.0910	-.0724	.0947	.8239	-.0066	-.0889	-.1206
	4.5	.3925	-.0680	.2244	.1683	-.0771	.0790	-.0915
	6.5	.7447	.7772	.3450	1.0091	.3185	.6680	.2447

Range: $-1.0000 \sim 8.4608$

Symbols explained in Table I.

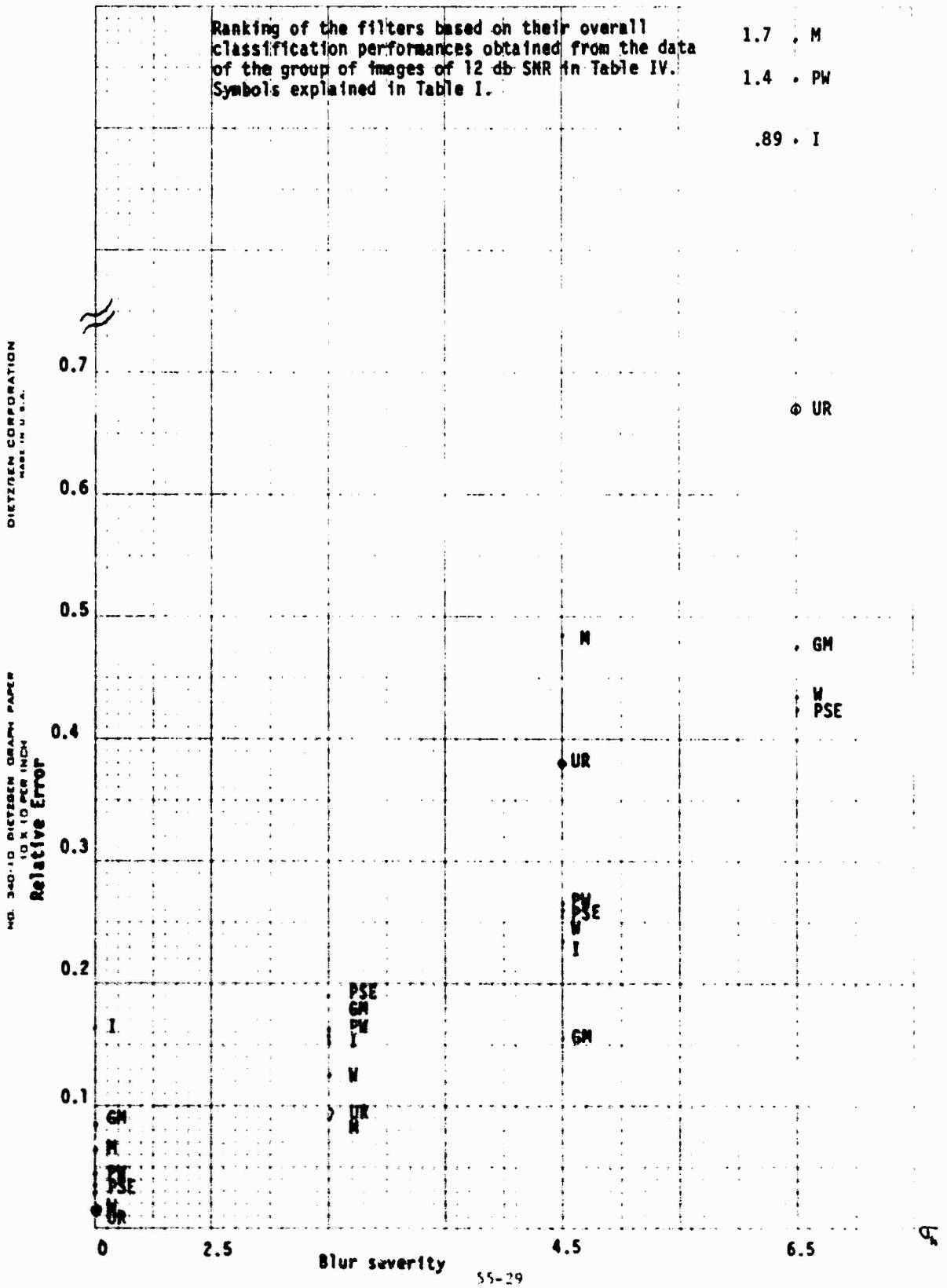
TABLE IV RELATIVE CLASSIFICATION ERRORS for ALL THREE CLASSES

SNR	σ_h	UR	I	W	PW	GM	PSE	M
Original	2.5	.0850	.0798	.0977	.0996	.1144	.0779	.0889
	4.5	.2382	.2135	.1100	.2741	.3060	.1966	.2495
	6.5	.5832	.2323	.1604	1.2466	1.2162	3.5845	.2358
12 dB	0	.0145	.1641	.0286	.0427	.0857	.0387	.0635
	2.5	.0960	.1507	.1270	.1538	.1559	.1881	.0958
	4.5	.3801	.2355	.2577	.2652	.1553	.2593	.4844
	6.5	.6663	.8891	.4435	1.4031	.4741	.4237	1.7420
5 dB	0	.0337	.0292	.1789	.1031	.0706	.0708	.0929
	2.5	.1393	.0903	.1155	.1093	.0197	.1040	.1558
	4.5	.5539	.1532	.3366	.2533	.3183	.1643	.1388
	6.5	1.0049	1.0678	.5123	1.1885	.4427	.9602	.3809

Range: 0 $\sim$ 10.4608

Symbols explained in Table I.

FIG. 2


Ranking of the filters based on their overall classification performances obtained from the data of the group of images of 12 db SNR in Table IV. Symbols explained in Table I.
1.7 . M
1.4 . PW
.89 . I
Relative Error
Blur severity
0
2.5
4.5
6.5
0.1
0.2
0.3
0.4
0.5
0.6
0.7
UR
GM
W
PSE
M
I
PW

REFERENCES

1. H.C. Andrews and B. R. Hunt, Digital Image Restoration, (Prentice-Hall, Englewood Cliffs, N.J. 1977), pp. 113 - 208.

2. W. K. Pratt, Digital Image Processing, (J. Wiley, New York , 1978) pp. 378-446.

3. A. Rosenfeld and A.C. Kak, Digital Picture Processing, (Academic Press, New York, 1976), pp. 203 - 255.

4. R. C. Gonzalez and P. Wintz, Digital Image Processing, (Addison-Wesley, Reading, MA 1977), pp. 184 - 227.

5. E. L. Hall, Computer Image Processing and Recognition, (Academic Press, New York, 1979) pp. 214-247.

6. T. S. Huang (Ed), Picture Procesing and Digital Filtering, 2nd Ed. (Springer-Verlag, New York, 1979), pp. 7-18, pp. 70-246, pp. 284-288.

7. T. S. Huang (Ed), Two-Dimensional Digital Signal Processing I., (Springer-Verlag, New York, 1981), pp. 4-7, pp. 11-40, pp. 155-205.

8. T. S. Huang (Ed), Two-Dimensional Digital Signal Processing II,, (Springer-Verlag, New York 1981), pp. 161-217.

9. Op Cit Ref 1, pp. 159-160.

10. T.M. Cannon, H.J. Trussell, and B. R. Hunt , "Comparison of Image Restoration Methods," Applied Optics,, Vol 17, No. 21, pp. 3384-3390, Nov 1978.

11. H. L. Snyder, M.E. Maddox, D.I. Shedivy, J.A. Turpin, J. J. Burke and R.N. Strickland, "Digital Image Quality and Interpretability: Database and Hardcopy Studies" Optical Engineering, Vol 21, No. 1, pp. 14-22, 1982.

12. G. Yang and T. S. Huang, "Median Filters and Their Applications to Image Processing", Tech Report., School of Electrical Engineering, Purdue Univ., 1980.

13. Op Cit Ref 4, pp. 215-217.

14. Op Cit Ref 1, pp. 160-161.

15. M. Cannon, "Blind Deconvolution of Spatially Invariant Image Blurs with Phase", IEEE Trans. ASSP24, No. 1, pp 58-63, Feb 1976.

16. J. W. Tukey, Exploratory Data Analysis, (Addison-Wesley, Reading, Ma 1977).

17. T.S. Huang, G.J. Yang, G.Y. Tang., "A Fast Two-Dimensional Median Filtering Algorithm" IEEE Trans. ASSP-27, No. 1 pp. 13-18, Feb 1979.

18. E. Ataman, V.K. Aatre, K.M. Wong, "A Fast Method for Real-Time Median Filtering", IEEE Trans. ASSP-28, No. 4, pp 415-421, Aug 1980.

19. G. P. Wolfe, J.L. Mannos, "A Fast Median Filter Implementation", Proc. SPIE Seminar on Image Procesing, Sept 1979, San Deigo, CA

20. W.L. Evensole, D.J. Mayer, F.B. Frazee, T.F. Cheek, Jr., "Investigation of VLSI Technologies for Image Procesing", Proc. Image Understanding Workshop, Pittsburgh PA pp. 191-195, Nov 1978.

21. P.E. Hart, "The Condensed Nearest Neighbor Rule," IEEE Trsns. IT-14, No. 3. pp. 515-516, May 1968.

22. Op Cit Ref 1, pp. 124-125.

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

EVALUATION OF PROPOSED INTEGRATED COMMUNICATION NAVIGATION IDENTIFICATION AVIONICS (ICNIA) ARCHITECTURES FOR THEIR FAULT TOLERANT CHARACTERISTICS AND POTENTIALS

Prepared by:	Stella Lawrence
Academic Rank:	Professor
Department and University:	Department of Engineering Technologies Bronx Community College City University of New York
Research Location:	AIR FORCE WRIGHT AERONAUTICAL LABORATORIES AVIONICS LABORATORY SYSTEMS AVIONICS DIVISION INFORMATION TRANSMISSION BRANCH RADIO SYSTEMS GROUP
USAF Research Colleague:	Robert Harris
Date:	August 12, 1983
Contract No:	F49620-82-C-0035

EVALUATION OF PROPOSED INTEGRATED COMMUNICATION NAVIGATION IDENTIFICATION AVIONICS (ICNIA) ARCHITECTURES FOR THEIR FAULT TOLERANCE CHARACTERISTICS AND POTENTIALS

By

Stella Lawrence

ABSTRACT

An assessment of the fault tolerant merits of the hardware and software designs of two proposed Integrated Communication Navigation Identification Avionics (ICNIA) terminals was made. The ITT candidate system is a "software radio" utilizing transversal filters to convert from carrier to baseband frequencies. The TRW candidate system design is a conventional superhetrodyne radio system utilizing mixing and downconversion. Classical fault tolerance implies component redundancy for the entire system; however in ICNIA-fault-tolerance the communication navigation identification (CNI) system functions are prioritized and the most critical functions are preserved at the expense of less critical ones. The ITT and TRW architectures represent a first step in the application of the principles of fault tolerance to avionics CNI system design. Suggestions are offered for further research in this area.

Acknowledgement

The author would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for providing her with the opportunity of spending a worthwhile and interesting summer at the AFWAL Avionics Laboratory, Wright Patterson AFB, Ohio. She would like to thank the laboratory, in particular the Information Transmission Branch, without whose assistance this effort would not have been possible, for its hospitality and excellent working conditions. In particular, Major Charles R. Register, Mr. Darlow Botha and Mr. Robert Harris have been very supportive and helpful throughout the ten weeks. A special thanks to Ms. Tammy Sizemore and Mrs. Sabrina Furlow for their excellent secretarial support.

I. INTRODUCTION

The Integrated Communication, Navigation, Identification Avionics (ICNIA) program is an Avionics Laboratory effort aimed at reducing total size, weight, and cost of CNI radio equipment. Although space limitations of tactical aircraft provide the driving ICNIA design requirements, the resultant ICNIA size, weight, and cost savings also are applicable to larger aircraft.

ICNIA now offers a single integrated CNI system, compatible with the multiplexed interface of today's military avionics while retaining the present standard CNI waveforms. ICNIA will contain many "building block stages," which by means of processor controlled matrix switches can be configured to receive or transmit standard CNI signals in the 2-MHz to 2-GHz portion of the radio spectrum.

The control processor accomplishes I/O with the remainder of the avionics system via a multiplex bus. It selects the configuration of the building blocks thereby controlling which CNI functions are active. In the event of a function failure the control processor has the capability of substituting redundant equipment or, if necessary, of reconfiguring the building blocks to regain the use of a function at the expense of a lower priority function. The reliability of each CNI function is thus increased by reconfiguration so that essential functions are maintained.

ICNIA possesses the following advantages over discrete function CNI subsystems:

a. A single system that eliminates the requirement for many dedicated CNI functions with duplicative hardware,

b. Increased CNI functional reliability due to reconfigurability,

c. Space, weight, power, and cooling savings brought about by a reduction of hardware.

Fault tolerance is the unique attribute of a system which makes it possible for the system to continue with its program-specified behavior after the occurrence of faults (malfunctions or failures of physical parameters)[1]. Fault tolerance is a vitally important and desirable property that ICNIA should possess, since it is the survival attribute that will enable the ICNIA system to maintain its specified behavior despite hardware or software failures.

Two conceptual design architectures for an integrated communication, navigation, identification radio system have been developed; one by ITT Avionics, Inc., and the other by TRW, Inc. Both include some degree of fault tolerance in their hardware and software designs; but prior to this paper no comparative analysis of each had been performed to determine their relative fault tolerant merits.

An assessment of the fault tolerant merits of the hardware and software design in each proposed ICNIA was the goal of this work.

II. OBJECTIVES

The project had three objectives:

A. To determine general desirable standards for fault tolerant hardware and software designs,

B. To assess the planned fault tolerant features of the current system definition phase for the ITT and TRW designs,

C. To make recommendations regarding the fault tolerant features to be implemented in the hardware and software of the Advanced Development Models.

III. GENERAL FAULT TOLERANT SYSTEMS AND STANDARDS

The criteria for avionics fault tolerance are that it be cost effective, not add to size or weight, be as simple as possible consistent with meeting desired reliability goals, and not introduce additional faults and errors.

Reliability should first be attained by prior elimination of all possible causes of failure. The first step consists of fault intolerance. Fault intolerance mandates the use of high reliability components and extensive testing. The second step is to provide redundancy to raise system reliability.

Effective fault-tolerant systems may be designed following a systematic design methodology that can be summarized in five steps: 1) specify system architecture without fault tolerance; 2) identify classes of faults that are to be tolerated; 3) select cost-effective forms of protective redundancy to cover the identified faults; 4) estimate the effectiveness of the redundancy by analytic, simulation, or experimental techniques; 5) devise checkout methods to test the redundant features of the system. In practice, several iterations of steps (3) and (4) are usually needed to arrive at a satisfactory design.

Reliability, maintainability, and fault tolerance must be designed into the basic system. After the fact, "add-on" concepts rarely accomplish the desired goals.

Two basic methods of recovery are utilized. Hardware-controlled recovery is implemented by means of modular replacements or reconfiguration. Software-controlled recovery depends on special programs to carry out recovery after faults are detected.

Faults should be detected, isolated, and eliminated. The system must then be rechecked to confirm that the fault is no longer present.

Critical points and critical equipment must be identified and duplicated.

The choice of failure unit is crucial[2]. Its size and complexity must be neither too large nor too small. The replaceable units or modules should be self-testing, if possible. The levels of redundancy of the various elements must be chosen to be consistent with required levels of reliability, maintainability, and performance.

The data-processing, control, and internal communications configuration should be extremely reliable, maintainable, and flexible. The system must establish a hierarchial organization. The levels of redundancy of the various elements must be chosen consistent with required levels of reliability and maintainability. The data processing system control language used, and the software employed should be fault tolerant[3].

IV. FAULT TOLERANT FEATURES OF THE ITT ARCHITECTURE

The ITT architecture[4] consists of six LRU's (Line Replaceable Units). There are four categories of LRUs: the RF (radio frequency) LRU, the digital LRU, the HF transmitter, and the HF antenna coupler.

The RF LRU is not entirely duplicated. Its two receiver-transmitter units are similar but not identical. Two-thirds of the RF modules in each RF LRU are duplicated to provide redundancy for the highest priority functions. The L-Band transmitter and front ends are duplicated for IFF and TACAN only, however VHF/UHF reception and transmission are duplicated.

The digital LRU is duplicated. Within the digital LRU's the data processor, vector processor, programmable controller, and programmable digital matched filter modules are duplicated.

The system is a "software radio" and its design proposes initially the use of the J73 language and later Ada. It uses the MIL-STD-1750 instruction set, and the MIL-STD-1553 digital data bus interface. It provides immediate

A/D conversion from baseband and is based upon all digital signal processing. It utilizes a digitally controlled transversal filter for direct tuning.

The fault detection and reconfiguration (FDR) software design is similar to TIES[7] (Navy Tactical System Information Exchange). The allocation of hardware system resources will use a pre-stored linkage table or set of chain rules which specify the various resources necessary to provide a CNI function. The FDR software is to be stored in each data processor unit. When an error condition is detected control will be transferred to unaffected FDR programs.

The FDR software should be designed to isolate the problem and select the optimum configuration for the remaining resources based on mission priorities.

Fault tolerance and fail-soft capability are very important aspects of data processor design. Each DPU (data processing unit) is independent. When faults occur the DPUs should be capable of supporting the top priority functions of the system (possibly with a reduction in throughput below normal maximum load). Within each DPU, duplicate devices and alternate bus paths should provide fail-soft capability. If any device shuts down, part of its work load should be absorbed by the spare capacity of a matching counterpart. After a hard failure, system capacity will have to be evaluated. The system should then be reconfigured by reassigning tasks to the DPUs.

The design includes vector processed, high speed computations on structured data streams. An example of this is, digital-filtered functions.

The ITT system plans extensive implementation of VLSI/VHSIC (Very High Speed Integrated Circuit) technology. The vector and data processors will utilize VHSIC modules.

V. FAULT TOLERANT FEATURES OF THE TRW ARCHITECTURE

The TRW design is a conventional superheterodyne radio. Its architecture[5] is organized as a collection of resource pools, which are dynamically allocated resources. It consists of 7 LRUs.

Each resource pool contains a number of identical elements that are available on a highly flexible basis. The pooled elements are line replaceable units or modules, within the avionics system. All internal fault isolation is carried to the level of these units. Maintenance and logistics procedures are designed around them as basic system elements. In addition, the pooled units or basic building blocks are hierarchically interconnected to perform the necessary avionics system functions.

Three types of maintenance modes are envisioned[6]. Non-disruptive self-testing involves the real-time monitoring of appropriate test points to determine subsystem health dynamically. This represents the built-in-test (BIT). These tests are non-disruptive in that they involve no changes in configuration and will not interfere with ongoing ICNIA signal and data processing. A second test is disruptive self-testing. It provides more detailed and diagnostic information. It is a closed-loop test and can require some ICNIA reconfiguration and the preemption of other processing tasks. A third supplementary test mode consists of a final inventory of system health information. This is part of the data processing tasks for each relevent ICNIA system.

Non-disruptive and disruptive self-testing are conducted by maintenance control algorithms[7]. This information must include mode select commands to enable reconfiguration. Failures are to be detected and isolated by extensive built-in computer-controlled performance monitoring, designed for real time fault detection, isolation, and system reconfiguration. The data processing

tasks are to be partitioned between the ICNIA subsystem and the core avionics[9]. Of the various options, desirable partitionings will be determined in the course of the ADM design. The terminal will be designed to interface with a MIL-STD-1553B multiplex bus, and to be compatible with future high speed digital or analog buses.

It is desirable for some degree of system autonomy to support stand-alone operation and test. The physical location of the redundant backup data processor shold be considered.

The TRW system proposes extensive use of VLSI/VHSIC. In particular, the design incorporates the use of VHSIC technology in the matrix switch signal processors[10].

VI. RECOMMENDATIONS REGARDING FAULT TOLERANT FEATURES TO BE IMPLEMENTED IN THE ITT AND TRW ARCHITECTURES

The terminals outlined in Sections IV and V represent interim designs. Careful consideration should be given to desirable fault tolerant criteria and practices for the next design stage. In the present designs these are outlined but should be further developed. Fault tolerance should be an integral part of the designs. Despite size, weight, and cost constraints, basic fault tolerant criteria can be applied in the following sense. ICNIA is a complex, time varying, integrated, reconfigurable system. It involves several types of aircraft, different missions, and different functions. During the various mission segments, the equipment for noncritical functions can be considered redundant backup for the critical functions. With this viewpoint in mind, standard fault tolerant practices can be applied. A careful analysis should be made via probability models to insure that sufficient equipment redundancy is provided for the critical functions in each

mission segment. Levels of redundancy for the various elements must be chosen consistent with required levels of reliability, maintainability, and performance.

The first step in fault tolerance design consists of fault intolerance. There should be prior elimination of the causes of failure by selection of reliable components and appropriate testing.

Since fault tolerance is an integral part of the ICNIA system design, the entire system including the fault tolerant implementation, should be subjected to hardware, computer, software, and human factor fault tolerance tests. The tests should consist of reliability and performance tests using mathematical models, simulation models, and actual hardware.

When it is certain that the system is one hundred percent healthy, it should be subjected to representative scenario testing, as determined by operational doctrine, to develop confidence that the ICNIA terminal will satisfy the criterion of 10,000 hours mean time between critical failure (MTBCF). A critical failure is the inability to perform a CNI function during any part of a mission that is necessary for mission success.

A careful study should be made of critical points in the ICNIA terminal. These include the transversal filters for ITT, the matrix switches for TRW, the control computer, the control executive, control algorithms, and software for both designs. Critical parts should be duplicated, or at the very least, dual contacts should be provided for the interconnections. Probability analysis is essential to prevent switch overloading. Control algorithm(s) must be carefully analyzed, designed, and tested. Artificial intelligence (adaptive learning networks) might be incorporated into the control algorithm. Fault detection and reconfiguration (FDR) software should be fail-safe, incorporating error coding, and perhaps interleaving. The control computer

should be fault tolerant, or at least duplicated. Provisions should also be made for memory storage in case of intermittent or longer time power failures.

Fault isolation should be at the LRM (Line Replacement Module) rather than the LRU level. Several LRMs could reside in an LRU. The LRMs should be self-testing. Careful thought should be given to the choice of complexity and LRM size in terms of functions. Failure of one module should not cause any impediment to reconfiguration. Failure of a module during self-test should set a flag. The module should then be retried. This will help locate intermittent faults.

As a result of VHSIC technology which makes possible the use of self-testing, fault tolerant modules, and embedded processors, pooled architecture systems can be used to obtain economical dependability[12].

Tasks should be properly apportioned in a hierarchial fashion to match control and data processor capability. In addition, redundant units should not be "cold spares". Equipment should be rotated periodically and all equipment should be subject to periodic test.

It is highly recommended that a further, careful study be made of the fault tolerant features of the next design iteration of ICNIA.

REFERENCES

1. Avizienis, A., "Fault-Tolerance: The Survival Attribute of Digital Systems" Proceedings of the IEEE, Vol 66, pp. 1109-1125, 1978.

2. Charles Stark Draper Laboratory, Inc., An Integrated Fault-tolerant Avionics System Concept for Advanced Aircraft, Contract N00019-78-C-0572, February 1979.

3. International Business Machines Corporation, Fault Tolerant Computer Network Study, Contract AFWAL-TR-80-1046, April 1980.

4. ITT Avionics Division, Interim Technical Report of Integrated Communication, Navigation, Identification Avionics (ICNIA), Contract F33615-80-C-1176, January 1982.

5. TRW Military Electronics Division, Interim Technical Report, ICNIA Architecture CNI Functional Description, Contract F33615-81-C-1512, March 26, 1982.

6. TRW Military Electronics Division, Interim Technical Report, ASA Integrated CNI Interface, Contract F33615-81-C-1512, February 26, 1982.

7. TRW Defense and Space Systems Group, Final Report, Multifunction Multiband Airborne Radio Architecture Study, Contract AFWAL-TR-81-1113, January 1982.

8. Communications, Command and Control Division, Communications and Navigation Technology Directorate, Naval Air Development Center, Tactical Information Exchange System, (Draft Report), March 7,1983.

9. TRW Military Electronics Division, Interim Technical Report, Functional Interface Control Document: ICNIA User Equipment/Advanced Aircraft Avionics Mission Software, Contract F33615-81-C-1512, April 20, 1982.

10. TRW Electronic Systems Group, Requirements Specification for ICNIA System Architecture Simulation, April 1, 1982.

11. General Dynamis, Fort Division, Report of Integrated Communication Navigation and Identification Avionics (ICNIA), Contract FZM-7092, December 10, 1977.

12. SRI International, Fault Tolerant Architecture fo VHSIC, Contract F0602-80-C-0303, March 1982.

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

Decision Aids for Selecting Air Force Manufacturing Technology Projects

Prepared by:	David R. Lee, Ph.D.
Academic Rank:	Associate Professor
Department and University	Chairman, Management Department University of Dayton
Research Location:	Air Force Business Research Management Center
USAF Research Colleague:	Captain Mike Tankersley
Date:	22 July 1983
Contract No.:	F49620-82-C-0035

DECISION AIDS FOR SELECTING AIR FORCE MANUFACTURING TECHNOLOGY PROJECTS

By

DAVID R. LEE

ABSTRACT

The availability and level of development of specific manufacturing technologies directly impact the U.S. Government's ability to achieve certain national defense objectives. The Air Force Manufacturing Technology (MANTECH) Program, begun some 35 years ago, supports the development of advanced manufacturing technologies for application at contractor and Air Force facilities.

In a resource constrained environment, the situation involves deciding which MANTECH projects to select and how to allocate resources. Both of these decisions are complex, partially defined processes involving many variables. This research, however, only focuses on the project selection process, with specific attention on evaluation of the decision variables.

With a semi-structured problem, intuitive judgement by the decision maker is an essential part of the decision process. Thus, the objective of this research was to develop decision aids, or management tools, to assist the decision maker in selecting projects to receive MANTECH support. This report describes the methodology employed, describes the major factor categories, and presents a method for summarizing these decision variables.

Acknowledgement

The author wishes to thank the Air Force Business Research Management Center, the Air Force Office of Scientific Research, and the Southeastern Center for Electrical Engineering Education for providing him and his SCEEE graduate student assistant, Mr. Daniel T. Fetsko, with the opportunity to spend a very worthwhile and interesting summer at Wright-Patterson AFB, Ohio. He wishes to acknowledge the excellent support and hospitality of Captain Michael Tankersley, Major Robert Golden, and and Colonel Ronald Deep of the Air Force Business Research Management Center.

Table of Contents

		Page
I.	Introduction	5
II.	Objective of the Research Effort	7
III.	Research Approach	7
IV.	Research Results	9
	A. Evaluation Scheme	9
	B. Factor Categories	11
	C. Project Evaluation Summary Chart	15
V.	Conclusions	17
VI.	Recommendations	17
	Attachment	18

I. INTRODUCTION:

The term manufacturing technology, in the more broadly used sense, pertains to technologies employed in the production, inspection and repair of an item. The availability and level of development of manufacturing technologies affect the cost, quality, quantity, physical characteristics, and/or useability of an item.

Because of its specific needs and resource limitations, the Department of Defense (DOD) is very concerned about the development of certain manufacturing technologies which impact its ability to obtain, maintain, and effectively utilize military-related items. World War II highlighted the close relationship that must exist between the national industrial base and the DOD. Since 1947, the US Air Force has been involved in fostering the development and growth of specific manufacturing technologies which, without its participation, would not be advanced or perhaps not advanced in a timely fashion.

The USAF currently conducts their activities in this area under AF Regulation 800-23, 25 April 1982, entitled Manufacturing Technology Program. This progam is therein identified as MANTECH and described as:

> "A program which consists of all actions taken by the Air Force to develop and carry out new or significantly improved production systems, processes, techniques, or equipment for use (near or long range) in contractor facilities or ALCs in support of Air Force systems, subsystems, or equipment."

As noted, MANTECH interests are centered not only on Air Force contractor activities but also on those at the Air Logistics Centers (ALCs). This expansion, from looking at just contractor activities to now considering both internal and external activities, necessitates evaluating additional candidate technologies for possible MANTECH support.

The Manufacturing Technology Division (MLT) located in the Air Force Systems Command Materials Laboratory at Wright-Patterson AFB is responsible for planning, implementing, managing and controlling the Air Force MANTECH program. An important element in the MANTECH program is that the payback and improvements in military capability generated as a result of the proper MANTECH support can be awesome. For instance, a MANTECH investment of $267,550 on just one aircraft engine project is estimated to produce a savings of $195.8 million over an eleven-year period.

In essence, there is a sizeable MANTECH budget (approximately $60 million for FY 1983), a large number of candidate projects, and opportunities for significant payback. Proper project evaluation, as well as the accompanying selection and resource allocation decisions involve many factors and are not well defined processes.

Topic N82-01, "Assessment of Productivity Potential of Manufacturing Technologies," contained in the Air Force Business Research Management Center's publication, Business Research Topics for FY-83, addressed the need for developing a methodology or planning model in the MANTECH area. This topic was the generating source for this research effort. The objective of N82-01, as submitted by AFSC/PMDP, was to:

> "Develop a methodology or planning model for identifying manufacturing technologies which generate the highest marginal increases in productivity growth, potential areas of application, and the availability of such technologies."

Early investigation of that topic showed that productivity alone, is just one of the many factors entering into the selection and resource allocation decision processes. Since there are so many factors to consider, preliminary attention needed to be directed toward identifying the influencing factors, assessing their relevancy, and then bringing these factors together in a comprehensible way to aid the decision makers. This study focuses on these elements.

II. OBJECTIVE OF THE RESEARCH EFFORT:

The main objective of this research effort was to develop decision aids, or management tools, to be used in evaluating and selecting which manufacturing technology projects should be supported under the USAF MANTECH Program. The major supporting goals were: (1) determine what factors influence MANTECH selection decisions, (2) categorize these factors into major groups, (3) develop ways to assess the strength of the factors, their interrelationships, and/or ensure that they were not overlooked, and (4) to summarize the relevancy of these factors in a comprehensible format for use by the key decision-makers.

III. RESEARCH APPROACH:

The first step in this research effort centered on collecting background information. This was accomplished by performing literature searches, reviewing several reports written on MANTECH activities, and conducting initial interviews. These interview results suggested that before the topic of productivity itself should be addressed, we needed to look deeper into the current decision process.

From the onset, it is helpful to realize that the selection and funding of MANTECH projects is a comprehensive and complex decison process involving a variety of factors and is compounded by many future uncertainties. Pressures are on MANTECH projects that they produce useful results in a timely manner.

The generation of MANTECH ideas come from numerous sources. Many ideas originate from discussions with contractors regarding problems, potential solutions, and promising opportunities. Projects are also generated by Systems Program Offices (SPOs), higher Headquarters, other Air Force Major Air Commands, other DOD and Government agencies, professional and academic associations, and from the basic research laboratories. Obviously, there are other input sources, but the point of this is to demonstrate that there are numerous activities attempting to impose their thoughts and projects into the MANTECH selection process.

In an effort to manage and evaluate this multiplicity of inputs, the Materials Laboratory has developed a Focal Point System. There are currently ten Focal Points such as thermal protection materials, metalic structural materials, non-metalic structural evaluation, CAM/manufacturing R&D, systems support, etc. These Focal Point managers employ a matrix structured organizational approach in prioritizing MANTECH candidate projects within their specific area. That is, representatives from other interested Air Force offices are charged with providing information to and confering with the Focal Point manager. At the appropriate time, these managers present their projects to a Materials Laboratory Executive Group which determines what projects will be selected for MANTECH funding, and at what level of funding. Not all annual project funds are for new project starts: a portion of the MANTECH fiscal year monies are directed toward projects that began in previous years but which are not yet completed. The annual MANTECH program is then approved at the appropriate Air Force level as outlined in AFR 800-33. While the MANTECH Divison manages and controls the program, this occurs primarily after the projects are selected and funding is determined.

Following collection of initial background information, we could see that while the basic flow of the selection decision process seemed reasonable, perhaps a meaningful contribution could be made by adding more structure to assessment of the evaluation factors. Our research objective was re-formulated to that wh'ch appeared earlier in Section II. Since an expanded data base was required to achieve the objective, further interviews were then conducted with people inside and outside the Materials Laboratory. This included interviewing people in the AFSC Aerospace Industrial Modernization office, ASD Plans and Programs office, former MANTECH personnel, several Focal Effort managers, project engineers, the current MANTECH Branch Chiefs , the Materials Laboratory Director, and other personnel within the Materials Laboratory. From these discussions, numerous considerations were collected and then categorized into appropriate evaluation groups.

IV. RESEARCH RESULTS:

A. Evaluation Scheme:

The proposed project evaluation scheme involves 16 factor categories as listed below:

I. Title

II. Objective

III. Description

IV. Nature of Need

V. Support

VI. Funding for Development

VII. Non-Direct Financial Requirements

VIII. System Integration

IX. Interlock with Other Thrust Areas

X. Technical Success

XI. Future Investment Requirements for Implementation

XII. Federal Agency Implementation

XIII. Industrial Implementation

XIV. Benefits

XV. Global Relationships

XVI. Return on Investment

Except for the first three factors (Title, Objective and Description), the remaining 13 factors are to be scored on a Project Evaluation Summary Sheet, which follows later in this report. There are a variety of elements in each of these major factor categories. In reaching a rating in each category, a systemized thought process is necessary to ensure that the essential elements are not inadvertently excluded. The categories are not totally independent or mutually exclusive of one another. The proposed method does attempt to minimize duplication and places most elements (or evaluation considerations) in only one factor group.

Mathematically directed formats and straight checklists were considered for arranging the elements, but were dismissed in favor of a decision stimulation format involving questions, fill in the blanks, and comparisons. The elements are included in a 38-page document entitled "MANTECH Project Evaluation Back-Up Sheets." Space limitations, however prevent their being published as an attachment to this report. Copies of the evaluation Back-Up Sheets have been provided to the AFBRMC and to various offices in the Materials Laboratory. Attachment 1 is a portion of the Back-Up Sheets which support a rating for the Need factor, Category IV.

Categories are arranged so as to start with a general view of the project and then become more specific as one progresses through the evaluation. The Back-Up Sheets are not designed to automatically generate a final conclusion on the rating of a particular factor. Their purpose is to function as an assessment guide and assist individuals in reaching their conclusions. Effective use of the Back-Up Sheets rely heavily upon experienced intuitive judgement. As with the final Project Evaluation Summary Sheet, both tools can be "fudged" to meet a particular person's individual needs and objectives. Project selection in MANTECH or in any other institutions is an unstructured, or at best a semi-structured, problem. The key is to reduce the subjectivity by discretely applying models or methodologies and wisely using available information to its full potential.

B. Factor Categories:

This section provides the philosophy and rationale for the various categories.

1. Title: Self Explanatory.

2. Objective: The objective of the project sets the benchmark for the rest of the evaluation. It should be concise, quantifiable, and measureable.

3. Description: In this section, the project is classified as generic, repair and manufacture, integrating, or quality technology. More so, it stresses the element of time by highlighting the required completion date, time for development, time for integration or time for implementation into a manufacturing line, and the recommended project start time. Worth noting is that once a MANTECH project is completed, there is still some implementation time until an industrial output or impact is actually realized. It is not uncommon to be talking of several years for implementation, following completion of the project development phase. This section also summarizes the relative funding requirement.

4. Nature of Need: Who is the principal sponsor of the project? Is it a project directed by higher Headquarters? Where did the need originate? Who will the project serve and is the need a current one or a projected future need? These elements begin to put the project into perspective. Applicability and project timing are critical issues. Thus, key questions have to be whether or not this project will be directly applicable to a weapon system or some support system and will it be done in time to have an impact. The dates in this section can be compared to the project dates in the Description category. The last part of this category centers on what is the real driving factor for the project. A project will generally serve multiple purposes, which are addressed later under Benefits. The question here is to select the single most important issue. This section also brings out the idea of technology push/pull. Projects which are need driven and pulling technology for the development often have a higher probability for implementation success than technology push-type projects.

5. Support: Whether or not the project is need-driven is important, but who supports the project is critical. This section provides a matrix where organizations can be listed vertically with their interest level indicated on the horizontal axis, on a scale from -3 to +3. It is important to record the priority which the organizations place on the project within their programs, if applicable. Comparisons of their priority ratings against the interest ratings help validate their support ratings. A potentially good project without the proper interest may have a tough time "weathering the storm." Similarily there have to be a sufficient number of engineers in MANTECH interested in the project. Also an assessment of foreign interest in the project can frequently change opinions and thinking about the project.

6. Funding For Development: In an earlier section, under Description, the funding picture was summarized. Here, the funding requirement is detailed in a matrix arrangement where it shows who is expected to pay for the development. Is another Federal agency going to help back the project? Are we expecting too much investment from industry? Are the projected funds really adequate for completing the project? What sort of potential cost overrun is possible? Answering these questions and displaying the data in a useful manner is a definite aid to reaching an overall category rating.

7. Non-Direct Financial Requirements: Government and private contractors may have to provide other support or aid to the project which is sometimes not "costed-out." For instance, there may be a need for government equipment and facilities, an interruption of production, and other time demands which could dwarf the direct funding requirement. Even if it is not that significant, decision makers should know what the total requirements are before supporting the project.

8. Systems Integration: Essentially there are two types of projects, new requirements and those previously identified. Questions in this area seek to determine how the particular project fits into the larger system or major thrust area's plans. The elements of substitution, completeness of other supporting projects and timing are emphasized.

9. Interlock with Other Thrust Areas. Although a project may be managed by one Focal Point manager, it may support other major thrust areas. Identification of those other areas is exhibited in this area via an interlock table which shows the criticality of the project to those other areas. A comparison of information in this section to that earlier displayed under the Support category provides a means for validating the interest-level rating.

10. Technical Success: Before an output or impact can be realized, regardless of the need, the project must be technically feasible. MANTECH projects are not basic research projects, rather they are developmental projects which carry technically feasible projects up to the point of showing that they can be implemented on a production basis. This section ensures that the evaluators consider such items as: the rationale for success; technical assumptions; views of proponents, moderates, and skeptics; past successes in this area; potential contractors; boundary conditions; need for more basic research; and the ability to overcome technical problems. In all, there are 20 elements in this category, some of which require establishing probability estimates.

11. Future Investment Requirements For Implementation: If project results are favorable, there may be costs on the part of the government, private activities, or perhaps both in order to achieve actual implementation. In some instances costs could be stagered over several years if results are implemented piecemeal or at multiple locations. Return On Investment is considered separately in a later section. The emphasis here is on defining the facility, equipment, material, personnel and overhead requirements for implementation, plus any other ownership considerations.

12. Federal Agency Implementation: There can be immediate potential users and future possible users of the project's results. This section discusses various aspects concerning the immediate potential users. In those instances where a Federal Agency(s) needs to fund at least a portion of the implementation: 1) Do they know the cost requirements? and 2) Can they accommodate the funding and lead-times? The bottomline centers on, will the project have a meaningful impact on a specific system's acquisition, manufacturing or repair cycles? Arranged in a tabular form, the evaluator can follow down through 13 data points to arrive at a final rating of this factor.

13. Industrial Implementation: For a large number of projects, private industry will be required to fund the implementation. Some evaluators may balk at this area and plead ignorance of what the private sector may or may not do. Without some understanding of the inclinations of the implementation thinkings by industry, the possibility exists that good project results may just "sit on the shelf." Information needs to be obtained on how industry perceives the future investment costs (Section XI) and the probability that funds will be available. Project results may support two or more different market sectors. Naturally for MANTECH projects, the obvious concern is focused on private investments to meet DOD/Federal requirements. However, potential for technology transfer to the commercial non-DOD marketplace can influence the project's chances for success and implementation to meet DOD requirements. Thus, intelligence on industries' perceptions of commerical applicability, market risk, total private developmental funding (independent of MANTECH), and estimates of transfer time are useful. Also, the results of the section can be compared to the Support ratings (Section V).

14. Benefits: In the Need category, (Section IV), the single most important driving element was identified. This section is structured so that the primary and secondary benefits can be rank ordered and benefit start-dates estimated. So as not to be overlooked, a critical question is asked: "What is the probability estimate that the government will incur lower procurement or other life cycle costs?" For some projects, cost is not a basic issue. However, for most projects, lower costs is at least a secondary benefit.

15. Global Relationships: This category includes 14 elements which draw attention to possible duplicating efforts, vertical integration, automation enhancements, sub-tier impact, political considerations and impact on competition, to name but a few. The thrust of this area is to stimulate a macro-perspective and an appreciation of the potential far-reaching implications of the project.

16. Return On Investment: The term, ROI, is gaining in popularity. The methodology for calculating ROI is relatively straight forward, but the estimates of future costs and savings can be quite inaccurate. This section is rather detailed; however, only applicable portions need be completed. Two paths through this section exist. The first deals with projects designed to improve or replace an existing/projected method, process or item. As with the next path, procurement cost, process cost and life cycle cost are candidate analysis items. When there is a basis of comparison, potential financial savings can be computed. On the other hand, in the second path where the project generates a new method, process or item and no basis of comparison exists, only costs can be projected. The section is well outlined for ease of completion and includes guidelines for reaching a final rating.

C. Project Evaluation Summary Chart: The final goal in this research effort was to summarize the factors into a comprehensible format for use by key decision makers. Various schemes are contained in the literature which attempt to weight factors and/or mathematically manipulate the data in numerous ways so as to derive a final project index number. (See "R&D Project Selection: An Annotated Bibliography," Daniel T. Fetsko, 1983 SCEEE SFRP Report.) Generally, these schemes are never really used or are quickly abandoned. With the risk that someone will again add up numbers, the final summary chart (Figure 1) contains a rating scale ranging from to +2.

These ratings or as some may call them, "Figures of Merit", are not intended at this stage of design to undergo any mathematical operation The appropriate blocks in the chart should be marked and presented to the key decision makers along with the first three sections of the evaluation Back-Up Sheets (Title, Objective and Description). The rest of the Back-Up Sheets support more detailed discussion. Final decisions on project selection and funding still rest with the key decision makers. With the suggested approach, their view can now be centered at least initially on assimilating the major factors, rather than being inundated with all the sub-elements.

PROJECT EVALUATION SUMMARY SHEET

	EVALUATION RATING:				
	-2	-1	0	+1	+2
1. Nature of Need					
2. Support					
3. Funding For Development					
4. Non-Direct Financial Requirements					
5. System Integration					
6. Interlock With Other Thrust Areas					
7. Technical Success					
8. Future Investment Requirements For Implementation					
9. Federal Agency Implementation					
10. Industrial Implementation					
11. Benefits					
12. Global Relationships					
13. Return on Investment					

FIGURE 1

V. CONCLUSIONS

MANTECH project selection and resource allocation decisions are complex mental processes involving numerous considerations. Without a systematic evaluation methodology, some considerations can be overlooked, while others may be overemphasized. The Project Evaluation Back-Up Sheets and the Summary Chart are management tools designed to support, not supplant, the intuitive judgement of the decision makers. The use of these evaluation tools can also assist in structuring discussion of candidate projects. Furthermore, these devices can be used to orient new employees regarding project assessment considerations. Lastly, they can also be applied to temper external pressures when there are obvious project shortcomings.

VI. RECOMMENDATIONS:

A. Copies of this report and the Project Evaluation Back-Up Sheets should be provided to interested DOD organizations

B. The products of this research effort need to be used with the view that they are preliminary documents. Local efforts should be made to modify these products as needed.

C. Recommend that an evaluation be conducted within one year to determine if this research was of value and if future refinement of the methodology is appropriate.

ATTACHMENT 1

EXTRACTED SECTION FROM MANTECH PROJECT EVALUATION BACK-UP SHEETS

IV. NATURE OF NEED:

A. Principal Sponsor: _____

B. Directed Project: _____ (Yes/No) Office: _____ (As Applicable) Date: _____

C. Source of Need Identification: (Specify)

_____ AF _____ Other DOD _____ Industry _____ Other U.S.

_____ Report _____ Conference _____ Other Date: _____

D. Required to meet:

	AF	DOD	Industry	Other U.S.
Current Need:	_____	_____	_____	_____
Future Need:	_____	_____	_____	_____

E. Applicability to Weapon System(s):

Is this project directly related to a weapon system? _____ Yes _____ No

Is the weapon system: _____ Current _____ Future

Name of System(s): _____ _____ _____

Projected Obsolescence Date: _____ _____ _____

F. Applicability to Support System(s):

Is project directly related to support system? _____ Yes _____ No

Is the support system: _____ Current _____ Future

Name of System(s): _____ _____ _____

Projected Obsolescence Date: _____ _____ _____

G. The Single Most Important driving factor for this project is:
Note: Other factors are listed in Section XIV, Benefits.

_____ System will not function without this project
_____ Counter a potential hostile threat
_____ Improve system performance
_____ Major focus of a thrust area
_____ User demand
_____ Basic technological problem preventing general use
_____ Speed up development for DOD use
_____ Lead time reduction
_____ Stimulate capital Investment
_____ Improve readiness
_____ Meet surge requirements
_____ Increase production throughput/yields
_____ Satisfy health requirement
_____ Improve safety features
_____ Minimize an environmental problem
_____ Achieve MILSPEC standards
_____ Reduce change-out problems
_____ Enhance sustainability
_____ Improve power efficiency
_____ Decrease energy consumption
_____ Improve engineering properties
_____ Gain an improved understanding of a production problem
_____ Develop reliable domestic sources
_____ Minimize the need for critical materials
_____ Reduce/eliminate foreign dependency
_____ Correct a spares problem
_____ External organizational pressure
_____ Achieve a high return on investment (ROI)
_____ Accomplish a key step in a thrust area
_____ Complete a thrust area
_____ Reduce maintenance/repair costs
_____ Reduce production costs in the area of:
  _____ direct labor
  _____ indirect labor
  _____ scrap rate
  _____ raw materials
  _____ capital investment
_____ Foster more rapid technology transfer
_____ Need to scale up
_____ Highly promising product/process
_____ Industry capability is low or lacking
_____ Correct obvious inefficiencies
_____ Improve productivity (specify:_____)
_____ Other (specify:_____)

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

AN ASSESSMENT OF WARTIME AVAILABILITY OF RECOVERABLE ITEMS

Prepared by:	E. Stanley Lee, Ph.D.
Academic Rank:	Professor
Department:	Industrial Engineering Department
University:	Kansas State University
Research Location:	Air Force Logistics Command, Directorate Of Plans and Programs, Management Science Department, Wright-Patterson Air Force Base, Ohio
USAF Research:	Mr. Victor Presutti
Date:	August 26, 1983
Contract No.:	F49620-82-C-0035

AN ASSESSMENT OF WARTIME AVAILABILITY OF RECOVERABLE ITEMS

by

E. Stanley Lee, Ph.D.

ABSTRACT

The most critical test of the adequacy of the logistics systems of the Air Force is during the worst scenario such as the initial war surge period or any other critical period when the demands far exceed the anticipated repair or transportation capabilities. In order to assess the war time readiness and sustainability and in order to obtain the best logistics policies during the worst scenario, the true dynamics during this period must be analyzed and accurately modeled based on war time emergency criteria and characteristics which are completely different from peace time.

The true dynamics with limited repair and transportation capabilities for recoverable items during the worst scenario are analyzed by the use of simple examples. Due to its simplicity in computation, the deterministic fluid approximation is analyzed and appears to be ideally suited to model the complicated and large Air Force logistics system during the worst scenario. To further up-grade the model to include the stochastic characteristics, diffusion approximation of the queueing theory can be used. Another possible approach is the formulation of new stochastic processes which may be easier to handle than the traditionally stochastic processes encountered in queueing theory.

ACKNOWLEDGMENT

The author would like to thank the Air Force Systems Command, the Air Force Logistics Command, the Air Force Office of Scientific Research, and the Southeastern Center for Electrical Engineering Education for providing me the opportunity to spend the summer in the XRX(MS) department of the Air Force Logistics Command at Wright-Patterson Air Force Base, Ohio. The experience was very educational and rewarding.

I would like to thank all of the people of XRX(MS) for their untiring help, hospitality, and friendship. In particular, I would like to thank the division chiefs, Mr. Jack Hill, Mr. Curt Neumann, and Mr. John Madden for their continued help through the summer. and Colonel Gerald F. Saxton for the very helpful briefings. I would also like to thank Mr. Victor Presutti for providing a good overview of the Logistics area, for providing a good working environment, and for his guidance and help throughout this summer experience.

I. INTRODUCTION

The most critical logistics problem in a war is during the initial surge period after D-Day. Both because of the sudden surge of war activities and the need of rapid transition from peace time to war time environment, the logistics policies and resource levels for sustaining a war are subjected to the most severe test during this surge period. Thus, the preparation of logistics support of a war should be primarily based upon this initial period which is characterized by:

(1) Because of the very high demands on repair facilities there is a high probability of a queue of repairable items waiting to be repaired.

(2) Because of the very large volume of items required to be transported, there is a high probability of a queue waiting for transportation.

(3) Because of the war urgency, cannibalization and lateral resupply may become an important part of the logistics system.

In addition to the above characteristics, the recoverable items are generally multi-indenture in structure and multi-echelon in repairable stations. Thus, in order to assess the war time materiel readiness and sustainability, all these factors based on the rapidly changing war dynamics must be analyzed. These assessments are important not only for fighting and sustaining a war, they also allow the Air Force to determine, for a given level of support, where funds would best be spent in providing logistics support such that the Air Force's ability to fight a war is maximized.

The various important factors for the general resupply process for recoverable items during the initial surge of war are illustrated in Figure 1. A two echelon system with base and depot repair is assumed. Because of the limited facilities in repair and transportation, tandem queues or queueing networks are needed for the two echelon repair stations and transportation queues are assumed between the base and the depot.

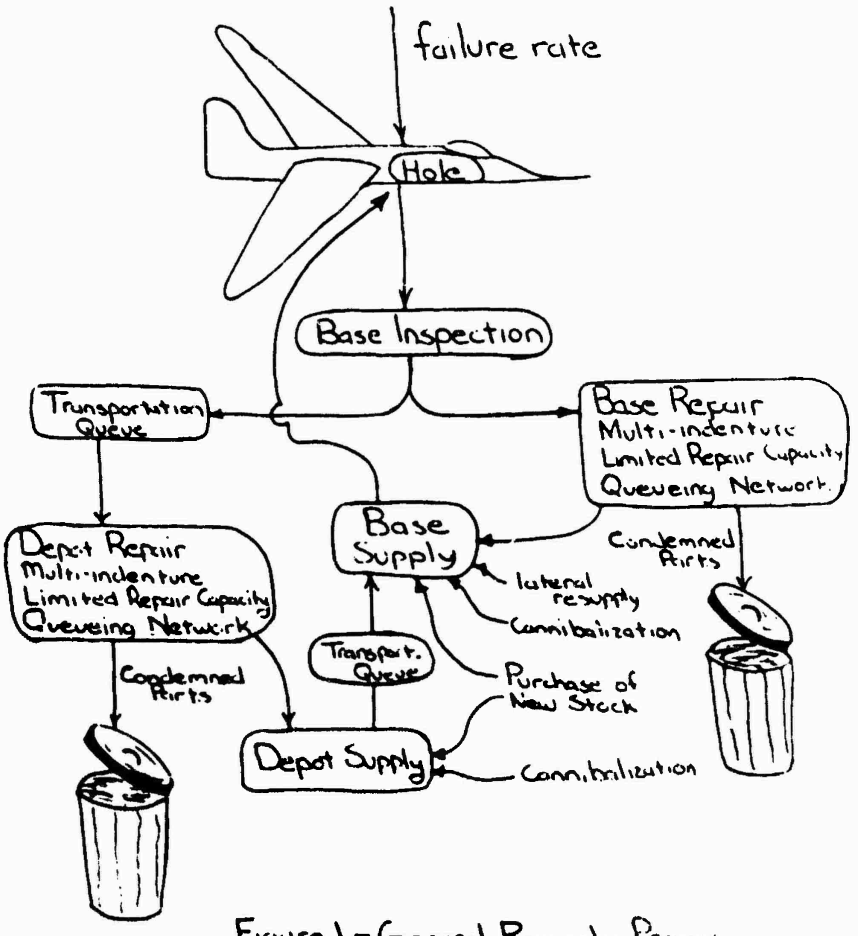


Figure 1 - General Resupply Process During the Initial Surge of a War

Models. Two basic approaches have been used to analyze the logistics of recoverable items: (1) Monte Carlo simulation models, and (2) analytical models. Monte Carlo simulation models emulate a real world process through the use of probabilistic descriptions of process activities and events. Although his type of model can represent the process in great detail, it suffers from the enormous computation requirements. The true dynamics of the process with the consideration of queue has been modeled successfully in JEMS (Jet Engine Management Simulator) [13].

Analytical models can abstract the essentials of the logistics system and represent it mathematically. Thus, it can be used to handle large complex systems such as the logistics of recoverable items. However, in order to consider the true dynamics of repairable items during the initial war surge, the three characteristics listed above must be included. But, unfortunately, very little has been done to include these characteristics in the various existing analytical models such as Dyna-METRIC, LMI availability model, and WARS [3-10]. One reason is because the extreme difficulty encountered in solving the dynamic queueing equations. Notice that both of the first two charateristics require the solution of this type of equations.

Although the general dynamic queueing equations are very difficult to handle, the dynamic equations representing temporary over capacity where the arrival or demand rate temporarily exceeds the service rates can be handled reasonably easily. The reason for this is that we are not concerned over whether there is a queue, we are only concerned with the queue length. Because of the large number of arrivals and the certainty that there is a queue, the system behaves like fluid flow with diffusion. Thus, the fluid and diffusion approximations of queues can be used to handle the limited repair and transportation facilities during the initial war surge.

In this report, these approaches will be explored. Some simple examples will be given to explore the fluid approximation approach which is a deterministic approach. The stochastic diffusion approach of queues and its possible applications to the war surge period will also be examined.

Another approach to avoid the dynamic queueing equation is to reformulate the problem in terms of simpler stochastic process. Several approaches have been proposed in the literature. These approach also will be examined briefly for its suitability to model the Air Force logistics systems.

One of the most effective ways to represent stochastic processes is the Markov process. This approach combined with basic reliability analysis are also examined for the study of cannibalization in a separate report [17]. It should be noted that Markov decision process is ideally suited for treating the opportunistic engine maintenance

problem [18].

II OBJECTIVES OF THE RESEARCH EFFORT

The objectives of this research were:

(1) To study and to analyze the war dynamics during the initial surge period, or any critical period where the demands exceed repair capabilities, for recoverable items. The initial surge period after D-Day is the most critical portion of a war and its time dynamics are the most important factors in assessing logistics support capabilities. The influences of some of the assumptions in existing analytical models such as Dyna-METRIC on the war dynamics during this surge will be examined.

(2) To explore modeling methodologies which can be used to model the true fluctuating dynamics with limited repair and transport capabilities during the surge period for multi-echelon and multi-indenture recoverable items.

(3) To assess the effectiveness and applicability of these various methodologies for modeling recoverable items and for estimating war time materiel readiness and sustainability during the surge period. Because of the limited time and also because of the extreme complexity of the Air Force logistics systems, only an initial start can be made. More detailed study and modeling must be carried out during later investigations.

III DETERMINISTIC FLUID APPROXIMATION

During the surge of a war where the expected number of units of demand is very high, the queueing system behaves like a fluid and thus it is called the (deterministic) fluid approximation [14, 15]. Since there definitely has a queue and we are only concerned with the surge of queue length with a fixed and constant repair rate, the stochastic aspects such as whether or not there is a queue is no longer important. Therefore, deterministic approximation is ideally suited to model this

process.

To illustrate the true dynamics of a war during the initial surge after D-day, consider a very simplified example of a single base without depot support. In other words, a one-echelon repair system with no indenture.

During peace time, the base has an unlimited repair capacity and all repairs are performed at the base. Suppose that an item has a fixed daily demand rate of 0.8 units and a fixed base repair time of 5 days. Assume that the initial surge of a war lasts 15 days. The daily demand rate during this surge period is 3 units per day. After the first 15 days of the war, the demand rate gradually decreases following the decrease in flying activity which has the form:

$$\text{Demand} = \lambda(t) = 3.16e^{-0.1t} \tag{1}$$

This decrease in flying activity was used by Muckstadt [1] to demonstrate the differences between steady state and dynamic models. Thus, the expected demand rate in the second column of Table 1 is obtained by using:

$$d_k = 0.8, \qquad k < 1, \quad k > 28$$

$$d_k = 3.0, \qquad 1 \le k \le 14$$

$$d_k = \int_{k-15}^{k-14} \lambda(t)\, dt, \qquad k = 15,16,\ldots,28 \tag{2}$$

where d_k is the expected demand rate at day k.

It should be noted that these equations for daily demand are arbitrarily assumed. The daily demand rates are also shown in Figure 3. Since the repair time is 5 days and the repair capacity is assumed unlimited during both peace and war times, the third column in Table 1 is the summation of the demands of the previous 5 days. Assuming Poisson arrival, the last column in Table 1 is the smallest nonnegative integer, N_k, such that

TABLE 1

UNLIMITED REPAIR CAPACITY, 15 DAYS INTENSIVE WAR

Day	Expected Demand	Expected no. of units in resupply	Stock level required to achieve 0.8 probability of no back orders
0	0.8	4.0	6
1	3.0	6.2	8
2	3.0	8.4	11
3	3.0	10.6	13
4	3.0	12.8	16
5	3.0	15	18
10	3.0	15	18
15	3.0	15	18
16	2.7	14.7	18
17	2.5	14.2	17
18	2.2	13.4	17
19	2.0	12.4	15
20	1.8	11.2	14
21	1.7	10.2	13
22	1.5	9.2	12
23	1.4	8.4	11
24	1.2	7.6	10
25	1.1	6.9	9
26	1.0	6.2	8
27	0.9	5.6	8
28	0.8	5.0	7
29	0.8	4.6	6
30	0.8	4.3	6

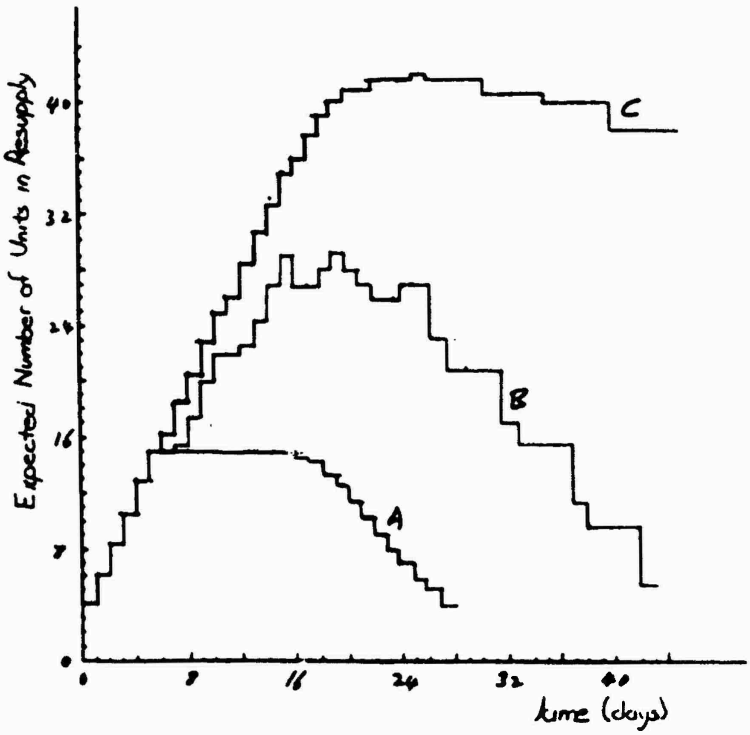


Figure 2 - Expected Number of Units in Resupply

Curve A: Unlimited Repair Capability
Curve B: Peacetime - 5 servers
Day 1 to Day 10 - 8 servers
Day 11 and later - 10 servers
Curve C: 5 servers

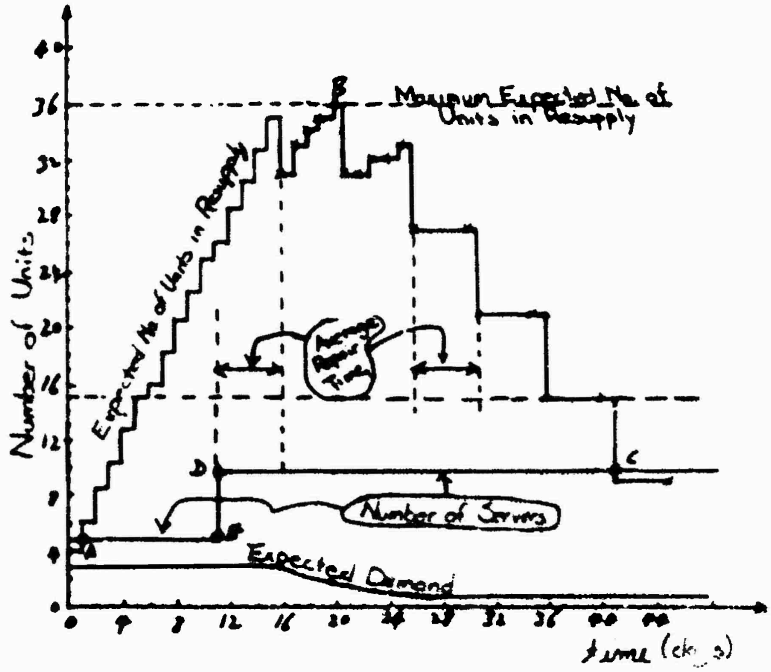


Figure 3 - Expected Number of Units in Resupply for 15 days Intensive War with Repair Capability doubled after 10 days

$$\sum_{s=0}^{N_k} \frac{e^{-M_k}(M_k)^s}{s!} \geq 0.8 \qquad (3)$$

where M_k is the number of units in resupply during the kth day. In other words, the last column of Table 1 represents the stock level on each day such that the probability of having one or more backorders is no greater than 0.2. This is the policy the Air Force uses to compute the stock level for spare aircraft engines [2]. Observe that the peak requirement of 18 units occurs on days 5 through 16. (See also Figure 2 where the third column in Table 1 is shown as Curve A.) The process returns to the minimum or peace time stock requirement of 6 on the 29th day after D-day.

Although Table 1 shows the dynamics of initial surge of a war, it does not show the true dynamics of demand due to one important unrealistic assumption, namely: the base has unlimited repair capacity. Since the war time demand is asssumed 3.0/0.8 = 3.75 times higher than peace time demand, it would be very unrealistic and very expensive to keep a repair facility with repair capability 3.75 times higher than required during peace time. Thus, in nearly all practical situations, the repair capacity is certainly limited during the initial surge of a war. Even if we assume a restricted war limited to a small number of bases, we still need time to increase the repair capacity by moving equipment and technical personnel from other bases to the bases concerned.

Suppose that during peace time only 80% of the repair capacity is utilized in this single base and 20% of the repair capacity remains idle. Since the item requires 5 days to repair and the peace time demand is 0.8 unit per day, the base must have 5 simultaneous repair facilities for this item. Table 2 is obtained by assuming 5 servers or 5 repair facilities in the base. The expected demand is the same as in Table 1 and is obtained by using the equations in (2). Column 3 shows the number of units in resupply which are in repair and Column 4 lists the number of units in resupply which are waiting to be repaired because of limited service capacity. The total

TABLE 2

LIMITED REPAIR CAPACITY, 15 DAYS INTENSIVE WAR

Day	Expected Demand	No. in service	No.in queue	Expected no. of units in resupply	Stock level required to to achieve 0.8 probability of no back orders*
0	0.8	4.0	0.0	4.0	6
1	3.0	5.0	1.2	6.2	8
2	3.0	5.0	3.4	8.4	11
5	3.0	5.0	10.0	15.0	18
6	3.0	5.0	11.2	16.2	20
10	3.0	5.0	20.0	25.0	29
15	3.0	5.0	30.0	35.0	40
20	1.8	5.0	36.2	41.2	47
25	1.1	5.0	38.1	43.1	49
30	0.8	5.0	37.4	42.4	48
35	0.8	5.0	36.4	41.4	47
51	0.8	5.0	32.4	37.4	43
101	0.8	5.0	22.4	27.4	32
201	0.8	5.0	2.4	7.4	10

* - Assumed Poisson distribution. This assumption is not completely correct.

number of units in resupply is the summation of Columns 3 and 4. The last column in Table 2 is obtained in the same way as those obtained in Table 1 by using Equation (3). However, for large values of the mean, M_k, the cumulative sums for Poisson distribution can be approximated by normal distribution. Thus, for $M_k > 20$, cumulative sums of normal distribution are used.

It should be emphasized that the last column in Table 2 is not correct. This is because Poisson distribution is assumed. Thus, the last column in Table 2 should only be used with great care.

Notice that, except in day zero, all the servers are busy in Table 2 for all the time until day 201. (See also Figure 2 where the fifth column in Table 2 is shown as Curve C) The peak requirement of 49 units occurs around day 25. The process did not return to the peace time stock requirement even on day 201.

If the required stock level is obtained based on unlimited repair capacity as those listed in Table 1, the stock level would have understated starting on the 6th day of war and lasting over 100 days. Furthermore, the actual maximum requirements would be understated by 49 - 18 = 31 units which is 172 percent understocked.

Suppose this is a limited war and the repair capacity at the base can be increased by moving repair equipment and repair technical personnel to the base fairly quickly. Furthermore, suppose that this movement required ten days. Let us assume that the repair capacity is doubled at the 11th day of war. In other words, we have ten servers starting on the 11th day. Using the same demand scenario, Table 3 is obtained. Observe that the maximum expected number of units in resupply (MENUR) is 36.2 and this peak demand occurs at around day 15 to day 20 (see also Figure 3). Compared to Table 1, the MENUR is understated by 36.2 - 15 = 21.2 units. This duration of understock started on day 6 and lasted until day 36.

Observe that by doubling the service capacity at the 11th day of war, the MENUR is reduced by 7 units or only approximately 20%. However, the duration of shortage is considerably reduced.

Another way to increase the repair capacity is to increase or add overtime for the workers. Obviously, there is a limit for this overtime because of repair and testing equipment and manpower limitations. Let us assume that on D-day we can increase our number of servers from 5 to 8 immediately by overtime and after 10 days of war we can add repair and testing equipments and manpower to 10 servers. The results in Table 4 (see also Figure 2) are obtained under these service capability assumptions with the same demand scenario. Notice that except peace time and the first day of war, all the servers are busy until the 36th day of war. Compared to

Table 1, the MENUR is understated by a maximum of 30.2 - 15 = 15.2 units.

TABLE 3

Repair capacity doubled after ten days, 15 days intensive war.

Day	Expected Demand	No. in service	No. in queue	Expected no. of units in resupply
0	0.8	4.0	0.0	4.0
[illegible]	3.0	5.0	1.2	6.2
5	3.0	5.0	10.0	15.0
10	3.0	5.0	20.0	25.0
11	3.0	10	16.2	26.2
15	3.0	10	25.0	35.0
20	1.8	10	26.2	36.2
21	1.7	10	21.1	31.1
31	0.8	10	11.4	21.4
41	0.8	9.4	0.0	9.4
46	0.8	3.4	0.0	3.4

If the required stock is obtained by assuming no repair capacity limitation, the duration of understock which would limit the war fighting capability starts on day 6 and lasts until day 35. In other words, our war fighting capability would be greatly reduced during most of the time of the war.

Thus, the dynamics of war can be influenced greatly by repair limitations. The expected number of units in resupply due to both in service and waiting to be served from Tables 1, 2 and 4 are plotted in Figure 2. Unlimited repair capacity assumption can cause the required stock level greatly underestimated even under the most optimistic

TABLE 4

Five servers in peace time, 8 servers during day 1 to day 10 and 10 servers after day 10.

Day	Expected Demand	No. in service	No. in queue	Total Expected no. of units in resupply
0	0.8	4.0	0.0	4.0
1	3.0	6.2	0.0	6.2
2	3.0	8.0	0.4	8.4
5	3.0	8.0	7.0	15.0
10	3.0	8.0	14.0	22.0
11	3.0	10.0	12.0	22.0
15	3.0	10.0	19.0	29.0
20	1.8	10.0	20.2	30.2
25	1.1	10.0	17.3	27.3
30	0.8	10.0	11.6	21.6
35	0.8	10.0	5.6	15.6
36	0.8	10.0	1.4	11.4
37	0.8	9.6	0	9.6
40	0.8	9.6	0.0	9.6
41	0.8	5.4	0.0	5.4

assumptions. The dynamics of war not only depends on the change in demand rate, it also depends heavily on the repair capacity of the system. Since nearly all of the currently available analytical models such as the dyna-METRIC and the LMI availability model do not consider the repair capability limitations, it is important to add this aspect into the models. This is especially true in view of the fact that because of the severe logistics demand during the initial surge of war, the required stock level is very sensitive to the intensity and duration of war, the repair capacity or limitations, the transportation capacity

and how quickly we can mobilize to wartime repair and transportation capabilities. Because of this severe demand on the various resources, a detailed study of the true dynamics of war is a necessity in order to prepare for war adequately during peace time.

It should be mentioned that Monte Carlo simulation models such as JEMS and TJEMS do consider repair capacity limitations [13].

To illustrate the influence of repair capacity on the dynamics of war, the results in Table 3 is plotted in Figure 3. The difference between the expected number of units in resupply and the number of servers in any given day gives the queue length or the number waiting for service. It is this queue length or the area inside the curves formed by the points ABCDE that is important from logistics standpoint. The maximum height of the curve at point B represents the time at which the maximum queue or maximum number of units in resupply which is not available for use and the duration from point A to point C represents the duration at which there are units waiting to be repaired or the duration during which the repair capacity is not adequate. Thus, both the height and the width (or duration) are important parameters. For example, if we assume the system fails any time there are 15 units or more in resupply, then, knowing the duration or width of the area ABCDE, one can estimate the mean time between system failures (MTBSF). For the scenario in Figure 3, this MTBSF is 36 - 5 = 31 days which is obviously too long for a war of only 15 days duration. The maximum height of the area can be used to estimate the maximum number of units which are in resupply. Notice that because of the fixed repair time of 5 days, there is always a delay of at least 5 days during which time the unit is not available. It is this delay that makes the dynamic so severe.

This type of analysis can be used to illustrate the influences or sensitivity on the dynamics of war under various different conditions such as the length of repair time, transportation time, the duration and intensity of war, repair capacity, and the various delays due to limitations on repair capacity, transportation capacity and shortage of parts.

IV. MULTI-ECHELON AND MULTI-INDENTURE SYSTEM

The above illustration was purposely made simple. The actual problem is complicated by both the structure of the system and the repair procedures. An airplane is structured multi-indenture. The repair operation in the Air Force logistics systems is multi-echelon in location and multi-station at a given location for any given repairable item.

Even for this complicated system, it is expected that the above procedure can still be used to study the dynamics of the initial war surge. This is because the basic simplicity in the approach and in the computation. However, it should be emphasized that only during the war surge period where demand rate is much higher than repair capabilities, this approach can be used. If the demand or failure rate is relatively low, the probability of formulating queue or no queue becomes important and the deterministic assumption can no longer be used.

Another disadvantage of this approach is that because this is a deterministic approach, all the stochastic properties are lost. For example, in obtaining the last column in Table 1, Poisson distribution can be assumed because of unlimited repair capacity. However, the last column in Table 2 is not completely correct. This is because of limited repair capacity with queue and thus the distribution of the expected number in resupply is unknown. This is why in Tables 2-4, only the expected numbers of units in resupply is used. The stock level required to achieve 0.8 probability of no back orders is unknown. One way to overcome this disadvantage is by the use of diffusion approximation which will be discussed in the next section. However, by the use of this approximation, the computational process becomes more complicated as compared to the simple procedure discussed in the previous section.

To illustrate the multi-station nature of the repair process, consider the "raw engine build-up" procedure after the base receives the engine from the depot [13]. The flow chart for this multi-station repair process is shown in Figure 4 [same as Reference 13, Chart 6]. As can be seen, a series of queues can be formed. A queue may be formed to wait for the stand, as a second queue may be formed to wait for crew, a third queue for parts if parts shortage occurs, and a fourth queue for

the test cell. Thus, instead of a simple queuing process as that discussed in the above illustrations, we have a series of four queues in tandem. The distribution of the output from this series of queues is obviously much more difficult to obtain. However, if we only consider the initial surge period, the dynamics of this tandem queue can still be obtained in essentially the same way as that discussed in the illustrations by assuming constant service times and by using deterministic approximations. The only difference from Tables 2 through 4 are that there are more than one queues waiting to be served, more than one service station in series, and the total expected number of units in resupply would be the sum of all those in services and in queues.

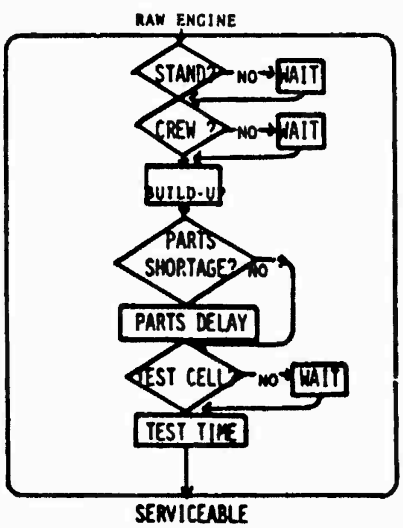


Figure 4
R.W Engine Built-up Procedure

To add the multi-indenture and multi-echelon property, the system is even more complicated. For example, in order to repair component j, sub-component i may be needed. But, sub-component i may be back-ordered and is not available. Thus, another queue can be formed for this sub-component. It is proposed that the extension of the above procedure to the multi-indenture and multi-echelon system be investigated during the mini-grant program.

V. DIFFUSION APPROXIMATION

Although the deterministic fluid approximation discussed in the previous sections is simple to calculate and appears to be ideally suited for study the dynamics of the initial war surge, the stochastic property of the system is completely lost. For example, we can calculate the mean of the back orders for certain item, but, we don't know the distributions of this back order. Obviously, it is desirable to know the distribution also.

Since the dynamic queuing equations with variable arrival rates or variable demand rates are too complicated to be of practical use, various other approaches to consider the stochastic property of the system have been devised. Two approaches which appear especially suited to model the Air Force logistics systems will be discussed briefly in this and the next section.

The diffusion approach [14-16] is based on the idea that the number of units in the system is so large that the integer nature of the count can be ignored. During the initial war surge period, this is exactly the situation where the number of units in resupply waiting to be served is large.

The behavior of queuing system is abstractly very similar to the behavior of fluid diffusion. In queueing theory, the first moment of the probability density is the analogue of the center of gravity of the mass distribution and the variance is the analogue of the moment of inertia about the center of gravity. Based on mass balance of fluid flow with diffusion and after some manipulations, the following differential equation representing the probability distribution function of the queue of an item with time-dependent arrival rate can be obtained

$$\frac{\partial}{\partial t}F(\ell,t)=\left[(\mu-\lambda(t))\frac{\partial}{\partial \ell}+\frac{1}{2}(I_A\lambda(t)+I_D\mu)\frac{\partial^2}{\partial \ell^2}\right]F(\ell,t) \quad (4)$$

where $F(\ell,t)$ is the probability distribution function of the queue at time t, $\lambda(t)$ is the time-dependent arrival rate at time t, μ is the time-independent or constant service rate, ℓ is the queue length, I_A is the variance to mean ratio for the arrival rate, and I_D is the variance to mean ratio for the departure or service rate.

Although Equation (4) is a very complicated partial differential equation, it is much easier to handle than the differential-difference classical dynamic queueing equation. Furthermore, since this is basically the same equation in fluid diffusion and heat transfer, it has been solved under many different assumptions in the literature.

Equation (4) can be simplified by the following transformation of the time coordinate:

$$t' = \int_{i}^{t} [I_A \lambda(\tau) + I_D] d\tau \qquad (5)$$

Equation (4) becomes:

$$\frac{\partial}{\partial t'} F'(\ell, t') = \left[\xi(t') \frac{\partial}{\partial \ell} + \frac{1}{2} \frac{\partial^2}{\partial \ell^2} \right] F'(\ell, t') \qquad (6)$$

with

$$\xi(t') = \frac{1 - \lambda(t(t'))/\mu}{I_A \lambda(t(t'))/\mu + I_D} \qquad (7)$$

Since $F'(\ell, t')$ is a distribution function, the boundary condition for Equation (6) must be

$$F'(0, t') = 0, \qquad F'(\infty, t') = 1 \qquad (8)$$

Once the arrival or demand rate, $\lambda(t)$, is given for a given problem, Equation (6) can be solved numerically with Equation (8) as the boundary conditions for the distribution function $F'(\ell, t')$.

Notice that Equation (6) only represents the queue of one item at one repair station. Thus, if the arrival or demand rate is different for different items, Equation (6) must be solved for each different arrival rates. Based on this distribution for every different arrival rate, a model can be established. This model should be established along the same lines as Dyna-METRIC except that repair capacity is limited.

Obviously, a lot of work needs to be done to establish this model. It is hoped that this approach can be further explored in the Mini-grant program.

VI. THE FORMULATION OF NEW STOCHASTIC PROCESS

Because of the simultaneous consideration of the interaction of two nonstationary stochastic variables (the arrival and departure stochastic variables), the dynamic queueing approach is very complicated. If one could extract a simpler stochastic process from the system based on the particular characteristics or assumptions of a given process, the problem may be easier to handle. Several different approaches, such as the busy and idle periods concept [15] and the MTBSF idea [12] are proposed in this direction. The advantages of these approaches, are that they are not limited to the war surge period. Due to space limitations, this possible approach will only be discussed briefly here.

As discussed in the previous sections, one of the principal difficulties in using existing analytical models to model the dynamics of the war surge is the assumption of unlimited repair capacity. This assumption can cause considerable under estimation of the required stock level. From Figure 3, it can be seen that the duration from A to C is an important variable. This duration depends on the relative rates of demand and repair and this is a partial measure of the repair limitation. The MTBSF defined in Section III represents this duration. For example, if we assume the system fails at the following different numbers of units in resupply: 10, 15, 25, 30, etc., the corresponding MTBSF's are 38, 31, 21, 13, etc. Thus, the area AEDCB can be obtained indirectly by using different values of MTBSF's.

In the LMI availability model or the recent version of Dyna-METRIC, the problem is:

$$\begin{aligned} &\text{Maximize:} && \text{(System Availability)} \\ &\text{Subject to:} && \text{(Total Logistic System Cost)} \leq C \end{aligned} \qquad (9)$$

However, the MTBSF is also an important variable. It measures the duraion during which the war fighting capacity is greatly limited. Thus, instead of Equation (9) if we solve the problem:

$$\begin{aligned} &\text{Maximize:} && \text{(System Availability)} \\ &\text{Subject to:} && \text{(Total Logistic System Cost)} \leq C \\ & && \text{(MTBSF)} \geq T \end{aligned} \tag{10}$$

The dynamics of war is also indirectly and partially considered. The reason that this is only a partial consideration is that the maximum queue is not considered. Once the queue is over certain length, the system is considered failed. It should be emphasized that this is a practical and useful approach for the Air Force logistics systems.

Graves and Keilson [12] proposed this approach which is similar to that of Goss et al [11]. The main advantage of this approach is that the stochastic variable, MTBSF, is much easier to describe than the stochastic variables traditionaly used in queueing theory. These authors reduced Equation (10) into a zero-one integer programming problem which can be solved approximately by linear programming. This approach appears promising although some computational difficulties must be first overcome.

VII. RECOMMENDATIONS

The deterministic fluid approximation has shown to be an effective tool for studying the logistics for recoverable items during the initial war surge or during any critical period where the demands extends the repair capacity. However, because of the complexity of the Air Force logistics systems and also because of the limited time during the summer research program, only a start has been made to investigate the logistics of this critical war period. Several areas need further research. The most fruitful areas for further investigations are:

(1) Expand the deterministic fluid approximation into one mathematical analytical model for the thousands of recoverable items in the Air Force logistics system. This model should at least include the multi-indenture, multi-echelon and multi-station repair features with limited transportation and repair capabilities. Cannibalization and lateral resupply will also be included. Because of this limited transportation and repair capacities, the model is much more complicated

due to the presence of various queues. However, because of the simplicity of the approach as shown in this report, the model should be within the available computer capacities.

(2) Expand the analytical model obtained above to cover the multi-indenture in structure and multi-echelon in storage consumable items with emphasis on the limited transportation capabilities.

(3) In order to further expand the above model to include more accurate stochastic distributions concerning the number of units in the repair pipeline, both the diffusion approximation and the formulation of new stochastic processes should be further investigated. Both approaches appear to be an effective technique at this stage of the research. However, the computational effort involved must be considerably reduced before it can be applied to the Air Force logistics systems. These approaches can be applied right now to a small sub-system of the Air Force logistics systems such as a module or an engine with only few most important modules.

I purpose to continue the investigation based on the above recommendations with emphasis on the particular characteristics during a war emergency such as the limited repair and transportation capabilities for the multi-echelon and multi-indenture recoverable items, the limited transportation and storage capabilities for consumable items, and the importance of cannibalization and lateral resupply.

REFERENCES

1. Muckstadt, J. A., "Comparative Adequacy of Steady-State versus Dynamic Models for Calculating Stockage Requirements," The Rand Corporation, R-2636-AF, November, 1980.
2. "Selective Management of Propulsion Units," AFLC Supplement 1 to AFM 400-1, Oct., 1978.
3. Sherbrooke, Craig C., "METRIC: A Multi-Echelon Technique for Recoverable Item Control," Operations Research, 16, 122 (1968), also The Rand Corp. RM-5078-PR, Nov., 1966.
4. Muckstadt, J. A., "A Model for a Multi-Item, Multi-Echelon Inventory System," Management Science, 20, 472 (1973).

5. Hillestad, R. J., "Dyna-METRIC: Dynamic Multi-Echelon Technique for Recoverable Item Control," The Rand Corp, R-2785-AF, July, 1982.
6. Demmy, W. S. and Presutti, Victor J., "Multi-echelon Inventory Theory in the Air Force Logistics Command," TIMS Studies in the Management Sciences, 16, 279 (1981).
7. Hill, J. M., "Marginal Analysis at AFLC for Recoverable Items," Directorate of Management Science, Report AFLC/XRX, April, 1978.
8. Pyles, Raymond, "The Dyna-metric Readiness Assessment Model: Motivation, Capabilities, and Uses," The Rand Corp., WD-892-AF, Feb. 1982.
9. Hillestad, R. J., and M. J. Carrillo, "Models and Techniques for Recoverable Item Stackage When Demand and the Repair Process are Nonstationary--Part I, Performance Measurement," The Rand Corp., N-1482-AF, May, 1980.
10. Bigelow, J. H., and K. Isaacson, "Models to Assess the Peacetime Material Readiness and Wartime Sustainability of U. S. Air Forces: A Progress Report," The Rand Corp., N-1896-MRAL, Oct., 1982.
11. Gross, D., H. D. Kahn, and J. D. Marsh, "Queueing Models for spares provisioning," Naval Research Logistics Quarterly, 24, 521 (1977).
12. Graves, S. C., and J. Keilson, "System Balance for Extended Logistic Systems," Operations Research, 31, 234 (1983).
13. Novak, R. A., and J. L. Madden, "TJEMS, Transport Jet Engine Management Simulator," Working Notes XRS-79-015, AFLC, DCS/Planes and Programs, September, 1982.
14. Newell, G. F., Application of Queueing Theory, 2nd edition, Chapman and Hall, 1982.
15. Kleinrock, L., Queueing Systems, Volume 2, John Wiley (1976).
16. Lee, E. S., and Richard Bellman, "Invariant Imbedding in the Diffusion Approximation of Queueing Theory," to be published.
17. Kramer, Bradley A., "Toward the Development of a Cannibalization Policy via Markov Chain Computations," 1983. USAF-SCEEE Graduate Student Program Report.
18. Williamson, V. L., and J. L. Madden, "Generalized Opportunistic Maintenance Engine Simulation," working paper XRS-81-121, AFLC/XRS.

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

AND GRADUATE STUDENT SUMMER SUPPORT PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

LASER DENSITOMETER DESIGN

Prepared by:	Dr. Mark E. Lewittes
Academic Rank:	Assistant Professor
Department and University:	Electrical Engineering Univ of Texas at San Antonio
and	Joanne M. Foley, BS Graduate Student
Department and University:	Biomedical Engineering Drexel University
Research Location:	USAF School of Aerospace Medicine Brooks AFB, Radiation Sciences Division, Vulnerability Assessment Branch
USAF Research	Major Dennis A. Maier Lt Byron P. Edmonds
Date:	August 15, 1983
Contract No:	F49620-82-C-0035

LASER DENSITOMETER DESIGN

by

Joanne M. Foley
and
Mark E. Lewittes

ABSTRACT

A design for an automated laser densitometer is presented. The design allows for the measurement of the optical density of samples with high attenuation factors. Such materials are used for or have the potential for use as laser protective eyewear.

The densitometer is designed to measure optical densities as high as 8±.1 OD. Both pulsed and continuous laser sources are incorporated in the densitometer. Wavelengths in the range 488 nm to 1064 nm are available. A microcomputer provides the automation of the laser densitometer, prompting the user through the measurement procedure, making the measurements, and ultimately performing the calculations to determine the optical density of the sample.

Acknowledgement

The authors would like to acknowledge the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for providing the research opportunity at the USAF School of Aerospace Medicine, Brooks AFB, Texas.

We appreciate the cooperation from the Radiation Sciences Division and Vulnerability Assessment Branch for providing a supportive working environment.

We offer special thanks to Lieutenant B. Edmonds who collaborated with us on every aspect of this project, and Major D. Maier whose illuminating discussions focused the project goals.

I. INTRODUCTION

The task of developing laser protective eyewear is especially difficult. This is due to the availability of many types of lasers with output wavelengths covering the UV, Visible and IR spectrum. No single protective filter can function at all wavelengths. Thus, the problem of evaluating protective eyewear materials is important for determining the best materials for any particular environment. The Air Force has had a continuing program for determining the suitability of laser protective eyewear. (1-5)

Much of the previous work used laser densitometry to measure the transmission of optically dense samples. (3-5) Commercially available spectrophotometers such as the Beckman ACTA MVII are capable of measuring a change in transmission of three orders of magnitude. A sample which can attenuate light by three orders of magnitude is said to have an optical density of three. More generally optical density (O.D.) is related to transmission (T) by the relation

$$OD = LOG_{10}\frac{1}{T}. \qquad (1)$$

The laser densitometer used by Murray (4) was capable of measuring the optical density of very dense samples. It employed a nitrogen pumped pulsed dye laser. The pulse widths of these lasers were less than 6 nano-seconds. The method employed required that this short optical pulse be split into two optical paths, one path being longer than the other. The light from both paths was detected by the same photomultiplier tube. Since one path was longer, the resulting optical

delay allowed the signals from the two paths to be observed as two sequential pulses on an oscilloscope. The optical sample being tested was placed in one path which changed the relative size of the two pulses. Manually changing the filters in the sample path allowed the experimenter to obtain a near balance of the pulse signals. The optical density of the sample was then calculated from the optical density of the filters removed from the sample path and the ratio of the sample and reference pulse amplitudes.

While this laser densitometer was effective it had two drawbacks. First, it required the use of a very short pulsed laser and would not function with either a long pulse or a continuous laser. Second, the manual filter adjustments were cumbersome. The accuracy of the optical density measurements depended on the skill of the operator in reading pulse heights on the oscilloscope and his willingness to perform the task.

Subsequent to this work, Major D. Maier of USAFSAM/RZ Laser Laboratory, constructed a prototype laser densitometer. This eliminated the need for an ultra short pulsed laser by detecting the reference beam with a second optical detector. This laser densitometer still required the operator to manually move neutral density filters into or out of the optical path. The fundamental concern of the present work is to design an automated laser densitometer capable of using either pulsed or continuous lasers as a light source.

II. OBJECTIVES

The design objectives of the laser densitometer were as follows:

(1) To have the measurement of optical density be independent of laser intensity instability.

(2) To measure optical density up to 8 OD with an accuracy better than .1 OD at wavelengths of common laser sources, and at different polarizations of the source.

(3) To automate the system and provide user interaction using a microcomputer.

(4) To accommodate the various laser types which differ in waveform, wavelength and power in a single experimental arrangement, avoiding realignment or replacement of optics.

(5) To select and order the appropriate electronic and optical components.

III. System Overview

As discussed above one of our design goals was to measure the transmission of samples over 8 orders of magnitude. This is difficult in a single beam absorption experiment because even the best optical detectors are only linear over six orders of magnitude in intensity.

By insuring that the intensity which impinges on the optical detector is independent of the optical density of the sample, the difficulty of nonlinearity can be overcome. This is realized by using a feedback system which alters the intensity impinging on the optical sample such that the transmitted light is a constant and thus independent of the sample's optical density. The measurement is made on the basis of how much the light intensity has to be altered.

A diagram of our implementation of this method is shown in Figure 1. The laser light source represents one of several lasers available as sources in this system. The laser beam itself is split into two beams by a beam splitter. The left optical path in Fig. 1 is considered the reference path. It provides a signal to the feedback system which is proportional to the instantaneous laser intensity. The right path is the sample path.

With the sample temporarily out of the optical path, and filters summing to an optical density of 8 in the sample path, the electronic signal generated by the beam in the sample path is matched to the voltage of the reference path's signal. When the sample is placed in the sample beam path the two signals are no longer equal in magnitude. The microcomputer provides the feedback mechanism in this system. It drives the linear actuators (stepping motors) to move the appropriate filters out of the sample path until the two beams are again equal. Since optical density is additive, the total optical density of the filters which were removed from the path is equal to the optical density of the sample itself. This value is displayed on the CRT to the operator.

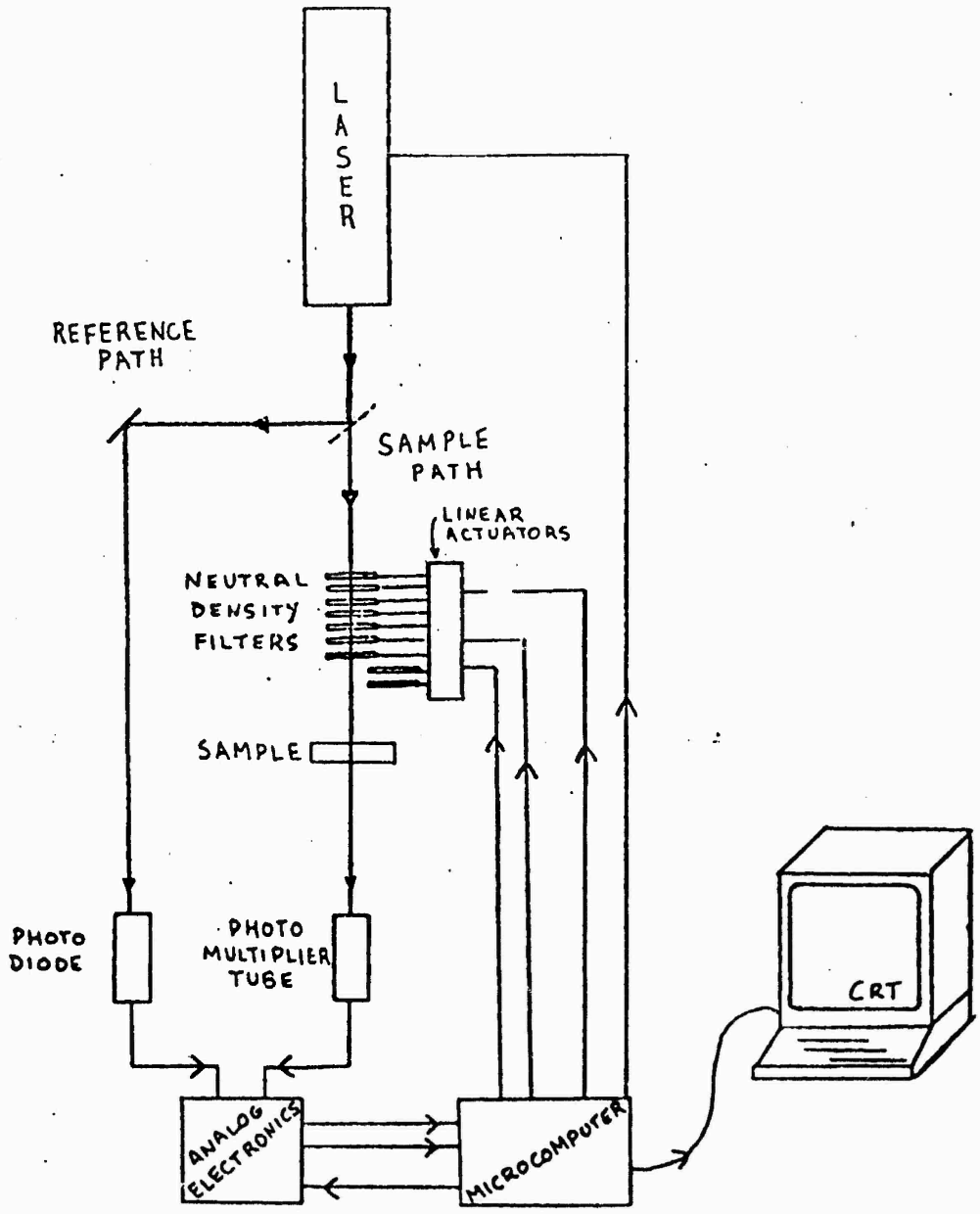


FIGURE 1. FUNCTIONAL SYSTEM OVERVIEW, LASER DENSITOMETER DESIGN

The densitometer contains three subsystems, the optics, the analog electronics, and the microcomputer. These subsystems will be discussed in detail here.

Optics

One objective of the design was to be able to test optical samples over a wide range of irradiation. Our optical design includes four separate lasers which provide radiation at six distinct wavelengths spanning the spectrum from 488 nm to 1060 nm. Two of the lasers provide continuous radiation, while the other two provide a pulsed source of radiation. The pulses are either of long (750 μsec) or short (15 nsec) duration depending on whether or not the laser is Q switched. Table 1 lists the lasers and their output capabilities including wavelength and pulse length.

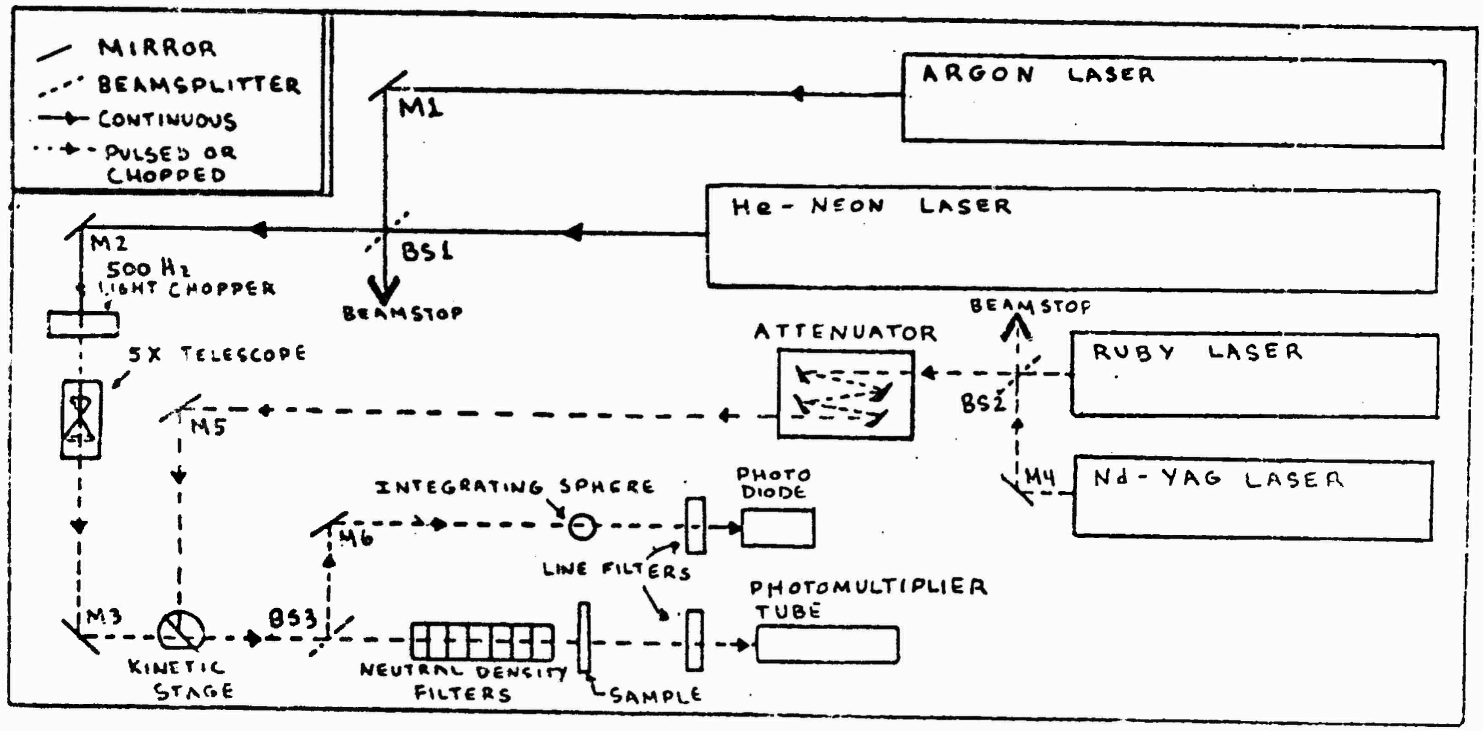
Laser Source	Wavelength (nm)	Waveform
Argon Ion	514.5 488.0	Continuous
Helium-Neon	632.8	Continuous
Ruby	694.0	Pulsed 15 ns or 750 μsec
Nd-Yag	1064 530	Pulsed 15 ns or 750 μsec

TABLE 1: LASER SOURCE OUTPUT CAPABILITIES

The optics table layout is shown in Figure 2. By using beam splitters and a single kinematic stage, all four of the laser beams become colinear before the optical beam path is split into the sample and reference paths. Thus, all of the lasers are permanently aligned so that any of the lasers can be used as the radiation source with a minimal effort on the part of the user. As any one of the four beams enters the sample path, it must be within a certain intensity range. The minimum intensity is limited by the minimum light required to produce a mid-range signal from the photomultiplier tube used as the detector in the sample path.

Remember that in the sample path, the light is attenuated by a factor of 10^8 by the combination of filters and the sample itself. The damage threshold of the neutral density places the upper limit on the intensity. Using the characteristics of 9808QB photomultiplier, a beam intensity in the range of 50 mwatts/cm^2 entering the sample path is optimal for a continuous laser. For a pulsed laser, the time response characteristics of the photomultiplier tube determine the optimum peak intensity in the sample path to be 100 watts/cm^2.

All lasers used in this system have sufficient power to provide the optimal intensities when the beams are 1 cm in diameter. In fact, the pulsed lasers must be attenuated considerably from their initial peak intensities which are on the order of 75 megawatts by a series of wedge prisms. The two pulsed laser beams propagate through the wedge prism attenuator and then are directed into the common optical path through a mirror mounted on a kinematic stage. A kinematic stage can be removed from



FICURE 2. OPTICS TABLE LAYOUT, LASER DENSITOMETER DESIGN

and replaced back into an optical system with high precision so that no optical alignment adjustments are necessary. The beam is then split by an uncoated beam splitter. A little over 90% of the beam goes into the sample path. The remainder of the light goes into the reference path where a photodiode monitors the laser's intensity.

The light in the sample path is attenuated by the series of neutral density filters mounted on the linear actuator shafts. The nominal values of the filters were chosen to minimize the total number of filters required to obtain any combination of optical density up to OD 8 with a resolution of ± .02 or better. The filters chosen have nominal optical densities of 4, 2, 1, .5, .3, .2, .1, .04, .03. The OD 4 filter is constructed out of two OD 2 filters.

Even though the filters nominally have a constant optical density over the entire spectrum, this is not true to the accuracy which is required in this system. To compensate for this fact, the optical density of each filter must be measured individually at each wavelength of interest. This table of measured densities vs. wavelength is then stored in an array within the microprocessor memory. It is used when calculating the total optical density which has been removed from the sample beam path.

The sample can be moved vertically and horizontally as well as rotated around both its vertical and horizontal axes. These position adjustments are available so that the optical density of the sample can be measured as a function of position and angle of incidence. The two

angular adjustments are required because the lasers are linearly polarized. The resulting asymmetry may cause an angular dependence of the samples optical density.

There are two minor differences between the paths of the continuous and pulsed lasers. The first difference is that the continuous beam is chopped by a 500 Hz mechanical chopper. Chopping the optical beam results in an electrical signal from the optical detectors which is a square wave of 500 Hz. We will later see that this is important in removing the detector's dark current and other noise components.

The next difference in the optical path is the passing of the chopped beam through a 5x expanding telescope to expand the beam diameter to 1 cm. In the pulsed source case the laser beam diameters are already 1cm in diameter upon exiting the laser, so no expansion is necessary.

Just before the beam gets to the detectors, the light passes through a laser line filter. Its transmission band is matched to the laser being used. The function of the filter is to pass the laser signal and prevent the room lights from entering the sensitive detectors. The last elements in the optical system are the detectors. A photomultiplier tube is used in the sample path and a photodiode in the reference path. Both the S-1 photocathode of the photomultiplier tube and the silicon photodiode detector have spectral sensitivity over the range of wavelengths of the laser sources used. (See Table 1)

A comparison of the sample and reference signal magnitudes will explain the choice of having different detectors. The sample path has an attenuation of 10^8 of the incoming light because of the neutral density filters and sample in its path. This is partially compensated for by allowing 95% of the source light to enter the sample path and only 5% to enter the reference path. The photomultiplier tube has a quantum efficiency of about 1/500 (1 photoelectron for every 500 photons), whereas the photodiode's quantum efficiency is close to 1. The photomultiplier has an internal current gain of 10^6, whereas the photodiode has unity gain. This would still leave a considerable mismatch in the two signals. To adjust for this, as well as limit the position sensitivity of the photodiode in the reference path, the reference beam passes through an intergrating sphere before the photodiode senses it. The integrating sphere reduces the intensity seen by the photodiode by a factor of 2000. Thus the ratio of the photomultiplier signal to photodiode signal is nominally:

$$\frac{\text{PMT Signal}}{\text{PD Signal}} = \frac{.95 \times 10^{-8} \times (1/500) \times 10^6}{.05 \times (1/2000)} = .76 \qquad (2)$$

Thus, the two signals are of comparable magnitudes. Any residual difference between the two signal amplitudes will be compensated for by variable gain amplifiers within the following analog electronics.

A different photodiode is used for the pulsed lasers than the one used for the continuous lasers. The one for the continuous laser detection has a surface area of 1 cm^2. This diode's surface area has considerable capacitance, making the detection of the pulsed lasers impractical. Therefore, the UDT PIN 040 photodiode with a .04 mm^2 surface area is used. This diode has a rise time less than 4×10^{-9} seconds and can be used to temporally resolve the pulsed laser intensity profile. The same photomultiplier tube is used for both the continuous and pulsed lasers. The photomultiplier tube has a rise time of 2×10^{-9} sec even though it has a large area photosensitive cathode (2 inch diameter).

The output signals from both optical transducers are currents. These currents are converted into voltage signals by transimpedance amplifiers. The expected magnitudes of the output voltages are on the order of .5 V for both the sample and the reference signals.

Electronics

We have used an analog circuit subsequent to the phototransducers and their transimpedance amplifiers. The function of this circuit is to determine which (if either) of the sample or reference signals is larger and to digitally encode this information for the microcomputer. The digital signal is then read by the microprocessor which will in turn alter the intensity in the sample path in an attempt to match the signals.

Because the temporal characteristics of the analog signals expected from the pulsed laser sources and continuous laser sources are significantly different, two separate analog signal processing circuits were designed. In both circuits care must be taken to remove the portion of the signal due to the dc dark current.

When using a continuous laser source, recall that a mechanical chopper interrupts the laser beam at a frequency of 500 Hz. Refer to Figure 3 for the schematic of continuous case electronics. The first stage of the analog circuit is a narrow bandpass filter with unity gain at the center frequency. The center frequency is tuned to 500 Hz with a bandwidth of ± 10 Hz. Thus, the dc dark current portion of the signal is rejected by this amplifier. In addition, the noise bandwidth is limited to 20 Hz, eliminating most of the shot noise of the detector. Additionally, by operating away from very low frequencies, 1/f noise generated in the preamplifier stage is minimized.

The second stage of the circuit is a precision full wave rectifier followed by an RC filter. Notice that there is an adjustable gain on the rectifier so that the signals in the sample and reference paths can be balanced prior to the insertion of the sample into the optical path. The RC filter converts the rectified signal into a dc level which is proportional to the 500 hertz signal which is in turn proportional to the light intensity seen by the phototransducer. The .1 second time constant of the RC filter limits the ripple of the 500 Hz fundamental fre-

quency to 1/100 of the dc level produced. The buffer amplifier following the RC filter isolates the filter from the input impedance of the following stage.

Up to this point, the circuitry of the reference and sample paths are identical. The dc signals from these two paths are now subtracted and the difference is multiplied by 10 for better resolution. The resulting output of the subtractor will be a positive dc level if the sample path voltage is greater than the reference path voltage and negative if the sample path voltage is less than the reference voltage. The result of the subtraction between the signals is compared to 5% of the reference signal. A table showing the encoding of the relative amplitudes of the sample and reference signals follows. L and M refer to the outputs of the two comparators in Figure 3.

Comparator Outputs

0 = low 1 = high

L	M	Significance
0	0	Sample signal is within 5% of reference
0	1	Sample signal is greater than reference by more than 5%, subtract OD from path
1	0	Sample signal is less than reference by more than 5%, add OD to path
1	1	Illogical and almost impossible

TABLE 2: COMPARATOR OUTPUT SIGNIFICANCE

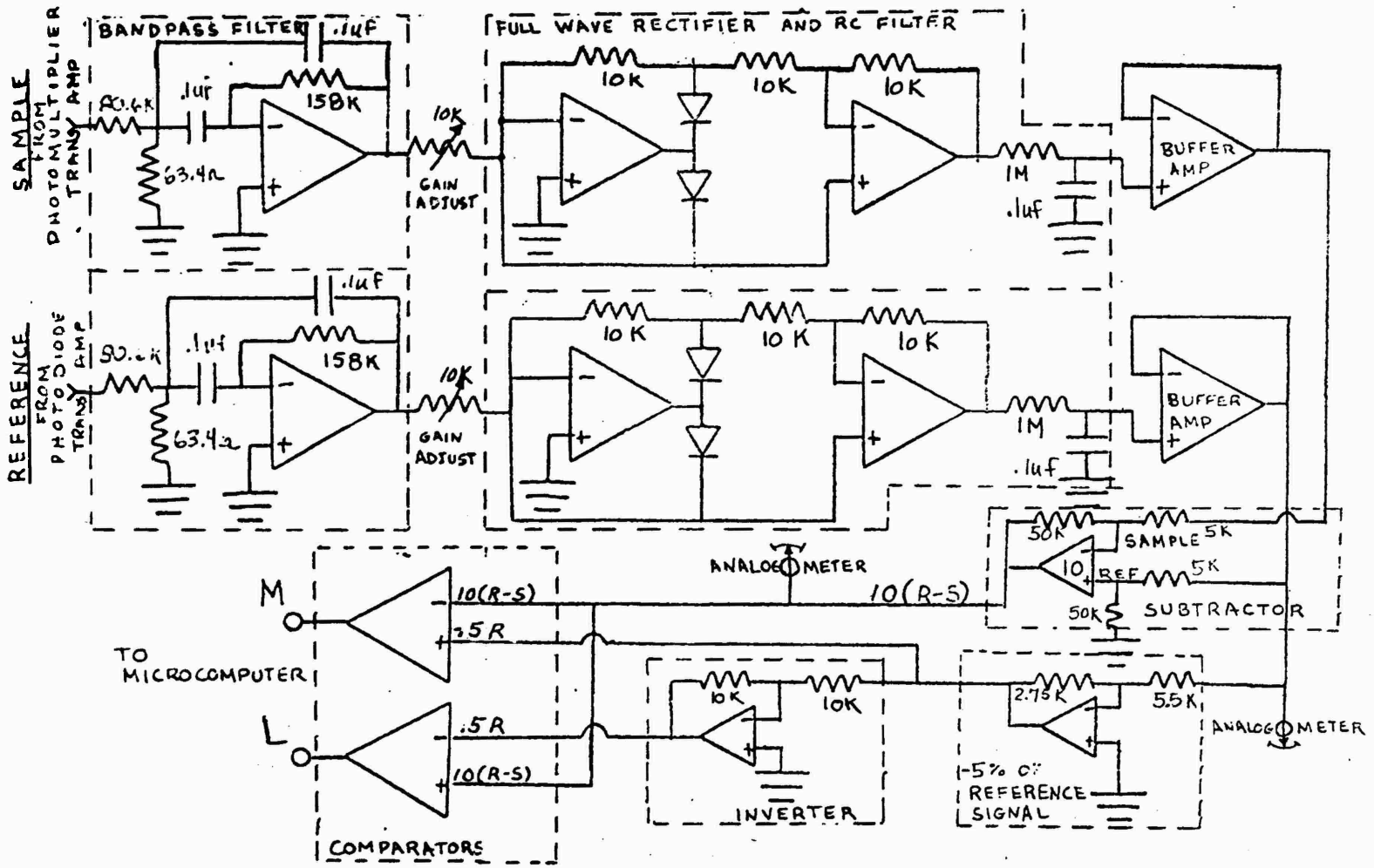


FIGURE 3. ANALOG ELECTRONICS SIGNAL PROCESSING, CONTINUOUS LASER SOURCE CASE

The logic of the subtraction and comparator functions is the same for the continuous and the pulsed circuits. The methods used to obtain the signal levels proportional to the peak intensities of the sample and reference optical signals are considerably different. See Figure 4 for the schematic diagram of the pulsed case.

The specifications of the OEI 5032 peak sample and hold device require that the input signal be positive. The output of both phototransducers are negative currents, and subsequently reslut in negative voltages from the transimpedance amplifiers. To provide a positive input to the peak sample and hold device, an inverting transistor amplifier is used. See Figure 4.

The duty cycle of the pulsed lasers is on the order of one pulse per second. It is deisrable that the intensity comparison be made on a single pulse rather than integrating the signal over 5 or 10 pulses. Thus, the information must be obtained from a single pulse of short duration. A peak sample hold device captures and holds the peak signal seen from the photransducer. While the total current in the pulse train is very small (a few picoamps) the peak current is in the milliamp range. Because the peak sample and hold device is sensitive to the peak current, the dark current can be neglected since it is orders of magnitude less than the peak current.

In order to obtain the long hold time required to allow the longer laser pulse of 750 μ s duration to finish before the reading is taken, two sequential peak sample and hold devices are used. The first peak sample and hold device has a fast rise time in order to catch the peak of the 15 ns pulse. There is a trade off, for it has a short holding time of

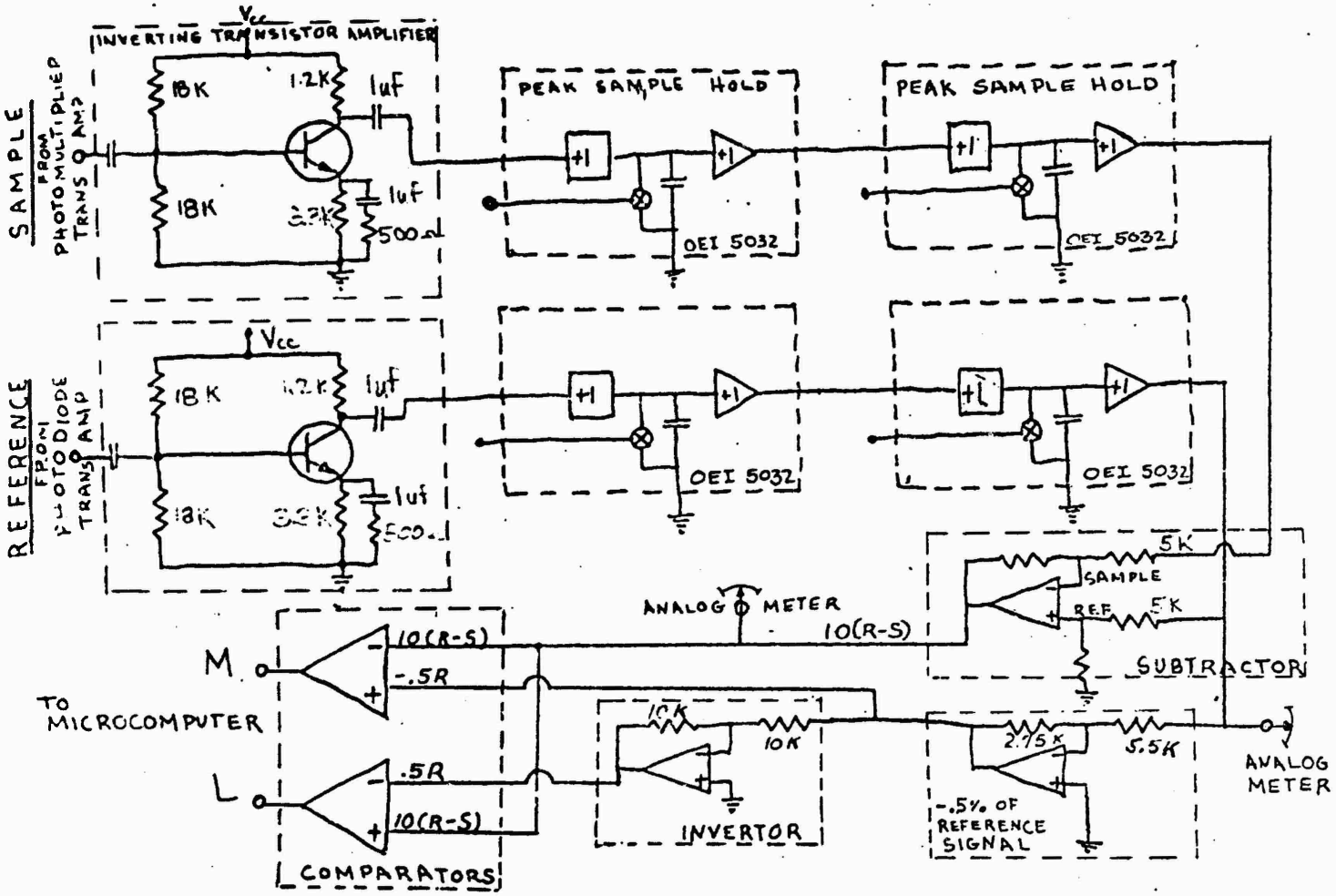


FIGURE 4. ANALOG ELECTONICS SIGNAL PROCESSING, PULSED WAVE LASER SOURCE CASE

less than 10 μsec. However, this holding time is sufficient to allow a slower rising peak sample and hold device in series with the first device sense the peak. Its decay rate is .01 volts/sec. The peak value is held for more than a millisecond. This is long enough for the microcomputer to read the resulting comparator outputs. The microprocessor requires a minimum of six machine cycles to perform the reading, with each cycle being one microsecond in duration.

All of the peak sample and hold devices are discharged by a signal from the microprocessor just prior to the microprocessor signal which fires the lasers. This insures that a reading from the present laser pulse is made rather than from the peak of a previous pulse which is still being held.

As mentioned above the dc levels from peak sample and hold devices are inputs to the indentical subtraction and comparison network as described for the continuous laser electronics. The resulting output signals from the comparators are TTL compatible and are read directly by the microcomputer.

MICROCOMPUTER

The microcomputer chosen for our application is a Synertek Systems SYM-1. It is based on the 6502 Central Processing Unit. By utilizing plug-in expansion capabilities, 4K of Random Access Memory (RAM) is available. To accomodate our application software the Read Only Memory (ROM) is expanded to 24K. The SY6522 Versatile Interface Adapter (VIA)

is used on the SYM-1 to handle peripheral interfaces. Three VIA's are used, each with 16 lines which are individually programmable to act as input or output lines.

The microcomputer system provides a degree of automation to the densitometer system by activating the linear actuators, reading the processed signals from the electronics, firing the laser in the pulsed case, and giving the user the opportunity to repeat the test. The main program, written in BASIC, provides user interaction and contains the decision making algorithm. Machine language subroutines communicate to and from the input/output ports. Information is passed from BASIC to machine language and vice versa. These programs are stored in Erasable programmable Read Only Memory (EPROM). The user is instructed to turn on the laser he wishes to use for this particular measurement and to input to the terminal the wavelength and wave form of the laser.

The difference between the pulsed and continuous cases is that for the pulsed case the microprocessor must provide the discharge capacitor signal to the peak sample and hold devices before each new pulse, and then fire the laser after an appropriate delay. For the continuous case, the optical signal is time invariant. Thus there is no need to exercise control as in the pulsed case.

Control is then transferred to the microporcessor where the filter positions are intialized to an optical density of 8 in the sample path. At this time the user is instructed to match the sample and refer-

ence beam electronic signals by adjusting the gains in the electronics. The comparators are repeatedly read and when both are low and remain low, then we are certain that the amplitude of the reference signal is equal to that of the sample signal with OD 8 in path. Once a match is made the user is instructed to position the sample and indicate when ready for measurement.

The insertion of the sample causes an inbalance between the sample and reference signals. Initially the microprocessor senses from the electronics connected to the input ports that there is less intensity in the sample path. The algorithm responds by indicating to the machine language linear actuator drive program to change filters in decrements of 1 OD. The machine language program then determines which filters need to be changed in the beam path and provides the triggering and direction input (protraction or retraction) to the motors. This continues until the imbalance condition is reversed, that is until the sample path has just less than an optical density of 8. A binary search is initiated until the combined OD of the known filters and the sample is as close to OD 8 as the smallest filter's resolution allows. The optical density of the sample is thus found by subtracting the known filter values at the wavelength under test from the total OD used at null (nominally 8 OD). This value is printed on the CRT and the user has the choice of repeating the measurement, changing the laser, repeating the electronics balance routine or stopping the program.

IV. RECOMMENDATIONS

When the design of this densitometer is implemented, the operator must be aware of its capabilities and limitations. The densitometer will measure optical densities over a range from 0 to 8 OD with a resolution of ± .02 OD. The densitometer may give erroneous measurements when the density of the material is either time or intensity dependent. Glass embedded with a bleachable absorber is an example of such material.

Other efforts in the USAFSAM Laser Laboratory have been in exploring the spectral emission generated by the plasma created by a high power infrared laser impinging on airplane windscreen material. The plasma emission contains light of wavelengths throughout the visible spectrum. It would be necessary to modify the densitometer design so as to test samples with a continuous spectrum of visible light rather than at a few wavelengths. An organic dye laser would be a suitable source for it can provide either a pulsed or continuous wave, tunable throughout the visible and near infrared spectra (.4 μm to .9 μm). Thus, incorporating the dye laser into the design would expand the capabilities of the laser densitometer.

REFERENCES

1. Fodor,W.J Laser-Protection Eyewear: An Evaluation Procedure. SAM-TR-76-19. USAF School of Aerospace Medicine, Brooks AFB, Texas (May, 1976).

2. Fodor,W.J Laser-Protection Eyewear: Preliminary Results from Colorfastness Testing. SAM TR-76-45. USAF School of Aerospace Medicine, Brooks AFB, Texas (Dec, 1976).

3. Taboada J. and W.J. Fodor. Pulsed Dye Laser Densitometry Using An Optical Delay. Appl. Opt. 16(5) 1132-1133 (1977).

4. Murray R. et al. Evaluation of Laser-Protection Eyewear. SAM TR-78-30. USAF School of Aerospace Medicine, Brooks AFB, Texas (Oct, 1978).

5. Envall K. R. and Murray R. Evaluation of Commercially Available Laser Protective Eyewear. FDA 79-8086. US Dept. of Health Education, and Welfare (May, 1979).

6. Budak A., Passive and Active Network Analysis and Synthesis, Houghton Mifflin Co.,Boston, 1974.

7. SYM-1 Reference Manual, Synertek Systems Corporation, Santa Clara, CA 1980.

REAL-TIME DATA QUALITY ASSESSMENT OF DISTRIBUTED DATA ACQUISITION SYSTEMS

by

Sigurd L. Lillevik

Christopher T. W. Rentola

ABSTRACT

This report describes a 10-week research effort to study the data quality of the Avionics Laboratory's TSCF distributed data acquisition system. This facility generates data for validating Air Force models and for sensor evaluation. A second-order, alpha-beta tracker was designed to predict the value of a parameter one sample in advance. Thus, whenever the difference of the predicted and actual data exceed a threshold value, erroneous data are detected. To demonstrate this principle and to develop an off-line aid for detecting erroneous data, a software package called DBQA was written, debugged, and verified. This software package was used to evaluate raw data and obvious data excursions were correctly detected as erroneous data. These results indicate that the method will work very well in a real-time environment.

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

REAL-TIME DATA QUALITY ASSESSMENT OF DISTRIBUTED DATA ACQUISITION SYSTEMS

Prepared by:	Dr. Sigurd L. Lillevik
	Mr. Christopher T. W. Rentola
Academic Rank:	Assistant Professor
	Graduate Student
Department and University:	Electrical and Computer Engineering
	Oregon State University
Research Location:	Air Force Avionics Laboratory, Mission Avionics Division, Electro-Optics Branch
USAF Research:	Major James R. Johnson

Date:	August 26, 1983
Contract Number:	F49620-82-C-0035

ACKNOWLEDGEMENTS

The authors wish to express their appreciation to the Air Force Systems Command, the Air Force Office of Scientific Research, and the Southeastern Center for Electrical Engineering Education for the opportunity to conduct research at the Air Force Avionics Laboratory, Wright-Patterson AFB, OH. Specifically, we would like to thank the Electro-Optics Branch for the use of their facilities and the hospitality extended to us during the summer. In addition, we wish to thank Major James R. Johnson for suggesting the research topic and for his collaboration. Finally, we appreciate the help provided by Mr. Leo Vroombout in reviewing our work, and we acknowledge the assistance of Sergeant Sammy Jiles in answering the many questions we asked about the TSCF computer system.

I. INTRODUCTION

The Department of the Air Force conducts research in modeling of many naturally occurring phenomenon. For example, models of the atmosphere assist pilots in determining optimal flight plans. Other examples include Tactical Decision Aides (TDA) which provide fighters with information on when, where, and how to attack a particular target. All models approximate a complicated system and require validation to determine their accuracy and limitations. To meet this need, the Air Force Avionics Laboratory developed a Targeting Systems Characterization Facility (TSCF) composed of several minicomputers, a four-tank scene or target, and an array of sensors. Besides validating several models, the TSCF laboratory allows for extensive evaluation of sensor requirement, precisions, and ranges. Both the Air Force Armament and Geophysics Laboratories use the data collected by the TSCF and its instruments.

The TSCF laboratory consists of a network of loosely-coupled minicomputers configured in a star architecture. Each remote node is physically separated from the central node but linked to the central node with a serial communications line.

Building 622 houses the central node which contains HP 2113 and 2112 minicomputers. As the central node, these

minicomputers coordinate all data acquisition activities and enter collected data into a data base management system. This software package provides for efficient storage, manipulation, and retrieval of data.

Located in a tower attached to Building 622, the first remote node uses an HP 6942 Multiprogrammer to monitor several meteorological sensors. In addition, a forward looking infrared (FLIR) sensor scans the target scene and an additional hot-patch. Also, this tower holds a calibrated infrared source monitored by a transmissometer at the tank site. Approximately 2.25 km separates the remote node and the target scene.

Two additional remote nodes are located at the tank site. One of them contains various meteorological sensors plus several temperature sensors mounted on a tank. These sensors and the transmissometer connect to an HP 6942 Multiprogrammer which serves as the computer for this node.

Finally, the last remote node, also located at the tank site, contains an Inframetrics 210 radiometer which monitors the tank scene. The radiometer sits on a cherry picker which swings a 180 degree half-circle to obtain different tank profiles. Under control of a DEC 11/23, this remote node determines average radiance values and sends them onto the central node for insertion into the data base.

Typical data acquisition periods last from fifty-two (52) to ninety-seven (97) hours. During this time, there

exists the possibility of collecting erroneous data. Since the TSCF collects data for model verification and sensor evaluation, the quality of data must be assured. Three factors influence the quality of data: completeness, out-of-bounds, and correlations. Completeness implies contiguous data for the duration of the sample period. Out-of-bounds refers to the magnitude of data approaching a value not consistant with existing conditions. And data correlations relate to the interaction of one parameter with respect to another. Together, these three features dictate the quality of collected data.

With this brief background, the problem studied in this report may now be stated:

HOW CAN THE QUALITY OF COLLECTED DATA BE ASSURED?

Actually, this broad question consists of several parts as follows:

1. How can erroneous data be detected?
2. What should be the response?
3. How can the detection and response be implemented and evaluated?

Although these questions are dependent on each other, they must all be answered to assure the data quality of the TSCF and the integrity of the models designed by Air Force organizations.

II. RESEARCH OBJECTIVES

To arrive at solutions to the above questions, several goals were established. During the 10-week research period, it became obvious that some of the goals required more time than was available. None-the-less, to solve the above problem requires completion of all of the following goals. Specifically, the goals are:

1. To survey the scientific literature for methods to detect erroneous data.
2. To determine the best method for the given application.
3. To demonstrate that the selected method will detect erroneous data.
4. To implement the method.
5. To evaluate the effectiveness of the implemented solution.

The first three goals represent exploratory research while the last two entail less basic, more applied research. Thus, only the first three objectives have been considered during the 10-week research period.

Meeting these goals greatly impacts the future research conducted at the TSCF laboratory. But moreover, the results of this research (a methodology to detect and respond to erroneous data) may possibly impact and benefit a much broader range of scientists. In fact, anyone performing

data acquisition and/or control requires quality data, and this research will be of interest to them.

III. DETECTING ERRONEOUS DATA

From the beginning of computer-aided data acquisition, designers developed methods to check and sometimes correct for erroneous data. The overall approach involves partitioning the system into several sections or components, and then looking for the presence of erroneous data within each section. Thus, designers may view a distributed data acquisition system from various levels.

To detect erroneous data, two methods are usually considered: direct, and indirect. The direct method uses hardware redundancy to compare two samples of the same parameter. If the samples do not agree, then erroneous data have been found. Conversely, the second method, indirect, relies on a mathematical treatment of the present and past data. When the data fail to exhibit the expected characteristics, then erroneous data have been found. The direct method represents a simple but expensive approach while the indirect method involves a sophisticated but inexpensive technique. As with most decisions, each alternative presents some advantages and some disadvantages, and the specific application dictates the correct choice.

Typical direct methods use Triple Modular Redundancy (TMR) to both detect and isolate the problem. For example,

Evans and Price[3] reported on a reliable computer using three central processing units. To minimize costs, designers often place redundancy only in critical sections and Hertel and Clark[5] developed a technique to assist in location of the TMR. All TMR schemes require a perfect voting mechanism to decide which modules agree and to isolate the faulty module.

Indirect methods usually fall under one of two techniques: statistical inference or digital signal processing. Often, designers use estimation theory to find an approximate form for the distribution and density functions of the mean time before failure (MTBF). This becomes difficult when dealing with limited sample sizes and Dey[2] proposed a solution to this problem. Using MTBF information, equipment may be removed from service prior to their expected failure for assured data quality.

In addition to MTBF, other statistics such as means, standard deviations, and correlations may be estimated using past data. Whenever the new data exhibit statistics beyond that which is expected, then the presence of erroneous data have been detected. For some parameters, the data excursions follow a repeatable and predictable pattern. Applying regression techniques, past data may be fit to curves using, for example, the method of least squares. During data acquisition, should the data not follow their expected curves, then the presence of erroneous data have

been detected. Good, et al.[4], suggested that each sensor periodically read a calibrated source. Whenever these readings deviate too far, then the presence of erroneous data have been detected. All of the above methods utilize statistical inference theory to detect erroneous data and will sometimes miss bad data and flag good data.

Digital signal processing may be viewed as the mathematical transformation of a sequence of numbers. Algorithms may be used to estimate various parameters associated with a sequence, or sequences may be modified into other more useful forms. When sequences represent signals such as meteorological data, they may be processed in the time or frequency domain.

Data quality assessment in the frequency domain may take several approaches. One possibility involves spectral density estimation of the signal and noise bandwidths. If the noise bandwidth is narrow compared to the signal bandwidth, a digital filter may be designed to simply remove the noise. More often, however, the noise bandwith is wide compared to the signal bandwith. According to Papoulis[8], if the signal and noise are jointly stationary with known spectra, a Wiener filter can be designed to produce the best estimate of the signal.

Another possibility for data quality assessment in the frequency domain involves fitting the data spectral samples to a finite impulse response (FIR), linear prediction

filter. Oppenheim and Schaffer[7] have shown that the coefficients of the FIR filter can be fit to the data spectral samples using the least squares technique or equiripple approximation.

If the data are nonstationary, a problem common to all these solutions is that filter parameters require constant updating. Since the analysis proceeds in the frequency domain, updating the parameters requires considerable computation for conversion from the time domain to the frequency domain.

To reduce the required computation, a digital filter can be designed in the time domain. Here, filter coefficients are selected to minimize the mean-squared error criterion. For the research discussed in this report, a second-order, infinite impulse response (IIR) linear prediction filter was selected for two reasons. First, the incoming data are assumed to be nonstationary. Kikuchi, et al.[6], have stated that higher-order filters may not respond to quick changes because they tend to fit long time-data. Secondly, a second-order filter requires fewer coefficients to optimize. This is particularly attractive since the minimization of the mean-squared error, according to Rabiner and Gold[9], is a nonlinear function usually solved with iterative techniques.

The second-order, linear prediction filter selected for this application is the alpha-beta tracker developed by Benedict and Bordner.[1] Originally developed for a radar track-while-scan system, the basic difference equations are:

$$y_p(k) = y(k-1) + T\dot{y}(k-1) \tag{1}$$

$$y(k) = y_p(k) + \alpha[u(k) - y_p(k)] \tag{2}$$

$$\dot{y}(k) = \dot{y}(k-1) + \frac{\beta}{T}[u(k) - y_p(k)] \tag{3}$$

where:

$u(k)$ = noise contaminated input signal,

$y(k)$ = estimated (smoothed) input signal,

$\dot{y}(k)$ = estimated (smoothed) rate of change,

$y_p(k)$ = predicted next value, and

T = sample interval.

After taking the Z-transform and decoupling Equations (1)-(3), the transfer functions are:

$$\frac{Y_p(Z)}{U(Z)} = \frac{Z(\alpha + \beta) - \alpha}{Z^2 + (\beta + \alpha - 2)Z + 1 - \alpha} \tag{4}$$

$$\frac{Y(Z)}{U(Z)} = \frac{Z(\alpha Z + \beta - \alpha)}{Z^2 + (\beta + \alpha - 2)Z + 1 - \alpha} \tag{5}$$

$$\frac{\dot{Y}(Z)}{U(Z)} = \frac{\beta}{T}\,\frac{Z(Z-1)}{Z^2 + (\beta + \alpha - 2)Z + 1 - \alpha} \tag{6}$$

Benedict and Vordner have shown that for optimal position and tracking,

$$\beta = \frac{\alpha^2}{2-\alpha} \tag{7}$$

which corresponds to a slightly underdamped system with

$$0 \leq \alpha \leq 1.236$$

the stability constraint.

Applying the optimally constraint, Equation (7), and transforming Equations (4)-(6) to the time domain yields:

$$y_p(k) = \frac{2\alpha}{2-\alpha} u(k-1) - \alpha u(k-2) - \frac{4(\alpha-1)}{2-\alpha} y_p(k-1) - (1-\alpha) y_p(k-2) \tag{8}$$

$$y(k) = \alpha u(k) + \frac{2\alpha(\alpha-1)}{2-\alpha} u(k-1) - \frac{4(\alpha-1)}{2-\alpha} y(k-1) - (1-\alpha) y(k-2) \tag{9}$$

$$\dot{y}(k) = \frac{\alpha^2}{T(2-\alpha)} u(k) - \frac{\alpha^2}{T(2-\alpha)} u(k-1) - \frac{4(\alpha-1)}{2-\alpha} \dot{y}(k-1) - (1-\alpha) \dot{y}(k-2) \tag{10}$$

These equations were implemented in the data base quality assurance (DBQA) software package to detect

erroneous data. This was accomplished by comparison of predicted values generated by Equation (8), with actual data. While Equations (9)-(10) are not required for data prediction, they were implemented to observe the basis for the system's prediction.

IV. DBQA PROGRAM

The DBQA software package was written in Fortran 4 for an HP 1000 series minicomputer with an RTE 4 operating system. This program was used to verify the above concept, aid in parameter selection, and analyze data not previously checked for erroneous values. Due to the large size of the DBQA software package, it was necessary to segment the program into nine overlays. Each overlay contains a user-interactive command scheduled by the DBQA root. A description of the commands follows.

The HELP command is divided into brief and detailed explanations of the other eight commands. An overview of each command is provided by the brief HELP command. Conversely, a description of each command's input prompts and helpful suggestions for choosing proper responses is furnished by the detailed HELP command. Also provided are explanations of the output produced by each command.

Access to the IMAGE 1000 data base by year, day, start and stop times is provided by the READ command. The output is placed in the DBQA buffer whose capacity is 44 by 360.

Positions 1 through 40 contain meteorological data while positions 41 through 44 are reserved for calculations.

Numeric values of the DBQA buffer may be displayed by the use of the PRINT command. Up to four items may be displayed at once.

The PLOT command allows the DBQA buffer to be illustrated graphically on the user's CRT or a hardcopy may be generated on the HP 7210A digital plotter. Up to four items may be plotted at once.

To evaluate arbitrary filter parameters, the PREDICT command has been provided. It keeps statistics on the error so different parameters may be compared.

The CALIBRATE command was originally designed for program development. It generates three deterministic signals; a sine, step, and ramp function which may be used to verify predictor operation.

With the OPTIMIZE command, filter parameters may be located which minimize the mean-squared error of the predicted and actual data values. The number of data points used in the optimization depends upon the requested READ command start and stop times along with the sampling interval used during data collection. Using an initial guess and step size, the optimizer determines a new error and then compares it with the old error. When the error does not improve, the optimizer cuts the step size in half and reverses direction.

Evaluating data with the predictor is performed with the TEST command. If the test data differ from the predicted values more than a predetermined limit, the data are flagged. Limits to flag data default to 3 standard deviations but may be changed as required.

The DISTRIBUTION command provides for a plot or list of the probability density and distribution function of the error of the predicted and actual data. This command is useful in determining limits to flag erroneous data.

Together, the above nine commands provide a comprehensive tool to evaluate past data and demonstrate the real-time capabilities of the alpha-beta tracker.

V. OBSERVATIONS

Since the DBQA software package operates in an off-line mode, all observations used past data held in the data base. These include samples not previously checked for erroneous values. When the DBQA software evaluated the raw data, it correctly detected obvious data excursions as erroneous data. For some parameters, proper detection required an error bound of 2.5 standard deviations instead of the defaults value of 3.

Both the density and distribution of the error between the actual and predicted values were determined and plotted for several parameters. In all cases, the error resembled a

normal curve within the limitations of a small sample size. This fact verified our use of standard deviations as a bound on acceptable data.

While using the optimizer, we found that nearly all alpha values ranged from 0 to 1. In fact, the optimizer works quickly with an initial guess of 0.5 and step size of 0.25. After 20 to 30 iterations, the improvement in error becomes quite small. Because of the algorithm, the optimizer functions more efficiently when the step size equals one-half of the initial alpha value.

A moving time-window of samples was used to optimize for new alpha values. Using these results, the DBQA software correctly predicted in advance and properly detected erroneous data. From these experiments, we have confidence that a moving time-window of samples to optimize for new alpha constants will function well in a real-time environment.

VI. RECOMMENDATIONS

Now that exploratory research has demonstrated the potential for success of the above method to detect erroneous data, some suggestions are appropriate to aid in real-time implementation and for topics of further research. First, a real-time implementation involves modification of the software which coordinates the TSCF distributed data

acquisition system. Since the DBQA software package was written primarily to demonstrate a principle, its architecture requires major changes for real-time execution. We recommend that a new software package be developed with an architecture that considers real-time constraints. Furthermore, this package should use a moving time-window to optimize for new alpha values prior to predicting the next samples. In fact, the old alpha values will serve as excellent initial guesses for the optimizer, and 20 to 30 iterations should provide 4 to 5 digits of precision.

For a response to detection of erroneous data, we suggest that the suspect data not be removed from the data base until an expert evaluates the data for consistencies. Instead, one response is to notify the appropriate computer operator or other responsible person that a specific parameter requires checking. Then, the person can log onto the computer system what corrective action was taken for later reference.

As for topics of further research, one concerns the alpha-beta tracker. Higher-order equations may prove superior prediction than the alpha-beta method, but they may also require greater computational time. Perhaps a tradeoff exists here between accuracy and computational requirements. Another topic of study involves the moving time-window for optimization of the alpha constants. A longer interval improves accuracy and nullifies the effects of erroneous

data, but since the data are probably not stationary a smaller time interval improves accuracy. In addition, the time-window shape may impact accuracy. Again, tradeoffs exist here between accuracy, computational requirements, effects of nonstationary data, and length and shape of the moving time-window. Even though the DBQA optimizer functions well, other schemes may converge quicker and require fewer computations. Thus, the optimizer should also be studied. Finally, the last topic for further investigation is to determine the data characteristics in the frequency domain. Few researchers have considered the frequency domain to detect erroneous data primarily because of the real-time computational constraints. But with today's high-performance computers and the increased understanding of multiple processor architectures, real-time frequency analysis becomes feasible. Investigation of these topics will impact research conducted at the Avionics Laboratory's TSCF by providing incresead data quality assurance. A minigrant proposal will be submitted to AFOSR for follow-on support of these research topics.

REFERENCES

1. T. R. Benedict and G. W. Bordner, "Synthesis of an Optimal Set of Radar Track-While-Scan Smoothing Equations," IRE Trans. on Auto. Control, AC-7, pp. 27-32, July 1962.

2. K. A. Dey, "Statistical Analysis of Noisy and Incomplete Failure Data," 1982 Proc. Annual Reliab. and Maintain. Symp., pp. 246-251, Jan. 1982.

3. R. J. Evans and S. H. Price, "Fault Tolerant Microprocessor System Design," 11th Annual Intl. Symp. on Fault Tolerant Comput., pp. 248-250, June 1981.

4. R. R. Good, et al., "Automated Data Qualification," Proc. 27th Intl. Instrum. Symp., vol. 27, pp. 467-471, April 1981.

5. J. E. Hertel and R. N. Clark, "Instrument Failure Detection in Partially Observable Systems," IEEE Trans. on Aerospace and elect. Syst., vol AES-18, pp. 310-317, May 1982.

6. A. Kikuchi, et al., "Application of Adaptive Digital Filtering to the Data Processing for the Environmental System," IEEE Trans. on Acoustics, Speech, and Sig. Proc., vol. ASSP-27. pp. 790-803, Dec. 1979.

7. A. V. Oppenheim and R. W. Schaffer, Digital Signal Processing. Prentice-Hall, Inc.: New York, N.Y., pp. 263-270, 1975.

8. A. Papoulis, Signal Analysis. McGraw-Hill: New York, N.Y., pp. 324-350, 1977.

9. L. W. Rabiner and B. Gold, Theory and Application of Digital Signal Processing. Prentice-Hall, Inc.: New York, N.Y., pp. 263-270, 1975.

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

PHOTOCHEMICAL REACTIONS IN A SMALL INDOOR SMOG CHAMBER

Prepared by:	Dr. Stephen F. Lin and Burton D. Price
Academic Rank:	Associate Professor and Graduate Student
Department and University:	Department of Chemistry North Carolina Central University
Research Location:	Air Force Engineering and Services Center, Engineering and Services Laboratory, Environics Division, Environmental Sciences Branch
USAF Research:	Dr. Daniel A. Stone
Date:	August 12, 1983
Contract No:	F49620-82-C-0035

PHOTOCHEMICAL REACTIONS

IN A SMALL INDOOR SMOG CHAMBER

BY

Stephen F. Lin and Burton D. Price

ABSTRACT

Photchemical reactions were studied in a small indoor smog chamber using long-path Fourier transform infrared spectroscopy. Three types of experiments were carried out to characterize the smog chamber: the irradiation of background air, ozone decay in the absence of light, and photooxidation reactions of the propene-NO_x-air system. The absorbance-time profiles of major products were determined in the background air system. The ozone decay rate was determined. The concentration-time profiles of reactants and major products were determined and important reaction mechanisms were mentioned for the propene-NO_x-air system.

Acknowledgement

The authors would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research, and the Southeastern Center for Electrical Engineering Education for providing them with the opportunity to work on a very interesting project at the Air Force Engineering and Services Laboratory, Tyndall AFB, FL. They would like to acknowledge the laboratory, in particular the Environmental Sciences branch, for its hospitality and excellent working conditions.

They would like to express their appreciation especially to Dr. Daniel A. Stone for introducing them to the area of atmospheric photochemistry and for his collaboration and guidance, and to Mr. Michael V. Henley for assisting in the experiments and data analyses.

I. INTRODUCTION:

Photochemical smog was recognized more than three decades ago in Los Angeles. Since then a large number of studies have been carried out trying to understand the photochemistry of air pollution. Unfortunately, the photochemistry of the polluted atmosphere is exceedingly complex; many gaseous pollutants exist in the atmosphere and hundreds of chemical reactions may take place simultaneously. One approach to solve this problem is to conduct laboratory experiments in which factors affecting the photochemical reactions can be carefully controlled. Laboratory experiments usually involve only one or two hydrocarbon and typical amounts of nitrogen oxides, carbon monoxide, and water vapor. Such experiments have been carried out since the 1950's[1] in reaction cells commonly referred to as "smog chambers" and are still being done in many laboratories[2-7]. A great deal of highly useful information has been generated in such chamber studies and many kinetic mechanisms have been proposed which have helped elucidate the fundamental chemistry of photooxidiation processes. Another approach of current air pollution research is the development of computer models for photochemical smog formation[8-12]. Computer models are based upon detailed reaction mechanisms and the rate constants of all elementary processes involved, and they should be validated by data collected in smog chamber experiments. Therefore, generation of accurate data from laboratory and smog chamber experiments is absolutely necessary for the success of the computer modeling and the understanding of the photochemistry of air pollution.

Many military operations involve large quantities of aircraft fuels. In the presence of NO_x and sunlight, the hydrocarbon components of such fuels can be photo-oxidized to yield ozone and other oxidants. The Air Force is interested in the investigation of the atmospheric fates and impacts of these jet fuels. Smog chamber experiments involving major components of these fuels are required in order to understand the consequences of these fuels in the atmosphere. However, before any experiments can be performed in a smog chamber, the chamber has to be well-characterized. The project was undertaken with the goal of characterizing the smog chamber, so that subsequent experiments involving the major components of jet fuels can be carried out.

II. OBJECTIVE:

Although smog chambers have been used since the late 1950's, the quantitative agreement among the results of different systems has been poor. Many factors have contributed to this problem, such as different sizes, shapes, and the nature of the chambers, different light sources, and different analytical techniques and instruments[13]. It is, therefore, very important that each smog chamber be well characterized. For this reason, the main objective of this project was the characterization of the smog chamber.

The approach to smog chamber characterization involved three types of experiments. The first type was irradiations of pure air to determine background reactivity levels. The second was the determination of ozone decay rate in the absence of light. The third was photooxidation experiments of propene in combination with nitrogene oxides. The time history of reactants, and important intermediates and products was followed with long-path Fourier transform infrared spectroscopy.

III. EXPERIMENTAL:

Long-path Fourier transform infrared spectroscopy was used as the analytical monitoring technique for our studies. The LP-FTIR technique has been applied to the measurement of low level pollutants in the atmosphere for many years and has proved to be among the more specific, sensitive, and accurate methods for the detection of many molecular species[5,7,14-17].

The schematic diagram of the long-path infrared photochemical reaction chamber is shown in Figure 1. The infrared instrument is a Nicolet 6000C Fourier transform infrared spectroscopy system. The system uses a globar as the light source with optimized emissivity for the 400 to 4000 cm^{-1} spectral range. The interferometer is a continuous scan Michelson interferometer with germanium-coated KBr beam splitter and a liquid-nitrogen cooled mercury-cadmium-telluride detector for the region 700-4000 cm^{-1}. It can provide a variable scan speed from 0.07 to 5 cm/sec. and a variable resolution from 0.06 to 8 cm^{-1}. Data is processed by the Nicolet IR-80 data system. The data system consists of 64K memory, analog-to-digital converter, color raster-scan for alphanumerics or

interactive graphics display, keyboard terminal, digital plotter, and dual disk drive. The system also has a software package allowing many kinds of data manipulation such as spectral subtraction, baseline correction, integration, peak picking, spectral library searching, etc. During a scan of the interferometer mirror, the detector signal is digitized at the number of points required for the desired spectral resolution and is stored in the computer. A chosen number of successive scans are added together in the computer, thus averaging the noise but building up the signal. The signal digitized sum interferogram is then transformed by the computer to a spectrum. The spectra can be stored, retrieved, divided into each other, subtracted from each other, and plotted in various modes.

The long-path cell was constructed of a 0.152-meter by 3.05-meter section of Pyrex® pipe[18]. Three mirror White-type[19] multiple-pass optics were installed in Plexiglass® mounts in the cell at 2.75 meter separation. A helium-neon laser was mounted under the optical bench and its beam directed to follow the infrared beam optical path for use in the alignment and adjustment of the multiple pass cell. With the aid of this laser, mirrors II were adjusted to give two rows of images on mirror I corresponding to the desired number of passes of the laser beam through the cell. During the course of the experiments, 19 laser dots were obtained corresponding to 4x19 or 76 passes through the cell for a total path length of 209 meters. The entire length of the cell can be irradiated by banks of fluorescent lamps. The fluorescent lamps installed in the system are GE® F40BLB black lights and Westinghouse® FS40 sun lamps with an output in the region from 300 to 400 nm. The light intensity of the lamps used in the irradiation experiments was determined to be k_1=0.33 $min.^{-1}$ by the rate of photolysis of NO_2[20].

A typical experimental run is described below. First, the cell was evacuated. Then pre-measured amounts of materials were introduced into the long-path cell via an attached vacuum manifold, either through a silicon septum on the manifold with a syringe or by attaching a glass bulb with Teflon® valves at both ends to the manifold and flushing the contents into the cell with fill gas, which was either nitrogen for reference gases or air for sample gases. The gas mixture was circulated with a mechanical diaphram pump for a few minutes before an experiment started. All experiments were run at room temperature and atmospheric pressure. An air conditioning unit was used to keep the temperature from rising when

the fluorescent lamps were turned on. A background spectrum was always taken before the sample spectrum so that the single beam sample spectrum could be ratioed against it and converted to a transmittance spectrum or an absorption spectrum. Reference spectra of O_3, NO_2, HCHO, HCOOH, CH_3COOH, PAN, HNO_3, propene, and n-butane were taken and stored on the monitor disk for use in subtraction mode or other purposes. The reference spectra and the background spectra were taken at 512 scans and the sample spectra at 128 scans, all at 1 cm^{-1} resolution. In the irradiation experiments, spectra were taken at appropriate time intervals, stored on a disk, and analyzed at a later time.

IV. DETERMINATION OF THE CONCENTRATION:

In order to follow the chemical kinetics of reactions, an accurate knowledge of the concentrations of reactants and products as functions of time is required. The infrared method uses a characteristic peak of a compound in the spectrum for identification and applies the Beer-Lambert law for the determination of the concentration.

According to the Beer-Lambert Law:

$$A = \ln \frac{I_o}{I} = \epsilon LC \qquad (1)$$

where A = absorbance

ln = natural logarithm

I_o = incident intensity of infrared frequency of the characteristic peak

I = transmitted intensity

ϵ = absorption coefficient, cm^{-1} atm^{-1}

L = optical path length, cm

C = concentration in terms of partial pressure, atm

In practice, the absorption coefficient ϵ was either determined experimentally or taken from the literature. The path length L was

determined from the number of laser dots as explained in Section III, and the absorbance A was determined from the absorption spectrum. The concentration of the gas in terms of its partial pressure was then calculated from eq. (1).

The characteric frequencies chosen and the absorption coefficients of the molecular species involved in the experiments are listed in Table 1. Incorrect absorption coefficients are the most likely source of error in atmospheric infrared studies. They depend on the shapes of the infrared bands, the experimental conditions, and the instruments. There is a large variation in absorption coefficients found in the literature.

Table 1. Absorption Coefficients

Compound	Frequency cm^{-1}	Coefficient cm^{-1} atm^{-1}	Reference
Acetaldehyde (CH_3CHO)	2729	3.8	22
Acetic Acid (CH_3COOH)	1180		
Carbon Monoxide (CO)	2170	Varies between 5.1 and 7.2	5
Formaldehyde (HCHO)	2779	16	14
Formic Acid (HCOOH)	1105	13	5
Nitric Acid (HNO_3)	1323	10.5	7
Nitric Oxide (NO)	1876	1.5	a
Nitrogen Dioxide (NO_2)	1629	16.6	a
Nitrogen Pentoxide (N_2O_5)	1248	40	5
Ozone (O_3)	1055	9.7	14
Peroxyacetyl Nitrate (PAN)	1162	13.9	21
Propylene glycol 1,2 - dinitrate (PG DN)	1280	26.6	7
Propene (C_3H_6)	912	6.8	a

a. Values from earlier RDVC studies.

V. THE BACKGROUND REACTIVITY LEVELS OF AIR:

Smog chamber contamination is one of the problems encountered in the laboratory experiments[23,24]. The contamination can arise from the wall material itself or from residual compounds sticking to the chamber wall from previous runs. For this reason, it is important to know the background reactivity levels of the air used in the experiments.

The zero air was introduced into the reaction chamber at room temperature and a pressure of 760 torr. After a background spectrum was taken, the air was irradiated with the fluorescent lamps described in Section III. Spectra were taken every 30 minutes for 450 minutes and analyzed. The following products were identified by their characteristic peaks: HCOOH at 1105 cm^{-1}, CO at 2170 cm^{-1}, CH_3COOH at 1180 cm^{-1}, and CH_3CHO at 1760 cm^{-1}. The absorbances at the frequencies chosen were measured and plotted against the irradiation time as shown in Figure 2. The absorbances, instead of the concentrations, were plotted against the irradiation time because the absorption coefficients of CH_3COOH at 1180 cm^{-1} and CH_3CHO at 1760 cm^{-1} were not available. However, the concentrations of HCOOH and CO after 450 minutes of irradiation were calculated to be 0.093 ppm and 0.525 ppm, respectively, using the absorption coefficients listed in Table 1.

VI. THE OZONE DARK DECAY:

An experiment often needed to characterize a reaction chamber is the determination of the ozone decay rate in the absence of light and other reactants. In the absence of other reactants, ozone is removed by the following process:

$$O_3 + \text{wall} \text{------>} \text{loss of } O_3 \qquad (2)$$

and hence

$$-\frac{d[O_3]}{dt} = k_1 [O_3] \qquad (3)$$

where k_1 is the rate constant for reaction (2).

Solving eq. (3) yields:

$$-\ln [O_3]_t = k_1 t - \ln [O_3]_o \qquad (4)$$

where $[O_3]_o$ and $[O_3]_t$ are the concentrations of ozone at time 0 and time t, respectively. A plot of $-\ln [O_3]_t$ versus t will give a straight line and the slope is the rate constant k_1.

A small amount of ozone was introduced into the reaction chamber and the chamber filled with the zero air at 760 torr and room temperature. Spectra was taken at t=0 and every 30 minutes after the lamps were turned on. The absorbance at 1055 cm^{-1} was measured at various irradiation times. A plot of $-\ln [O_3]_t$ versus t was made and the slope calculated as shown in Figure 3. The rate constant k_1 was determined to be 0.2369 hr^{-1} or 6.6×10^{-5} $sec.^{-1}$.

Background ozone decay rate constants have been determined elsewhere and usually were in the range of $(0.5-3.2) \times 10^{-5}$ $sec.^{-1}$ [25-27]. Our constant was higher due to the fact that our chamber is smaller and thus has a higher surface to volume ratio. The background ozone decay rate is negligible compared to the ozone decay rate in the presence of hydrocarbon reactants, which is greater than 1×10^{-3} $sec.^{-1}$ [25-27].

VII. PHOTOOXIDATION OF THE PROPENE-NO_x-AIR SYSTEM:

The photooxidation of the propene-NO_x-air-system has been studied most frequently and most extensively in both smog chamber experiments and computer simulations.[7-9, 12, 28, 29]. It is, therefore, an important model reaction of photochemical air pollution. We chose to study this system as a reference baseline to characterize our smog chamber.

Experiments with different initial concentrations of gases were carried out in the smog chamber: (1) 3 ppm C_3H_6+1.5 ppm NO_2-rich NO_x in dry air, (2) 10 ppm C_3H_6 + 3 ppm NO_2-rich NO_x in dry air, (3) 3 ppm C_3H_6 + 1.5 ppm NO-rich NO_x in dry air, and (4) 10 ppm C_3H_6 + 3 ppm NO-rich NO_x in dry air. Important products positively identified were CH_3CHO, CO, HCHO, HCOOH, O_3, and PAN. The concentration time profiles of these products

along with the reactants are shown in Figures 4-7. Small absorption bands showed up in some spectra at the characteristic frequencies of HNO_3, N_2O_5, and PGDN. The absorption bands were so small that their presence could not be absolutely certain, at least within the sensitivity of the instrument.

The general shapes of the profiles of each gas in Figures 4 and 5 (NO_2 -rich NO_x) and in Figures 6 and 7 (NO-rich NO_x) are about the same, except that the rise times for the products were shorter when the intitial concentrations of the reactants were higher. When Figures 4 and 6 (or 5 and 7) are compared, it is obvious that when the initial NO_x was rich in NO, C_3H_6 was depleted and all the products began to show up after a longer irradiation time. The familiar profiles of NO and NO_2 were clearly observed in Figures 6 and 7.

Many mechanisms have been proposed and computer simulations performed on the propene-NO_x-air system to explain and predict the concentration-time behavior of the reactants and major products.[8-12,29] Although the proposed mechanisms and the computer models are different in details, they seem to contain the following important steps:

1. The initial reactions of NO, NO_2, O_3, explaining the depletion of NO, the rise and fall of NO_2, and the formation of O_3:

$$2\ NO + O_2 \longrightarrow 2\ NO_2 \qquad (5)$$

$$NO_2 + h \longrightarrow NO + O \qquad (6)$$

$$O + O_2 + M \longrightarrow O_3 + M \qquad (7)$$

$$O_3 + NO \longrightarrow NO_2 + O_2 \qquad (8)$$

2. The reactions for the formation of N_2O_5 and HNO_3:

$$O_3 + NO_2 \longrightarrow NO_3 + O_2 \qquad (9)$$

$$NO_3 + NO_2 \longrightarrow N_2O_5 \qquad (10)$$

$$N_2O_5 + H_2O \longrightarrow 2\ HNO_3 \qquad (11)$$

3. It has generally been recognized that OH attack on the propene is the most important pathway in the propene consumption. There may be as many as five possible pathways for the reaction of OH with propene, but the following pathway is believed to be the most important:

$$CH_3CH{=}CH_2 + \dot{O}H \rightarrow CH_3\dot{C}HCH_2OH \quad (12)$$

$$CH_3\dot{C}HCH_2OH + O_2 \rightarrow CH_3CH\dot{O}_2CH_2OH \quad (13)$$

$$CH_3CH\dot{O}_2CH_2OH + NO \rightarrow CH_3CH\dot{O}CH_2OH + NO_2 \quad (14)$$

$$CH_3CH\dot{O}CH_2OH \rightarrow CH_3CHO + \dot{C}H_2OH \quad (15)$$

$$\dot{C}H_2OH + O_2 \rightarrow HCHO + HO_2\cdot \quad (16)$$

4. Another important pathway for propene consumption is via ozone attack. Addition and fragmentation reactions are possible. The addition reaction is known as the Criegee pathway:

$$CH_3CH{=}CH_2 + O_3 \rightarrow CH_3CHO + H\dot{C}HO\dot{O} \quad (17)$$

$$H\dot{C}HO\dot{O} + NO \rightarrow HCHO + NO_2 \quad (18)$$

$$H\dot{C}HO\dot{O} \rightarrow HCOOH \quad (19)$$

$$CH_3CH{=}CH_2 + O_3 \rightarrow CH_3\dot{C}HO\dot{O} + HCHO \quad (20)$$

$$CH_3\dot{C}HO\dot{O} + NO \rightarrow CH_3CHO + NO_2 \quad (21)$$

The following reactions show the fragmentation reactions:

$$CH_3CH{=}CH_2 + O_3 \rightarrow HCHO + CH_3\cdot + HO_2\cdot + CO_2 \quad (22)$$

$$CH_3CH{=}CH_2 + O_3 \rightarrow CH_3CHO + 2\ HO_2\cdot + CO_2 \quad (23)$$

5. The third pathway for the propene consumption is the attack of O atoms on the propene:

$$CH_3CH{=}CH_2 + O \longrightarrow CH_3\overline{CHCH_2O} \longrightarrow C_2H_5CHO \quad (24)$$

$$CH_3CH{=}CH_2 + O \longrightarrow CH_3CO\cdot + CH_3\cdot \quad (25)$$

$$CH_3CO\cdot + O_2 \longrightarrow CH_3CO_3\cdot \quad (26)$$

$$CH_3CO_3\cdot + NO \longrightarrow CH_3CO_2\cdot + NO_2 \quad (27)$$

$$CH_3CO_3\cdot + NO_2 \longrightarrow CH_3CO_3NO_2\,(PAN) \quad (28)$$

VIII. RECOMMENDATIONS:

The results from the three types of experiments performed in the smog chamber generally were in agreement with the results obtained from other smog chambers in terms of the products found as well as the concentration-time behavior of the major species. The smog chamber system, therefore, has the capability and sensitivity to be used to perform most types of photochemical smog simulations. The experiences we have gained from these experiments will allow us to continue working on research in the area of atmospheric photochemistry.

Our recommendations and suggestions for follow-on research are as follows:

1. Incorrect absorption coefficients used to calculate concentration are the most likely source of error in atmospheric infrared studies. Different values were obtained from different systems and they may differ by as much as an order of magnitude. Therefore, absorption coefficients should be determined experimentally for the RDVC smog chamber.

2. It has been generally agreed that OH· radicals initiate photochemical reactions of hydrocarbons and are the major attacking species in these reactions. Therefore, we suggest that photochemical reactions be carried out with OH· radicals. OH· radicals can be generated most conveniently by the photolysis of gaseous nitrous acid, HONO.[30].
3. For the follow-on research, we plan to continue working in the area of atmospheric photochemistry. We would propose specific research projects in which the Air Force has an interest. This would be done through a mini-grant proposal using the facilities at the Environmental Sciences Research Laboratory of the EPA in Research Triangle Park, N.C., or at the Department of Environmental Sciences and Engineering of the University of North Carolina at Chapel Hill, N.C.

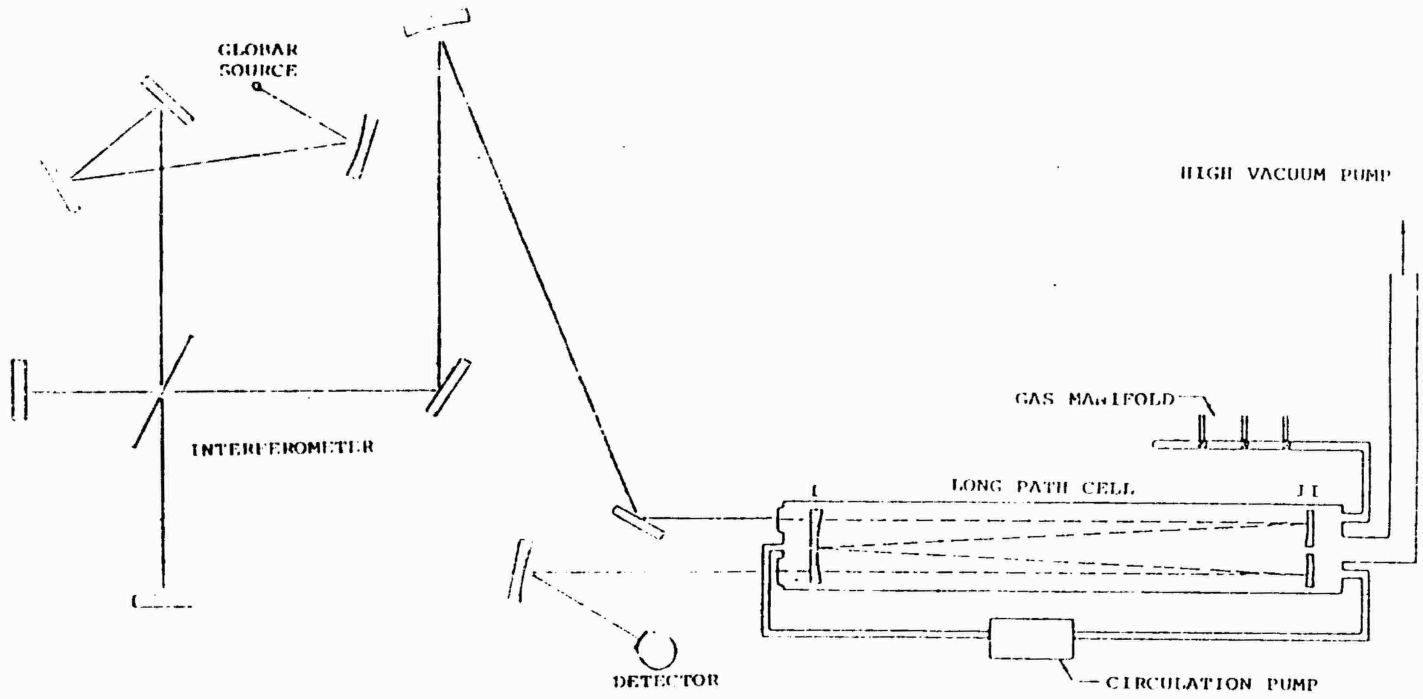


Figure 1. Schematic Diagram of the Smog Chamber System.

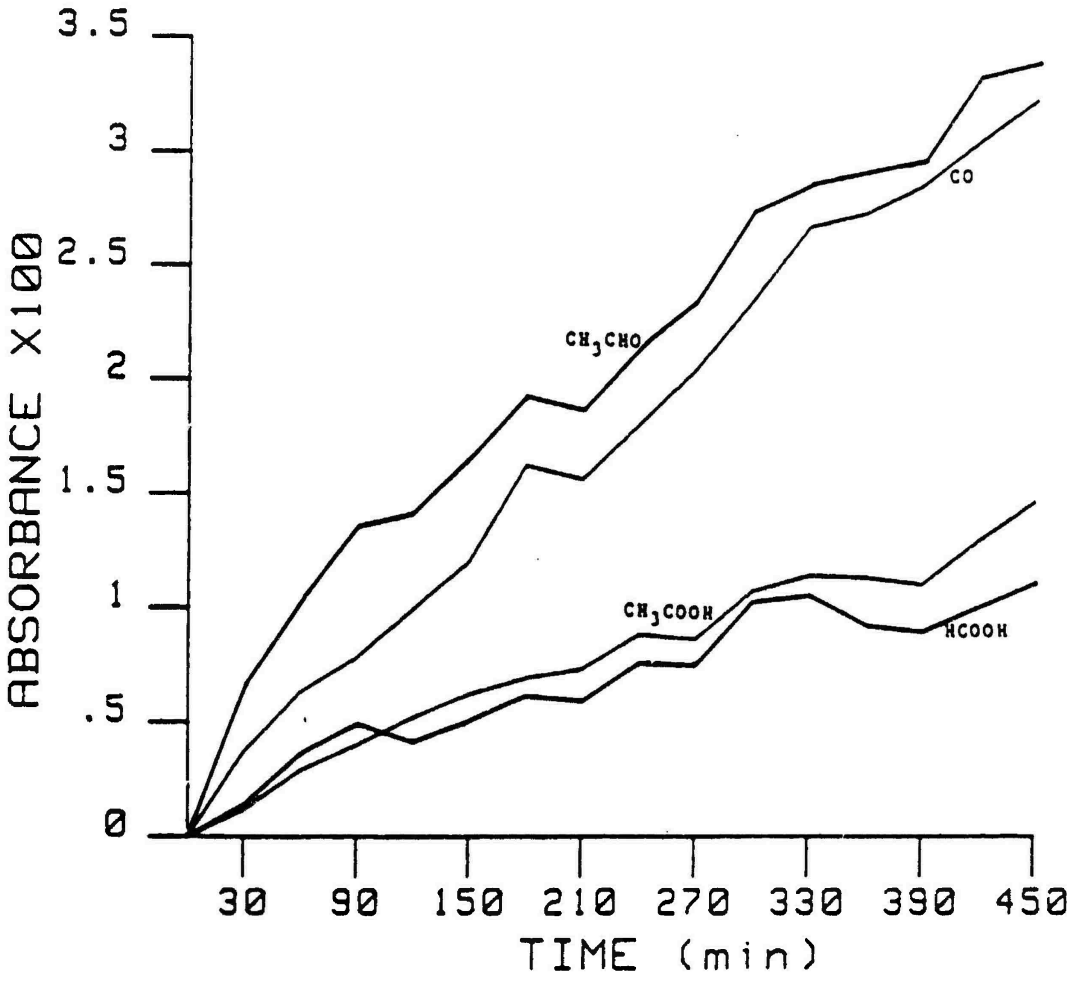


Figure 2. Absorbances of Products vs. Irradiation Time for the Background Air.

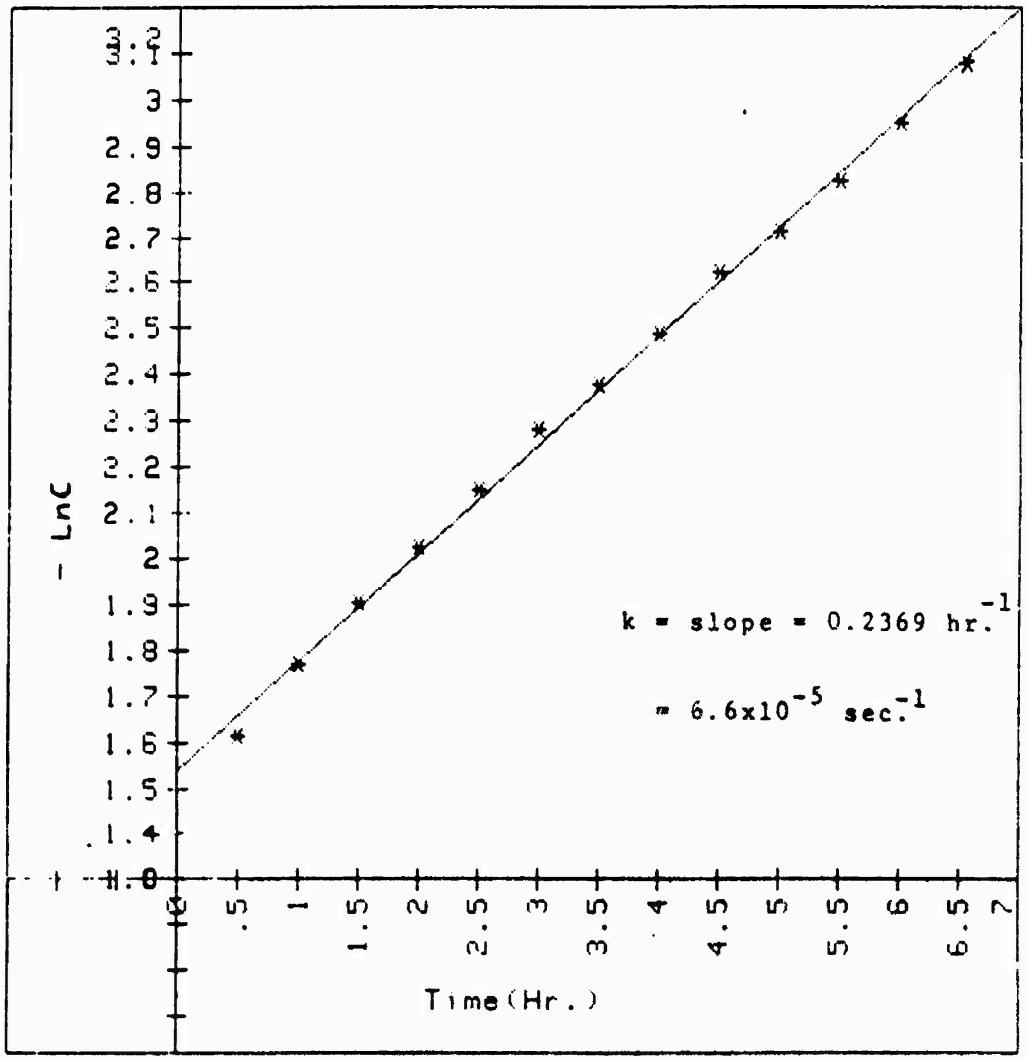


Figure 3. Plot of $-\ln[O_3]$ vs. Irradiation Time for Ozone Dark Decay.

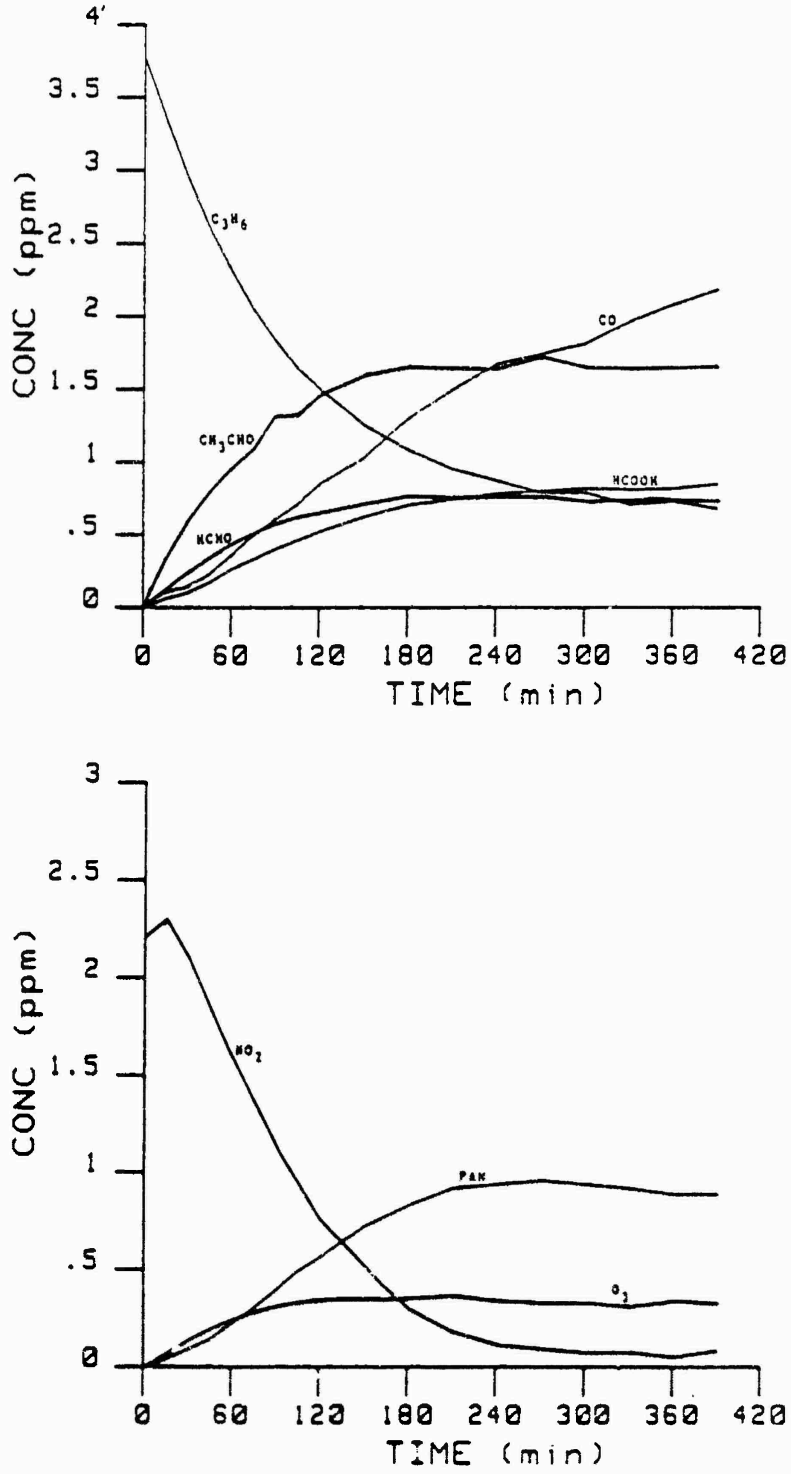


Figure 4. Concentrations of Reactants and Products vs. Irradiation Time for 3 ppm C_3H_6-1.5 ppm NO_x (NO_2 rich)-Air System.

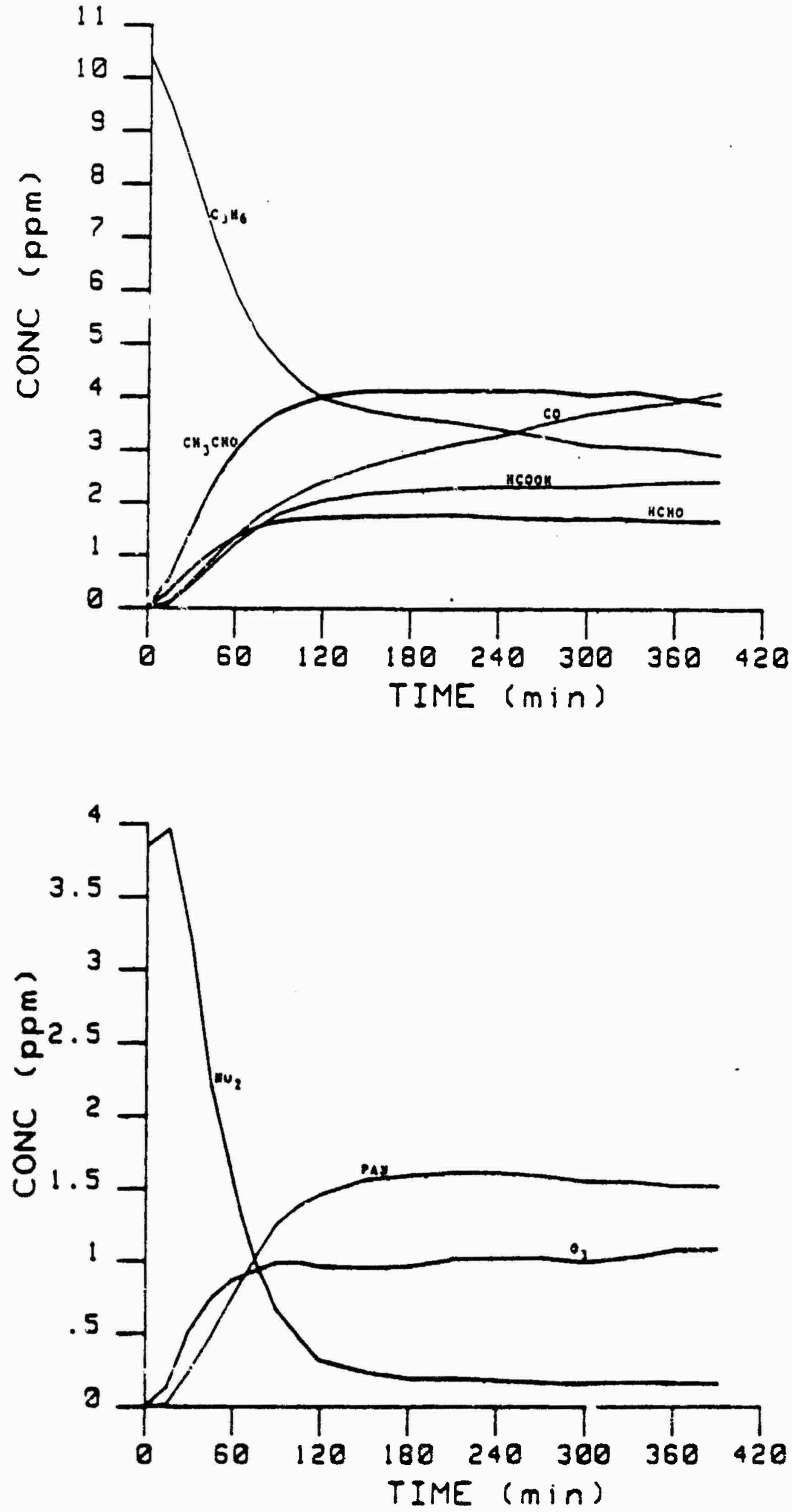


Figure 5. Concentrations of Reactants and Products vs. Irradiation Time for 10 ppm C_3H_6-3 ppm NO_x (NO_2 rich)-Air System.

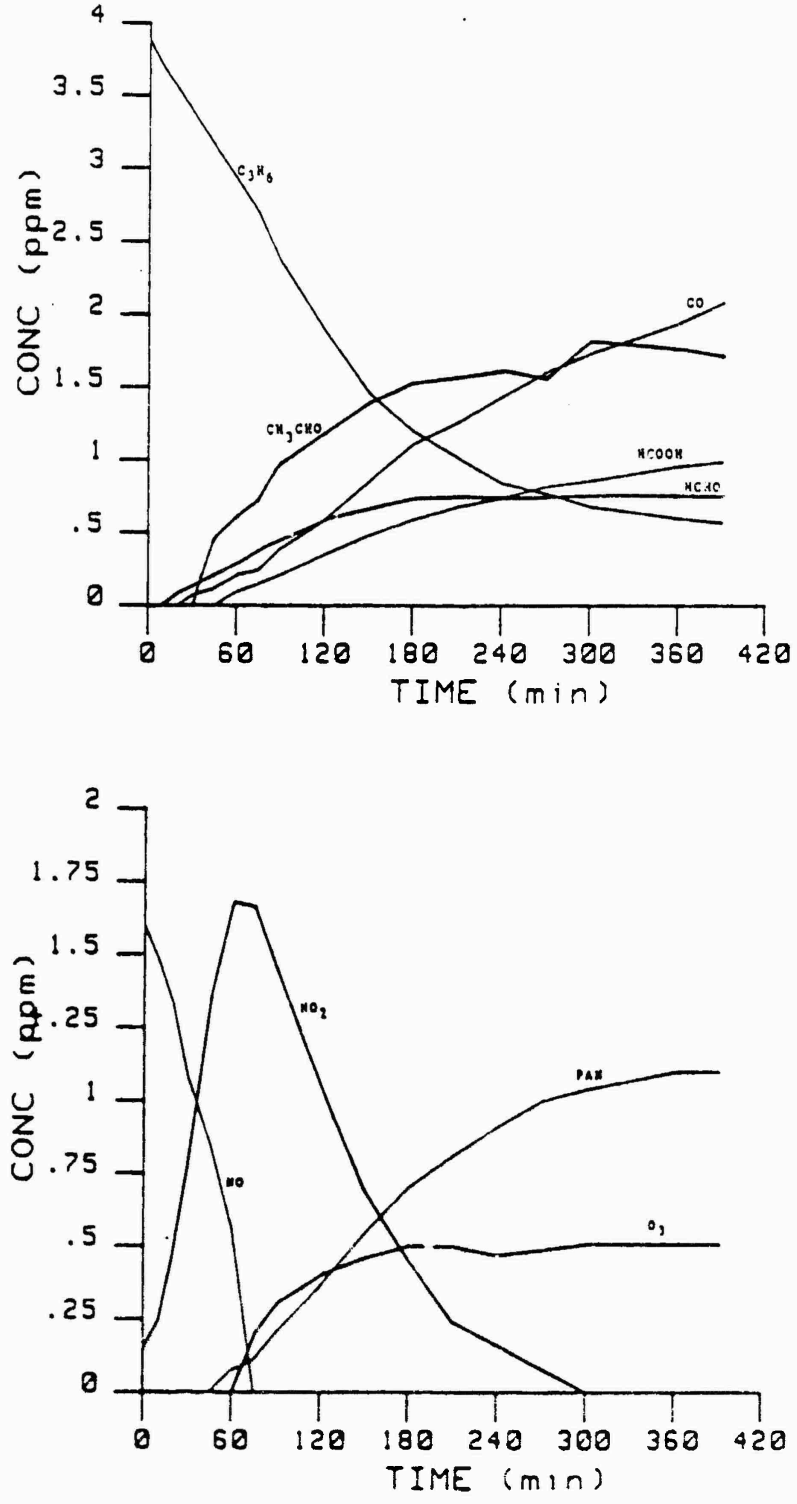


Figure 6. Concentrations of Reactants and Products vs. Irradiation Time for 3 ppm C_3H_6-1.5 ppm NO_x (NO rich)-Air System.

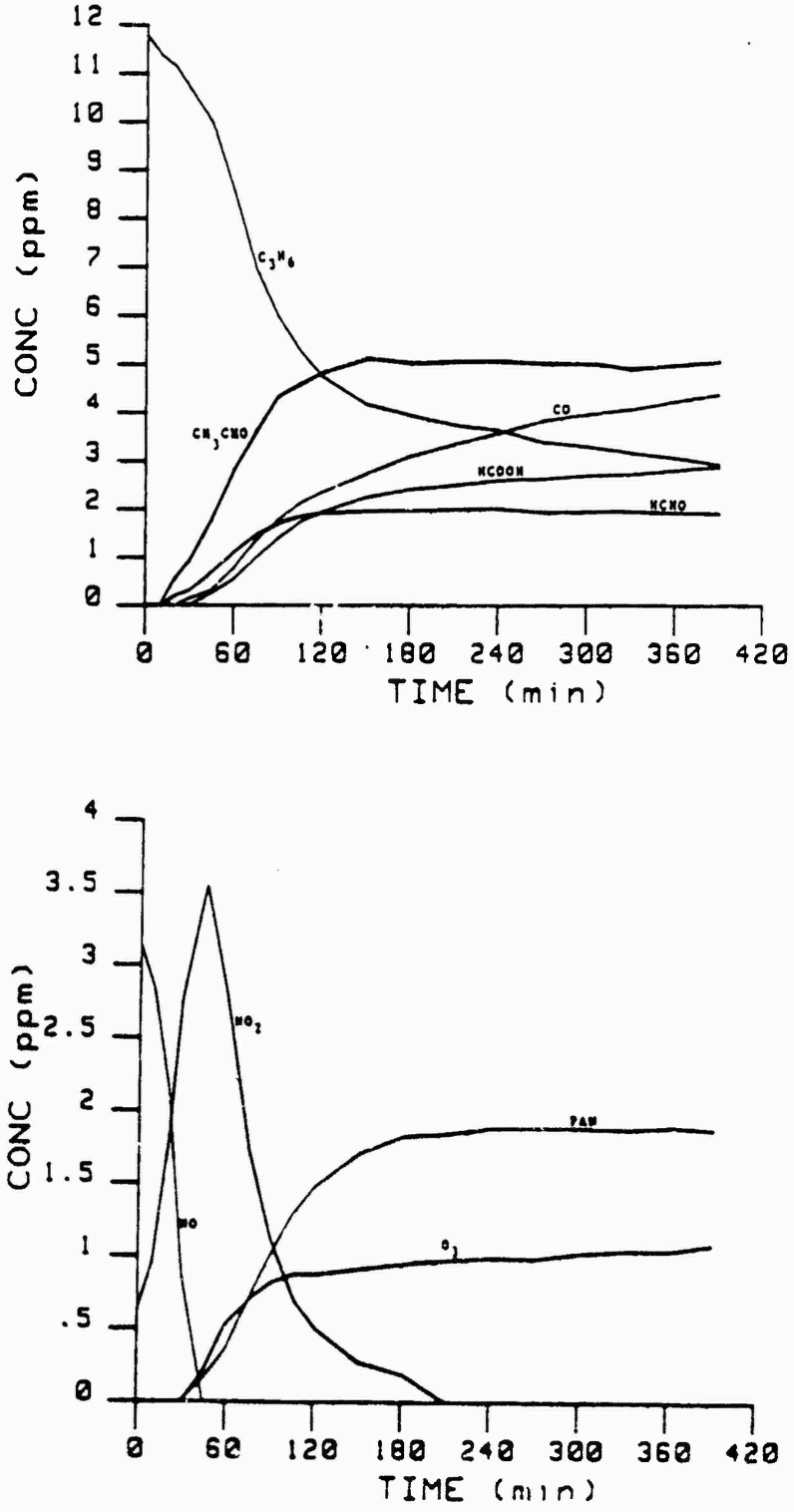


Figure 7. Concentrations of Reactants and Products vs. Irradiation Time for 10 ppm C_3H_6-3 ppm NO_x (NO rich)-Air System.

REFERENCES

1. Schuck, E.A. and G.J. Doyle, Photooxidation of Hydrocarbons in Mixtures Containing Oxides of Nitrogen and Sulfur Dioxide, Report No. 29, Air Pollution Foundation, San Marino, California, 1959.
2. Leighton, P.A., Photochemistry of Air Pollution, Academic Press, New York, 1961.
3. Altshuller, A.P. and J.J. Bufalini, "Photochemical Aspects of Air Pollution: A Review", Environ. Sci. Technol., Vol 5, pp. 39-64, 1971.
4. Finlayson-Pitts, B.J. and J.N. Pitts, Jr., "The Chemical Basis of Air Quality: Kinetics and Mechanisms of Photochemical Air Pollution and Application to Control Strategies", in Advances In Environmental Science and Technology, Vol 7, pp. 25-162, J.N. Pitts, Jr. and R. Metcalf, Eds., Wiley-Interscience, New York, 1977.
5. Hanst, P.L., W.E. Wilson, R.K. Patterson, B.W. Gay, Jr., L.W. Chaney, and C.S. Burton, A Spectroscopic Study of California Smog, EPA Publication No. 650/4-75-006, Research Triangle Park, N.C., 1975.
6. Cox, R.A., K.F. Patrick, and S.A. Chant, "Mechanism of Atmospheric Photooxidation of Organic Compounds. Reactions of Alkoxy Radicals in Oxidation of n-Butane and Simple Ketones", Environ. Sci. Technol., Vol. 15, pp. 587-592, 1981.
7. Akimoto, H., H. Bandow, F. Sakamaki, G. Inoue, M. Hoshino, and M. Okuda, "Photooxidation of the Propylene-NO_x-Air System Studied by Long-Path Fourier Transform Infrared Spectrometry", Environ. Sci. Technol., Vol. 14, pp. 172-178, 1980.
8. Carter, W.P.L., A.C. Lloyd, J.L. Sprung, and J.N. Pitts, Jr., "Computer Modeling of Smog Chamber Data: Progress in Validation of a Detailed Mechanism for the Photooxidation of Propene and n-Butane in Photochemical Smog", Int. J. Chem. Kinet., Vol. 11, pp. 45-101, 1979.

9. Demerjian, K.L., J.A. Kerr, and J.G. Calvert, "The Mechanism of Photochemical Smog Formation" in Advances in Environmental Science and Technology, Vol. 4, pp. 1-262, J.N. Pitts, Jr. and R. Metcalf, Eds., Wiley-Interscience, New York, 1974.

10. Whitten, G.Z., Killus, J.P., and H. Hogo, Modeling of Simulated Photochemical Smog with Kinetic Mechanisms, Vol. 1, Final Report, EPA Publication No. 600/3-80-028a, 1980.

11. Calvert, J.G. and R.D. McQuigg, "Computer Simulation of the Rates and Mechanisms of Photochemical Smog Formation", Int. J. Chem. Kinet, Symp., No. 1, pp. 113-154, 1975.

12. Hecht, T.A. and J.H. Seinfield, "Development and Validation of a Generalized Mechanism for Photochemical Smog", Environ. Sci. Technology, Vol. 6, pp. 47-57, 1972.

13. Bufalini, J.J., T.A. Walter, and M. M. Bufalini, "Contamination Effects on Ozone Formation in Smog Chambers", Environ, Sci. Technol., Vol.11, pp. 1181-1185, 1977.

14. Tuazon, E.C., A.M. Winer, R.A. Graham, and J.N. Pitts, Jr., "Atmospheric Measurements of Trace Pollutants by Kilometer-Pathlength FT-IR Spectroscopy", in Advances In Environmental Science and Technology, Vol. 10,. pp. 259-299,J.N. Pitts, Jr. and R.L. Metcalf, E ., Wiley-Interscience, New York,1980.

15. Hanst, P.L., A.S. Lefohn, and B.W. Gay, Jr., "Detection of Atmospheric Pollutants at Part-per-Billion Levels by Infrared Spectroscopy", Appl. Spectroscopy, Vol. 27, pp. 188-198, 1973.

16. Gay, B.W., Jr., P.L. Hanst, J.J. Bafalini, and R.C. Noonan, "Atmospheric Oxidation of Chlorinated Ethylenes", Environ. Sci. Technology, Vol. 10, pp. 58-67, 1976.

17. Hanst, P.L., N.W. Wong, and J. Bragin, "A Long-Path Infrared Study of Los Angeles Smog", Atmospheric Environmental, Vol. 16, pp. 969-981, 1982.

18. Stone, D.A., The Autooxidation of Monomethylhydrazine Vapor, Engineering and Services Laboratory Report No. ESL-TR-79-10, Air Force Engineering and Services Center, Tyndall AFB, FL, April, 1979.

19. White, J.L., "Long Optical Paths at Large Aperture", J. Opt. Soc. Amer., Vol. 32, p. 285, 1942.

20. Tuesday, C.S., "The Atmospheric Photooxidation of Trans-2-Butane and Nitric Acid", in Chemical Reactions in the Lower and Upper Atmosphere, pp. 1-49, Interscience, New York, 1961.

21. Stephens, E.R., "Absorptivities for the Infrared Determination of Peroxyacyl Nitrates", Analyt. Chem., Vol. 36, pp. 928-929, 1964.

22. Private Communication with B.W. Gray, Jr., of the Environmental Sciences Research Laboratory, Environmental Protection Agency, Research Triangle Park, N.C.

23. Bufalini, J.J., S.L. Kopczynski, and M.C. Dodge, "Contaminated Smog Chambers in Air Pollution Research", Environmental Letters, Vol. 3, pp. 101-109, 1972.

24. Bufalini, J.J., T.A. Walter, and M.M. Bufalini, "Contaminated Effects On Ozone Formation in Smog Chambers", Environ. Sci. Technol., Vol. 11, pp. 1181-1185, 1977.

25. Atkinson, R., S.M. Aschmann, A.M. Winer, and J.N. Pitts, Jr., "Rate Constants for the Gas Phase Reactions of O_3 with a Series of Carbonyls at 296 K", Int. J. Chem. Kinet, Vol. 13, p. 1133, 1981.

26. Atkinson, R., S.M. Aschmann, D.R. Fitz, A.M. Winer, and J.N. Pitts, Jr., "Rate Constants for the Gas-Phase Reactions of O_3 With Selected Organics at 296 K", Int. J. Chem. Kinet., Vol. 14, pp. 13-18, 1982.

27. Atkinson, R., A.M. Winer, and J.N. Pitts, Jr., "Rate Constants for the Gas-Phase Reactions of O_3 with the Natural Hydrocarbons Isoprene and α-and β-Pinene", Atmospheric Environmental, Vol. 16, pp. 1017-1020, 1982.

28. Altshuller, A.P., S.L. Kopczynski, W.A. Lonneman, T.L. Becker, and R. Slater, "Chemical Aspects of the Photooxidation of the Propylene-Nitrogen Oxide System", Environ. Sci. Technol., Vol. 1, pp. 899-914, 1967.

29. Sakamaki, F., M. Okuda, H. Akimoto, and H. Yamazaki, "Computer Modeling Study of Photochemical Ozone Formation in the Propene-Nitrogen Oxides-Dry Air System. Generalized Maximum Ozone Isopleth", Environ. Sci. Technol., Vol. 16, pp. 45-52, 1982.

30. Cox, R.A., "The Photolysis of Gaseous Nitrous Acid," J. Photochem., Vol.3, pp. 175-188, 1974.

1983 USAF - SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

PARTITIONING EQUILIBRIA OF VOLATILE POLLUTANTS IN THREE PHASE SYSTEMS

Prepared By: LEONARD W. LION / DOUG GARBARINI

Academic Rank: ASSISTANT PROFESSOR / RESEARCH ASSISTANT

Department and University: DEPARTMENT OF ENVIRONMENTAL ENGINEERING
CORNELL UNIVERSITY

Research Location: AIR FORCE ENGINEERING AND SERVICES CENTER
ENGINEERING AND SERVICES LABORATORY, ENVIRONICS DIVISION,
ENVIRONMENTAL CHEMISTRY BRANCH
TYNDALL AFB, FL.

USAF Research Colleague: MR. THOMAS B. STAUFFER

Date: August 8, 1983

Contract NO: F49620-82-C-0035

PARTITIONING EQUILIBRIA OF VOLATILE POLLUTANTS IN THREE PHASE SYSTEMS

by

Leonard W. Lion and Doug Garbarini

ABSTRACT

Volatile pollutants in unsaturated ground water systems are partitioned between the gaseous, aqueous, and solid phases. Equilibria between aqueous solution, the atmosphere, and a solid adsorbent are evaluated in this research. An equilibrium head space technique is employed to determine Henry's Law constants and activity coefficients for components of JP-4 Fuels and for trichloroethylene. Sorption equilibria for these same compounds are also determined on aluminum oxide (Al_2O_3) with and without humic acid coating materials, on humic acids in the absence of Al_2O_3, and on naturally occurring Air Force base soils.

Acknowledgement

The authors thank the Air Force Systems Command, the Air Force Office of Scientific Research and the Southeastern Center For Electrical Engineering Education for providing both the opportunity and financial support for our research on a challenging problem of environmental concern. In particular we would like to thank Mr. Thomas Stauffer of the Environics Laboratory at Tyndall Air Force Base for lending his valuable assistance and technical expertise in the separation and analysis of the compounds of concern. We also thank Don Wickman and Sergeants Stacey Brown, Chuck Manikas, Bill Nemeier and Dan Stork and the rest of the personnel at the Environics Laboratory who went out of their way to assist us in this research effort.

I. INTRODUCTION:

The fate of various pollutant organic compounds introduced into groundwater systems is a pressing and ubiquitous environmental problem in many areas of the United States. This problem is of concern to the U.S. Air Force since groundwater contamination has been detected on or near several Air Force Bases.

The U.S. Air Force is currently conducting an active research program designed to both investigate feasible treatment strategies for groundwater clean up and to study the reactions of groundwater pollutants which may control their transport. Two classes of nonionic organic pollutants are currently among those of concern to the Air Force: (1) those typically found in cleaning solvents and degreasers such as trichloroethylene (TCE) and (2) those found in jet fuels, such as toluene and naphthalene. These materials may be introduced into soils and ultimately groundwater aquifers in a number of different ways including through spills, and rupture or leaking of underground storage containers.

In unsaturated aquifer systems contaminant compounds may be removed from solution by several mechanisms including volatilization, sorption and biodegradation. The first two of these processes were examined in the research which was conducted. The reactions are illustrated in schematic fashion in figure 1 and are briefly discussed below.

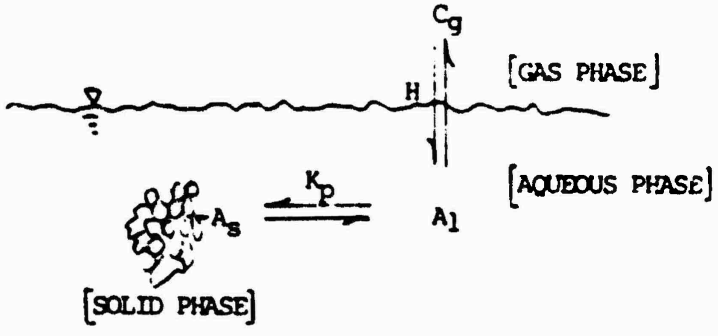


Figure 1: Potential reactions for a nonionic organic pollutant compound in a three phase abiotic system.

A. Volatilization

Gas exchange equalibria for compounds are governed by Henry's Law:

$$H = C_g/A_l \quad (1)$$

where H is the Henry's Law constant

C_g is gas concentration

and A_l is the activity of the compound in solution.

Compound activity and concentration in solution are interrelated through the activity coefficient:

$$A_l = \gamma C_l \quad (2)$$

where γ is the activity coefficient

and C_l is the aqueous solution concentration.

Activity coefficients for neutral molecules, such as those considered in this study, are commonly related to solution ionic strength by the empirical equation:

$$\log \gamma = kI \quad (3)$$

where k is referred to as the salting coefficient

and I is the ionic strength of solution[1].

To the authors knowledge, activity coefficients and salting coefficients for many of the constituents of Air Force jet fuels and degreasing compounds have not been determined, however it is clear that these values may need to be known in order to precisely describe both gas exchange and sorption equilbria.

Henry's Law Constants for compounds are often reported as either dimensionless numbers (this convention has been adopted in this report) or with units of atm-l/mole; the two values being interrelated by the universal gas constant and temperature (°K) as follows:

$$H \text{ (dimensionless)} = H \text{ (atm-l/mole)}/RT \quad (4)$$

Henry's Law Constants are sometimes estimated from reported values for compound solubility in water and vapor pressure.

$$H \text{ (dimensionless)} = 16.04\ PM\ /\ S\ T \quad (5)$$

where P = vapor pressure in mm Hg

M = molecular weight (grams)

S = solubility (mg/l)

T = absolute temperature (°K)

Unfortunately, reported values for solubility of many organic compounds can often vary over an order of magnitude with a concomitant uncertainty in the calculated value of H. Accurate values for Henry's Law Constants are essential to the description of the fate of volatile compounds in unsaturated groundwater aquifers.

B. Sorption

Available data for the sorption of nonionic organic compounds typically reveal linear sorption isotherms[2,3,4] and therefore the equilibrium between solution and the adsorbing solid may be described by a simple partitioning coefficient:

$$Kp = \Gamma / A_1 \tag{6}$$

where Kp is the partitioning coefficient and Γ is adsorption density (mg adsorbed/g adsorbent).

It is noteworthy that such a linear relationship results from the Langmuir adsorption isotherm:

$$\Gamma = \frac{\Gamma_{max} A_1}{K + A_1} \tag{7}$$

at low solution concentrations (ie when $A_1 >> K$). A linear relationship also results from the Freundlich adsorption isotherm:

$$\Gamma = K A_1^n \tag{8}$$

when $n = 1$.

At soil ogranic carbon concentrations > 1% organic solute equilibria appear to be controlled by the organic content of the adsorbent[5] and therefore the observed partitioning coefficient may be normalized to the weight fraction of organic carbon as:

$$Koc = Kp/f \tag{9}$$

where f is the weight fraction of organic carbon on the soil

and Koc is the normalized partitioning coefficient.

Although investigators are in agreement on the importance of soil organic matter, the mechanism of the binding reaction between the organic solutes and the solid phase is uncertain. Both the terms "sorption," which is frequently used to refer to a surface reaction, and "partitioning," which refers to an extraction phenomena, have been employed to describe the binding process[2,4,6,7].

Observed Koc values have been related to compound solubilities in water and to compound octanol-water partitioning coefficients (Kow)[2,4,6]. An implication of these results is that sorption of nonionic organic pollutants on a given aquifer material may be predicted solely on the basis of the sorbent's organic content. In addition, these results suggest that the specific nature and form of the soil organic material may not be relevant to sorption equilibria. These implications have not been rigorously tested.

II. OBJECTIVES OF THE RESEARCH EFFORT

Major research objectives were as follows:

1. To evaluate the applicability of an equilibrium head space technique to the determination of Henry's Law Constants for individual compounds typical of JP-4 fuels. Initial method verification was performed using TCE, a compound for which the experimental procedure had previously been successfully employed. TCE was also included as a compound of interest in subsequent experiments in which compound activity coefficients and sorption equilibria were evaluated.

2. To determine the effect of solution composition on compound activity in solution. Variables examined included the effect of compound concentration, the effect of the presence of additional compounds in solution - both polar and nonpolar, the effect of ionic strength, and the effect of electrolyte composition.

3. To examine the applicability of a modification of the equilibrium headspace technique to the measurement of compound sorption from solution and to study the influence of organic carbon on compound sorption behavior. Variables considered included the effect of the presence of an organic (e.g. humic acid) coating on compound sorption by a hydrous oxide (Al_2O_3), the effect of the form of the organic material (e.g. humic acids coated on Al_2O_3 *vs.* humic acids in the absence of Al_2O_3) and the effect of the nature of the organic material (varied by carrying out sorption experiments on both synthetic and naturally occurring soils).

III. GAS EXCHANGE EQUILIBRIA AND MEASUREMENT OF ACTIVITY COEFFICIENTS

A. Introduction

Recently equilibrium partitioning in closed systems (EPICS) has been shown by Lincoff and Gossett[8] to be a reliable precedure for the determination of Henry's Law Constants for compounds with dimensionless Henry's constants < 3. For two identical bottles (1 and 2) containing the same compound mass but with different liquid and gas volumes, Lincoff and Gossett[8] show that:

$$H = \frac{(C_{g_1}/C_{g_2})V_{l_1} - V_{l_2}}{V_{g_2} - (C_{g_1}/C_{g_2})V_{g_1}} \tag{10}$$

where C_{g_1} and C_{g_2} are compound concentrations in the bottle headspaces V_{l_1}, and V_{l_2} are the liquid volumes in the bottles. and V_{g1} and V_{g2} are the gas volumes in each bottle.

Thus, Henry's Law Constant may be determined by G.C. headspace analysis of gas concentration, and is independent of the compound mass employed in the experiment.

For compounds of low volatility or high molecular weight, G.C. headspace analysis may not be feasible. However, liquid phase analysis may also be employed since for the same two identical bottles, it may be shown that:

$$H = \frac{(C_{l_1}/C_{l_2})V_{l_1} - V_{l_2}}{V_{g_2} - (C_{l_1}/C_{l_2})V_{g_1}} \tag{11}$$

Equations 10 and 11 are written assuming the aqueous phase behaves as an ideal solution and therefore $\gamma = 1$. However, the EPICS technique may also be employed to evaluate compound activity coefficients.

For two identical bottles containing the same compound mass and the same liquid and gas volume; but in only one of which (bottle 1) the compound behaves ideally, the following relationship applies:

$$M_T = C_{g_1}V_{g_1} + C_{l_1}V_{l_1} = C_{g_1} + \frac{C_{g1}}{H} V_{l_1} \tag{12}$$

and

$$M_T = C_{g_2}V_{g_2} + C_{l_2}V_{l_2} = C_{g_2} + \frac{C_{g_2}}{\gamma H} V_{l_2} \tag{13}$$

Combining equations 12 and 13 and solving for γ gives:

$$\gamma = (V_1/H)\left[(C_{g_1}/C_{g_2})(V_g + V_1/H)\right]^{-1} \quad (14)$$

Therefore γ may be determined using the ratio of gas concentrations determined in the two bottles.

B. MATERIALS AND METHODS

The chemical constituents of JP-4 jet fuels have been analyzed and reported by Smith et al;[9] 77 major components were identified. A subset of 15 components of JP-4 were utilized by MacIntyre et al[10] for investigation of sorption equilibria on soils and are listed in table 1. Compounds from this list which were available at the USAF Environics Laboratory at Tyndall AFB were utilized in this study.

Table 1. Representative Components of JP-4 Fuels

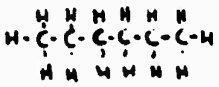
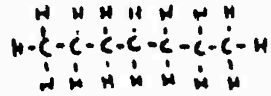
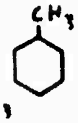
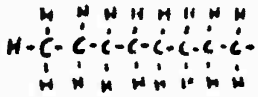
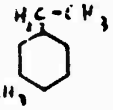
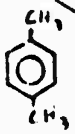
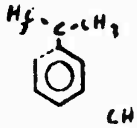
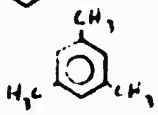
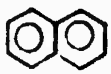
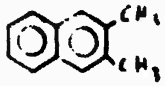
Compound	Structure	Molecular Weight (gm)	Vapor Pressure* (mm Hg)	Solubility (mg/l)	H** (dimensionless)
n-hexane		86.17	107[11] 120[12]	13[12]	41.2
cyclohexane		84.16	66.5[11] 77[12]	55[12]	6.0
n-heptane		100.20	32.8[11] 35[12]	2.4[12]	77.5
methylcyclohexane		98.19	34.25[11]	14[13]	13.2
toluene		92.13	20.9[11] 22.[11]	515[12]	0.27[13] (25°C)
n-octane		114.22	10.4[11] 11[12]	0.66[12]	101.5

Table 1 (Continued)

Compound	Structure	Molecular Weight (gm)	Vapor Pressure* (mm Hg)	Solubility (mg/l)	H** (dimensionless)
ethylcylohexane		112.2	9.4[11]		
p-xylene		106.2	6.5[12]	198[12]	0.19
isopropyl-benzene		120.2	3.4[11] 3.2[12]	50[12]	.42
1,3,5 trimethyl-benzene		120.2	2.1[11]	97.7[14]	.14
indan		118.17			
naphthalene		128.16	.053[11]	31.7[2]	.020[13] (25°C)
2-methylnaphthalene		142.2		25.4[2]	
n-tetradecane	$C_{14}H_{20}$	198.4	.024[11] .0022[12]		
2,3 dimethyl-naphthalene		156.2			

* vapor pressures reported at 20°C

** H calculated from average reported vapor pressure and solubility unless otherwise noted.

Henry's Law Constants were measured using glass hypo vials sealed with teflon rubber laminated discs (Pierce Chemical Co.). Discs and bottles were precleaned by washing with distilled water and heating overnight at 105°C. Two bottle sizes were employed. Bottle volumes were determined by weighing when full of water and averaged 156.10± 0.24 ml (n=9) and 120.98 ±0.78 ml (n=12). Liquid volumes normally employed in the 156.1 ml bottle size were 10.0 and 125 ml. Typically, five or more replicate bottles of each liquid volume were equilibrated over at least a 24 hour period in a circulating water bath at 25°C. Other experimental procedures were comparable to those described by Lincoff and Gossett[8]. Henry's Law Constants and activity coefficients for TCE and toluene were determined by removing of 1.00 ml of bottle headspace with a gas tight syringe and analysis with a Perkin Elmer model 900 gas chromatograph. Concentrations were measured isothermally at temperatures of approximately 135°C for TCE and 175°C for toluene using a Carbopack (60/80) with 1% SP-1000 column (Supelco, Inc). Compound concentration was measured as peak area determined by interface of the GC signal with a Hewlett Packard model 3388A Integrator. A linear relationship between gas concentration and GC peak area was established by analysis over a range of gas concentrations. A typical calibration curve is shown in figure 2.

Henry's Law Constants for naphthalene, indan, 2-3 dimethylnaphthalene, and 1,2,3 trimethylbenzene were measured by UV analysis of the aqueous solution using a Cary 219 spectrophotometer. Analytical wavelengths were 219.2nm for napthalene, 247nm for indan, 201.5 nm for 2,3 dimethylnaphthalene, and 196nm for 1,2,3 trimethylbenzene. With the exception of 1,2,3 trimethylbenzene, wavelengths employed correspond to those at which maximum absorbance was determined to occur by scan of the UV region. UV absorbance for napthalene and indan was determined to be directly proportional to solution concentration by analysis of standard decane solutions, an example calibration is shown in Figure 3.

FIGURE 2. G.C. Calibration Curve For Tolnene Headspace Analysis

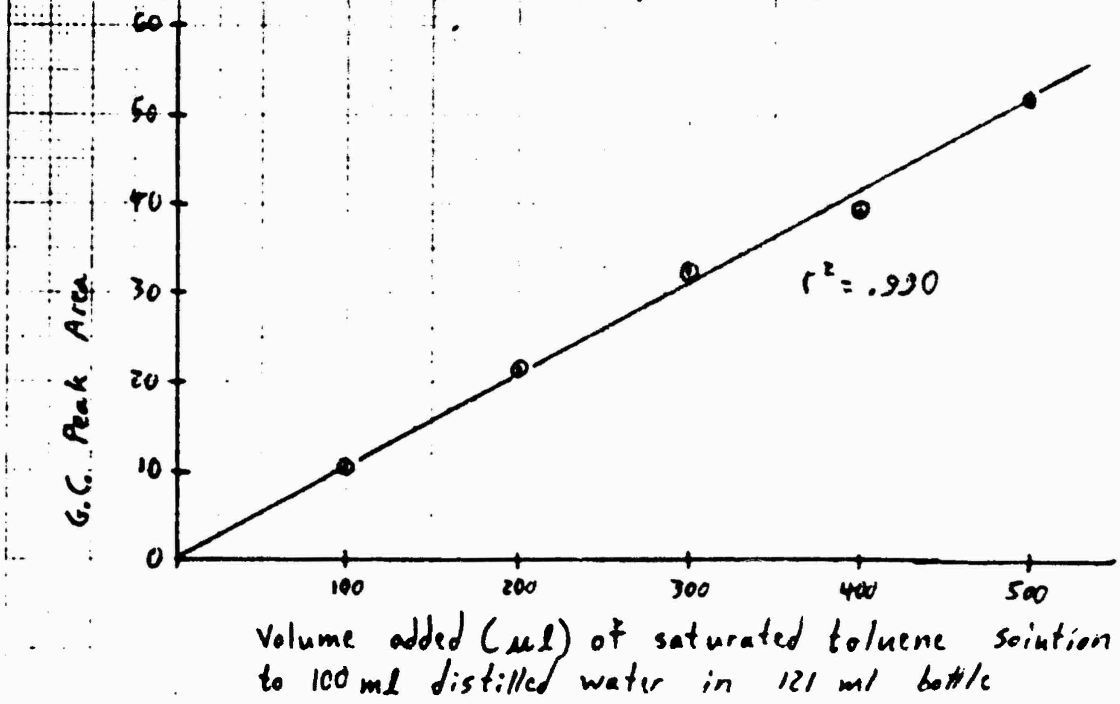
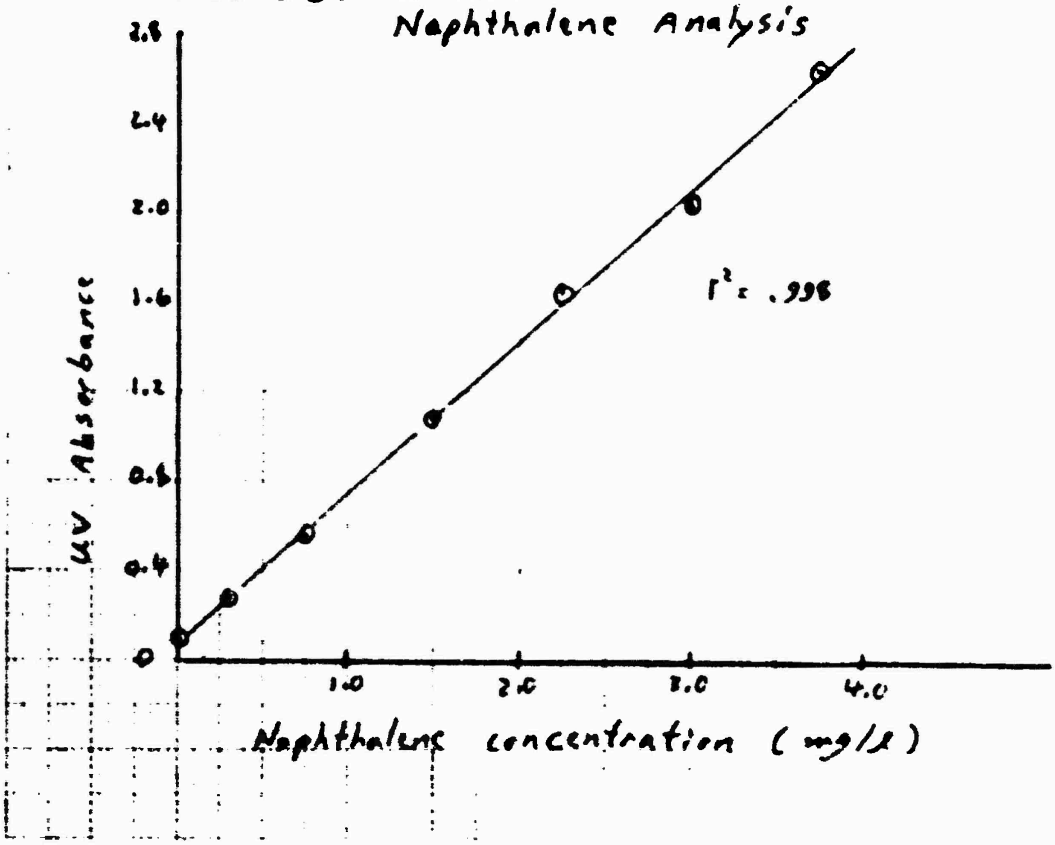


FIGURE 3. UV Calibration Curve For Naphthalene Analysis



After equilibration, samples for aqueous phase analysis were uncapped, and an aliquot was rapidly transferred to a stoppered quartz cuvette for measurement of UV absorbance. The compounds handled in this manner all had dimensionless H values $\leq$ .04 and therefore aqueous concentration changes caused by volatilization are likely to have been minimal.

C. RESULTS AND DISCUSSION

The Henry's Law Constants determined in this study are given in table 2.

Table 2. Henry's Law Constants for TCE and JP-4 Fuel Components

Compound	Measured Value @ 25°C	Literature Value	Reference	Remarks
TCE	.397	.411	8	reference value calculated for T=25°C from regression equation for H vs. temperature
toluene	.261	.270	13	
naphthalene	.0209	.0196	13	
2,3, dimethyl-napthalene	.0376			
indan	.0082			
1,2,3 trimethyl-benzene*	.086	.141		reference value calculated from reported values for vapor pressure[11] and solubility[14] of 1,3,5 trimethylbenzene

* the experimental compound used was technical grade (90% purity) and therefore the measured value for H is suspect.

The measured H value for TCE, is in good agreement with the value previously determined using the EPICS technique. H values measured for toluene and napthalene also agree closely with reported literature values. Neither water solubility and vapor pressure data from which H could be calculated nor measured H values for 2,3 dimethylnapthalene; indan; or 1,2,3 trimethylbenzene could be located by the authors for comparison with the measured values.

A Henry's Law constant was also measured for methylcyclohexane. The measured value (H = 9.81) was beyond the range in which the EPICS technique is sensitive (ie $H \leq 3$) and therefore is not considered to be reliable. Henry's Law constants calculated from reported values for vapor pressure and solubility for hexane, cyclohexane, heptane and octane are all greater than 6.0 (see table 1). It is therefore likely that Henry's Law Constants for these components of JP-4 Fuels can not be accurately determined using the EPICS procedure.

Activity coefficients obtained for TCE and toluene at four ionic strengths are given in table 3.

Table 3 - Activity Coefficients for TCE and Toluene*

compound	γ			
	I = .1	I = .3	I = .6	I = 1.0
TCE	1.055	1.16	1.32	1.58
toluene	1.054	1.14	1.31	1.62

* values obtained in NaCl electrolyte at 25°C

Based on these data, the calculated salting coefficients (k) for TCE and toluene are .197 and .207 respectively. Considering the precision of replicate GC analyses ($\simeq \pm$ 2%) the difference in these values is not considered to be statistically significant. Gossett[15] has previously reported a salting coefficient of .207 for TCE from KCl electrolyte which agrees closely with the measured value.

The activity of toluene in solution was measured over a threefold range in solution concentration and was unaffected over the toluene concentration range of 0.5 to 1.5 mg/l.

The activity of toluene in solution was also measured in the presence of other nonpolar (cyclohexane) and polar (acetone) solutes and was unaffected by the presence of these compounds.

Finally the influence of electrolyte composition on toluene activity was evaluated by measuring activity coefficients in $CaCl_2$ and $AlCl_3$ electrolytes with ionic strengths of .1 and .3; results are shown in table 4.

Table 4. Effect of Electrolyte Composition on Toluene Activity

Ionic Strength	Toluene Activity Coefficient		
	NaCl	$CaCl_2$	$AlCl_3$
0.10	1.054	1.022	1.043
0.30	1.136	1.077	1.057

Activity coefficients for toluene in both $AlCl_3$ and $CaCl_2$ were significantly lower than that determined for toluene in NaCl at I = 0.3. These data are consistent with reported values for other nonionic compounds which also vary with electrolyte composition[1].

D. SUMMARY

The EPICS technique has been demonstrated to be a reliable method for measurement of Henry's Law Constants and activity coefficients for many constituents of JP-4 Fuels. The procedure loses sensitivity for compounds with dimensionless H values $\geq$ 3 and is therefore not applicable to compounds such as hexanes, heptanes and octanes. For unsaturated compounds of low volatility UV analysis of the equilibrated aqueous solution is seen to be a viable analytical procedure. Gas exchange equilibria of volatile compounds is a function of their activity in the aqueous phase (see equation 1) and solution activity has been shown to be influenced by both electrolyte composition and concentration. Activity coefficients determined for TCE and toluene suggest that corrections for nonideality are likely to be negligible in most fresh water aquifers.

IV. SORPTION EQUILIBRIA

A. INTRODUCTION

At equilibrium, the gas concentration of a volatile compound in a closed bottle serves as a direct measure of the aqueous solution concentration providing the compound's Henry's Law Constant and activity coefficient are known (see equations 1 and 2). It is therefore reasonable to extend the EPICS technique to the study of sorption equilibria of volatile compounds such as toluene and TCE. A major advantage of this procedure is that separation of the solid from solution is no longer required in order to measure the equilibrium concentration in solution.

For compounds of low volatility, such as naphthalene, direct analysis of the solution is required, necessitating separation of solid from the aqueous phase.

B. MATERIALS AND METHODS

Sorption equilibria for toluene, naphthalene and TCE were determined on an artificial soil created by coating humic acids onto a hydrous aluminum oxide (Al_2O_3). Sorption equilibria were also measured for these compounds onto humic acids in the absence of Al_2O_3. Finally naphthalene and TCE sorption equilibria were evaluated using naturally occurring Air Force Base soils.

H 10. Al_2O_3 was added (approximately 450gm) and the pH lowered to approximately 4.0. After equilibration, excess humic acids and water were decanted. Solids were rinsed once with 1 liter distilled water acidified to pH = 4 and dried overnight at 105°C.

Air Force Base soil samples were obtained from cores, and shipped to the Environics Laboratory at Tyndall AFB wrapped in foil sealed in paraffin. Moisture contents for the soils were measured by weight loss after drying 108 hours at 105°C. The dried soils were ground and material passing through an 850u sieve was used in the sorption experiments.

The organic content of all adsorbents was characterized by dichromate oxidation using the Walkley-Black method for analysis of soil organic carbon[16]. Samples were analyzed in triplicate and the mean result is shown in table 5.

Table 5. Adsorbent Characteristics

Solid Phase	% organic carbon	density (g/cm^3)	moisture content	depth of core
Al_2O_3 (Fisher Scientific Co. adsorption alumina; 80-200 mesh)	varied from 0 to 0.54%	2.49 ± .055 (n = 5)		
Whiteman AFB soil Knobnoster, MO	0.40%	2.27 ± 0.148 (n = 6)	21.4%	3'-4.7'
Lowry AFB soil Denver, CO	0.11%	2.28 ± .049 (n = 3)	15.3%	5.3'-6.9'
Offutt AFB soil Omaha, NB	0.23%	2.50 ± .041 (n = 3)	22.2%	11'-13'
Minot AFB soil Minot, ND	0.12%	2.36 ± .078 (n = 3)	14.5%	10'-12'

Table 5. Adsorbent Characteristics (Continued)

Solid Phase	% organic carbon	density (g/cm^3)	moisture content	core depth
Humic Acid, ICN Pharmaceuticals Inc.	33.46%			
Humic Acid, Aldrich Chemical Co.	32.05%			

Since sorption equilibria were conducted in sealed bottles with a gas headspace (see discussion below) knowledge of adsorbent densities was required to determine the volume displaced by the solid. Solid densities were determined by measuring the water volume displaced by a weighed amount of soil in a 10 ml graduated cylinder and are summarized in table 5.

The protocol used for sorption experiments was as follows:

1. The calculation of compound mass remaining at equilibrium necessitated precise knowledge of the headspace in each bottle. The individual bottle volumes were determined by weighing when full with water. Results for the gas and/or liquid concentrations in individual bottle were normalized to a uniform standard bottle size using the following relationships:

$$C_{g(\text{normalized})} = C_{g(\text{observed})} \frac{(V_l + \Upsilon HV_g)_{\text{actual}}}{(V_l + \Upsilon HV_g)_{\text{std.}}} \qquad (15)$$

$$C_{l(\text{normalized})} = C_{l(\text{observed})} \frac{(V_l + \Upsilon HV_g)_{\text{actual}}}{(V_l + \Upsilon HV_g)_{\text{std.}}} \qquad (16)$$

2. A stock solution of the compound of interest was prepared. For toluene this was accomplished by delivering 500ml of toluene with an Eppendorf pipette to a 2 liter volume of water. The mass delivered was determined by weight to be .422 $\pm$ .001 gm (n = 4) giving a stock solution concentration of 211 mg/l. The toluene concentration in the stock solution was verified by micro-extraction as follows:

a) 1ml methylene chloride (MeCl) and 8.5 ml distilled water were delivered to a 10 ml glass bottle and sealed with a teflon disc.

b) 250μl of the toluene stock solution was delivered through the disk using a syringe while venting the bottle headspace gas through a second syringe needle.

c) The remaining bottle headspace was displaced by injecting distilled water into the bottle and venting the headspace gas as in (b).

d) bottle contents were shaken to extract toluene into MeCl and the MeCl phase analyzed for toluene by gas chromatography.

e) G.C. peak areas were calibrated against standard toluene solutions in MeCl prepared by delivering a weighed amount of toluene into MeCl and serial dilution.

The toluene concentration in the stock solution determined by micro-extraction was 214.1 mg/l $\pm$ 12.4 mg/l (n = 4). Based on this data it was assumed that toluene loss in preparation of the stock solution was negligible and a stock concentration of 211 mg/l was assumed. Aliquots of the stock solution were added to bottles containing solids for sorption experiments.

For naphthalene, a saturated aqueous stock solution was prepared by equilibrating an excess of naphthalene with water. Prior to each experiment an aliquot of stock solution was filtered through No. 5 Whatman filter paper and additions were made to experimental bottles containing solids. Identical additions were made to bottles without solids and used to establish the stock solution concentration. This aqueous solution was analyzed for naphthalene concentration by UV at λ = 219.2 nm. UV absorbance was calibrated against decane solutions of known naphthalene concentration made by dilution of a stock solution consisting of a weighed amount of naphthalene in decane.

For TCE, aliquots of a saturated aqueous stock solution were delivered to experimental bottles containing solids for equilibration. Aliquots delivered to bottles without solids were used to determine TCE concentration. Solution concentration was measured by rapidly transferring a portion of the aqueous phase to a stoppered cuvette and measuring its UV absorbance at 198nm. UV absorbance was calibrated against decane solutions of known TCE concentration made by dilution of a stock solution prepared by adding 500 ul of TCE (.693 $\pm$ .013 gm (n = 10)) to 250 ml decane.

3. Sorption equilibria were measured by adding aliquots of toluene, naphthalene or TCE stock solutions to glass bottles containing variable amounts of adsorbent and distilled water or 0.1m NaCl electrolyte. Bottles were rapidly capped with teflon laminated discs and continously rotated at room temperature for at least 12 hours. The bottles were then placed in a 25°C water bath overnight prior to analysis of the liquid or gas phases.

4. After equilibration vapor phase concentrations of TCE and toluene were determined by gas chromatography using the methods described in section III. B.

The liquid phase from bottles containing naphthalene was analyzed by high performance liquid chromatography (HPLC). Sample volumes of 25ul were removed through the bottle disc by syringe and injected into a Waters liquid chromatograph fitted with a Bondpack C-18 column (Waters, Inc.). An in-line prefilter was used to remove any colloidal solids sampled by the syringe. A mobile phase of 75/25 acetonitrile - H_2O was employed at a flow rate of 2.0 ml/min. Peak detection was by UV at λ = 254nm. Concentrations were quantified by comparison of sample peak heights to those of naphthalene standards in decane.

5. The sorption of solute in each experiment was calculated as follows:

   a) Measured bottle concentrations were normalized to those in a standard bottle of uniform size using equation 15 or 16 and replicate sample results were averaged. In bottles containing solids, gas volumes were corrected for the volume displaced by the adsorbent.

b) The fraction of solute remaining (f) was calculated as the ratio of normalized concentration in bottles with adsorbent to those in bottles containing the same addition of solute but with no solid. the percent solute sorbed is therefore:

$$\% \text{ Sorbed} = (1-f)100 \quad (17)$$

It should be noted that the percent solute sorbed (a ratio) may be accurately determined without knowing the precise mass of solute added in the experiment.

c) The mass sorbed was calculated as the product of the total mass added to the experimental bottles and the calculated percent sorbed (see (b) above). Total solute mass added was computed from the product of the solute stock solution concentration (measured as described in (2) above) and the volume of stock solution added to the experimental bottles.

C. RESULTS AND DISCUSSION

Several experiments were performed in which toluene sorption by Al_2O_3 was evaluated. In the absence of humic acid coatings, toluene concentration in the bottle headspace indicated that it was consistently excluded from solution by the oxide adsorbent. Exclusion was observed at adsorbent additions ranging from 5 to 60gm to a 50ml volume of distilled water. The specific conductance of the equilibrated solution was measured and results are shown in figure 4. The conductance in the equilibrated solution containing humic coated Al_2O_3 was substantially lower than that in solution equilibrated with the pure oxide.

Approximate ionic strengths may be computed from specific conductance using the relationship developed by Langlier[17]:

$$I = 1.6 \times 10^{-5} \times \text{specific conductance (umho/cm)} \quad (18)$$

Calculated ionic strengths and corresponding toluene activity coefficients are shown in figure 4. The calculated values are insufficient to explain the observed toluene exclusion from solution. As an example, the 15gm addition of Al_2O_3 produced a head space concentration of toluene which would require a toluene activity coefficient in solution of 1.05 (conservatively estimated assuming no toluene was sorbed) while the calculated toluene activity coefficient based on measured specific

FIGURE 4. CONDUCTANCE OF EQUILIBRATED SOLUTIONS

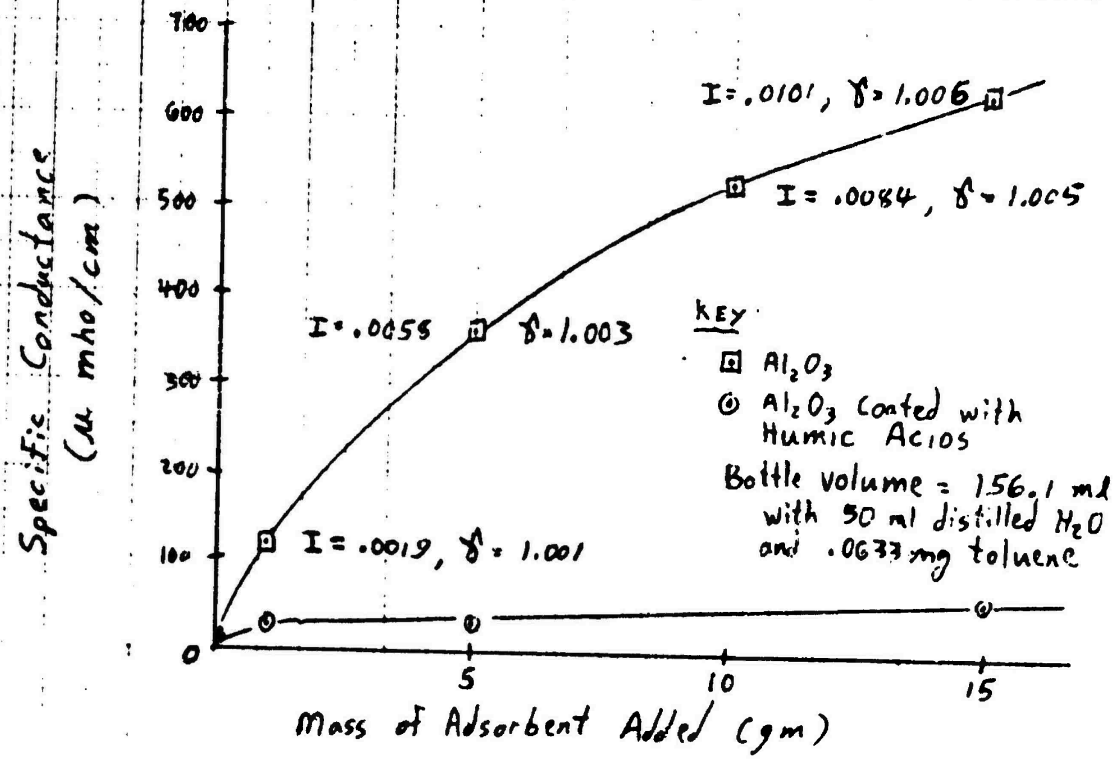
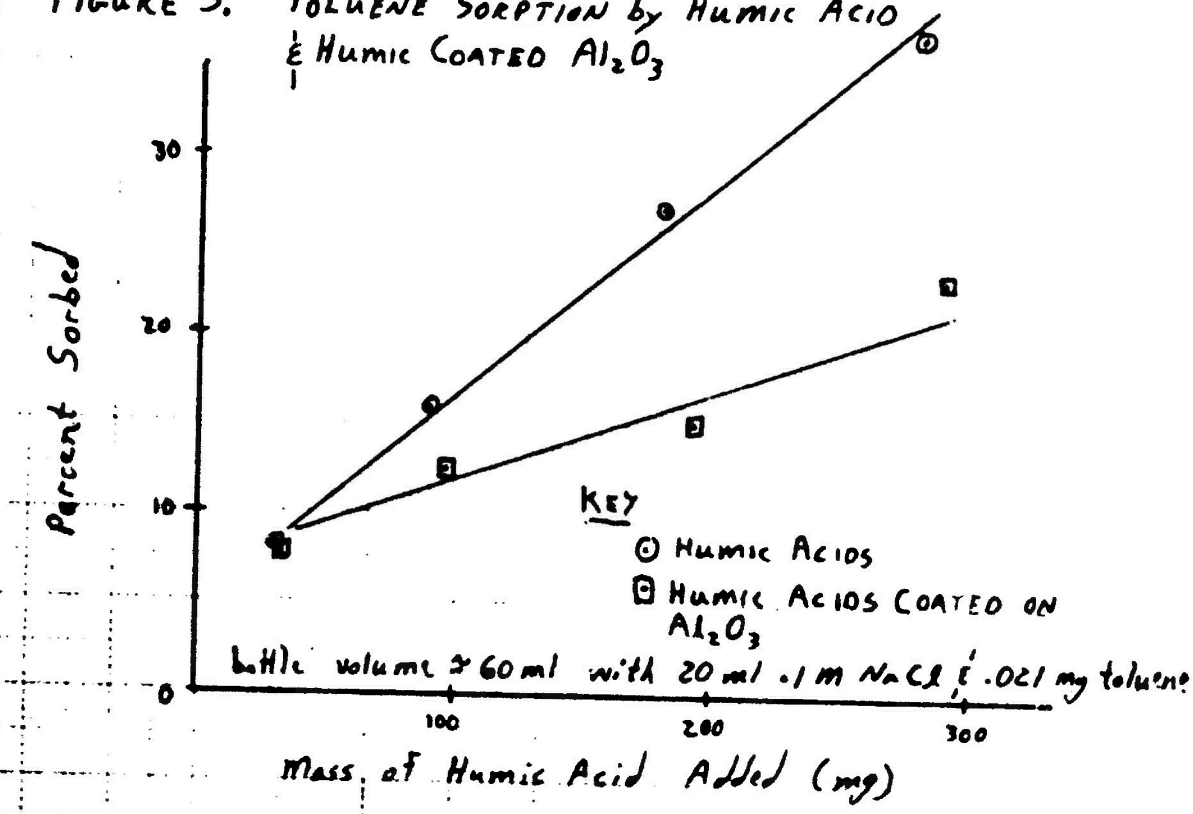


FIGURE 5. TOLUENE SORPTION by HUMIC ACID & HUMIC COATED Al_2O_3



conductivity was 1.006. Although no explanation for these results has been established, we speculate that colloidal Al_2O_3 carrying a surface charge may be affecting toluene activity. Such charged colloids are unlikely to be sufficiently mobile to influence the measured solution conductance which would explain the disparity between the calculated toluene activity coefficients and those which were observed.

Toluene was found to be consistently sorbed from aqeuous solution by Al_2O_3 coated with humic acids. The control ·ption by humic acids was tested by performing parallel experiments in which one set of bottles contained variable amounts of Al_2O_3 coated with humic acids and a second set contained the same mass of humic acids without Al_2O_3. A swamping NaCl electrolyte (0.1M) was used in all bottles to prevent toluene activity changes from influencing the results. Four or more replicate bottles of each solid addition were analyzed; mean results are shown in figure 5.

Sorption of toluene was observed to be greater by humic acids in the absence of Al_2O_3. This difference suggests that the form of the humic acid in solution (eg as a coating material vs. free humic acid), influences its sorptive properties.

It is noteworthy that the toluene sorption isotherm on humic acids without Al_2O_3 is a plateau (ie mg of toluene sorbed/g humic acid is independent of the equilibrium toluene concentration). This suggests that toluene sorption on the humics obeys a Langmuir isotherm and that full sorption site occupancy has occurred.

Sorption kinetics for naphthalene on humic coated Al_2O_3 were studied over a 72 hour time period. Results show that equilibrium conditions were obtained within the 18 minutes which elapsed prior to analysis of the first sample.

Naphthalene sorption on humic coated Al_2O_3 and naturally occurring Air Force Base soils was evaluated under conditions in which each experimental bottle contained the same mass of organic carbon. Triplicate bottles of each solid were analyzed, mean results are shown in table 6.

Table 6. Naphthalene Sorption Natural and Artificial Soils*

Soil	solid mass(g)	mass of organic carbon (mg)	% naphthalene sorbed
Al_2O_3 with humic acid	13.9	75	48.0%
Offutt AFB soil	32.8	75	100%
Whiteman AFB soil	18.3	75	39.8%
Lowry AFB soil	68.2	75	32.9%
Minot AFB soil	62.0	75	100%

* bottle volumes were 121 ml, containing soil, 50 ml 0.1M NaCl and 5ml of naphthalene stock solution.

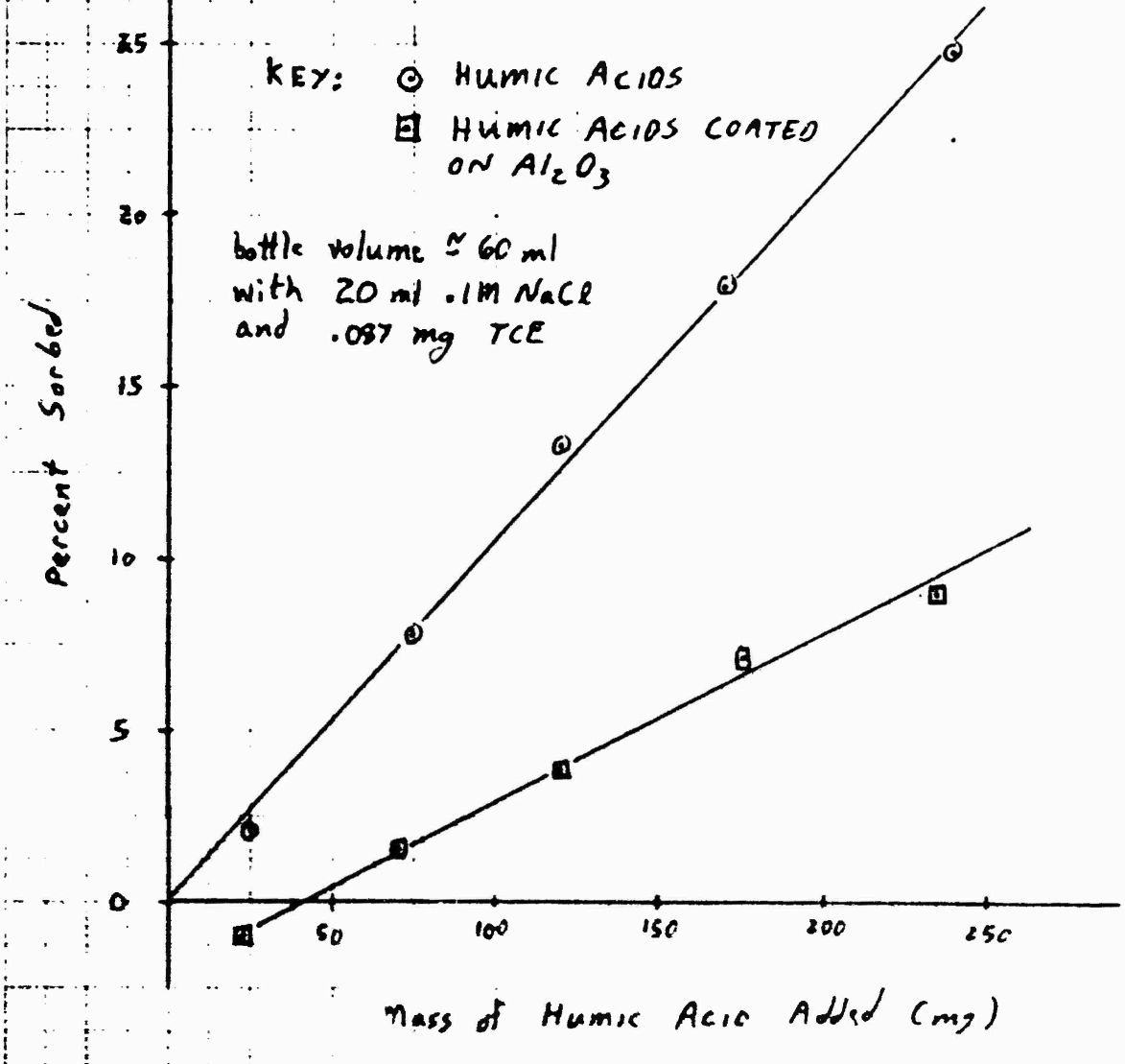
These results clearly show that, at the soil carbon concentrations employed (.11 to .54%), soil sorption capacity is not controlled solely by the mass of organic carbon present. In addition, sorption did not covary with the total mass of solid added, suggesting that the specific nature of the adsorbent and/or organic matter on the adsorbent must influence sorption equilibria.

The sorption of TCE onto Al_2O_3 coated with humic acids and onto humic acids was in the absence of Al_2O_3 was evaluated in an experiment comparable to that previously described for toluene. Results are shown in Figure 6. As was the case with toluene, the sorption of TCE was enhanced in the absence of Al_2O_3. The observed sorption density of TCE (mg sorbed/g sorbent) onto the humic material was independent of the equilbrium concentration of TCE in solution suggesting that the humic materials have finite binding capacity of approximately $7x10^{-7}$ moles TCE/gm (vs $3x10^{-7}$ moles toluene/gm) and that humic binding sites were saturated in the experiment performed.

D. Summary

The headspace technique has been demonstrated to be a reliable method for measuring the sorption of volatile components of JP-4 fuels and degreasers. High performance liquid chromatography permits separation and analysis of compunds of low volatility which absorb in the UV region.

FIGURE 6. SORPTION OF TCE BY HUMIC ACID & HUMIC COATED Al_2O_3



Sorption of toluene and TCE has been shown to be suppressed by the presence of a hydrous Al_2O_3 oxide. Sorption of these compounds by humic acids was enhanced when this sorbent was not present as a coating material on Al_2O_3. This data suggests that the physical state of the organic sorbent may influence the binding of nonionic solutes. The sorption density (mg sorbed/g sorbent) of toluene and TCE on humic acid was observed to be independent of the equilibrium concentration in solution indicating that the humic acid sorbent had a limited number of binding sites and that full site occupancy was achieved.

Sorption of naphthalene by Al_2O_3 coated with humic adids and naturally occurring Air Force base soils was not controlled solely by the organic content of these sorbents or by the mass of sorbent present. These data establish that the sorption behavior of nonionic soluter on some naturally occurring aquifer materials is likely to be a complex phenomena in which the nature of both the inorganic soil matrix and the soil organic matter may play a role. Additional research is required to establish the interdependency between sorption of nonionic solutes and the detailed physical and chemical characteristics of the adsorbent.

V. Recommendations

The research performed was designed to evaluate the equilibria for pertinent reactions of volatile nonionic organic compounds in unsaturated groundwater systems.

Gas exchange equilibria may be evalutated through knowledge of a compound's Henry's Law constant and it's activity coefficient in solution. An equilibrium head space technique has been demonstrated to be applicable to the determination of these parameters for many of the componets of distillate fuels and degreasers.

Sorption equilibria may also be evaluated using head space measurements for volatile compounds for which Henry's Law constants and solution activity coefficients are known. The observed sorption of organic solutes typical of degreasers (TCE) and distillate fuels (toulene and naphthalene)was seen to be a complex phenomena. Although our results

confirm the importance of organic materials in the binding of nonionic solutes, the data indicate that the physical and perhaps chemical nature of the soil organics are likely to influence results. Additional investigation is required to explore these interdependencies.

Specific recommendations for further research are as follows:

1. Experiments should be conducted to determine those characteristics of soil organics which influence their binding with nonionic organic solutes. Pertinent variables include but are not limited to (a) elemental composition (percent carbon, nitrogen, etc.) (b) molecular size (c) physical state, e.g., particulate vs. dissolved vs. coating materials and (d) exchangeable acidity.

2. Experiments are needed to evaluate the sorption of volatile compounds from the vapor phase onto soils. The importance of soil organic carbon in these sorption reactions requires evaluation.

3. The equilibria of mixed solutes with soil organic materials should be evaluated. Our data suggest humic acids have a finite binding capacity and compound competition for binding sites may therefore occur.

4. Henry's Law constants and activity coefficients for all of the identified components of distillate fuels should be evaluated as a prerequisite for the prediction of their gas exchange and sorption equilibra.

REFERENCES

1. Garrels, R.M. and C.L. Christ, Solutions, Minerals, and Equilibria, Harper and Row Publishers (1965).

2. Karickhoff, S.W.; D.S. Brown and T.A. Scott, "Sorption of Hydrophobic Pollutants on Natural Sediments," Water Research, 13, pp 241-248 (1979).

3. Means, J.C.; S.G. Wood; J.J. Hassett and W.H. Banwart "Sorption of Polynuclear Aromatic Hydrocarbons by Sediments and Soils"; Env. Sci. and Tech. 14(12); pp 1524-1528 (1980).

4. Schwarzenbach, R.P. and J. Westall; "Transport of Nonpolar Organic Compounds from Surface Water to Groundwater. Laboratory Sorption Studies" Env. Sci and Tech.; 15(11) pp 1360-1367 (1981).

5. McCarty, P.L; M. Reinhard and B.E. Rittmann "Trace Organics in Groundwater"; Env. Sci. and Tech.; 15(1) pp 40-51 (1981).

6. Chiou, C.T.; L.J. Peters and V.H. Freed; "A Physical Concept of Soil-Water Equilibria for Nonionic Organic Compounds" Sci. 206 pp 831-832 (1979).

7. Chiou, C.T.; P.E. Porter and D.W. Schmedding "Partition Equilibria of Nonionic Organic Compounds between Soil Organic Matter and Water" Env. Sci and Tech. 17(14) pp 227-231 (1983).

8. Lincoff, A.H. and J.M. Gossett; "The Determination of Henry's Constant for Volatile Organics by Equilibrium Partitioning in Closed Systems" Proc. Int. Symp. on Gas Transfer at Water Surfaces; Cornell Univ. (1983).

9. Smith, J.H., J.C. Harper and H. Jaber; "Analysis and Environmental Fate of Air Force Distillate and High Density Fuels"; Report No. ESL-TR-81-54 prepared by SRI International for Air Force ESC/RD (1981).

10. MacIntryre, W.G.; C.L. Smith, P.O. Defur and C.W. Su "Hydrocarbon Fuel Chemistry: Sediment water Interaction" Report No ESL-TR-82-06 prepared by Virginia Inst. of Mar. Sci for Air Force ESC/RD (1981).

11. Schlessinger, G.G. "Vapor Pressures, Critical Temperatures and Critical Pressures of Organic Compounds" in Handbook of Chemistry and Physics 52nd Ed., R.C. West (ed); CRC Press (1971).

12. Verschueren, K. Handbook of Environmental Data on Organic Chemicals; Van Norstraud Reinhold Co. (1977).

13. Mackay, D; W.Y. Shiu and R.P. Sutherland "Determination of Air-water Henry's Law Constants for Hydrophobic Pollutants" Env. Sci. and Tech 13(3) pp 333-337 (1979).

14. Chiou, C.T.; D.W. Schmedding; and M. Manes; "Partitioning of Organic Compounds in Octanol Water Systems" Env. Sci and Tech 16(1)pp 4-10 (1982).

15. Gossett, J.M., "Packed Tower Air Stripping of Trichloroethylene from Dilute Aqueous Solution" Report No ESL-TR-81-83 for Air Force ESC/RD (1980).

16. Allison, L.E., "Organic Carbon" in Methods of Soil Analysis; Part 2. Chemical and Microbiological Properties; C.A. Black et al. (eds). Am. Soc. of Agronomy (1965).

17. Snoeyink, V.L. and D. Jenkins; Water Chemistry John Wiley & Sons (1980).

1983 USAF - SCEEE SUMMER FACULTY RESEARCH PROGRAM

SPONSORED BY THE

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

CONDUCTED BY THE

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

EVALUATION OF PROJECTILE IMPACT ON EARTH COVERED STRUCTURES

Prepared By:	DR. DARYL L. LOGAN
Academic Rank:	ASSOCIATE PROFESSOR
Department and University:	DEPARTMENT OF CIVIL ENGINEERING ROSE-HULMAN INSTITUTE OF TECHNOLOGY
Research Location:	AIR FORCE ENGINEERING AND SERVICE CENTER ENGINEERING RESEARCH DIVISION, AIR BASE SURVIVABILITY BRANCH TYNDALL AFB, FL.
USAF Research Colleague:	CAPTAIN PAUL L. ROSENGREN
DATE:	AUGUST 1, 1983
CONTRACT NO:	F49620-82-C-0035

EVALUATION OF PROJECTILE IMPACT ON EARTH COVERED STRUCTURES

by

Daryl L. Logan

ABSTRACT

The capability of earth covered reinforced concrete structures to withstand the local response of impacting projectiles is investigated . Soil and concrete penetration, as well as thickness of structure concrete wall necessary to prevent perforation and scabbing, are predicted. The influence of parameters soil penetrablity index, soil cover thickness, concrete compressive strength, concrete thickness and impact velocity on the local response are examined. It is shown that for ranges of parameter values of interest soil penetrability index and soil cover thickness have the greatest influence on allowable missle weight (frontal pressure) at which incipient backface scabbing of the concrete wall occurs. The implementation of the results is demonstrated for data recorded from misale debris resulting from an actual aircraft shelter explosion experiment. The results indicate that nearly all data falls within safe limits of realizable parameter values. Implications for siting earth covered structures with respect to aircraft structures are evident. Suggestions for further research are indicated.

ACKNOWLEDGEMENT

The author thanks the Air Force Systems Command, the Air Force Office of Scientific Research, the Air Force Engineering and Services Center, and the Southeastern Center for Electrical Engineering Education for a truly rewarding summer at Tyndall AFB, Florida. He particularly thanks all his colleagues for their friendly hospitality and professional attitude.

In particular he especially thanks Captain Paul Rosengren for his friendly guidance and encouragement, and Dr. Chester Canada for many formal and informal discussions. Finally, he appreciates the help of Major John Centrone, Lt. Tom Hilferty and Major David Flaa for making his stay even more comfortable.

I. INTRODUCTION:

A primary concern of the Air Force is the present siting restriction placed on all inhabited buildings, including semi-hardened facilities and earth covered structures (1). Recent studies of the aftermath of aircraft shelter debris from bomb detonations within the aircraft shelter (2) have resulted presently in a 300 feet minimum spacing requirement between semihardened aircraft shelters and inhabited buildings regardless of the protective capabilities of these inhabited buildings. This 300 feet siting requirement could be relaxed if tests or analysis were available to demonstrate the added protection from shelter debris provided by these protective structures.

This paper is the result of a study undertaken to determine the capability of earth covered structures to withstand the debris threat from a most probable detonation within a nearby aircraft shelter. The study concerned itself with the local response due to projectiles (missles) impacting earth covered structures.

II. OBJECTIVE:

The primary objective of this research effort was to examine survivability capabilities of earth covered structures when such structures are subjected to debris missles resulting from a most probable sized bomb explosion within a nearby aircraft shelter.

Specific goals of the research were:

(1) To determine appropriate procedures to adequately predict the local response of an earth covered structure to missle impact.

(2) To illustrate use of this procedure for an Air Force structure of interest subjected to missle debris of interest (2).

(3) To recommend ways of increasing the survivability capabilities of the structure of interest.

III. METHOD OF ANALYSIS AND EQUATIONS USED:

Herein we will analyze the local response (as opposed to overall structural response) of an earth covered structure due to impact resulting from aircraft shelter debris (missles). The missles considered are those created during an explosion occurring within an aircraft shelter as

recorded in (2). The local response refers to analysis of the earth covered structure in the vicinity of the impact. The phenomena to be analyzed are penetration depth of a missle into the structure wall, perforation (a missle passing entirely through the wall thickness), and backface scabbing (scabbing of concrete off the inside face of the wall). The structures are assumed to be of reinforced concrete and to have an earth material overlying them.

Although analytical attempts have been under study (3) to predict local impact phenomena, these methods have not been fully developed. Therefore the analysis in this paper is based upon a series of recently assessed empirical equations (4-8) which are used to predict penetration, perforation and backface scabbing as a missle impacts the soil cover associated with the structure.

The analysis procedure is as follows:

(1) Calculate the depth of penetration, X_s, (in feet) of the tip of the missle into the earth overburden by

$$X_s = 0.53SN\left(\frac{W}{A}\right)^{1/2} \ln(1 + 2V^2 10^{-5}) \qquad (1)$$

where S = Soil penetrability index (soil constant)
N = missle nose shape performance coefficient
W = missle weight, in lb
A = missle impact cross-sectional area, in sq. in. and
V = missle impact velocity, in ft/sec

(2) Calculate the residual velocity, Vc, (in ft/sec) by

$$Vc = V\left(1 - \frac{t_s}{X_s}\right)^{1/2} \qquad (2)$$

where complete penetration of the overburden by the missle is assumed and t_s = the soil overburden thickness, in units of feet.

(3) Calculate the depth of penetration, x, (in inches) of the tip of the missle into the concrete wall as

$$x = 2d\sqrt{F} \quad ; \quad \text{for } \frac{x}{d} \leq 2.0 \qquad (3)$$

or $x = d(F + 1)$; for $\frac{x}{d} \geq 2.0$

where $$F = \frac{180}{\sqrt{f'_c}} N_2 \left(\frac{E}{E_m}\right)^{1.25} \frac{W}{(d^{2.80})} \left(\frac{V_c}{1000}\right)^{1.80} \quad (3)$$

(4)

and

f'_c = concrete compressive strength, in psi

N_2 = missle nose shape coefficient

E = modulus of elasticity of missle material, in psi

E_m = modulus of elasticity of mild steel, in psi and

d = effective diameter of a missle, which has same contact area as that of actual contact area, in inches

(4) Determine thickness, p, (in inches) of concrete wall to prevent perforation by

$$\frac{P}{d} = 1.32 + 1.24 \frac{x}{d} \quad ; \quad \text{for } 1.35 \leq \frac{x}{d} \leq 13.5$$

or $$\frac{P}{d} = 3.19 \frac{x}{d} - 0.718\left(\frac{x}{d}\right)^2 \quad ; \quad \text{for } \frac{x}{d} \leq 1.35 \quad (5)$$

(5) Determine thickness, s, (in inches) of concrete wall to prevent backface scabbing by

$$\frac{s}{d} = 2.12 + 1.36 \frac{x}{d} \quad ; \text{ for } 0.65 \leq \frac{x}{d} \leq 11.75$$

or $$\frac{s}{d} = 7.91 \frac{x}{d} - 5.06\left(\frac{x}{d}\right)^2 \quad ; \quad \text{for } \frac{x}{d} \leq 0.65 \quad (6)$$

In Eqs. (1) through (6), the impact is assumed to be normal to the surface. It can be observed that local impact is a function of many parameters including soil cover thickness, soil penetrability index, missle weight, missle contact cross-sectional area and impact velocity, and missle nose shape, and concrete compressive strength of structure.

IV. SOLUTION PROCEDURE:

To satisfy the goals critical frontal pressures, W/A, were determined for various distances (ranges) from the aircraft shelter for parameters of soil penetrability index, soil cover thickness, structure wall thickness and compressive strength, and missle impact velocity.

The solution procedure used was as follows. First, the initial velocity of a typical missle was calculated using particle projectile motion equations (where the range of interest and an assumed launch angle were substituted into the equations). Here range is defined to be the horizontal distance between where the missle is launched and where it lands. A representative missle contact area and nose shape coefficients ($N = 0.56$ and $N_2 = 0.72$) for blunt ended missles were assumed. Then the parametric study was undertaken using variations in soil penetrability indices, soil cover, and concrete wall thickness and compressive strength. A trial-and-error process of selecting a missle weight and subsequent solution of Eqs. (1-6) was used. The process was stopped when a missle weight determined by the minimum concrete thickness to prevent scabbing, from Eq.(6), was obtained. That is, a weight resulting in an s, from Eq. (6), equal to the concrete wall thickness was obtained. This weight is defined to be the critical weight resulting in incipient backface scabbing. To facilitate the parametric study, a FORTRAN computer program, based upon Eqs. (1-6), was written to determine the critical missle weight as well as missle penetration depth into the wall and minimum concrete thickness to prevent missle perforation.

V. NUMERICAL RESULTS:

Numerical results are now presented for typical parameters of interest to the Air Force. Some of the values of parameters used include,

1) Soil penetrability indices (constants) given by

Soil Constant	Typical Soil Description
5.2	Clayey silt, silty clay, dense, hard, dry
7.0	Sand, loose to medium, moist
10.5	Clay, moist, stiff
30.0	Loose, moist topsoil with humus material, mostly sand and silt. Moist to wet clay, soft, low shear strength.
40.0	Clay, silty, wet

2) Soil cover thicknesses of 3, 4 and 5 feet.

3) Concrete wall thicknesses of 9 and 12 inches

4) Concrete compressive strengths of 4000 and 5500 pounds per square inch.

It can be observed from Eqs. (1) and (4) that more meaningful results are obtained by expressing W and A as a single parameter. This possibility was verified as shown by Table 1 where W/A at incipient scabbing for different cross sectional areas at different ranges are given for values of S = 10.5, t_s = 3 ft, f_c' = 5500 psi and concrete thickness = 9 in.

Table 1. W/A at Incipient Scabbing for Different A's

	W/A	
Range, R, FT.	A = 1963 IN^2	A = 78.54 IN^2
50	687.5	700.1
100	178.2	184.4
150	86.7	85.9
200	48.4	49.0
250	32.1	33.1
300	23.1	23.5

The ratio W/A is called frontal pressure. In this paper critical frontal pressure is defined to be that frontal pressure causing incipient scabbing.

Figures 1 through 5 illustrate the influence of various parameters on critical frontal pressure. Throughout, the missle is taken to be of steel. Figure 1 shows critical frontal pressure, W/A, for various ranges, R, for different soil penetrability indices. (Actually $(W/A)^{1/2}$ is used in order to present the data in a more meaningful graphical form). The launch angle of the missle is 30 degrees from the horizontal, soil cover thickness is 3 feet and concrete wall thickness t_c is 9 inches. It can be observed from Figure 1 that as soil penetrability index decreases the critical frontal pressure increases. That is for a dense, hard, dry silty clay (S = 5.2), critical W/A is larger than for a loose to medium moist sand (S = 7.0). Comparisons of results for the soil descriptions

corresponding to each S indicate that in general dense, hard, dry soils resist penetration noticeably better than loose, soft, wet soils. Further it can be observed that as the range increases the critical W/A decreases. This is reasonable, based on the fact that it takes a larger initial velocity to project a missle a longer range. The resulting impact velocity is equal to the initial velocity based on projectile motion equations.

Figure 2 shows the results of W/A for various R for different soil cover thicknesses t_s for a given S and t_c. Here the greater t_s, the greater critical W/A. For S = 10.5, 4 feet of soil cover may increase the critical W/A by as much as 1.75 times compared to 3 feet of cover. Again W/A decreases with increasing R for reasons explained in the previous paragraph.

Figures 3 and 4 demonstrate critical W/A for various R for different concrete compressive strengths and concrete wall thickness, respectively. It can be observed that critical W/A are negligibly influenced by concrete compressive strength and wall thicknesses of usual interest.

Finally Figure 5 shows the influence of missle initial velocity of impact on the critical W/A at a range of 100 feet. As the initial velocity of impact increases the critical W/A decreases.

In summary, a parametric study, based on a series of empirical equations used to predict soil penetration, concrete penetration, perforation, and scabbing, was undertaken to determine the most important factors influencing local missle impact response for a typical Air Force earth covered structure. For ranges of parameters of interest, it was determined that soil penetrability index and soil cover thickness have the greatest influence on allowable frontal pressures at which incipient scabbing occurs. Fortunately these two parameters are quite easily controlled and their required values relatively economically achieved.

VI. RECOMMENDATIONS:

a. Implementation of Results

The results of this research have immediate application to a soil-concrete layered medium in predicting missle penetration into the medium and the associated concrete thickness needed to prevent backface

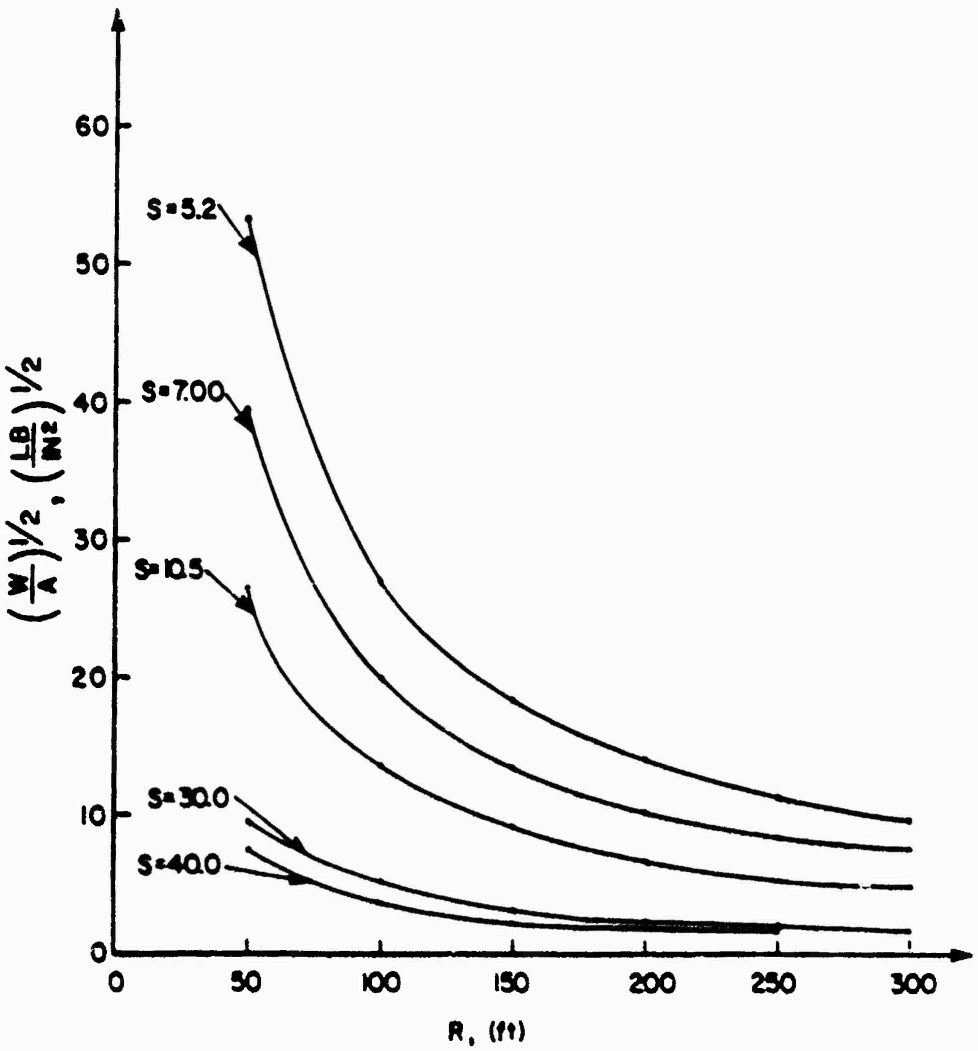


Figure 1 - Range versus (frontal pressure)$^{\frac{1}{2}}$ at Incipient Scabbing for Different Soil Penetrability Indices (For t_s=3 ft, t_c=9 in, f'_c = 5500 psi)

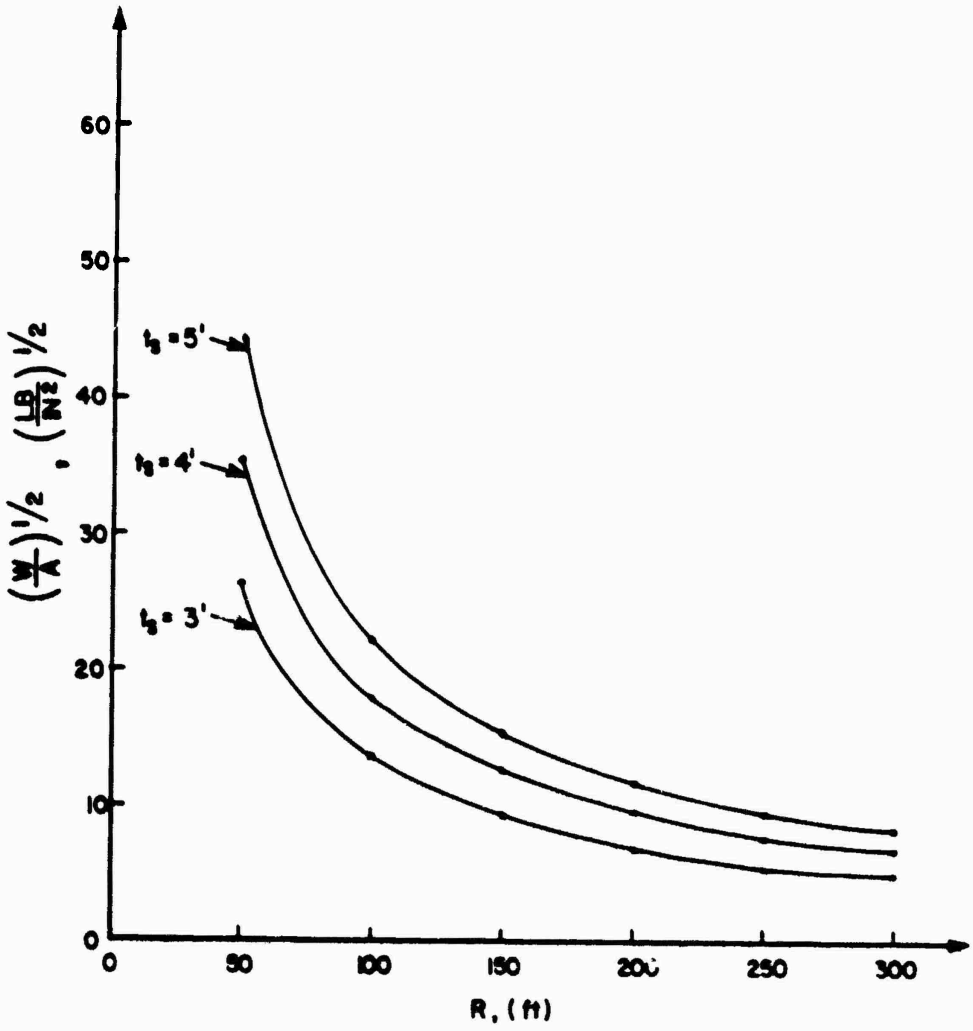


Figure 2 - Range versus (frontal pressure)$^{\frac{1}{2}}$ at Incipient Scabbing for Different Earth Cover Thicknesses (for S=10.5, t_c=9 in. f'_c=5500 psi)

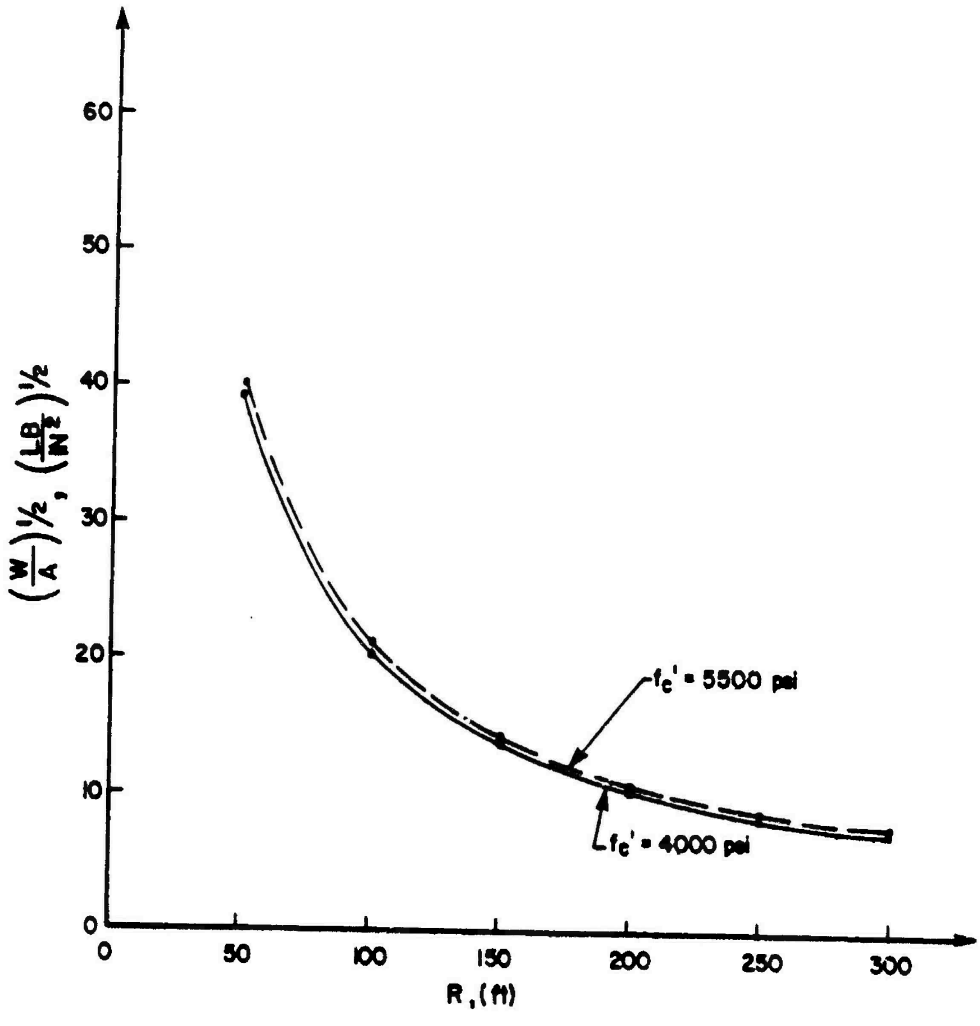


Figure 3 - Range versus (frontal pressure)$^{\frac{1}{3}}$ at Incipient Scabbing for Different Concrete Compressive Strengths (For S=7.0, t_s=3 ft, t_c=9 in.)

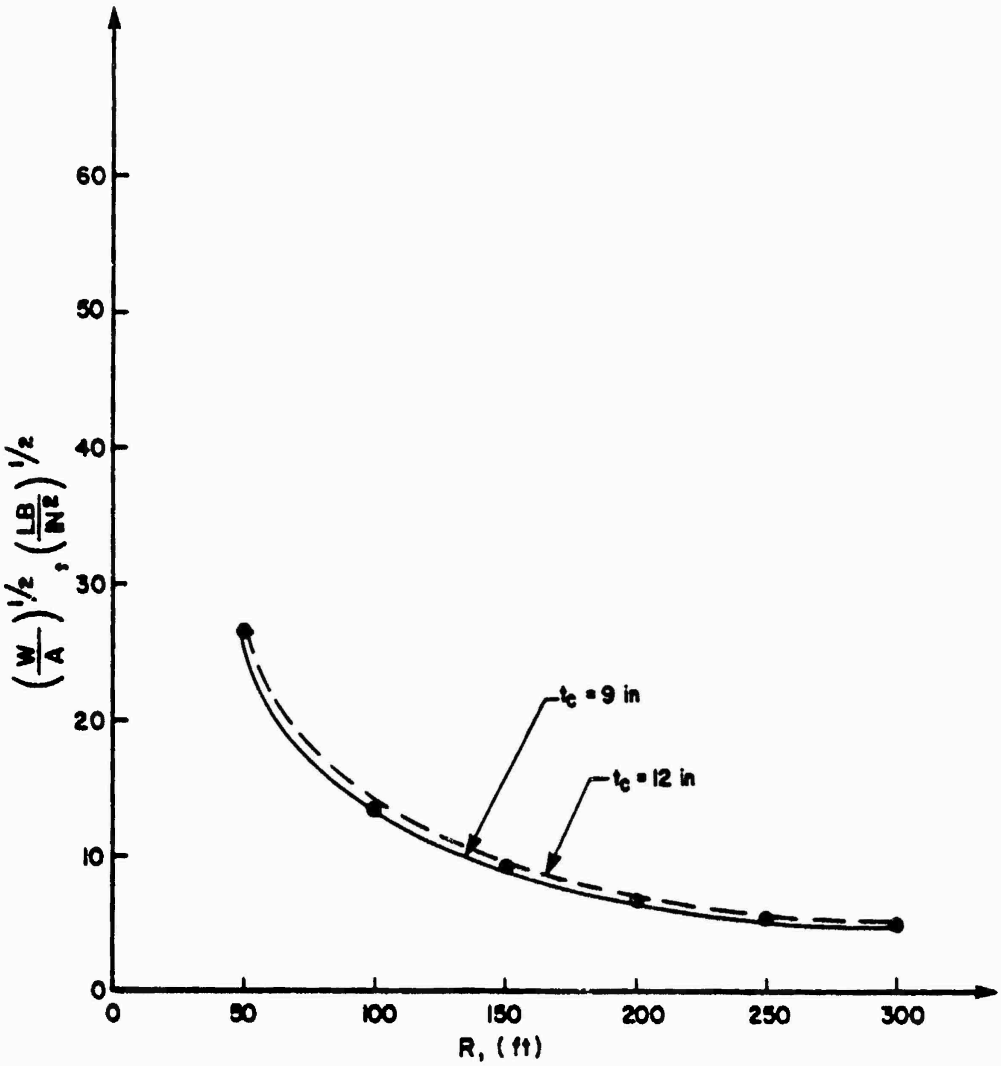


Figure 4 - Range versus (frontal pressure)$^{\frac{1}{2}}$ at Incipient Scabbing for Different Concrete Wall Thicknesses (For S=10.5, t_s=3 ft, f'_c=5500 psi)

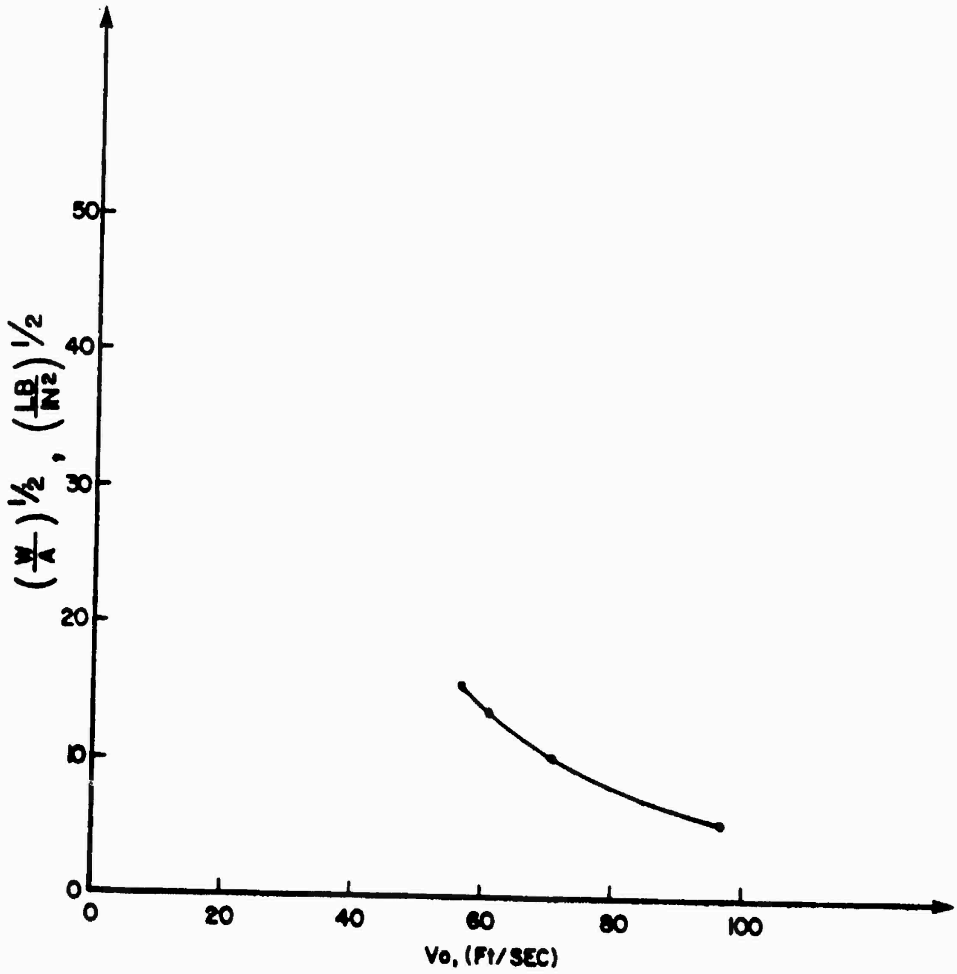


Figure 5 - Impact Velocity versus (frontal pressure)$^{\frac{1}{2}}$ at Incipient Scabbing (For R=100 ft, S=10.5, t_s=3 ft, t_c=9 in, f'_c=5500 psi)

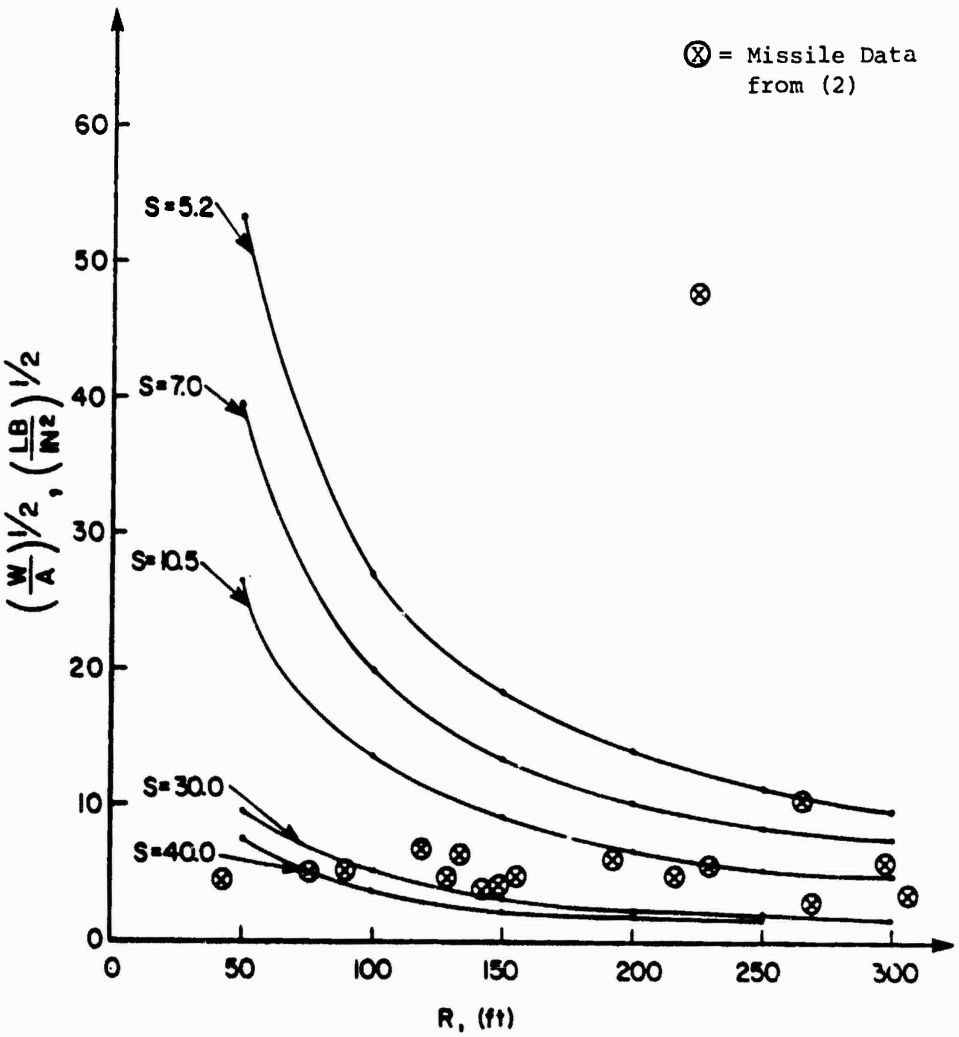


Figure 6 - Comparison of Range versus (frontal pressure)$^{\frac{1}{2}}$ at Incipient Scabbing for Different Soil Penetrability Indices and Missile Data from (2) (For t_s=3 ft, t_c=9 in, f'_c =5500 psi)

scabbing and perforation. An example of a military application is for earth covered concrete structures subjected to debris resulting from an aircraft shelter explosion(2). The implementation of results is demonstrated in Figure 6 where the large debris data from (2) is expressed as $(W/A)^{1/2}$ and plotted for their ranges. These results are compared to critical $(W/A)^{1/2}$ versus R for various soil penetrability indices for 3 feet of soil cover. As can be seen nearly all data falls within safe limits of realizable soil parameters. Implications for siting earth covered structures with respect to aircraft shelters are indicated.

Another application would be to predict the depth of penetration (or the burster layer thickness necessary to "catch" a bomb) into a concrete burster layer from a bomb. This information is necessary in order to define the ground shock load used for underground shelter design.

b. Suggestions for Follow-On-Research

This research was based on local response behavior from a missle impacting an earth covered structure. The local response equations were computer programmed in a user friendly manner for a soil-concrete medium. To expand the use of these equations, the computer program should be made more versatile, including capability of analysis for any combination of different materials (applications for composite construction barriers such as concrete-sand-concrete and for soil-burster layer-soil penetration predictions) and capability to automatically converge to a critical weight. (This option would be obtainable by programming a numerical method into the existing program.)

The research should also be extended to utilize the results from the penetration equations in a model to predict overall structural response from missles. This phase would include a method for determination of the force-time function(s) to be applied to the structure. This is a necessary phase of analysis in the determination of survivability of earth covered systems. A finite element program, including the force-time function developed and soil interaction, would be used to complete the analysis.

REFERENCES

1. HQ USAFE/DEX Message, P081429Z June 1983, Subject "Siting Impact of Distant Runner Test Series".

2. Ward, J.M. "Proceedings of the Distant Runner Symposium, 27-28 April 1982, Proceedings, Sept. 1982, pp 325-400.

3. Thigpen, L., "Projectile Penetration of Elastic-Plastic Earth Media," Journal of the Geotechnical Engineering Division, ASCE, Vol. 100, No. GT3, March 1974, pp 279-294.

4. Sliter, G.E., "Assessment of Empirical Concrete Impact Formulas," Journal of the Structural Division, ASCE, Vol. 106, No. ST5, May 1980, pp. 1023-1045.

5. Triandafilidis, G.E., "State-of-the Art of Earth Penetration Technology", DNA 001-74-C-0198 Defense Nuclear Agency, 1976, pp. 53-70.

6. Young, C.W., "Depth Prediction for Earth-Penetration Projectiles", Journal of the Soil Mechanics and Foundations Division, ASCE, vol. 95, No. SM3, Proc. Paper 6558, May, 1969, pp. 803-817.

7. Effects of Impact and Explosion, Summary Technical Report of Division 2, National Defense Research Committee, Vol. 1, Washington, D.C., 1946.

8. Kar, A.K., "Projectile Penetration into Buried Structures", Journal of the Structural Division, ASCE, Vol. 104, No. ST1, Jan. 1978, pp. 125-139.

1983 USAF-SCEEE FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

NUMERICAL SOLUTION OF THE EULER EQUATIONS ON DYNAMIC GRIDS

Prepared by:	C. Wayne Mastin Department of Mathematics and Statistics Mississippi State University
Research Location:	Air Force Armament Laboratory Aeromechanics Division Aerodynamics Branch
USAF Research Colleague:	Lawrence E. Lijewski
Date:	August 19, 1983
Contractor Number:	F49620-82-C-0035

NUMERICAL SOLUTION OF THE EULER EQUATIONS ON DYNAMIC GRIDS

by

C. Wayne Mastin

ABSTRACT

A method is developed for solving the Euler equations on two-dimensional regions with moving boundary components. An algorithm is presented for moving the interior grid points to follow the moving boundary component, and the modifications necessary for implementing the finite-volume algorithm for solving the Euler equations on a moving grid are stated. These concepts have been incorporated in a computer program and preliminary numerical results are presented.

ACKNOWLEDGEMENTS

I wish to express my gratitude to the Air Force Systems Command, the Air Force Office of Scientific Research, and the Southeastern Center for Electrical Engineering Education for the opportunity of spending a most enjoyable and rewarding summer of study and research at the Air Force Armament Laboratory, Eglin AFB, FL. I would like to thank the members of the Aerodynamics Branch for their cooperation and hospitality. I am especially grateful for the support and encouragement of Dr. Donald Daniel, Captain Gerry Harbaugh, and Dr. Lawrence Lijewski. Ms. Sallye Jackson also deserves a special thanks for typing this report.

I. INTRODUCTION

The field of computational fluid dynamics is continuously expanding to deal with more and more complex flow regimes. Improvements are being made in both hardware and software. As the new super computers are coming on-line, algorithms are being improved to solve new problems with greater accuracy and efficiency. An integral part of any numerical algorithm is the construction of a computational grid. Not only must the grid fit the boundary of a complicated three-dimensional region, it may also be required to dynamically adapt to the movement of the boundary components.

This report will examine several aspects in the solution of fluid flow problems on moving grids. The first consideration will be the modification needed in a typical numerical algorithm. The popular finite-volume method for solving the two-dimensional Euler equations has been selected as a model algorithm. It will be obvious that the techniques developed here can be easily extended to three-dimensions. The same techniques are also applicable to the finite-volume solution of the Navier-Stokes equations. The intent is to go beyond the superficial modifications, such as the recomputation of changing lengths and areas used in the algorithm, and to investigate the effect of grid movement on the accuracy and stability of the algorithm.

A secondary effort during this period has been development of an efficient grid moving scheme. For complicated regions, like the region about a complete aircraft, the construction of a grid may consume a significant portion of the overall effort. Therefore, a reconstruction of the entire grid each time a boundary component moves may be very expensive. This observation has motivated the development of an algorithm for advancing the interior grid points to follow the movements of a boundary component. It should be noted that the preservation of grid spacing at the boundary would not be as critical in the present solution of the Euler equations as it would be in the solution of the Navier-Stokes equations. In the latter case, a fine grid spacing near the boundary must be maintained due to the presence of a boundary layer. Hopefully, these concepts will be of value in other problems, such as resolving shocks or material interfaces, where moving meshes are advantageous.

An existing computer program for solving the Euler equations has been modified to solve a problem on a region with a moving boundary component. Numerical results are presented to validate the concepts presented in this report.

II. FINITE VOLUME METHOD

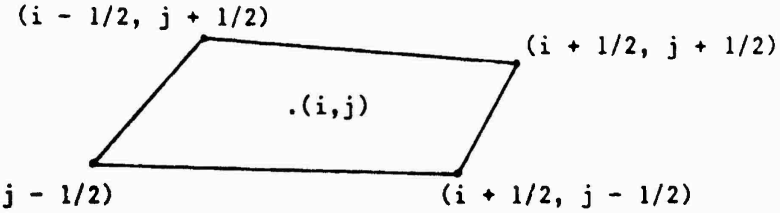
Finite-volume methods may be used to solve conservation laws of the form

$$w_t + f_x + g_y = 0 \tag{1}$$

where f and g are functions of w. If the above equation (1) is integrated over any region R with boundary ∂R, and the Gauss divergence theorem is applied, then

$$\iint_R w_t \, dxdy = -\int_{\partial R} f dy - g dx \tag{2}$$

Difference equations are derived by defining a grid on R and numerically approximating the integrals in (2) over each grid cell. The following indexing scheme will be used for function values at the centroid and vertices of a grid cell.



The area of the cell will be denoted by $A_{i,j}$ and the increment of a coordinate value along a side will be denoted, for example, by

$$\Delta x_{i+1/2,j} = x_{i+1/2,\, j+1/2} - x_{i+1/2,\, j-1/2}.$$

The time level of a variable will be indicated by a superscript. In order to maintain the second order accuracy of the numerical algorithm, a trapezoidal time differencing will be used in (2). Average values will also be used for the incremental values of x and y which will be used to approximate the differentials on the right hand side of (2). This is indicated notationally by

$$\Delta x^{n+1/2}_{i+1/2,j} = 1/2(\Delta x^{n+1}_{i+1/2,j} + \Delta x^{n}_{i+1/2,j}).$$

There are many variations of the finite-volume method in current use. The finite-volume method due to MacCormack [2] can be implemented on a moving grid using the following predictor-corrector equations.

$$\begin{aligned} \tilde{w}^{n+1}_{i,j} \tilde{A}^{n+1}_{i,j} = {} & w^{n}_{i,j} A^{n}_{i,j} - \Delta t\{ f^{n}_{i+1,j} \Delta y^{n+1/2}_{i+1/2,j} - g^{n}_{i+1,j} \Delta x^{n+1/2}_{i+1/2,\, j} \\ & + f^{n}_{i,j+1} \Delta y^{n+1/2}_{i,j+1/2} - g^{n}_{i,j+1} \Delta x^{n+1/2}_{i,j+1/2} \\ & + f^{n}_{i,j} \Delta y^{n+1/2}_{i-1/2,j} - g^{n}_{i,j} \Delta x^{n+1/2}_{i-1/2,j} \\ & + f^{n}_{i,j} \Delta y^{n+1/2}_{i,j-1/2} - g^{n}_{i,j} \Delta x^{n+1/2}_{i,j-1/2}) \end{aligned} \tag{3}$$

$$\begin{aligned} w^{n+1}_{i,j} A^{n+1}_{i,j} = {} & w^{n}_{i,j} A^{n}_{i,j} - \Delta t/2\{(\tilde{f}^{n+1}_{i,j} + f^{n}_{i+1,j}) \Delta y^{n+1/2}_{i+1/2,j} \\ & - (\tilde{g}^{n+1}_{i,j} + g^{n}_{i+1,j}) \Delta x^{n+1/2}_{i+1/2,j} \\ & + (\tilde{f}^{n+1}_{i,j} + f^{n}_{i,j+1}) \Delta y^{n+1/2}_{i,j+1/2} - (\tilde{g}^{n+1}_{i,j} + g^{n}_{i,j+1}) \Delta x^{n+1/2}_{i,j+1/2} \\ & + (\tilde{f}^{n+1}_{i-1,j} + f^{n}_{i,j}) \Delta y^{n+1/2}_{i-1/2,j} - (\tilde{g}^{n+1}_{i-1,j} + g^{n}_{i,j}) \Delta x^{n+1/2}_{i-1/2,j} \\ & + (\tilde{f}^{n+1}_{i,j-1} + f^{n}_{i,j}) \Delta y^{n+1/2}_{i,j-1/2} - (\tilde{g}^{n+1}_{i,j-1} + g^{n}_{i,j}) \Delta x^{n+1/2}_{i,j-1/2}) \end{aligned} \tag{4}$$

Since equation (4) includes values at the previous time step, it is computationally more convenient to eliminate these values by solving for $w^n_{i,j}\ A^n_{i,j}$ in (3) and substituting in (4). Thus the corrector step may be replaced by

$$
\begin{aligned}
w^{n+1}_{i,j}\ A^{n+1}_{i,j} = \tilde{w}^{n+1}_{i,j}\ A^{n+1}_{i,j} - \Delta t/2 \{ & (\tilde{f}^{n+1}_{i,j} - f^{n}_{i+1,j})\ \Delta y^{n+1/2}_{i+1/2,j} \\
& - (\tilde{g}^{n+1}_{i,j} - g^{n}_{i+1,j})\ \Delta x^{n+1/2}_{i+1/2,j} \\
& + (\tilde{f}^{n+1}_{i,j} - f^{n}_{i,j+1})\ \Delta y^{n+1/2}_{i,j+1/2} - (\tilde{g}^{n+1}_{i,j} - g^{n}_{i,j+1})\ \Delta x^{n+1/2}_{i,j+1/2} \\
& + (\tilde{f}^{n+1}_{i-1,j} - f^{n}_{i,j})\ \Delta y^{n+1/2}_{i-1/2,j} - (\tilde{g}^{n+1}_{i-1,j} - g^{n}_{i,j})\ \Delta x^{n+1/2}_{i-1/2,j} \\
& + (\tilde{f}^{n+1}_{i,j-1} - f^{n}_{i,j})\ \Delta y^{n+1/2}_{i,j-1/2} - (\tilde{g}^{n+1}_{i,j-1} - g^{n}_{i,j})\ \Delta x^{n+1/2}_{i,j-1/2} \}
\end{aligned}
\tag{5}
$$

Equations (3) and (5) differ from the usual predictor-corrector equations on a fixed grid only in the dependence of the cell areas and increments of x and y on the time step number. From the form of these equations, it is noted that the increments Δx and Δy should be computed as soon as the grid points at level n + 1 are defined and before any computations in (2) are performed. The cell areas at level n + 1 are not computed until after the computation of the right-hand side of (3) is completed. In this manner, previous values may be overwritten and no additional storage is required for computing on a moving grid. It should be noted that the computation of cell areas at the same time level on both sides of (3) would be equivalent to the approximation of the area integral in (2) by a rectangular rule. This would result in a method with only first order accuracy in time.

III. EULER EQUATIONS

The Euler equations for inviscid compressible flow are equations of the form (1) where

$$
w = \begin{pmatrix} \rho \\ \rho u \\ \rho v \\ e \end{pmatrix}, \quad f = \begin{pmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ (e + p)u \end{pmatrix}, \quad g = \begin{pmatrix} \rho v \\ \rho uv \\ \rho v^2 + p \\ (e + p)v \end{pmatrix} \tag{6}
$$

The pressure is given by

$$p = (\gamma - 1)(e - 1/2\rho\,(u^2 + v^2)).$$

The numerical solution of these equations can be computed using the finite-volume method discussed in the previous section. Along with the equations (3) and (5) one must impose appropriate boundary conditions. On a solid wall, the velocity component normal to the boundary vanishes, or

$$u\,\Delta y - v\,\Delta x = 0.$$

In this report the normal derivative of the pressure has also been set to zero at a solid wall. Previous experience of Jacocks and Kneile [1] indicate that

this boundary condition will give satisfactory results even though the curvature of the wall is neglected. The number and type of inflow and outflow boundary conditions depend on whether the free stream flow is subsonic or supersonic. In the computations reported here, supersonic free stream conditions, were used. Thus all variables were prescribed at the inflow boundary. A one-sided differencing of the Euler equations was used to determine the values of the variables at the outflow boundary.

The Courant-Friedrichs-Lewy (CFL) criteria is a necessary condition for stability of the MacCormack finite-volume method. A rigorous determination of the CFL criteria would require the computation of the eigenvalues of the Jacobian matrices of f and g in (6). However, the intuitive approach of Jacocks and Kneile [1] can also be used to estimate the CFL criteria on a moving grid. The local speed of a sound wave is given by

$$c = \left(\frac{\gamma p}{\rho}\right)^{1/2}.$$

Consider the time interval required for a sound wave to pass from one side of a typical grid cell to the other. If side s is defined by the vector

$$s = (\Delta x, \Delta y)$$

and the area of the cell is A, then the distance from s to the opposite side can be approximated as

$$d = \frac{A}{|s|}$$

If $q = (u,v)$ denotes the velocity of the fluid and $r = (x_t, y_t)$ is the velocity of the cell, then the net rate of speed of a sound wave moving normal to s is

$$|(q - r) \cdot n| + c$$

where

$$n = \frac{(-\Delta y, \Delta x)}{|s|}$$

Thus the sound wave will reach the other side of the cell in a time interval of approximately

$$\delta = \frac{A}{(|(q - r) \cdot n| + c)\,|s|}$$

It is therefore intuitively clear that the CFL criteria will be satisfied when Δt is chosen to be bounded above by the minimum of the quantity δ over all sides of all cells. The form of δ reveals that the stabiity of the scheme is enhanced when the grid cells move with the fluid since this tends to decrease the vector q-r.

IV. EQUATIONS FOR GRID DYNAMICS

There are numerous reasons why one would be faced with the problem of computing the solution of fluid flow equations on moving grids. The specific problem which is considered here is the case where a double-connected region has one stationary and one moving boundary component. This type of region might be encountered in computing the relative trajectory of one body with respect to another body. More specifically, one might be interested in determining the trajectory of a bomb or external fuel tank after it is released from an aircraft. Regardless of the specific problem, there are two grid

related tasks which must be addressed. The first is to determine the position of the moving boundary component after each time interval. The second task is to redefine the grid on the region with boundary consisting of the stationary component and the moving component in its new position.

Since the Euler equations are being solved, no viscous forces will be considered. Consider the case of a cylindrical body whose cross-section defines a boundary component of the region. The only forces that will be included are those due to surface pressure and gravity. If C denotes the boundary component, then the x and y components of the force per unit length on the cylindrical body are given by

$$Fx = -\int_C (p - p_\infty)\, y_s\, ds$$

$$Fy = \int_C (p - p_\infty)\, x_s\, ds - mg,$$

where p_∞ denotes the free-stream pressure. The constant m is the mass per unit length and g is the gravitational constant. The total pitching moment about the center of gravity of the body can also be computed. Let (xc, yc) be the center of gravity. Then the pitching moment is given by

$$M = \int_C (p - p_\infty)\, \{(y - yc)\, y_s + (x - xc)\, x_s\}\{drc/ds\}\, ds$$

where

$$rc = \{(x - xc)^2 + (y - yc)^2\}^{1/2}$$

These forces and moments can be used to determine the acceleration of (xc, yc) in the x and y directions and the angular acceleration of the body about (xc, yc). Specifically, the three accelerations are given by

$$Ax = Fx/m$$
$$Ay = Fy/m$$
$$A\theta = M/I$$

where I is the moment of inertia of the body. The implementation of these formulae require a numerical integration scheme. If the grid points on C are indexed by (1/2, j - 1/2), j = 1,2, ---,k, and the notation of Section III is used, the following approximations may be used.

$$Ax = -1/m \sum_{j=1}^{k} (p_{1/2,j} - p_\infty)\, \Delta y_{1/2,j}$$

$$Ay = 1/m \sum_{j=1}^{k} (p_{1/2,j} - p_\infty)\, \Delta y_{1/2,j} - g \qquad (7)$$

$$A\theta = 1/I \sum_{j=1}^{k} (p_{1/2,j} - p_\infty)\, \{\Delta y_{1/2,j}\, (y_{1/2,j} - yc) + \Delta x_{1/2,j}\, (x_{1/2,j} - xc)\}$$

$$\cdot \{\Delta rc_{1/2,j}\}$$

Note that interpolation is needed to find the values of x, y, and p indexed by (1/2,j). Once an initial position and velocity of the body is prescribed,

the computed accelerations determine the position at each time step. If the pressure at time level n + 1 has been computed, then the accelerations at level n + 1 can also be computed by equations (7). The velocity of the center of gravity and the angular velocity are now approximated as

$$\begin{aligned} Vx^{n+1} &= Vx^{n} + \Delta t Ax^{n+1} \\ Vy^{n+1} &= Vy^{n} + \Delta t Ay^{n+1} \\ V\theta^{n+1} &= V_{\theta}^{n} + \Delta t A\theta^{n+1} \end{aligned}$$

Finally, the coordinates of the center of gravity and the rotational displacement about the center of gravity can be computed as

$$xc^{n+2} = xc^{n+1} + \Delta t\, Vx^{n+1}$$

$$yc^{n+2} = yc^{n+1} + \Delta t\, Vy^{n+1},$$

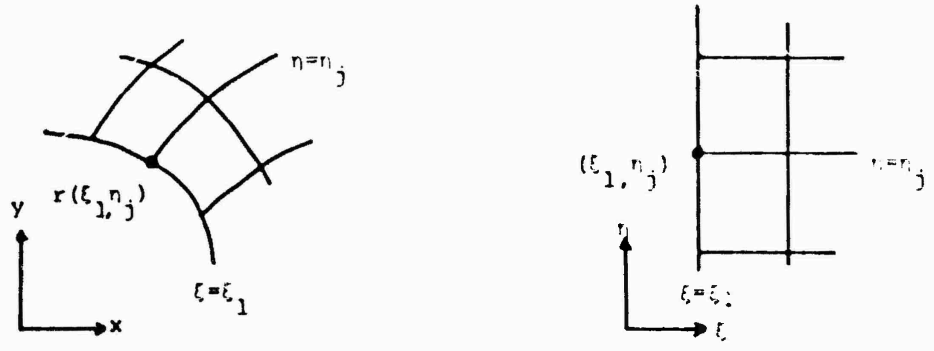
$$\Delta\theta = \Delta t\, V\theta^{n+1}$$

The translation and rotation of the body has been computed. Before the solution process can continue, a new grid is needed for time level n + 2. It will be assumed that the grid points on C move with the body and retain their same relative positions. While this may be a rational assumption for translations, it may not be advisable for large rotational displacements. The aim here is to derive an efficient method for redistributing the interior grid points so that they do not fall outside of the region when the boundary component C is moved.

Elliptic systems of partial differential equations have been very successful in generating grids for complicated two and three-dimensional regions. As discussed in the report by Thompson et. al. [3], an iterative procedure is required to solve these systems. This makes such procedures inefficient when frequent regridding is necessary. However, many of the advantages of elliptic systems are enjoyed by parabolic systems of the form

$$r_t + ar_{\xi} + br_{\eta} = cr_{\xi\xi} + dr_{\xi\eta} + er_{\eta\eta}. \qquad (8)$$

where $r = (x,y)$ and ξ and η are curvilinear coordinates as indicated in the following sketch.



The grid points are defined by solving this system on a rectangular region in the $\xi\eta$ - plane. If the initial grid is given, then succeeding grids can be

generated by any of the marching procedures for solving parabolic equations. The object then is to pick the coefficients so that the grid points follow the boundary component C. Suppose the grid motion near C is governed by the convective (first order) terms in equation (8) while the diffusion (second order) terms are much smaller and serve only to smooth the solution. Let C be the image of the line $\xi = \xi_1$ and let $\eta = \eta_j$ denote grid lines intersecting C as above. If the diffusion terms are neglected, then the trajectories of the grid points can be specified by choosing the coefficients a and b to be solutions of the system

$$r_t + ar_\xi + br_\eta = 0. \tag{9}$$

Now suppose the motion of the grid points along $\eta = \eta_j$ is the same as the point $r(\xi_1, \eta_j)$ on C. Then a and b are defined in terms of grid velocities on C by

$$a(\xi,\eta) = 1/J\ (x_\eta\ (\xi,\eta)\ y_t(\xi_1,\eta) - x_t(\xi_1,\eta)\ y_\eta\ (\xi,\eta))$$

$$b(\xi,\eta) = 1/J\ (x_t(\xi_1,\eta)\ y_\xi\ (\xi,\eta) - x_\xi\ (\xi,\eta)\ y_t\ (\xi_1,\eta))$$

$$J = x_\xi\ (\xi,\eta)\ y_\eta\ (\xi,\eta) - x_\eta\ (\xi,\eta)\ y_\xi\ (\xi,\eta).$$

The coefficients c, d, and e in (8) are quite arbitrary and need only satisfy the condition $d^2 - 4ce < 0$.

A well-posed initial-boundary-value problem for the coordinates of the moving grid has now been defined. An explicit numerical algorithm has been proposed for solving this problem concurrent with the solution of the Euler equations. Since only coordinate values are considered in this Section, and not flow variables, we will let $r(\xi_i, \eta_j) = r_{i,j}$ rather then revert to the fractional indexing used in Section II. Central differences are used for aproximating the second order derivatives. However, due to stability considerations, one-sided upwind differences must be used for the first order spatial derivatives. The complete difference equation which governs the movement of the interior grid points at time level n + 1 can be written as

$$r^{n+1}_{i,j} = r^n_{i,j} + \Delta t \{ \varepsilon a\ (r^n_{i,j} - r^n_{i+\varepsilon,j}) + \delta b\ (r^n_{i,j} - r^n_{i,j+\delta})$$

$$+ c\ (r^n_{i+1,j} + r^n_{i-1,j} - 2r^n_{i,j})$$

$$+ d\ (r^n_{i+1,j+1} + r^n_{i-1,j-1} - r^n_{i+1,j-1} - r^n_{i-1,j+1})/4$$

$$+ e\ (r^n_{i,j+1} + r^n_{i,j-1} - 2r^n_{i,j})\}$$

where

$$\varepsilon = \begin{cases} 1 & \text{if } a \leq 0 \\ -1 & \text{if } a > 0, \end{cases}$$

$$\delta = \begin{cases} 1 & \text{if } b \leq 0 \\ -1 & \text{if } b > 0. \end{cases}$$

Based on the coefficients used by Thompson et. al. [3] in elliptic systems, the following values were selected for c, d, and e.

$$c = |r_\eta|^2/f$$

$$d = r_\xi \cdot r_\eta/f$$

$$e = |r_\xi|^2/f$$

where

$$f = 2\,\Delta t(|r_\xi|^2 + |r_\eta|^2)$$

Since the factors multiplying a and b in (10) are the order of the grid spacing and the factors multiplying c, d, and e are the order of the square of the grid spacing, these coefficients work especially well when a fine boundary layer grid is to move with a body and the smoothing is to occur in the region of the coarse grid. This is exactly the situation encountered in the examples of the next Section.

V. EXAMPLES

The computational results of this report serve only to validate the numerical algorithms and are not intended to be precise fluid dynamic simulations. It must be admitted that the reason for this deficiency lies in the inability of the original Euler code, which was developed to solve a nozzle flow problem, to generate accurate pressure values near stagnation points. Therefore, the numerical results will be presented from a qualitative rather than a quantitative point of view. All grid movement was started in what was roughly a steady state flow field.

The first series of computations were performed on the grid illustrated in Figure 1. The interior boundary component is the crossection of an elliptic cylinder and the outer boundary component is a nearly circular ellipse. The free steam flow is at Mach number 1.1 and zero angle of attack. If the cylinder is released from its fixed position, and gravitational forces are negligible, the cylinder will move to the right with the center remaining on the major axis. Figures 2 and 3 are of the same region of the xy-plane. Figure 2 is a closeup view of the elliptic cylinder in its original position and Figure 3 is the grid about the cylinder after the acceleration due to drag has caused the cylinder to move a distance of nearly half of its major axis. The noticeable rotational movement was due to a slight nonsymmetry in the numerical solution. If the acceleration of the body is primarily due to gravity, the body translates along the minor axis. Figure 4 indicates the position of the cylinder after translation by a large gravitational force.

In both Figures 3 and 4, the grid cells near the body retain the same size and shape as in the original position indicated in Figure 2. The remaining part of the grid, which is not plotted, would appear as is Figure 1. A graphical

representation of the flow field generated by solving the Euler equations reveals the expected perturbations caused by movement of the body. Velocity vectors and mach contours for the body in Figure 3 are depicted in Figures 5 and 6. The velocity vectors have remained tangent to the body. Since the numerical solution was highly smoothed, the Mach contours are somewhat unrealistic. However, the position of the contours is characteristic of the angle of attack of the body at that particular point in time.

The second configuration is motivated by practical problems in aerodynamics such as aircraft-store separation and ground effects on low flying aircraft. A portion of the outer boundary component is now a horizontal line segment on which solid wall boundary conditions will be imposed. The grid for this region appears in Figure 7. The grid motion in this case is due to both gravitational and drag forces. At a later time, the body and nearby grid appears as in Figure 8. Note that the presence of the wall causes a nonsymmetric pressure distrution which induces a larger rotational motion in this example. A modification of the grid moving scheme was necessary to keep the grid points near the wall in their original position. All that was required was simply to multiply the coefficents in (8) by an exponential decay factor which was one near the inner boundary component and nearly zero at the outer boundary component. The free stream flow field for this problem is at a Mach number of 1.5 and the velocity field and Mach contours are plotted in Figures 9 and 10.

Although no addition computer storage is required for moving grids, the execution time is increased. The above cases employed a 30 by 40 computational grid and were run on a CDC 6600. In the initial time interval, when no grid movement was allowed, the computation of each 100 time steps required about 5 seconds of computer time. After grid movement was allowed to commence, each 100 time steps consumed about 7 seconds of computer time.

VI CONCLUSIONS

A numerical algorithm for solving the Euler equations on a moving grid has been developed and coded. The numerical results represent the gross characteristics of the flow field. The accuracy of the results was limited by the original Euler code and not related to grid motion. No doubt improved results could have been obtained if more effort had been expended on optimizing the smoothing procedure for this particular geometric configuration. Rather it was decided to direct our efforts toward the grid moving procedure. Most likely, the potential user will want to incorporate the moving grid techniques into a highly developed production code. This report should make that task easier. All of the concepts can be easily applied to three-dimesional problems with no more than the usual complexity associated with an added dimension.

It is obvious from the above statements that much work remains to be done. The eventual goal is the solution of the three-dimensional Navier-Stokes equations for fluid dynamic problems requiring moving grids. The technique of allowing the interior grid to move with the boundary was developed with this goal in mind. For more complex geometric problems involving multiple bodies or large rotational motions, this technique would have to be augmented by an occasional redistribution of boundary points or even a regriding of the complete region whenever the grid became distorted. It should be noted that a regriding of the region would require an interpolation of the solution values

from the old grid to the new grid. One advantage of the moving grid algorithm of this report is that interpolation is not needed.

A rather heavy penalty must be paid if the solution of the fluid flow problem is to follow the grid dynamics. In the examples reported here, the execution time of the Euler code was increased by about 40 per cent. Clearly, the program would be more efficient if the grid were only updated when necessary rather than at each time step. This option was not tested since no simple procedure for determining when updating was necessary could be found.

REFERENCES

1. Jacocks, J.L. and Kneile, K.R., "Computation of Three-Dimensional Time-Dependent Flow Using the Euler Equations," AEDC-TR-80-49, Arnold Engineering Development Center, Arnold Air Force Station, Tenn., October 1980.

2. MacCormack, R.W., "The Effect of Viscosity in Hypervelocity Impact Cratering," AIAA Paper 69-354, May 1969.

3. Thompson, J.F, Thames, F.C., and Mastin, C.W., "Boundary-Fitted Curvilinear Coordinate Systems for Solution of Partial Differential Equations on Fields Containing any Number of Arbitrary Two-Dimensional Bodies," NASA CR-2729, National Aeronautics and Space Administration, Washington, DC, July 1977.

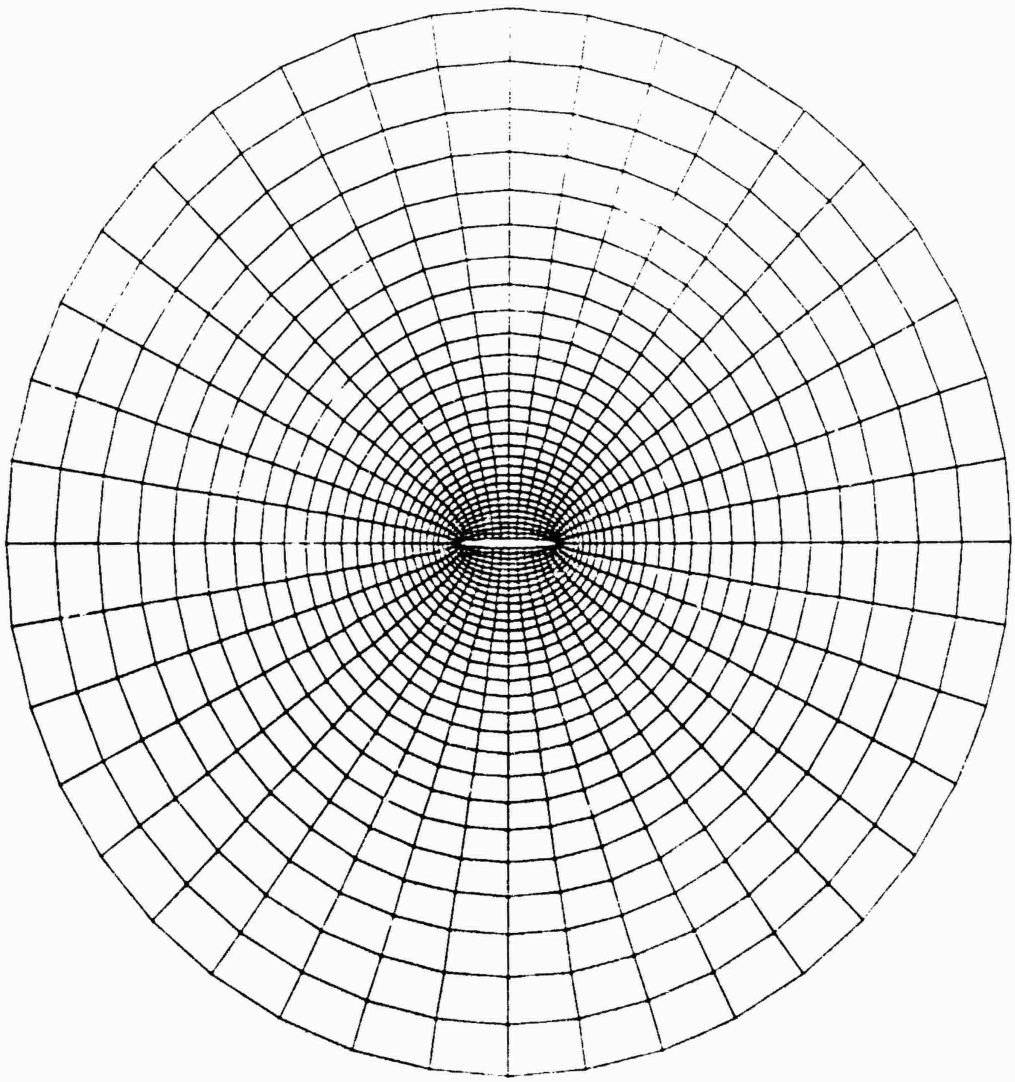


Figure 1. Complete grid about ellipse - initial position

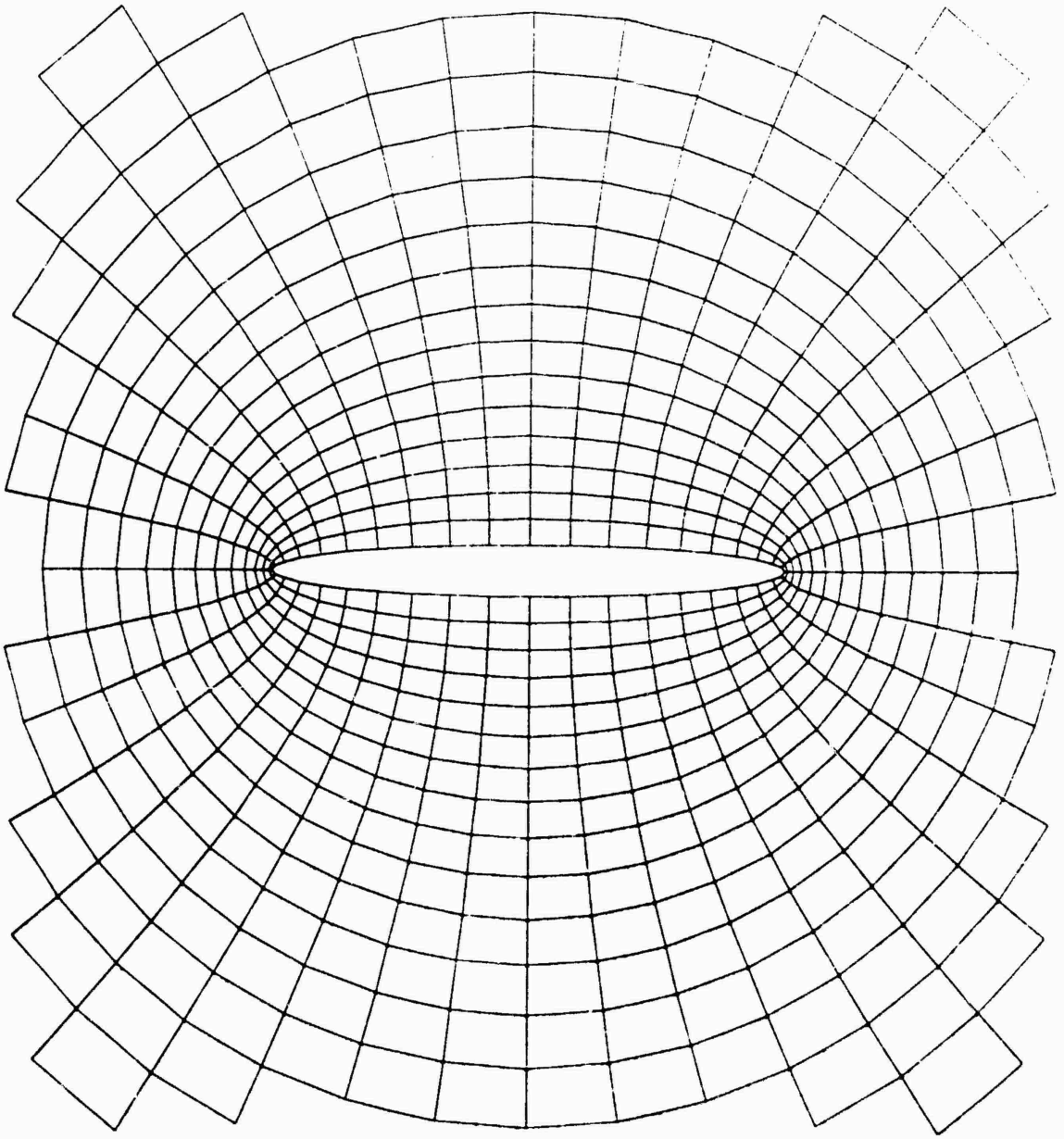


Figure 2. Partial grid about ellipse - initial position

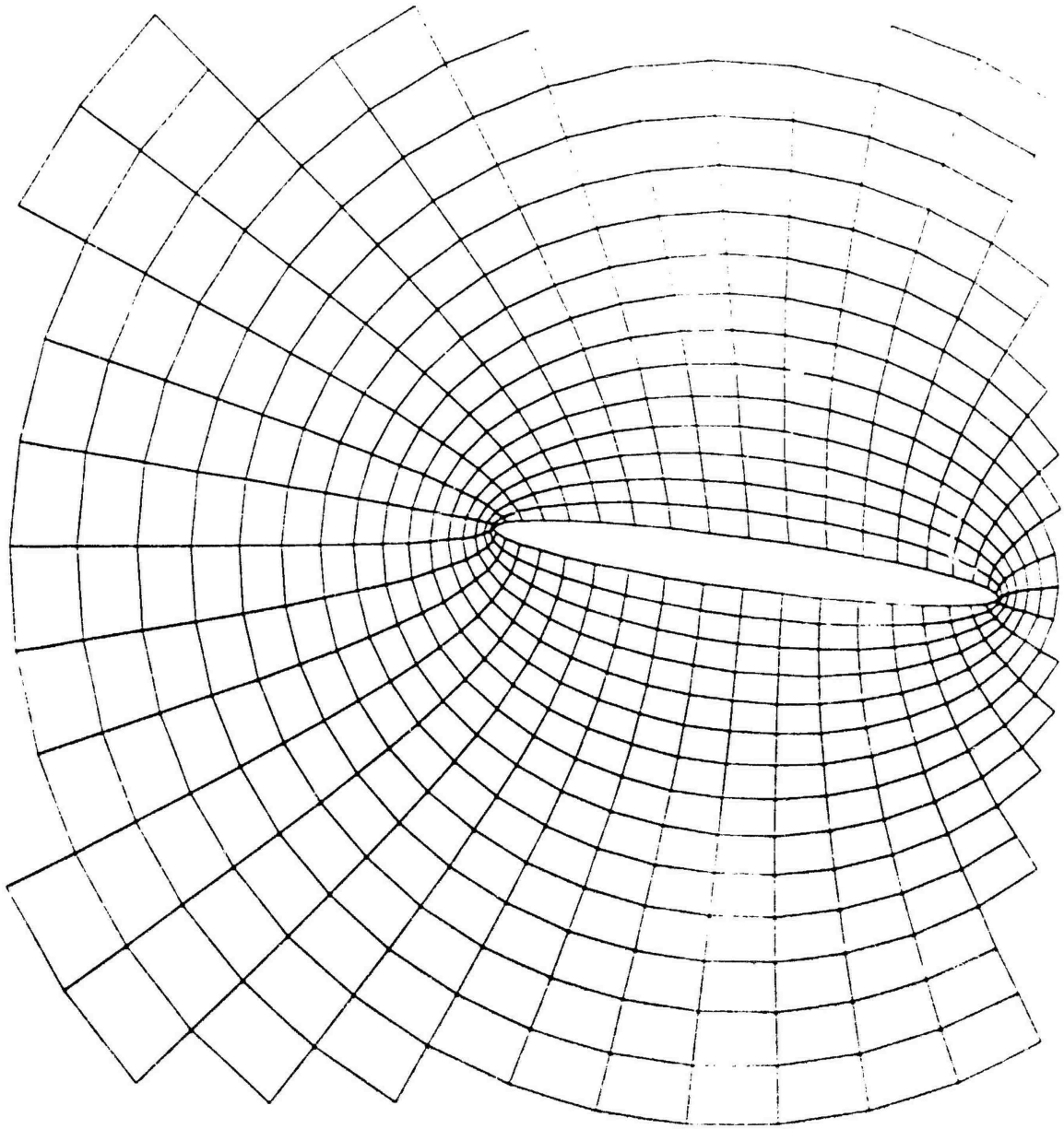


Figure 3. Partial grid about ellipse - after motion by drag forces

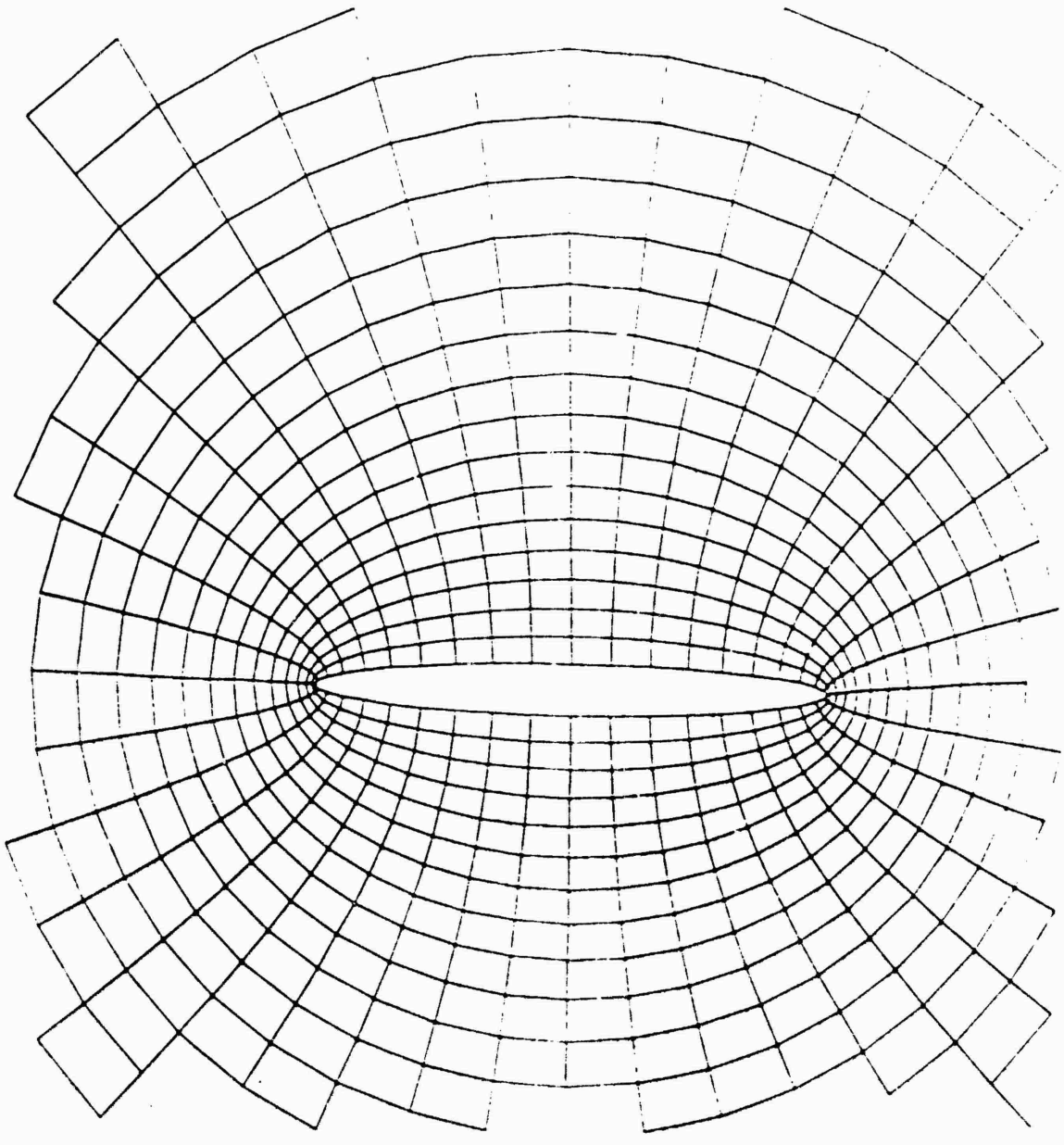


Figure 4. Partial grid about ellipse - after motion by gravitational forces

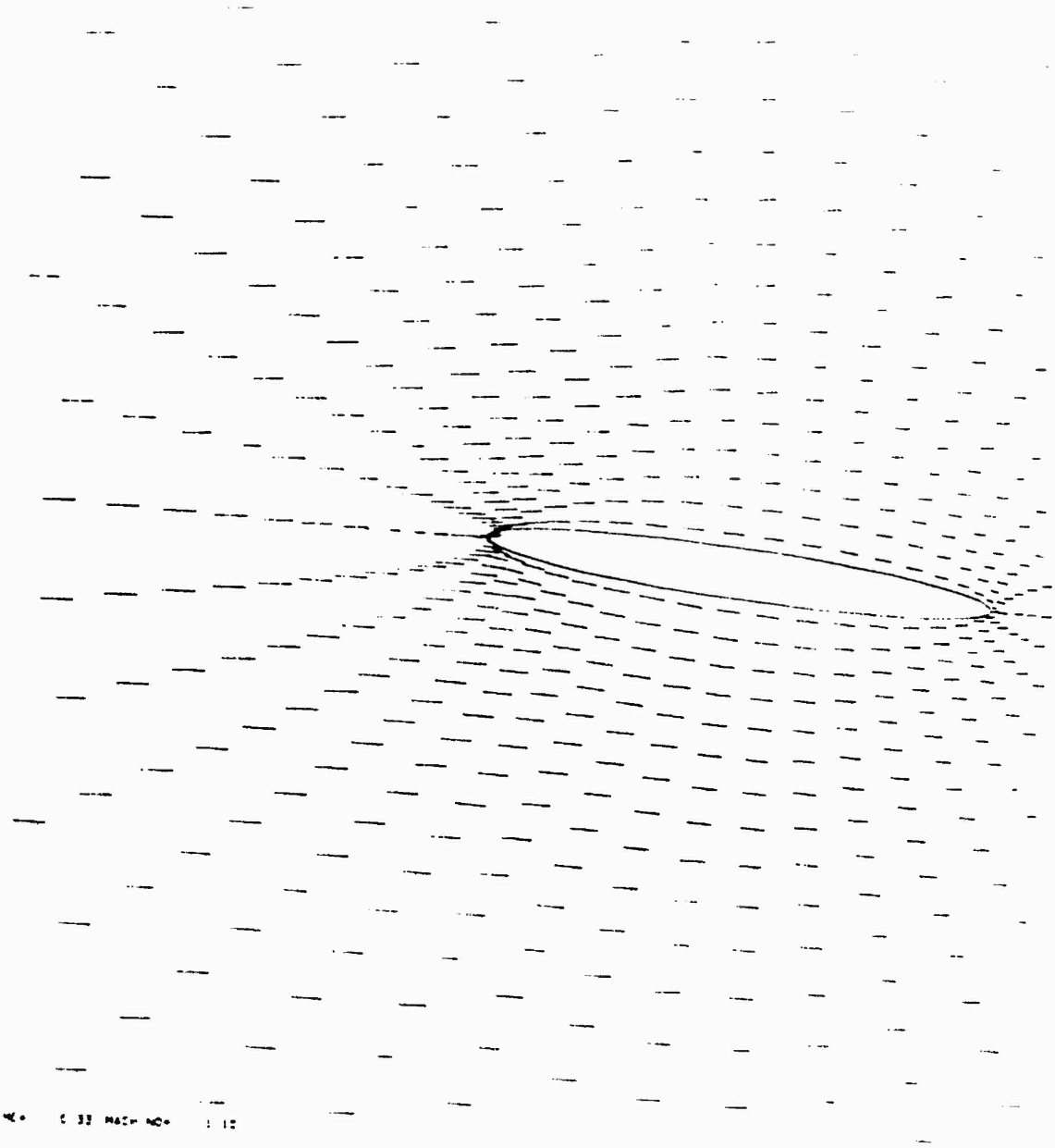


Figure 5. Velocity vectors

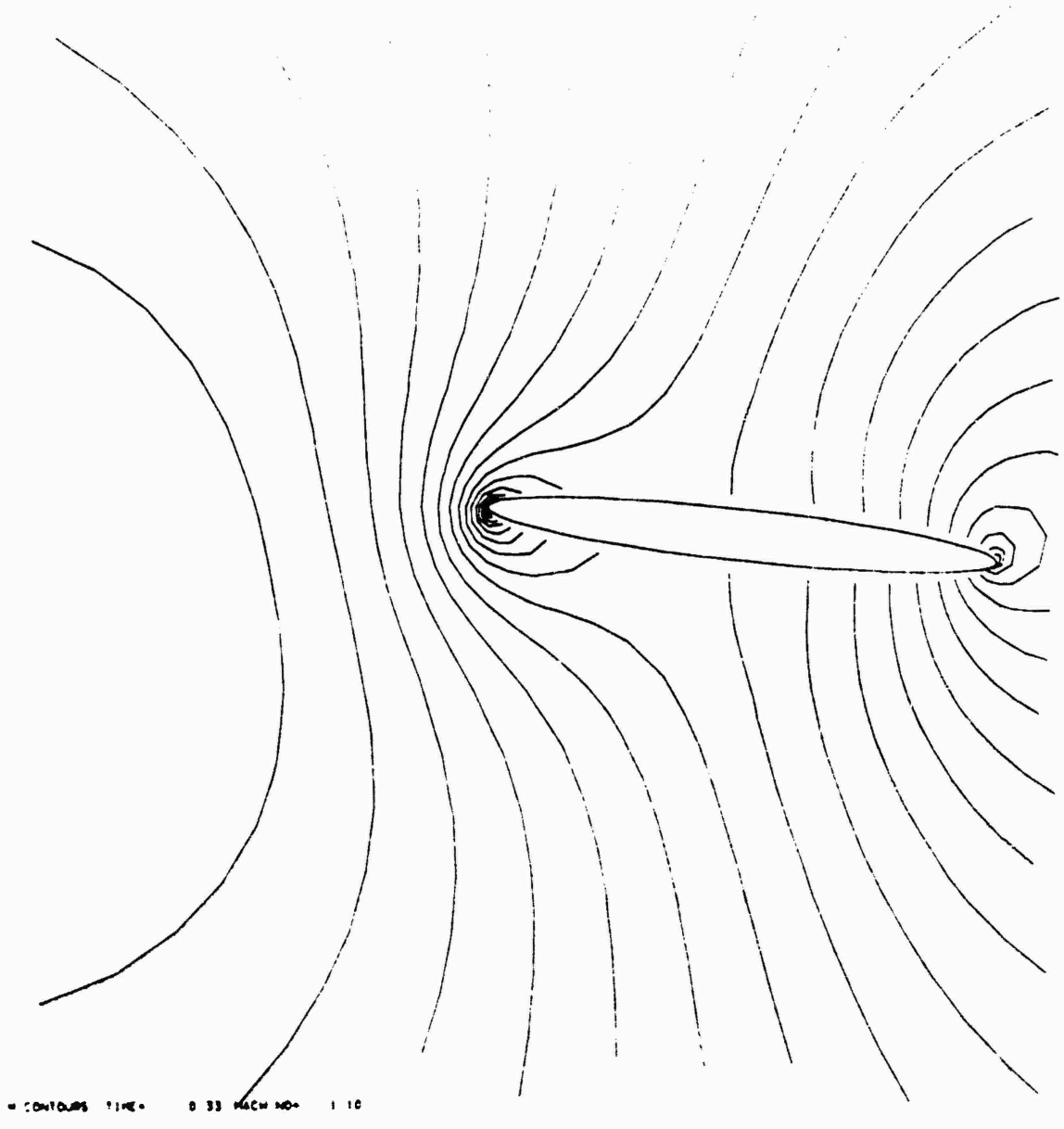


Figure 6. Contours of equal Mach number

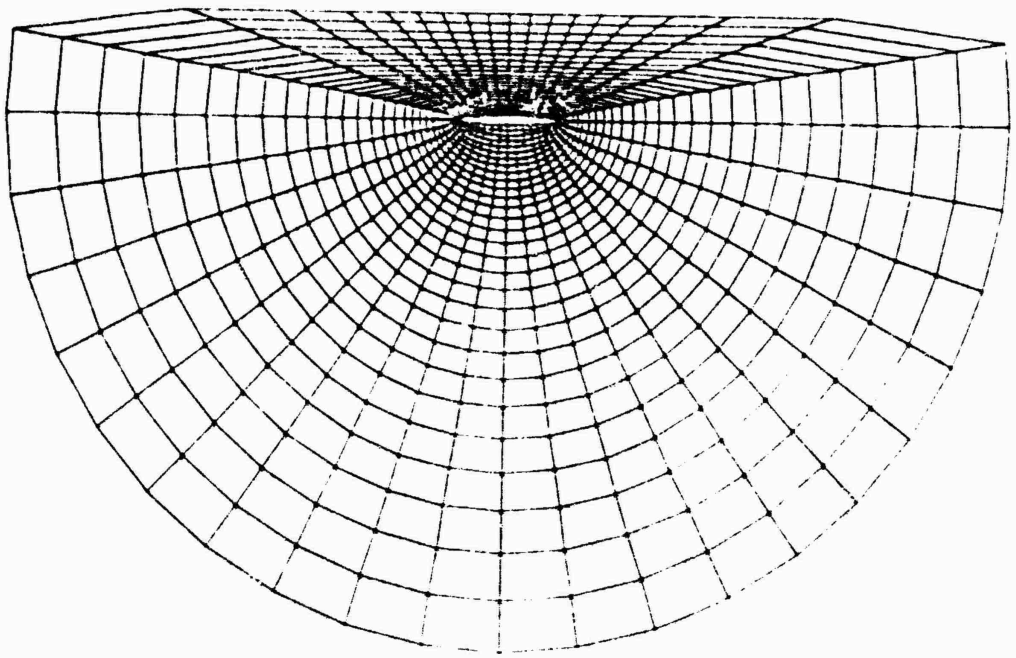


Figure 7. Complete grid for body/wall configuration - initial position

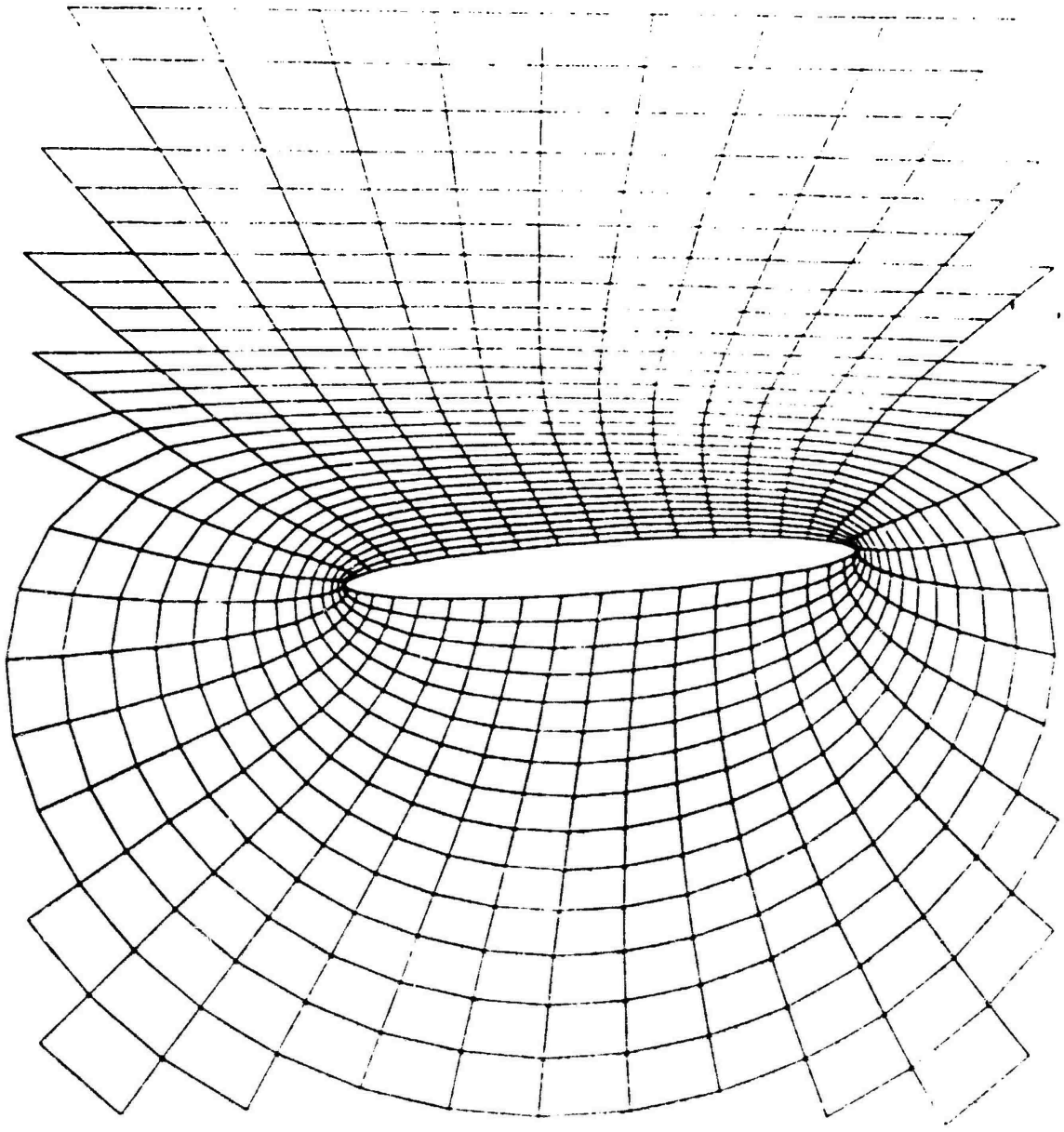


Figure 8. Partial grid for body/wall configuration - after grid movement

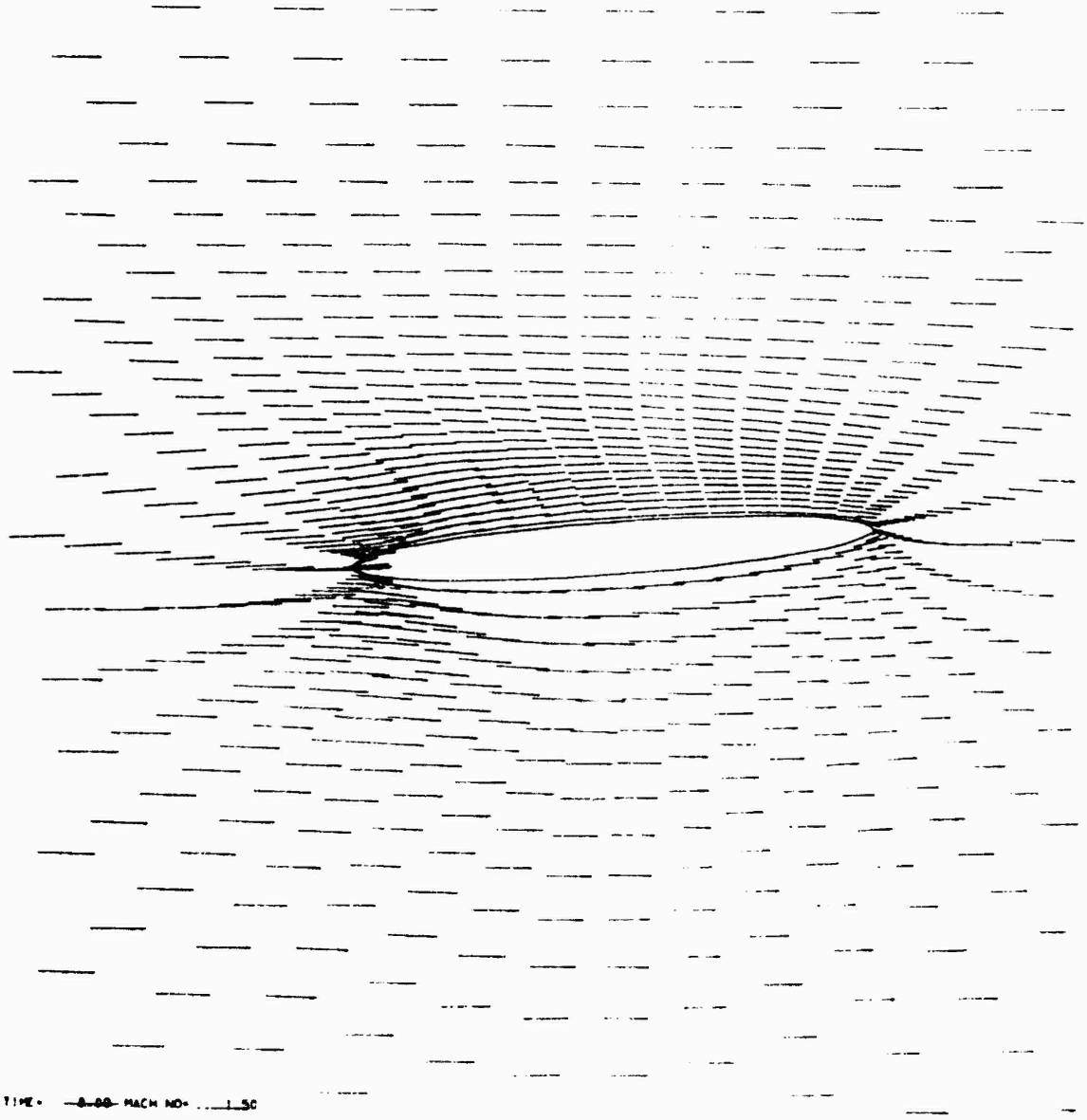


Figure 9. Velocity vectors

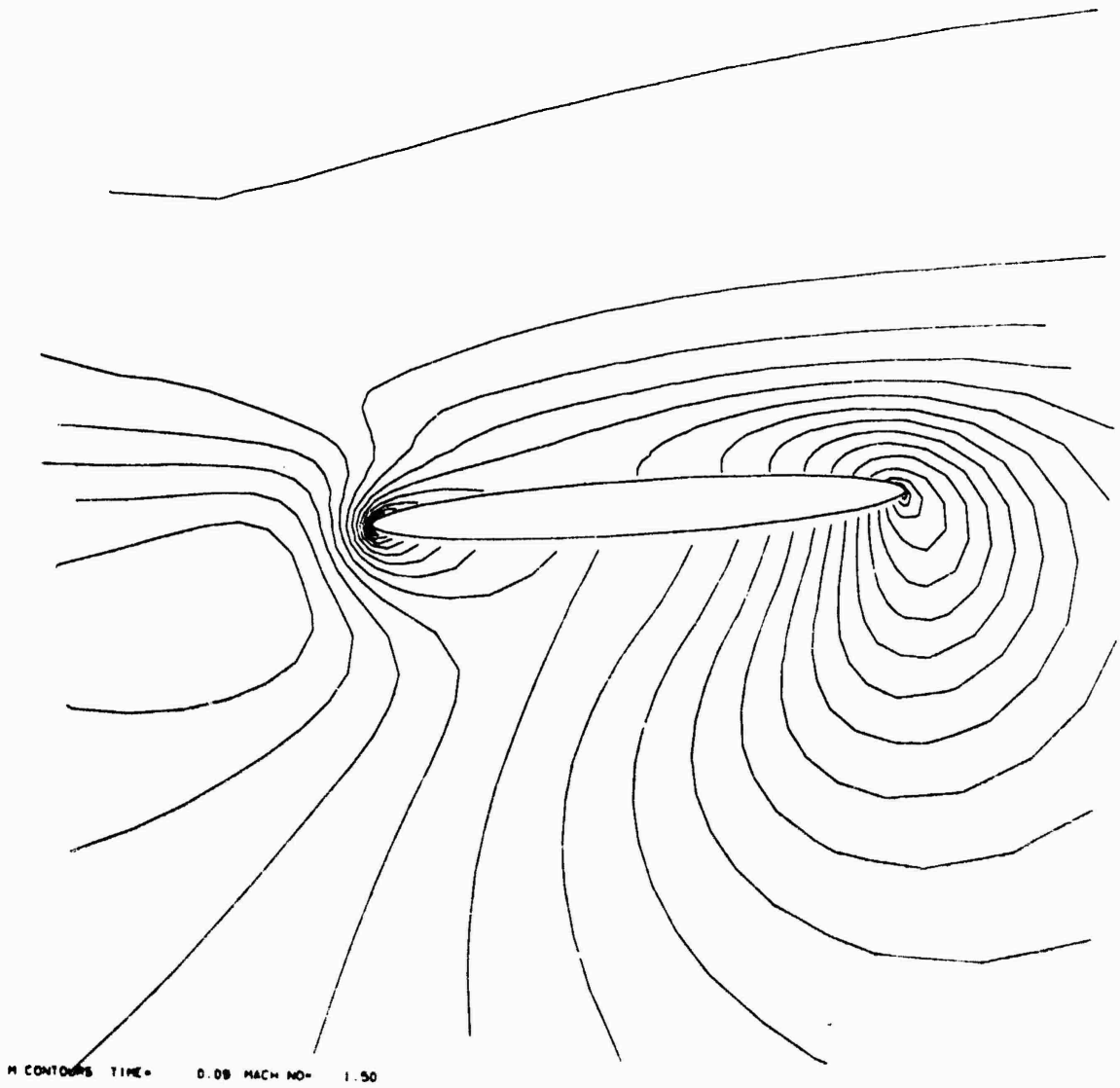


Figure 10. Contours of equal Mach number

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

APPROXIMATE EVALUATION OF OPTIMAL CONTROL MINIMIZING NONCOERCIVE COST-FUNCTIONALS OVER UNBOUNDED SETS; HYPERBOLIC SYSTEMS

Prepared By:	Dr. Tapas Mazumdar
Academic Rank:	Associate Professor
Department and University:	Department of Mathematics and Statistics, Wright State University, Dayton, Ohio 45435
Research Location:	Flight Dynamics Laboratory, Structures and Dynamics Division, Analysis and Optimization Branch, Wright Aeronautical Laboratories, Wright-Patterson Air Force Base
USAF Research:	Dr. Vipperla B. Venkayya
Date:	September 9, 1983
Contract No.:	F49620-82-C-0035

APPROXIMATE EVALUATION OF OPTIMAL CONTROL MINIMIZING NONCOERCIVE COST-FUNCTIONALS OVER UNBOUNDED SETS; HYPERBOLIC SYSTEMS

by

TAPAS MAZUMDAR

ABSTRACT

In this report we give a preliminary iterative approximation procedure to solve the optimal control scheme that leads to the minimization of a, possibly noncoercive, specified performance index of a system governed by the usual type of hyperbolic equation. The minimization is over specified types of closed convex sets that may possibly be unbounded.

ACKNOWLEDGEMENT

The author would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for providing him with the opportunity to spend a very worthwhile and educational summer at the Flight Dynamics Laboratory of the Wright-Patterson Air Force Base, Ohio.

Thanks are also due to Dr. Vipperla B. Venkayya for suggesting this area of research and for his constant encouragement.

I. INTRODUCTION:

Feedback optimal control problems related to partial differential equations of the hyperbolic type are of interest because it is known that vibrations of large flexible structures in space are governed by hyperbolic differential equations. To prevent a breakdown in the space-structure, these vibrations must be kept under control, near to a pre-determined desirable level. This is achieved by first sensing the magnitude of the vibrations by sensors which, then, trigger the controls or actuators to alter the external force-function that creates the vibrations in accordance with a law governed by a hyperbolic differential equation. So, one of the underlying problems is to arrive at a control energy that will alter the external forcing function to just the right extent with a view to keeping to a minimum level a certain "performance index" or "cost-functional" which measures, among other things, "the excess vibrations" i.e. the deviation of the state of the system from a desired state.

Lacking the knowledge of a most suitable expression for the performance index, we keep all possibilities in mind in doing the thoeretical work of minimizing it. Of several types of performance indices that may be employed, we have here worked with a general type that is not restricted by the "convexity" assumptions usual in literature (cf. [1]) when the set of admissible controls is unbounded. It has been shown in [2] that with this kind of performance index, in certain general types of unbounded sets of admissible controls, an optimal control exists that minimizes the performance index. So, under the stipulated conditions, this renders false the naive intuition: the

greater is our technological ability to develop more and more control energy, the greater is our success in controlling unwanted vibrations. Thus, the energy producing technology need not be developed beyond an optimum point, in our context.

We have maintained the "partial differential equations approach" to the optimal control problem for various reasons. Efforts are more needed to (i) devise ways of approximate numerical calculation of the optimal control whose existence and uniqueness is proved in [2], and (ii) study the various instability problems that arise because of the practical difficulty of controlling all the variables on which the optimal control depends.

II. OBJECTIVE AND THE PROBLEM SET-UP:

The basic Hilbert spaces with which we work here are $H = L^2(\Omega; \mathbb{C})$ and $\mathcal{H} = L^2([0,T]; H) \equiv L^2(\Omega'; \mathbb{C})$, where Ω is a bounded open subset of $\mathbb{R}^p$, p being a positive integer, $\mathbb{R}$ being the set of real numbers, $\mathbb{C}$ being the set of complex numbers, and $\Omega' = \Omega \times (0,T)$, $0<T<\infty$. The norms in H and $\mathcal{H}$ are denoted by $|\cdot|_H$ and $|\cdot|_{\mathcal{H}}$ respectively, with the corresponding inner products given by

$$(u,v)_H = \int_\Omega u(x)\,\overline{v(x)}\,dx \quad \text{for all} \quad u,v \in H,$$

and

$$(f,g)_{\mathcal{H}} = \int_0^T \int_\Omega f(x,t)\,\overline{g(x,t)}\,dx\,dt \quad \text{for all} \quad f,g \in \mathcal{H}.$$

Here $\overline{v(x)}$ denotes the complex conjugate of $v(x)$. H is identified with its anti-dual H', and $\mathcal{H}$ with its. In addition, we will deal with Hilbert spaces $(V, \|\cdot\|_V)$ and $(\mathcal{V}, \|\cdot\|_{\mathcal{V}})$ where V is algebraically and topologically contained in H, V is a dense subset of H, and $\mathcal{V} = L^2([0,T]; V)$. The inner products of V and $\mathcal{V}$ are denoted respectively by $((\cdot, \cdot))_V$ and $((\cdot, \cdot))_{\mathcal{V}}$.

Of concern in this report are certain aspects of a feedback control problem, as explained below.

We will begin with a space of controls, $\mathcal{U} = \Xi^{(n)} \equiv \Xi \times \Xi \times \cdots \times \Xi$ (n factors), where $(\Xi, |||\cdot|||_\Xi)$ is a complex separable Hilbert space, algebraically and topologically contained in $(\mathcal{H}, |\cdot|_{\mathcal{H}})$. $\mathcal{U}$ is then a product Hilbert space, with the usual Euclidean norm and inner product denoted by $|||\cdot|||_{\mathcal{U}}$ and $(((\cdot, \cdot)))_{\mathcal{U}}$ respectively. The control feedback operator $\mathcal{B}$ on $\mathcal{U}$ is assumed to be extendible to a continuous linear operator $: \mathcal{H}^{(n)} \to \mathcal{H}$, the extension being denoted again by $\mathcal{B}$. The set of admissible controls, in this report, is a set of the form

$$\mathcal{U}_{ad} = \left\{ \underline{v} \in \mathcal{U} : \left| \sum_{i=1}^{n} (\beta_i v_i - \mathcal{K}_i(v_i)) \right|_{\mathcal{H}} \leq \lambda \right\} \quad \cdots \quad (1)$$

where λ is a constant, $\underline{v} = (v_1, v_2, \cdots, v_n) \equiv (v_i)_{i=1}^n \in \mathcal{U}$, and for all i, $1 \leq i \leq n$, $\mathcal{K}_i : \Xi \to \mathcal{H}$ is a continuous, linear operator. Also, $\beta_i \in \mathcal{H}$ for all i, $1 \leq i \leq n$.

Assumption on each $\mathcal{K}_i$. There exists a constant $\gamma > 0$ such that

$$\left| \sum_{i=1}^{n} \mathcal{K}_i(v_i) \right|_{\mathcal{H}} \geq \nu \, |||\underline{v}|||_{\mathcal{U}} \quad \text{for all} \quad \underline{v} \in \mathcal{U}. \quad \cdots \quad (2)$$

Assumption on each f_i. There exist constants m_i, M_i such that

$$0 < m_i \leq f_i(x,t) \leq M_i < \infty \quad \cdots \quad \cdots \quad (3)$$

for almost all $(x,t) \in \Omega'$. (Some or all of the f_i's may satisfy the alternative requirement, $-\infty < -M_i \leq f_i(x,t) \leq -m_i < 0$, and our work goes through with only minor modifications).

Although our results are true for all $\mathcal{U}_{ad}$'s of the form (1), and, indeed, for somewhat more general forms of $\mathcal{U}_{ad}$'s, our interest obviously lies only on those $\mathcal{U}_{ad}$'s which are unbounded in $\mathcal{U}$. It is easy to show (cf. [2]) that the definition (1) of $\mathcal{U}_{ad}$ renders it closed and convex in $\mathcal{U}$.

In this report we will lay down the beginning of a theory that will eventually allow us to numerically approximate the optimal control $u_0 \in \mathcal{U}_{ad}$ that minimizes a certain performance index (described later in (6)) of a system whose state y is governed by a hyperbolic partial differential equation of the form:

$$\frac{\partial^2 y}{\partial t^2} + A(t)\, y = f. \quad \cdots \quad \cdots \quad \cdots \quad (4)$$

where, $\forall\, t \in [0,T]$, $A(t) : V \to V'$, V' being the anti-dual of V, is a continuous linear operator satisfying certain well-known

conditions of continuous weak-differentiability, symmetry and strong coercivity (or V-ellipticity). These conditions are needed for the existence and uniqueness of solution of (4) under very general types of initial and boundary conditions ([1]).

We are interested in a certain type of performance index, $J : \mathcal{U} \to \mathbb{R}$, and to completely describe this J, we introduce among the hypotheses, (i) a complex Hilbert space of "observations", $(\mathcal{W}, \|\cdot\|_{\mathcal{W}})$, (ii) a "desired observation" $z_d \in \mathcal{W}$, (iii) a continuous linear "observation" operator $C : \mathcal{H} \to \mathcal{W}$, and (iv) a continuous, linear "control law" operator $N : \mathcal{U} \to \mathcal{U}$ such that

$$\left.\begin{aligned} &(\!(\underline{v}, N\underline{v})\!)_{\mathcal{U}} \text{ is real, and} \\ &\sum_{i=1}^{n} \mathrm{Re}\,(v_i, \mathcal{K}_i(v_i))_{\mathcal{H}} + (\!(N\underline{v}, \underline{v})\!)_{\mathcal{U}} \geq 0 \\ &\qquad \text{for all } \underline{v} = (v_i)_{i=1}^{n} \in \mathcal{U}_{ad}. \end{aligned}\right\} \quad \cdots \quad (5)$$

The performance index or the cost-functional J is now given by

$$J(\underline{v}) = F_1(\underline{v}, \underline{v}) + \|C y(\cdot,\cdot;\underline{v}) - z_d\|_{\mathcal{W}} + \sum_{i=1}^{n} (v_i, \mathcal{K}_i(v_i))_{\mathcal{H}} + (\!(\underline{v}, N\underline{v})\!)_{\mathcal{U}} \quad \text{for all} \quad \underline{v} \in \mathcal{U}_{ad}, \quad \cdots \quad (6)$$

where $y(\cdot,\cdot;\underline{v})$ is the "state of the system", and is, in fact, the unique solution of the closed-loop system ($f \in \mathcal{H}$, $y_0 \in V$ and $y_1 \in H$ being given)

$$\left.\begin{aligned}
&\frac{\partial^2 y(x,t;v)}{\partial t^2} + A(t)\,y(x,t;v) \\
&\qquad = f(x,t) + (Bv)(x,t) \quad \text{a.e. in} \quad \Omega', \\
&y(\cdot,\cdot;v) \in \mathcal{Y}, \qquad \frac{\partial y(\cdot,\cdot;v)}{\partial t} \in \mathcal{H}, \\
&y(x,0;v) = y_0(x) \quad \text{a.e. in} \quad \Omega, \\
&\frac{\partial y(x,0;v)}{\partial t} = y_1(x) \quad \text{a.e. in} \quad \Omega,
\end{aligned}\right\} \tag{7}$$

and $F_1 : \mathcal{U} \times \mathcal{U} \to \mathbb{R}$ is a function satisfying the following conditions,

$$v \mapsto F_1(v,v) : \mathcal{U} \to \mathbb{R} \quad \text{is convex,} \quad \cdots \quad (8)$$

and

there exists a constant $\alpha > 0$ such that

$$\left.\begin{aligned}
&F_1(v,v) - \frac{1}{2}\sum_{i=1}^{n} \frac{1}{m_i} \left| B_i v_i - \mathcal{K}_i(v_i) \right|_{\mathcal{H}}^2 \\
&\qquad + \frac{1}{2}\sum_{i=1}^{n} \frac{1}{M_i} \left| \mathcal{K}_i(v_i) \right|_{\mathcal{H}}^2 \geq \alpha \sum_{i=1}^{n} |v_i|_{\mathcal{H}}^2
\end{aligned}\right\} \tag{9}$$

for all $v = (v_1, \cdots, v_n) \in \mathcal{U}_{ad}$.

Let us emphasize that the J given in (6) need not be coercive, i.e. J

need not satisfy the condition " $J(\underline{v}) \geq (\text{const}) \|\underline{v}\|_{\mathcal{U}}^2 \; \forall \, \underline{v} \in \mathcal{U}_{ad}$ ". With minor modifications in the definition(1) and the assumption (8), we may even supply examples in which J does not satisfy the condtion, " $J(\underline{v}_m) \to \infty$ whenever $\|\underline{v}_m\|_{\mathcal{U}} \to \infty$ in $\mathcal{U}_{ad}$ ". Of course, our theory holds even if these conditons were true.

As stated earlier, the objective of the present report is to put forward a theory which will eventually lead to methods of numerically approximating the optimal control $u_0 \in \mathcal{U}_{ad}$ that will minimize the J given in (6) subject to the other stated stipulations.

III. EXISTENCE AND UNIQUENESS OF OPTIMAL CONTROL - VARIATIONAL FORMULATION:

We continue with the notations and the assumptions laid out above.

Let $\Lambda_{\mathcal{W}} : \mathcal{W} \to \mathcal{W}'$ and $\Lambda_{\mathcal{U}} : \mathcal{U} \to \mathcal{U}'$ be the canonical isomorphisms where $\mathcal{W}'$ and $\mathcal{U}'$ are the anti-duals of $\mathcal{W}$ and $\mathcal{U}$ respectively. Let the operator $\mathcal{K} : \mathcal{U} \to \mathcal{H}^{(m)}$ be defined by $(\underline{\mu}, \mathcal{K}(\underline{v}))_{\mathcal{H}^{(m)}} = \sum_{i=1}^{m} (\mu_i, \mathcal{K}_i(v_i))_{\mathcal{H}}$ for all $\underline{v} \in (\mathcal{U}_i)_{i=1}^{m}$ and for all $\underline{\mu} = (\mu_i)_{i=1}^{m}$ in $\mathcal{U}$. Let $B' : \mathcal{H} \to \mathcal{U}'$, $N^* : \mathcal{U} \to \mathcal{U}$, $\mathcal{K}' : \mathcal{H}^{(m)} \to \mathcal{U}'$, $\mathcal{C}' : \mathcal{W}' \to \mathcal{H}$ be the appropriate adjoints (cf. [3]) of the continuous linear operators $B : \mathcal{U} \to \mathcal{H}$, $N : \mathcal{U} \to \mathcal{U}$, $\mathcal{K} : \mathcal{U} \to \mathcal{H}^{(m)}$ and $\mathcal{C} : \mathcal{H} \to \mathcal{W}$ respectively. Then, under all the assumptions made earlier, we can prove the following theorem ([2]),

Theorem. There exists an optimal control $\underline{u}_0 \in \mathcal{U}_{ad}$ (see (1)) minimizing the performance index given by (6). This optimal control $\underline{u}_0$

is a simultaneous solution $\underline{v}$ of the system (7), the following adjoint system (p is called the adjoint state),

$$\left.\begin{aligned} &\frac{\partial^2 p}{\partial t^2} + A(t)p = C'\Lambda_W(Cy - y_d) \quad \text{in} \quad \Sigma, \\ &p \in V, \quad \frac{\partial p}{\partial t} \in H, \quad p(T) = 0 = \frac{\partial p}{\partial t}(T). \end{aligned}\right\} \quad \cdots \quad (10)$$

(where y, being a solution of (7), is a function of $\underline{v} \in \mathcal{U}_{ad}$, and so p should in reality be represented as $p(\underline{v})$, and $p(T)$ as $p(T;\underline{v})$, etc.), and the following variational inequality,

$$\begin{aligned} &F_1(\underline{v},\underline{u}) - F_1(\underline{v},\underline{v}) \\ &+ \mathrm{Re}\,(((N+N^*)(\underline{u}) + \Lambda_{\mathcal{U}}^{-1}(\mathcal{K}+\mathcal{K}')(\underline{v}), \underline{u}-\underline{v}))_{\mathcal{U}} \\ &+ \mathrm{Re}\,((2\Lambda_{\mathcal{U}}^{-1}B'p(\underline{v}), \underline{u}-\underline{v}))_{\mathcal{U}} \geq 0 \quad \text{for all } \underline{u} \in \mathcal{U}_{ad}. \end{aligned} \quad (11)$$

Moreover, this optimal control is unique in $\mathcal{U}_{ad}$ if

$$\left.\begin{aligned} &\text{at least one of the } n+2 \text{ maps, } \underline{u} \mapsto F_1(\underline{v},\underline{u}), \\ &v_i \mapsto \mathrm{Re}\,(v_i, \mathcal{K}_i(v_i))_{H_i}, \quad \underline{v} \mapsto (((\underline{v}, N\underline{v})))_{\mathcal{U}} \\ &\text{is strictly convex.} \end{aligned}\right\} \quad (12)$$

This completes the statement of the theorem. Nontrivial examples applying this existence-uniqueness theorem are given in [2].

IV. A PRELIMINARY APPROXIMATION PROCEDURE TO EVALUATE THE OPTIMAL CONTROL $\underline{u}_0$:

We make all the assumptions made above, including (12). In addition, we assume that (see [3] for technical definitons)

$$B : \mathcal{U} \to H \text{ is a compact operator,} \quad \cdots \quad \cdots \quad (13)$$

and

$$\left.\begin{array}{l} \underline{v} \mapsto F_1(\underline{v}, \underline{v}) : \mathcal{U} \to \mathbb{R} \quad \text{is lower} \\ \text{semi-continuous in the weak topology of } \mathcal{U}. \end{array}\right\} \quad \cdots \quad \cdots \quad (14)$$

Results of the previous section yield existence and uniqueness of an optimal control $\underline{u}_0 \in \mathcal{U}_{ad}$. Fix a positive number $\rho \geq 2 |||\underline{u}_0|||_{\mathcal{U}}$, and let

$$\mathcal{U}_{ad}^{(\rho)} = \mathcal{U}_{ad} \cap \left\{ \underline{v} \in \mathcal{U} : |||\underline{v}|||_{\mathcal{U}} \leq \rho \right\}.$$

We now give below steps of a preliminary approximation procedure with the limited objective of approximating the solution of only the variational inequality (11). This procedure, as it stands here, is not directly adaptable to numerical treatment because at each step exact solution of an approximate variational inequality is called for. Development of the following ideas into a numerical analytic procedure is left for the future.

Step 1. Start with any conveniently chosen $\underline{u}^{(1)} \in \mathcal{U}_{ad}^{(\rho)}$.

Step 2. Obtain the exact solution of the system (7) in which $\underline{y}$ has been replaced by $\underline{u}^{(1)}$. Such a solution exists by a known theorem (cf. [1]). Call this solution $\underline{y}^{(1)}$.

Step 3. Replace $\underline{y}$ by $\underline{y}^{(1)}$ in (10), and then obtain its exact solution. Such a solution exists by a known theorem. Call this solution $p^{(1)}$.

Step 4. In (11) replace $p(\underline{v})$ by $p^{(1)}$ and $\mathcal{U}_{ad}$ by $\mathcal{U}_{ad}^{(\rho)}$. Obtain the exact solution of the resulting variational inequality. Since $\mathcal{U}_{ad}^{(\rho)}$ is a closed convex bounded set in $\mathcal{U}$, a

known theorem tells us that such a solution exists. Call this solution $\underline{u}^{(2)}$.

Step 5. Repeat Step 2 with $\underline{u}^{(2)}$ replacing the $\underline{v}$ of system (7) (instead of $\underline{u}^{(1)}$ replacing the $\underline{v}$). This will yield a $y^{(2)}$.

Step 6. Repeat Step 3 with $y^{(2)}$ replacing the y in (10). This will yield a $p^{(2)}$.

Step 7. Repeat Step 4 with $p^{(2)}$ replacing $p(\underline{v})$ in (11). This will yield a $\underline{u}^{(3)}$.

This procedure may be repeated indefinitely. We have thus inductively constructed an infinite sequence

$$\underline{u}^{(1)}, \underline{u}^{(2)}, \cdots\cdots, \underline{u}^{(i)}, \cdots\cdots, \underline{u}^{(m)}, \cdots\cdots \qquad \cdots \qquad (15)$$

of elements of the closed convex bounded set $\mathcal{U}_{ad}^{(\rho)}$. At the i^{th} stage we actually have

$$\begin{aligned} &F_1(\underline{v}, \underline{v}) - F_1(\underline{u}^{(i+1)}, \underline{u}^{(i+1)}) \\ &+ \mathcal{R}e\,(\!(\!((N+N^*)(\underline{u}^{(i+1)}) \\ &\qquad + \Lambda_u^{-1}(\mathcal{K}+\mathcal{K}')(\underline{u}^{(i+1)}),\ \underline{v} - \underline{u}^{(i+1)})\!)\!)_u \\ &+ \mathcal{R}e\,(\!(2\Lambda_u^{-1}B'p^{(i)},\ \underline{v} - \underline{u}^{(i+1)})\!)_u \geq 0 \quad \text{for all } \underline{v} \in \mathcal{U}_{ad}^{(\rho)}. \end{aligned} \qquad (16)$$

Since $\underline{u}^{(i)} \in \mathcal{U}_{ad}^{(\rho)}$, we have $\|\underline{u}^{(i)}\|_u \leq \rho$ for all i. Thus, the sequence (15) has a weakly convergent subsequence (for this weak sequential compactness property of Hilbert spaces, see [3]). We are now going to show that whenever $\{\underline{u}^{(\mu)}\}_{\mu=1}^{\infty}$ is a subsequence of (15) that converges weakly in $\mathcal{U}$, this subsequence converges weakly to $\underline{u}_0$.

So, suppose $\{\underline{u}^{(\mu)}\}_{\mu=1}^{\infty}$ converges, in the weak topology of $\mathcal{U}$, to $\underline{v}_0 \in \mathcal{U}$. Since $\mathcal{U}_{ad}^{(P)}$ is weakly closed in $\mathcal{U}$, we have, in fact, $\underline{v}_0 \in \mathcal{U}_{ad}^{(P)}$. By (13), $\{B(\underline{u}^{(\mu)})\}_{\mu=1}^{\infty}$ converges to $B\underline{v}_0$ in H. Since the solution of the system (7), and the derivative of the solution, depend continuously on the forcing function ([1]), we have

$$y^{(\mu)} \to y(\underline{v}_0) \text{ in } V \text{ as } \mu \to \infty,$$

$$\frac{\partial y^{(\mu)}}{\partial t} \to \frac{\partial y}{\partial t}(\underline{v}_0) \text{ in } H \text{ as } \mu \to \infty.$$

This yields, in turn,

$$C'\Lambda_W C y^{(\mu)} \to C'\Lambda_W C y(\underline{v}_0) \text{ in } H \text{ as } \mu \to \infty,$$

$$p^{(\mu)} \to p(\underline{v}_0) \text{ in } V \text{ as } \mu \to \infty,$$

$$\frac{\partial p^{(\mu)}}{\partial t} \to \frac{\partial p(\underline{v}_0)}{\partial t} \text{ in } H \text{ as } \mu \to \infty.$$

Let us note now that the inequality (16) implies that

$$F_1(\underline{w}, \underline{w}) - F_1(\underline{u}^{(\mu+1)}, \underline{u}^{(\mu+1)})$$
$$+ \mathrm{Re}\left(\left(\left(N+N^*\right)(\underline{w}) + \Lambda_{\mathcal{U}}^{-1}(\mathcal{K}+\mathcal{K}')(\underline{w}),\ \underline{w} - \underline{u}^{(\mu+1)}\right)\right)_{\mathcal{U}}$$
$$+ \mathrm{Re}\left(\left(2\Lambda_{\mathcal{U}}^{-1}B'p^{(\mu)},\ \underline{w} - \underline{u}^{(\mu+1)}\right)\right)_{\mathcal{U}} \geq 0$$

$$\text{for all } \underline{w} \in \mathcal{U}_{ad}^{(P)}. \tag{17}$$

because $\mathcal{R}_e \left(\!\!\left(\Lambda_{\mathcal{U}}^{-1}(\mathcal{K}+\mathcal{K}')(\underline{v}), \underline{v} \right)\!\!\right)_{\mathcal{U}} = 2\mathcal{R}_e \sum_{i=1}^{m} (\mathcal{K}_i(v_i), v_i)_{\mathcal{H}}$

and $\mathcal{R}_e \left(\!\!\left((N+N^*)(\underline{v}), \underline{v} \right)\!\!\right)_{\mathcal{U}} = 2\mathcal{R}_e \left(\!\!\left(\underline{v}, N\underline{v} \right)\!\!\right)_{\mathcal{U}}$, and the assumption (5) can be applied with $\underline{v}$ replaced by $\underline{v} - \underline{u}^{(\mu+1)}$. In the inequality (17) transpose the term $F_1(\underline{u}^{(\mu+1)}, \underline{u}^{(\mu+1)})$ to the right side, take the "lim inf" of both sides as $\mu \to \infty$, apply (14), and we eventually end up with

$$
\begin{aligned}
& F_1(\underline{v}, \underline{v}) - F_1(\underline{v}_0, \underline{v}_0) \\
& + \mathcal{R}_e \left(\!\!\left((N+N^*)(\underline{v}) + \Lambda_{\mathcal{U}}^{-1}(\mathcal{K}+\mathcal{K}')(\underline{v}), \underline{v} - \underline{v}_0 \right)\!\!\right)_{\mathcal{U}} \\
& + \mathcal{R}_e \left(\!\!\left(2\Lambda_{\mathcal{U}}^{-1} B' f'(\underline{v}_0), \underline{v} - \underline{v}_0 \right)\!\!\right)_{\mathcal{U}} \geq 0 \\
& \qquad \text{for all } \underline{v} \in \mathcal{U}_{ad}^{(\rho)}. \qquad (18)
\end{aligned}
$$

This yields that

$$
\begin{aligned}
& F_1(\underline{v}, \underline{v}) - F_1(\underline{v}_0, \underline{v}_0) \\
& + \mathcal{R}_e \left(\!\!\left((N+N^*)(\underline{v}_0) + \Lambda_{\mathcal{U}}^{-1}(\mathcal{K}+\mathcal{K}')(\underline{v}_0), \underline{v} - \underline{v}_0 \right)\!\!\right)_{\mathcal{U}} \\
& + \mathcal{R}_e \left(\!\!\left(2\Lambda_{\mathcal{U}}^{-1} B' f'(\underline{v}_0), \underline{v} - \underline{v}_0 \right)\!\!\right)_{\mathcal{U}} \geq 0 \\
& \qquad \text{for all } \underline{v} \in \mathcal{U}_{ad}^{(\rho)}, \qquad \ldots \qquad (19)
\end{aligned}
$$

as can be seen by replacing the $\underline{v}$ in (18) by $(1-\theta)\underline{v} + \theta\underline{v}_0$ for $0 < \theta < 1$, applying the convexity of $F_1(\underline{v}, \underline{v})$, dividing by $1-\theta$, and finally taking the limit as $\theta \to 1^-$.

The Theorem given in the previous section is true when $\mathcal{U}_{ad}$ is replaced by $\mathcal{U}_{ad}^{(\rho)}$. Using this theorem we see that the inequality (19) implies that $\underline{v}_0$ is the optimal control that minimizes J over

$\mathcal{U}_{ad}^{(P)}$. However, $\underline{u}_0$ minimizes J over $\mathcal{U}_{ad}$, and hence, over $\mathcal{U}_{ad}^{(P)}$. By the uniqueness part of the theorem just mentioned, we have $\underline{v}_0 = \underline{u}_0$.

Thus, $\underline{u}^{(\mu)} \to \underline{u}_0$ weakly in $\mathcal{U}$, as $\mu \to \infty$. But $\{\underline{u}^{(\mu)}\}_{\mu=1}^{\infty}$ was an arbitrary weakly convergent subsequence of the sequence (15). Hence, by standard methods of analysis, the sequence (15) itself converges weakly to $\underline{u}_0$ in $\mathcal{U}$.

Let $\{\underline{e}_i\}_{i=1}^{\infty}$ be an orthonormal basis of $\mathcal{U}$. We now have $(\!(\!(\underline{u}_0, \underline{e}_i)\!)\!)_{\mathcal{U}}$ = the usual limit (in $\mathbb{C}$) of $(\!(\!(\underline{u}^{(m)}, \underline{e}_i)\!)\!)_{\mathcal{U}}$ as $m \to \infty$. So, calculation of $\underline{u}_0 \equiv \sum_{i=1}^{\infty} (\!(\!(\underline{u}_0, \underline{e}_i)\!)\!)_{\mathcal{U}} \underline{e}_i$ to a desired degree of accuracy may be accomplished by calculating $(\!(\!(\underline{u}_0, \underline{e}_i)\!)\!)_{\mathcal{U}}$, for each i , to an appropriate degree of accuracy by representing $(\!(\!(\underline{u}_0, \underline{e}_i)\!)\!)_{\mathcal{U}}$ by $(\!(\!(\underline{u}^{(m)}, \underline{e}_i)\!)\!)_{\mathcal{U}}$ for an appropriate m. Since in numerical calculation, we are able to deal with only a finite number of i's by ignoring the rest, we may face here a different kind of "spillover effect" that depends on the i's that were ignored.

V. RECOMMENDATIONS:

Follow-up research may be performed

(i) to consider the spillover kind of effect mentioned in the last sentence of the preceding section,

(ii) to investigate under what conditions approximate solutions may be obtained in all the algorithmic steps given near the beginning of the previous section, and still come out with a sequence (15) that converges weakly to the exact solution $\underline{u}_0$,

(iii) to develop the numerical analysis needed to obtain the approximate solutions just mentioned,

(iv) to determine whether our theory may be successfully applied to systems governed by parabolic partial differential equations,

(v) to examine whether a hyperbolic equation more general than the one given in (4) may be dealt with,

(vi) on the effectiveness of our theory if the performance index J depends on $\frac{\partial y}{\partial t}$ or the space-time boundary values of y or the derivatives of these boundary values of y, or

(vii) on the problems of decoupling the systems (7), (10) and (11), and solving the resulting Riccati-type integrodifferential equations.

REFERENCES

1. J. L. Lions, Optimal Control of Systems Governed by Partial Differential Equations, translated in English by Dr. S. K. Mitter, Springer-Verlag, New York, 1971.

2. Tapas Mazumdar, "Minimization, over Possibly Unbounded Regions, of Certain Possibly Noncoercive Cost-Functionals in Hyperbolic Feedback Control Problems", in preparation.

3. Kôsaku Yosida, Functional Analysis, Fifth Edition, Springer-Verlag, New York, 1978.

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

A STUDY OF THE CH_2NO_2 RADICAL USING A MULTICONFIGUATIONAL SELF CONSISTANT FIELD APPROACH

Prepared by: Michael L. McKee

Academic Rank: Assistant Professor

Department and University: Department of Chemistry, Auburn University

Reasearch Location: Frank J. Seiler Labroatory, Air Force Academy, Colorado, Springs, Colorado

USAF Research: Dr. John S. Wilkes

Date: September 3, 1983

Contract No: F49620-82-C-0035

A STUDY OF THE CH_2NO_2 RADICAL USING A MULTICONFIGUATIONAL APPROACH

by

Micheal L. McKee

ABSTRACT

The CH_2NO_2 radical is used as a model for first likely reactive species in the decomposition of high energy materials such as TNT (1,3,5-trinitrotoluene) or HMX (octahydro-1,3,5,7-tetranitro-1,3,5,7-Tetrazocine). Radical species can be observed in an ESR cavity during the inductive phase of decomposition however their identity is not clearly established. Semiempirical, single configurational, and eventually multiconfigurational ab initio calculations were carried out to determined the ground electronic state and the electronic distribution. Ab initio calculations based on a single configuration plus correlation could not distinguish the ground state. The MCSCF results indicate that the ground state is a planar $^2A''$ state which results from the interaction of the planar $^2B_1$ and $^2A_2$ states that cross at a common C_{2v} geometry and which leads to a favorable asymmetric distortion from C_{2v} symmetry to a lower C_s symmetry. A similar distortion was observed by Davidson and coworkers for NO_2. The $^2A''$ state is 14.6 kcal/mol more stable than the $^2B_1$ state and 19.9 kcal/mol more stable than the $^2A_2$ state. The staggered $^2B_2$ state which is the C_{2v} ground state is 6.8 kcal/mol higher than $^2A''$ state. The planar $^2A''$ ground state has considerable spin density on the carbon and some on one oxygen in agreement with ESR results.

I. INTRODUCTION:

The nitromethyl radical has been postulated to be invloved in a number of different phenomena. In a study[1] of nitromethylation of aromatic compounds it was determined that the radical had appreciable electrophilic character. The only products formed were through reactions at the carbon suggesting that it was the more reactive center. Interestingly in the reaction with manganese (III) acetate

and nitromethane the authors[1] suggested that it might be the aci form (CH_3ONO) that underwent oxidation followed by tautomerization.

The radical could be produced in the following reaction:

$$CH_3 + CH_3NO_2 \rightarrow CH_4 + CH_2NO_2$$

which has been suggested[2] to explain the appearance of methane as a major product in the decompostion of nitromethane. Once formed the nitromethyl radical is believed to effect the thermal decomposition of nitromethane in condensed phase systems[3]. Such systems include propellant ignition and explosive initiation.

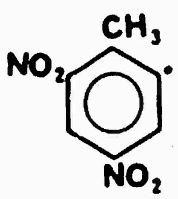
Nitromethyl radical may even play a role in the chemistry of the upper troposphere through the degradation of nitromethane formed from the possible reaction of CH_3 and NO_2[4].

Very recently the vibrational spectrum of CH_2NO_2 was reported by Jacox[4] who codeposited F atoms and nitromethane in an Argon matrix at 14 K and analyzed the changes in the IR spectra. Different isotopes of nitromethane were used in the study to help assign the bands.

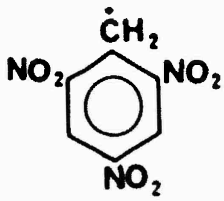
The heat of formation of gaseous nitromethyl radical has been estimated[5] to be 37.5 kcal/mol.

The ESR spectrum of the nitromethyl radical has been interpreted[6] as supporting a nonplanar geometry. However, the only reported theoretical calculation on CH_2NO_2 was an INDO calculation[7] which predicted a planar structure.

The motivation for this study was not only to gain an understanding of this particular system as a reactive intermediate, but also to view the system as a model for the radicals produced from high energy species such as TNT (1,3,5-trinitrotoulene) where radical intermediates[8] as 1 or 2 might occur.

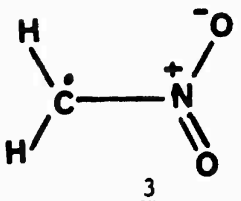


1

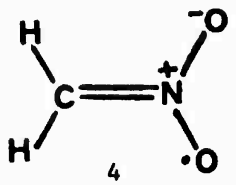


2

The two valence structures which can be written for planar CH_2NO_2 are:



3



4

A theoretical study[9] on radicals of this nature involving structures with comparable energies but inequivalent assignments of π bonds and the lone electron suggest that accurate results require substantial effort to correctly predict their relative energy. Indeed this could be predicted by the sophisticated basis sets needed[10] to predict the states of NO_2.

II. OBJECTIVES OF THE RESEARCH EFFORT:

The object of this work was to study the types of radicals that could be generated by thermal or photolytic decomposition of high energy materials. Specifically the radicals produced from polynitrated systems as TNT or HMX. To do this one must determine which is the most economical method that will produce results of sufficient quality or accuracy to answer the particular chemical problem. If empirical methods are not suffiently accurate and more sophisticated methods too costly, one may ask if there are trends which could still allow the empirical methods to be used predictively with some confidence. The objectives of this study were as follows:

(1) To study the electronic states of the CH_2NO_2 radical with the semiempirical method MNDO to determine the ground state and electronic configuration.

(2) To study the problem using ab initio methods with geometry optimization and correlation.

(3) To determine the extent that other configurations play in determining the properties of the electronic states by using a MCSCF method.

III. RESULTS AND DISCUSSION:

The first method used to study the electronic states of CH_2NO_2 was the MNDO method, a semiempirical method based on the NDDO approximation[11] and very successfully paramaterized by Dewar and Thiel[12]. All calculations were carried out using the unrestricted SCF formalism.

The heats of formation and geometries for C_{2v} species are given in Table 1. The most stable state is the planar $^2B_1$ followed very closely by the staggered $^2B_2$. These two states simply correspond to rotation of the singly occupied p orbital on carbon. The other states were obtained by starting with a bond order generated from an orbital guess in which selected orbitals were occupied. During the iteration to self consistency the switched orbital was montitored to insure that it remained occupied. Unfortunately in many instances the wavefunction lost the initial C_{2v} symmetry and it proved to be impossible to obtain those states.

The planar $^2A_2$ solution was found to be 18.3 kcal/mol higher in energy than the planar $^2B_1$ state. Roughly, these two states correspond to valence structures 3 and 4. The unrestricted solution obtained for $^2A_1$ was characterized by three singly occupied orbitals; a b_2 and a_2 orbital of alpha spin and a b_1 orbital of beta spin.

The assumption of C_{2v} symmetry for CH_2NO_2 radical seems intuitively obvious but due to the interaction of closely related electronic state this assumption may be wrong. It has previously been noted[10a,b] that under certain conditions electronic states in C_{2v} symmetry may mix by an asymmetric distortion and lead to a lower energy in C_s symmetry. The conditions which must be met are:

1) the two states in C_{2v} symmetry must transform as the same irreducible representation in the C_s point group;

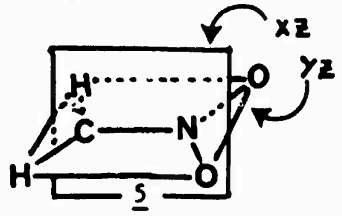
2) there must be a surface crossing at some nearby C_{2v} geometry;

3) the leading configurations must differ by only one spin orbital.

The $^2A_1$ and $^2B_2$ states of NO_2 meet the above conditions and it is found[10a,b] that an asymmetric distortion does lead to a

lower energy.

If we look at the C_{2v} point group table we see that the C_{2v} symmetry can be reduced to C_s by the loss of the sigma (yz) plane or the sigma (xz) plane (5)



5

If the xz symmetry plane is lost then the A_2 and B_1 states transform as A" while A_1 and B_2 states transform as A'. If the yz symmetry plane is lost then A_2 and B_2 transform as A" while the A_1 and B_1 transform as A'.

The MNDO wavefunctions were tested for stability under C_s distortions and the results are presented in Table 2 and 3.

Table 1 MNDO Heats of Formation (kcal/mol) and Geometries[a] of CH_2NO_2 radicals in C_{2v} symmetry

State	ΔH_f	CN	NO	CH	ONH	HCN
$^2B_1$[b]	35.2	1.476	1.214	1.084	119.3	118.3
$^2A_2$[b]	53.5	1.414	1.266	1.084	122.3	119.0
$^2B_2$[c]	35.6	1.482	1.219	1.082	119.4	117.6
$^2A_1$[c,d]	63.1	1.454	1.264	1.084	121.5	118.2

a) Distances are given in Angstroms and angles are given in degrees.

b) Planar geometry.

c) Staggered geometry.

d) The $^2A_1$ state has b_2, a_2, and b_1 orbitals singly occupied.

C_{2v} Symmetry to C_s Symmetry

Element of symmetry kept	Symmetry in C_{2v}	Symmetry in C_s	ΔH_f[a]	ΔH_f[b]
sigma(yz)	$^2A_2$, $^2B_1$	$^2A''$	33.8[c] (1.4)	71.0
	$^2A_1$, $^2B_2$	$^2A'$	---	d

sigma(xz)	$^2A_2$, $^2B_2$	$^2A''$	33.8[c] (1.8)	d	
	$^2A_1$, $^2B_1$	$^2A'$	d	61.6 (1.5)	

a) Planar geometry. Number in parenthesis refers to lowering in kcal/mol with respect to the lower of the two C_{2v} states.

b) Staggered geometry. Number in parenthesis refers to lowering in kcal/mol with respect to the lower of the two C_{2v} states.

c) The wavefunction exhibited only C_1 symmetry.

d) A distortion from C_{2v} symmetry was not energetically favorable.

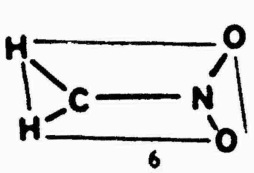
Table 3 MNDO Geometries[a]

of CH_2NO_2 radicals in C_s symmetry

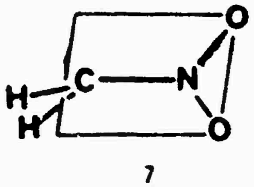
Geometry	ΔH_f	CN	NO	CH	ONC	HCN	Dihedral
6	33.8	1.466	1.221	1.083	120.5	118.3	
			1.227	1.083	119.1	118.5	
7	61.6	1.465	1.268	1.084	118.0	120.2	72.8
				1.082		116.4	
8	71.0	1.452	1.226	1.084	125.3	118.0	88.9
			1.306		116.6		

a) When two values for a variable are given, the first value corresponds to the bond distance or bond angle on the side of the CN bond denoted 'a' in figures 6, 7, and 8 while the second value corresponds to the side denoted 'b'.

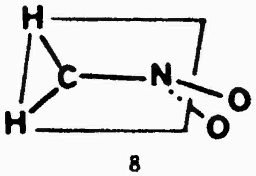
The global minimum was a distorted planar structure 6 but since the wavefunction did not possess C_s symmetry, the state could not be definitively assigned. Although the staggered $^2A_2$ and $^2B_1$ solutions could not be found, the MNDO value of 71.0 kcal/mol for the heat of formation of the distorted structure 7 is a lower bound for the lower of the C_{2v} states. When the sigma (yz) plane of symmetry is lost the $^2A_1$ and $^2B_1$ states interact and lead to a $^2A'$ state 8 1.5 kcal/mol lower in energy than the $^2A_1$ state.



6



7



8

Since the relative energies of NO_2 and related compounds are particularly sensitive to the basis set[10], ab initio calculations were performed on the CH_2NO_2 system[13]. Geometries were completely optimized within either the C_{2v} or C_s point group using the 3-21G basis[14] and single point calculations were ther made including correlation at the UMP3 level[15]. The energies relative to the staggered $^2B_2$ are presented in Table 4, absolute energies in Table 5, and geometries in Table 6. From Table 4 one can determine that the $^2A_2$ - $^2B_1$ (planar) separations calculated at the UHF/3-21G, UHF/6-31G, UMP2/6-31G, and UMP3/6-31G are respectively; -17.7, -4.6, 49.4, and 31.0 kcal/mol. This compares to the MNDO value of 18.3 kcal/mol from Table 1.

Table 4 Relative Energies (kcal/mol) for Different States of CH_2NO_2 Based on Single Configuration Calculation

		UMP3/6-31G	UMP2/6-31G	UHF/6-31G	UHF/3-21G
Planar	$^2A_2$	24	45.1	-12.7	-26.5
	$^2B_1$	-6.7	-4.3	-8.1	-8.8
	$^2B_2$	16.5	29.8	9.2	0.3
	$^2A_1$				122.9
	$^2A''$	5.7	23.3	-22.8	-32.6
	$^2A'$	37.5	57.9	9.0	-16.0
Stagg.	$^2A_2$	75.4	87.7	58.2	45.4
	$^2A_2$	110.4	133.0	76.0	53.1
	$^2B_1$	58.1	71.6	38.6	30.9
	$^2B_2$	0	0	0	0
	$^2A_1$				80.7
	$^2A''$	52.5	73.7	15.5	0.9
	$^2A'$	18.0	36.9	-14.4	-24.6

Table 5 Total Energies (Hartrees) for Different States of CH_2NO_2 Based on Single Configuration Calculation

State	config.[a]	UMP3/6-31G	UMP2/6-31G	UHF/6-31G	UHF/3-21G
Planar					
$^2B_1$	16/3/2/10	243.33871	243.35832	242.89561	241.62565
$^2B_2$	16/4/2/9	243.30166	243.30401	242.86808	241.61118
$^2A_2$	16/4/1/10	243.28929	243.27959	242.90298	241.65386
$^2A_1$	16/5/1/9				241.41585
$^2A''$[b]	26/5	243.31891	243.31428	242.91899	241.66365 (6.1)[c]
$^2A'$[d]	25/6	243.26825	243.25916	242.86835	241.63718 (16.3)[c]
Staggered					
$^2B_1$	16/5/2/8	243.23547	243.23744	242.82115	241.56241
$^2B_2$	16/4/2/9	243.32802	243.35148	242.88272	241.61165
$^2A_2$	15/5/2/9	243.20787	243.21167	242.78989	241.53924
$^2A_2$	16/4/1/10	243.15204	243.13955	242.76159	241.52701
$^2A_1$	16/5/1/9				241.48298
$^2A''$[e]	24/7	243.24441	243.23404	242.85805	241.61020 (30.0)[c]
$^2A'$[f]	25/6	243.29926	243.29263	242.90567	241.65085 (24.6)[c]

a) The notation refers to the number of electrons in the a_1, b_1, a_2, and b_2 orbitals respectively. The notation for $^2A''$ and $^2A'$ refers to the number of electrons in a' and a" orbitals respectively

b) Involves the interaction of planar states $^2B_1$ and $^2A_2$ and the loss of the sigma (xz) plane of symmetry.

c) The number in parenthesis refers to the lowering in kcal/mol with respect to the lower of the two interacting C_{2v} states for the 3-21G energies.

d) Involves the interaction of planar states $^2B_2$ and $^2A_1$ and the loss of the sigma (xz) plane of symmetry.

e) Involves the interaction of staggered states $^2B_1$ and $^2A_2$ and the loss of the sigma (xz) plane of symmetry.

f) Involves the interaction of staggered states $^2B_2$ and $^2A_1$ and the loss of the sigma (xz) plane of symmetry.

Table 6 Ab initio UHF/3-21G Geometries of CH_2NO_2 radicals in C_{2v} Symmetry and the Spin Squared Values (without projection)

	State	S^2	CN	NO	CH	ONC	HCN	Dihedral
Planar	$^2A_2$	1.59	1.377	1.360	1.065	122.5	117.2	
	$^2B_1$	0.79	1.408	1.257	1.065	117.1	116.6	
	$^2B_2$	1.08	1.281	1.369	1.066	129.2	118.3	
	$^2A_1$	1.80	1.384	1.428	1.064	138.6	116.1	
	$^2A''$[a]	1.47	1.391	1.288	1.065	124.5	117.1	
				1.362	1.064	115.1	117.1	
	$^2A'$[a]	1.37	1.341	1.458	1.065	115.0	118.9	
				1.284	1.066	129.8	116.5	
Stagg.	$^2A_2$[b]	1.54	1.336	1.454	1.068	128.4	117.2	
	$^2A_2$[b]	1.84	1.425	1.370	1.068	115.2	116.7	
	$^2B_1$	1.84	1.413	1.363	1.070	128.2	117.8	
	$^2B_2$	0.76	1.456	1.241	1.067	117.0	116.5	
	$^2A_1$	1.84	1.402	1.419	1.071	124.7	118.0	
	$^2A''$[c]	1.76	1.419	1.472	1.070	111.8	118.7	90.6
				1.299		128.9		
	$^2A'$[c]	1.63	1.431	1.319	1.070	120.8	118.0	96.1
				1.328		118.7		

a) See Table 3 and Figure 6.

b) The geometries for the two perpendicular states of $^2A_2$ symmetries differ in the orbital occupancy. The first entry was obtained with 3 open shells, and the second with one open shell(see text).

c) See Table 3 and Figure 8.

The rotational barrier of planar $^2B_1$ to staggered $^2B_2$ is fairly insensitive to the level of calculation (7-9 kcal/mol). The rotational barrier of planar $^2A_2$ to staggered $^2A_2$ is also insenstive to correlation if the same configuration is used. The barrier at the UHF/6-31G level is 79.6 kcal/mol compared to 86.1 kcal/mol at the UMP3/6-31G level. The high barrier reflects the multiple bond character of the C-N bond. At the single determinant level, however, the $[8a_1^2 2b_1^2 5b_2^2]1a_2^1$ configuration leads to a higher energy than the $[7a_1^2 2b_1^2 1a_2^2 4b_2^2]8a_1^1 3b_1^1 5b_2^1$ configuration[16]. The energy difference between the two solutions varied from 10-30 kcal/mol depending on level of calculation (see Table 4).

Of course the true $^2A_2$ staggered state will be some linear combination of these two states plus other configurations of the correct symmetry. The proper treatment of such states must be based on either an MCSCF approach or a full configuration interaction approach.

As seen in Table 6 the CN bond length of the planar $^2B_1$ states elongates 0.05 Å upon twisting to the staggered $^2B_2$ state. This suggests that the 6.7 kcal/mol required for rotation (UMP3/6-31G) may be due either to elimination of conjugation in the planar structure or increased repulsion of the electron in the singly occupied orbital with the oxygen atoms in the staggered structure.

It must be remembered that this is not an electronic transition between states but rather a conformational change. The states $^2B_1$ and $^2B_2$ are antisymmetric with respect to C_2 rotation while A_2 state is symmetric.
In the former case a rotation will split the $^2B_1$ and $^2B_2$ planar states into $1^2A''$ and $2^2A''$ which may interact. The rotation continues until at 180 degrees the states become noninteracting with new designations of $^2B_2$ and $^2B_1$. The interaction could lead to a weak minimum for the lower state or a slight maximum in the higher state. However in this paper rotational barrier refers to the energy difference in the two C_{2v} electronic states. For $^2A_2$ rotation the same applies to the interaction of $^2A_2$ and $^2A_1$ as $1^2A'$ and $2^2A'$. In this case the energy separation between the two states makes mixing unlikely.

The shortest CN bond occurs in the planar $^2B_2$ state, 0.10 Å shorter than the CN bond in the planar $^2A_2$ state 4 whose valence structure can be drawn with a double bond. The shorter CN bond is due to removing an electron (with respect to the $^2A_2$ configuration) from the CN antibonding $5b_2$ orbital and putting it in the CN bonding $1a_2$ orbital. The loss of a electron from the $5b_2$ orbital also causes the CNO bond angle to increase about 6 degrees when compared to the planar $^2A_2$ geometry.

The spin squared values for the 3-21G wavefunctions are given in Table 6. The only states which approach the ideal 0.75 value (S(S+1)) are the planar $^2B_1$ state and the staggered $^2B_2$ state. In these states the extra alpha spin electron is almost completely

localized in a p orbital perpendicular to the CH_2 plane. The other states indicate a large amount of spin polarization where alpha and beta spins localize in different regions of the molecule. This gives some indication that a restricted program must use additional configurations to achieve the same result.

It was found that the 3-21G optimized geometries are unstable with respect to distortions from C_{2v}. The results are also presented in Table 4, 5, and 6. The distortions and energy changes are larger than found for MNDO. When correlation is used the distortion becomes unfavorable.

A perturbational correlation treatment such as the UMP3 method becomes ineffective when the single Hartree-Fock configuration is not the dominant configuration in the wavefunction. Since it is known that there are several configurations of the same symmetry and of comparable energy, a MCSCF treatment was used to continue the study. The program ALIS[17] developed by Klaus Ruedenberg and coworkers and sponsored by the U.S. Department of Energy was used to determined MCSCF energies and wavefunctions.

In order to 'feel out' the potential surface and determine which configurations were important for the different states MCSCF calculations were made using a STO-3G basis[18]. In all calculations reported below the 3-21G geometries were used[19]. It was later confirmed that the largest contribution to the wavefunction was the configuration used to generate the geometry. It can be argued that using geometries optimized with a single configuration might 'favor' that configuration in a MCSCF scheme, however the extra expense of optimizing at this level was not feasible. MCSCF calculations were performed on the two staggered $^2A_2$ states obtained at the 3-21G level (vide supra) using different configurations. For the geometry obtained with 3 open shells, the leading contribution (one space orbital with two spin functions) also contained the 3 open shells (a_1, b_1, b_2) and accounted for 63% of the wavefunction. For the geometry obtained with one open shell (a_2), the same configuration accounted for 49% of the wavefunction.

The steps taken to obtain the MCSCF energy[20] are given below. An initial guess of orbitals is made by doing a restricted open shell calculation. For the STO-3G basis the energy normally converged but

for the 6-31G basis the energy diverged. In the latter case it was necessary to stop the SCF at the point of divergence and orbital improvement was made at the expense of extra MCSCF microiterations. Next a limited MCSCF was performed with a few intuitive configurations. Using this wavefunction a CI in a larger space (approximately 500 Spin Adapted Antisymmeterized Products) was made to determine which configurations made substantial contributions to the wavefunction. These configurations were then used in another MCSCF calculation with the CI orbitals as the initial guess. The process was repeated until the MCSCF and CI had the same dominant configurations and until no configuration in the CI which was not included in the MCSCF contributed 1% or more to the wavefunction. The final MCSCF was carried out with between 5 and 11 spin adapted configurations. This provides an approximation to the full 500 configuration MCSCF as the most important configurations have been iteratively selected.

The MCSCF/STO-3G results are presented in Table 7. The first column represents the restricted open shell energies while the second column gives the MCSCF energy with the number of configurations in parenthesis and the per cent contribution of the most important configuration. The third column gives the CI energy with the number of spin adapted configurations in parenthesis (8 orbitals and 7 electrons in C_{2v} symmetry). It has been pointed out that in some cases carrying out a calculation in C_{2v} symmetry may lead to a higher energy than if only C_s symmetry were used, but in the calculations[10a,b] for NO_2 it was found that the C_{2v} solution was the same as the C_s solution. For this reason plus economics, the C_{2v} point group was used when possible.

Relative energies are presented in Table 9. In contrast to MNDO and ab initio results, the planar $^2A_2$ state is predicted to be the most stable of the C_{2v} states followed by the planar $^2B_1$ state 8.0 kcal/mol higher. The rotational barrier for the $^2B_1$ to $^2B_2$ conformational change is 18.6 kcal/mol, considerablely higher than the MNDO (0.4 kcal/mol) or the UMP3/6-31G (6.7 kcal/mol) barrier. The barrier for the rotation of the $^2A_2$ state (45.1 kcal/mol) is larger than the UMP3/6-31G value (51.1 kcal/mol).

Table 7 Total Energies for Different States of CH_2NO_2 Based on Multiconfigurational STO-3G Wavefunction

	State	RSCF	MCSCF	#conf.	%[a]	CI	#conf.	%[b]
Planar	$^2A_2$	239.7296	239.79270	(4)		239.86545	(582)	11%
			239.88456	(6)	65%	239.90642	(582)	< 1%
	$^2B_1$	239.7699	239.86848	(6)	81%	239.89358	(588)	< 1%
	$^2B_2$	239.7512	239.82680	(5)	87%	239.83138	(594)	6%
			239.84988	(9)		239.85617	(594)	< 1%
Stagg.	$^2A_2$	239.5864	239.79005	(8)	52%	239.79205	(584)	< 1%
	$^2B_1$	No SCF	239.73403	(7)		239.79018	(588)	6%
			239.81702	(11)	30%	239.82255	(588)	< 1%
	$^2B_2$	239.7633	239.85302	(5)	83%	239.86399	(594)	< 1%

a) Largest contribution from a single configuration in the MCSCF wavefunction.

b) Largest contribution from a single configuration in the CI which was not optimized in the MCSCF wavefunction.

The STO-3G minimal basis set is known to be inferior to the near double zeta quality 6-31G basis set[21]. For this reason the calculations were repeated with the larger basis set. It was felt that most of the important configurations would already be identified from the STO-3G basis allowing rapid convergence to the most important configurations in the expanded basis.

The absolute energies for the MCSCF/6-31G calculations are presented in Table 8 and the relative energies are presented in Table 9. The restricted open shell energies are not given as the SCF diverged. The percent contribution of the single largest contributing configuration clearly indicates that a single determinant cannot be adequate. The planar $^2B_1$ and staggered $^2B_2$ states have one configuration comprising approximately

Table 8 Total Energies for Different States of CH_2NO_2 Based on Multiconfigurational 6-31G Wavefunction

	State	MCSCF	#conf.	%[a]	CI	#conf.	%[b]
Planar	$^2A_2$	242.86061	(6)		242.89856	(582)	less 1%
		242.9238	(8)	73%	242.94688	(582)	less 1%

$^2B_1$	242.8978	(7)		242.90731 (588)	1%
	242.94927	(7)	89%	242.95541 (588)	1%
$^2B_2$	242.86	(5)		242.90560 (594)	1%
	242.92690	(8)	91%	242.94086 (594)	1%
$^2B_1$[c]	242.92308	(7)	81%	242.92988 (588)	1%
$^2A_2$[c]	242.90617	(8)	70%	242.92207 (582)	1%
$^2A''$	242.92	(7)		242.92671 (68)	8%
	242.93088	(8)		242.95703 (1176)	1%
	242.96236	(8)	90%	242.97867 (1176)	< 1%
Stagg. $^2A_2$	242.6	(6)		242.62593 (584)	34%
	242.80496	(7)		242.81163 (584)	3%
	242.80498	(9)	24%	242.81155 (584)	1%
$^2A_2$	242.8238	(9)		242.83152 (584)	1%
	242.82383	(9)	52%	242.83165 (584)	1%
$^2B_1$	242.84862	(11)		242.86150 (588)	1%
	242.84806	(11)	49%	242.86847 (588)	1%
$^2B_2$	242.93351	(5)	90%	242.96780 (594)	1%

a) Largest contribution from a single configuration in the MCSCF wavefunction.

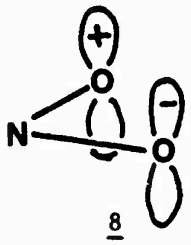
b) Largest contribution from a single configuration in the CI which was not optimized in the MCSCF wavefunction.

c) The $^2B_1$ calculation was at the $^2A_2$ geometry and the $^2A_2$ calcuation was at the $^2B_1$ geometry(see text).

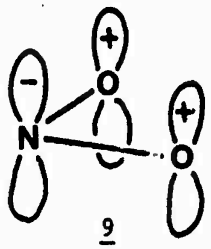
Table 9 Relative Energies (kcal/mol) for Different States of CH_2NO_2 Based on MCSCF/STO-3G and MCSCF/6-31G

State	MCSCF/STO-3G	+CI	MCSCF/6-31G	+CI
Planar $^2A_2$	-19.8	-26.6	6.1	13.1
$^2B_1$	-9.7	-18.6	-9.9	7.8
$^2B_2$	2.0	4.9	4.1	16.9
$^2A''$			-18.1	-6.8
Stagg. $^2A_2$	39.5	45.1	80.6	98.0
$^2A_2$			68.8	85.4
$^2B_1$	22.6	46.3	53.6	62.3
$^2B_2$	0	0	0	0

90% of the wavefunction. These corresponds to one electron in a p orbital perpendicular to the CH_2 plane and are the same states that produced 'normal' S^2 values. Even for these two states the orbitals occupations showed a significant population of a 'virtual' orbital. In Figure 8 and 9 the nonbonding oxygen b_2 orbital contained 1.8 electrons while the antibonding (bonding across oxygen) contained 0.2 electrons.



8



9

This suggests that there may be some stablization through pi overlap of the oxygen p orbitals.

The MCSCF/6-31G plus CI lead to a predicted staggered $^2B_2$ C_{2v} ground state for the nitromethyl radical. This is different from each of the previous methods. The planar $^2B_1$ state is 7.8 kcal/mol less stable and the planar $^2A_2$ state is 13.1 kcal/mol above the staggered $^2B_2$ state. The difference between the $^2A_2$ states is 72.3 kcal/mol which is almost identical to the STO-3G difference and greater than that predicted by UMP3/6-31G.

The relative energies have not converged with respect to increasing the sophistication of the basis set. Therefore it is premature to give the definitive electronic state ordering. In particular polarization functions as well as a more complete CI may make significant contributions to the relative energies.

Another possible influence on relative energies could be distortions from C_{2v}. This was found to be favorable in the MNDO approximation and at the UHF/3-21G and UHF/6-31G levels (though not at the UMP2/6-31G or UMP3/6-31G levels). In order to see if the planar $^2A_2$ and $^2B_1$ states might interact upon loss of the sigma (xz) plane (vide supra), MCSCF calculations were carried out on the planar $^2A_2$ state at the $^2B_1$ geometry and on the $^2B_1$ state at the $^2A_2$ geometry. The results shown in Figure 10 (see Table 8 for absolute energies) indicate that a crossing must occur somewhere between the two geometries along a C_{2v} reaction coordinate.

A MCSCF calculation was carried out on the distorted C_{2v}

geometry optimized with the UHF/3-21G basis (see Table 6). The results given in Table 8 and 9 indicate that the distortion is favorable and yields a $^2A''$ state 14.6 kcal/mol more stable than the planar $^2B_1$ state. This is 6.8 kcal/mol more stable than the staggered $^2B_2$ and thus becomes the ground state. It would be important to precisely determine the C_s minimum before structural predictions are made as the UHF/3-21G optimized geometry may not be accurate, but it is felt that such optimization would not qualitatively change this result.

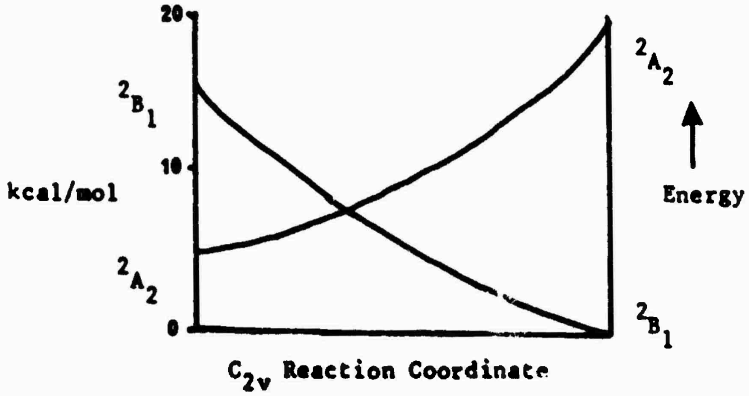


Figure 10. Illustration of C_{2v} Energy Crossing with the MCSCF calculations at the four labeled points.

IV. RECOMMENDATIONS:

On the basis of the above study the following recommendations can be made:

(1) That single determinant methods for the nitromethyl radical may neglect important contributions to the wavefunction and must be used with some caution in making structural and energetic predictions.

(2) That semi-empirical and ab initio methods provide similar qualitative predictions, however it is more difficult to obtain different electronic states with the MNDO method.

(3) That distortions from C_{2v} symmetry must be considered when states differing by only one spin orbital cross at some C_{2v} geometry.

(4) That unrestricted ab initio results without correlation probably overestimate such distortions.

(5) That a similar study be undertaken at the ab initio level for the radicals derived from 1-Nitropropene which may be a more realistic model for the radicals produced from TNT.

(6) That in order to study the decomposition mechanism for nitromethane the reaction pathway from the nitromethyl radical to CH_2 plus NO_2 and to CH_2O plus NO be studied.

ACKNOWLEDGEMENTS

I would like to thank the Air Force Office of Scientific Research for the opportunity to carry out research at their facilities. Financial support for the Summer Faculty Research Program Fellowship was administered through the Southeastern Center for Electrical Engineering Education. The hospitality shown by the members of the Frank J. Seiler Research Laboratory was greatly appreciated. In particular I would like to acknowledge helpful conversations with Capt. Larry P. Davis and Dr. Almon G. Turner.

The MCSCF program was provided by the Ames Laboratory, Iowa State University, Ames, Iowa. Stephen T. Elbert also provided technical assistance.

REFERENCES

(1) Kurz, M. E.; Chen, T.-Y. R. J. Org. Chem. 1978, 43, 239.

(2) Cottrell, T. L.; Graham, T. E.; Reid, T. J. Trans. Faraday Soc. 1951, 47, 584.

(3)(a) Nazin, G. B.; Manelis, G. B.; Dubovitskii, F. I. Russ. Chem. Rev. 1968, 37, 603. (b) Crawforth, C. G.; Waddington, D. J. Trans. Faraday Soc. 1969, 65, 1334. (c) Dubovitskii, F. I.; Korsunskii, B. L. Russ. Chem. Rev. 1981, 50, 958.

(4) Jacox, M. E. J. Phys. Chem. 1983, 87, 3126.

(5) Knobel, Yu. K.; Miroshnichenko, E. A.; Lebedev, Yu. A. Izv. Akad. Nauk. SSSR, Ser. Khim. 1971, 485.

(6)(a) Chachaty, C. J. Chim. Phys. 1965, 62 728. (b) Chachaty, C.; Rosilio, C. J. Chim. Phys. 1967, 64, 777.

(7) Bolsman, T. A. B. M.; Verhoeven, J. W.; De Boer, Th. J. Tetrahedron, 1975, 81, 898.

(8)(a) Guidry, R. M.; Davis, L. P. Themochimica Acta 1979, 32, 1. (b) Davis, L. P.; Wilkes, J. S.; Pugh, H. L.; Dorey, R. C. J. Phys. Chem. 1981, 85, 3505. (c) Davis, L. D.; Turner, A. G.; Carper, W. R.; Wilkes, J. S.; Dorey, R. C.; Pugh, H. L.; Siegenthaler, K. E. "Thermochemical Decomposition of TNT: Radical Identification and Theoretical Studies," FJSL TR 81-0002, April 1981.

(9) Baird, N. C.; Taylor, K. F. Can. J. Chem. 1980, 58, 733.

(10)(a) Jackels, C. F.; Davidson E. R. J. Chem. Phys. 1976, 64, 2908. (b) Jackels, C. F.; Davidson E. R. J. Chem. Phys. 1976, 65, 2941. (c) Handy, N. C.; Goddard, J. D.; Schaefer, H. F. III J. Chem. Phys. 1979, 71, 1979. (d) Benioff, P. A. J. Chem. Phys. 1978, 68, 3405. (e) Krauss, M.; Hopper, D. G.; Fortune, P. J.; Wahl, A. C.; Tiernan, T. O. "Potential Energy Surfaces for Air Triatomics, Vol I," June 1975, Final Report ARL TR 75-0202, Aerospace Research Laboratories, Wright-Patterson Air Force Base, Ohio.

(11) Pople, J. A.; Segal, G. A. J. Chem. Phys. 1965, 43, S129.

(12) Dewar, M. J. S.; Thiel, W. J. Am. Chem. Soc. 1977, 99 4899.

(13) The Gaussian 80 program was used; Binkley, J. S.; Whiteside, R. A.; Krishnan, R.; Seeger, R.; DeFrees, D. J.; Schlegel, H. B.; Topiol, S.; Kahn, L. R.; Pople, J. A. QCPE 1981, 13, 406.

(14) Binkley, J. S.; Pople, J. A.; Hehre, W. J. J. Am. Chem. Soc.

1980, 102, 939.

(15) Pople, J. A.; Binkley, J. S.; Seeger, R. Int. J. Quantum Chem. Symp. 1976, 10, 1.

(16) The orbitals in brackets indicate the highest doubly occupied orbital of each type.

(17) Elbert, S. T.; Cheung, L. M.; Ruedenberg, K. NRCC Software Catalog, Vol 1., Program No. QC01 (ALIS) (Ames Laboratory-Iowa State University). The MCSCF method used is described in: Ruedenberg, K.; Cheung, L. M.; Elbert, S. T. Int. J Quantum Chem. 1979, 16 1069.

(18) Hehre, W. J.; Stewart, R. F.; Pople, J. A. J. Chem. Phys. 1969, 51, 2657.

(19) The following paper finds that large spin contaminations is associated by poor geometry predictions. Farnell, L.; Pople, J. A.; Radom, L. J. Phys. Chem. 1983, 87, 79.

(20) Feller, D. F.; Schmidt, M. W.; Ruedenberg, K. J. Am. Chem. Soc. 1982, 104, 960.

(21) Ditchfield, R.; Hehre, W. J.; Pople, J. A. J. Chem Phys. 1971, 54, 724.

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

INVESTIGATION OF VIBRATION PROBLEMS WITH HETERODYNE HOLOGRAPHIC INTERFEROMETER

Prepared by:	Dr. Robert A. McLauchlan, P. E.
Academic Rank:	Assistant Professor
Department and University:	Department of Mechanical Engineering Texas Tech University
Research Location:	Air Force Weapons Laboratory, Advanced Radiation Technology Division, Advanced Beam Control Branch
USAF Research:	Dr. John R. Kenemuth, Captain Joseph T. Evans
Date:	1 September 1983
Contract No:	F49620-82-C-0035

INVESTIGATION OF VIBRATION PROBLEMS WITH HETERODYNE HOLOGRAPHIC INTERFEROMETER

by

Dr. Robert A. McLauchlan, P. E.

ABSTRACT

An investigation was made of the mechanical/acoustic vibration sources of the large ($\lambda/35$) RMS surface errors in 10 in. beam wavefronts measured with the AFWL/ARAA heterodyne holographic interferometer, wavefront sensor (HET) system. Seismometer, etc. vibration sensor measurements, HET optical system measurements and a simple undamped resonant frequency analysis were used to identify the attached acoustic/wind screen as a significant local "source" of nominal 30 and 40 Hz, pitch and yaw vibratory modulations of the HET optical beams. Supporting studies relevant: (a) rigid body motion of the HET optical table system, (b) flexural and torsional motion of HET optical table, and (c) impact of vibrating motion on differential hologram fringe modulation function were performed and are also discussed in this report.

EXECUTIVE SUMMARY

Consistently large mechanical/acoustic vibration induced RMS deviations ($\lambda/35$) in 10 in. beam wavefronts have been measured through late Spring 1983 with the AFWL/ARAA heterodyne holographic interferometer, wavefront sensor (HET) system. These large wavefront surface measurement errors motivated an investigation of the sources of these mechanical/acoustic vibrations and work to eliminate/minimize their impact on HET wavefront sensor performance. Measurements were made of the RMS acceleration level spectra at several locations on the floor in the HET lab and on the HET optical table with the acoustic/wind screen in place - i.e., mechanically attached to the table. These spectra showed two very significant local peaks in the vibration levels at nominal frequencies of 30 and 40 Hz. Independent spectral decompositions of the vertical and horizontal, two-point relative phase measurements made with the HET system indicate that these 30 and 40 Hz frequencies correspond to pitch and yaw vibratory modulations, respectively, of the HET (reference and object) optical beams. A simple analysis (using nominal mass and stiffness values) made of the motion of

the top member of the acoustic/wind screen with respect to the optical table gave translational (correspond to pitch, roll optical beam modulations) and yaw natural resonance frequencies of 29.03 Hz and 40.81 Hz, respectively. The close agreement between the analysis and measured RMS spectra results indicated that the acoustic/wind screen, when attached to the optical table, is a significant local "source" of vibratory motion and hence relative optical beam modulations which degrade the wavefront sensor performance of the HET system. Removal of the acoustic/wind screen gave significant reduction of these 30, 40 Hz vibration noise components for seismometer locations on the optical table. Use of a system of floor to ceiling curtains reduced the large-scale ($> \lambda/4$) random, low-frequency (0 to 2.5 Hz) oscillations--induced by random air movements and turbulence--which degraded HET system performance without the wind screen in place. Here ten-inch beam RMS wavefront surface errors were reduced to $\lambda/60$ with the curtains. In support of the above described measurement and analysis program, work was also done in three areas: (1) Rigid-body, time-domain analysis of the motion of the HET optical table, etc; (2) Tabulation of normalized, approximate-flexural and torsional frequencies for various free, pinned-corner and intermediate support leg conditions; (3) Impact of vibratory motion on differential hologram fringe modulation function. Based upon this work and the results of the measurement and analysis program, recommendations are made to further reduce or eliminate mechanical/acoustic vibration induced problems with the HET system.

ACKNOWLEDGMENTS

The author wishes to thank the Air Force Systems Command, the Air Force Office of Scientific Research, the Southeastern Center for Electrical Engineering Education and the Advanced Beam Control Branch/Air Force Weapons Laboratory, Kirtland AFB, NM for the opportunity to participate in the 1983 AFOSR/SCEEE Summer Faculty Research Program. Thanks are given to Dr. John R. Kenemuth for suggesting this area of research and to Captain Joseph T. Evans, USAF Project Officer for the HET System for his collaboration and help in the performance of the author's research work.

I. INTRODUCTION

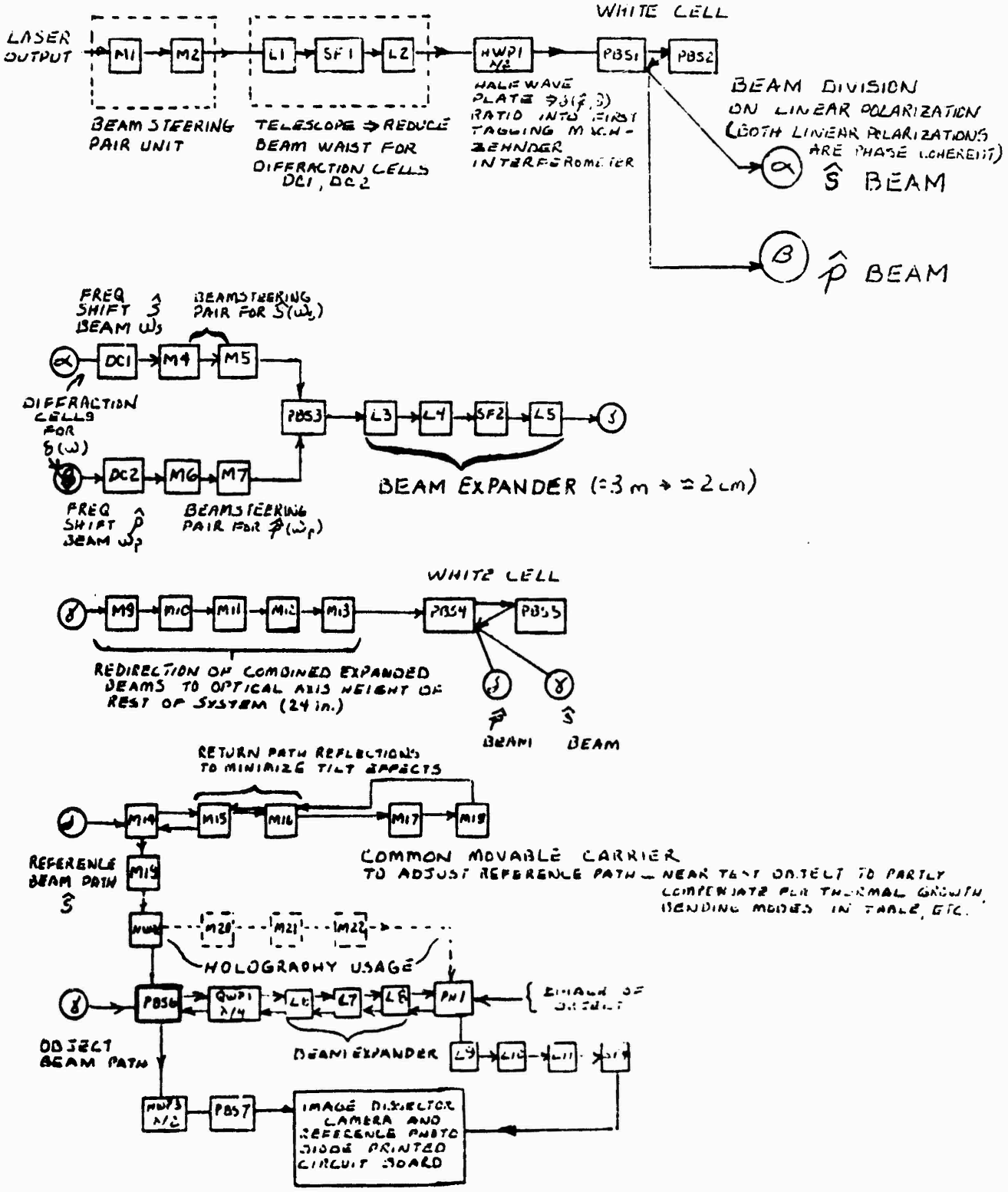
The AFWL/ARAA heterodyne holographic interferometer (HET) system is an extremely sensitive optical wavefront sensor system. The HET can operate in either a conventional hologram-interferometer or a heterodyning hologram-interferometer mode. In the conventional mode both the reference and object optical beams are at the same frequency. In the heterodyning mode the HET uses a constant 625 KHz frequency difference between the reference and the object beams in a heterodyning, frequency-modulation sense. Here the optical result is used to define the differential holographic fringes (difference between specified reference and desired wavefront surfaces) characterizing the spatial distortion of a nominal 10 in. diameter wavefront image of the transverse distortion of a static or a vibrating surface. The 10 in. diameter wavefront size is a nominal design value for the HET system. It can be varied from 1 in. to 16 in. by suitable choice of the beam expander lenses in the optical beam paths. Figure 1 gives a block diagram representation of the optical path components comprising the HET wavefront sensor system.

The HET system is mounted on an optical bench which is presently supported by six vibration isolation support mounts or legs. It originally (through 1 August 1983) was covered by an attached acoustic/wind screen to minimize the effects of random air movements and turbulence upon the performance of the HET system. Figure 2 depicts this HET configuration.

Significant mechanical and/or acoustic vibration level problems have been observed with the HET configuration in its present laboratory

FIG. 1

BLOCK DIAGRAM OF HETERODYNE INTERFEROMETER



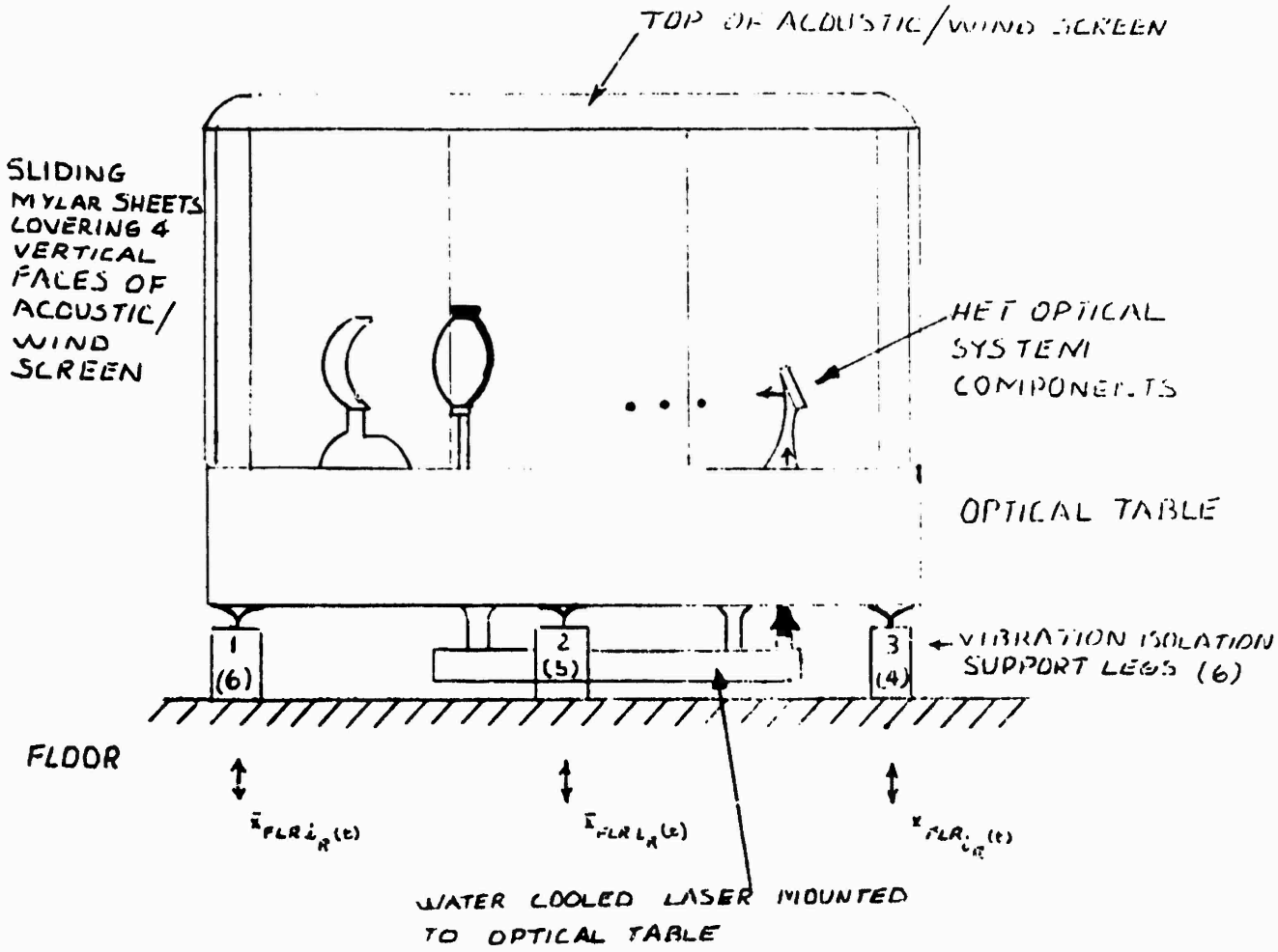


FIG. 2

HET OPTICAL TABLE PLUS ATTACHED ACOUSTIC/WIND SCREEN, LASER SUPPORT TABLE, VIBRATION ISOLATION SUPPORT LEGS, ETC. SYSTEM CONFIGURATION

location in the center of the first floor of Building 413 at AFWL/ARAA. This location may be a vibration antinode in an office building which was neither structurally nor acoustically designed to house high precision optical measurement laboratories.

The mechanical and acoustic vibration problems referred to above manifest themselves in real time as vibrating differential hologram, interferometer fringe patterns in the image formed of the transverse distortion of an illuminated object/wavefront surface. That is, the vibrations can cause modulations in the relative phase between the object and reference beams. After electronic detection at spatially discrete sampling locations and Zernicke orthogonal-modal-basis-function, least squares fit to the deformed surface, consistently large RMS deviations ($\lambda/35$, λ = 5145 Å) have been found in the 10 in. optical beam wavefronts through late Spring 1983. These large wavefront surface measurement errors as compared to an apparently obtainable ($\lambda/200$ to $\lambda/75$), have motivated the measurement and analysis work described in this report. The primary effort in this work consisted of an investigation of the sources of unwanted mechanical/acoustic vibrations and of system modifications to eliminate/minimize their impact on HET wavefront sensor performance. In support of the primary effort work was also accomplished in the following three areas:

- Rigid-body, undamped resonant frequency and time-domain analyses of the motion of the HET optical table, etc.

- Calculation of normalized, approximate-flexural and torsional frequencies of optical table for various free, pinned-corner and intermediate support leg conditions.

- Impact of vibrating motion on differential hologram fringe modulation function.

The remainder of this report is organized as follows. First, the objectives of the research project are stated. Next the experimental and analytical investigations of the sources of mechanical, etc. vibration are described and their results discussed. This is followed by sections describing modifications to the HET optical table configuration and the results of these modifications. Next the work in the three supporting areas are described. Their supporting analysis and results can be found in the Report Appendices. Conclusions are then made regarding the investigation of vibration sources and modifications tried to improve HET system performance, as well as the related support area work. Recommendations are made to further improve HET wavefront sensor performance at the end of the body of the report.

II. OBJECTIVES

The primary objective of this research project has been to investigate sources of mechanical and acoustically coupled vibrations (possibly parametric) to a heterodyne, holographic interferometer, optical wavefront sensor (HET) system. This investigation was made using:

- Limited number of available seismometers and accelerometers to measure velocity and acceleration levels on the floor in the HET lab and on the optical table.

- HET system itself to obtain horizontal and vertical, two-point, relative phase measurements corresponding to yaw and pitch vibratory modulations, respectively, of the HET optical beams.

- Simple resonant frequency analyses of the relative (to optical table) motion of the acoustic/wind screen and the laser support table.

- Limited time-domain computer modeling, analyses of rigid body motion of HET optical table, etc. system designs.

- Idealized temporal filtering of holograms consideration of the impact of vibratory motion upon wavefront sensor fringe modulation function.

Based upon the experimental and the computer, etc. analyses of the vibration problems; techniques and/or design modifications were to be recommended and, if possible, developed and actually tried out in the HET lab. Target goals of this vibration level reduction work were vibration levels permitting 10 in. beam wavefront interferometer system accuracy of $< \lambda/75$ (i.e., RMS wavefront surface errors $< \lambda/75$ with $\lambda = 5145$ Å).

III. HET PLUS ACCELEROMETER, ETC. SENSOR-MEASUREMENT AND DATA ANALYSIS INVESTIGATION OF VIBRATION PROBLEMS

III.1 FAMILIARIZATION WITH HET SYSTEM

A review was made of the relevant optics and of the laser and optical path system components comprising the HET system (Refs. 1-8). As indicated in the introduction, Figure 1 schematically depicts the component lineup and the optical path signal flow in the present HET

TABLE I Sources of Unwanted Noise

- Seismic/Structural Floor Vibration
- Acoustic Pressure Radiation Via Air Path (Fluidborne; Fluidborne, Structureborne)
- Acousto-optical Interaction Coupling
- Flow Induced Vibration Using Water Cooling for Laser (Monitor Transition, Mid and Laser Output Beam Sections of Flow-Line on Laser)
- Flow Induced Plasma Fluctuation and Beam Modal Drift
- Optical Separation of $\bar{p},\bar{s}$ Polarization Beams
- Electrical Line/Electronic Recording Processing and Radiation
- Laser-Detector Shot Noise <-> System Optical Gain Figures

system. In addition, the data collection electronics system and the heterodyning holographic interferometry/signal processing principles upon which the HET system is based were also reviewed (Refs. 4, 9-12).

Relevant all aspects of noise degrading HET wavefront sensor performance, the optical bench or table, etc. plus the vibration isolation legs and the complete optical path/optical, etc. component lineup in the HET were considered. This was done to place the mechanical/acoustic vibration problems in context with respect to other noise sources and to obtain an indication of the potential noise floor -- i.e., best wavefront sensor performance possible with the HET system. Table I summarizes the five mechanical/acoustic and the three-laser/optical polarization/detector -- potential sources or mechanisms of unwanted noise which were identified.

Newport Research Corporation was contacted relevant the design and dynamic mechanical response details of the optical bench, the attached acoustic screen, the bottom mounted laser bench, and the vibration isolation leg support system (Refs. 13,14). This was done to gain further insight from the equipment manufacturer into the mechanical/acoustic vibration frequencies and mode shapes possible in the HET optical table design. NRC indicated that the first free-free flexural bending and torsional resonance frequencies should be approximately 100 and 170 Hz, respectively for the HET's 12 ft x 4 ft x 1.5 ft optical table in an unloaded condition. Nominal vertical and transverse stiffnesses of 64 lb_f/in and 125 to 150 lb_f/in were indicated for each of the six pneumatic, vibration isolation mount support legs . . frequencies around 1 Hz.

It was also indicated these stiffnesses are frequency dependent. That is, above 1 Hz these stiffnesses are somewhat less than (nominal 50 to 60% of) the 1 Hz values. Vertical and transverse damping values were unavailable. They can, however, be inferred from the frequency response results given in the equipment catalog (Ref. 13).

III.2 HET INTERFEROMETER PLUS ACCELERATION, ETC. SENSOR MEASUREMENTS

Initially a complete set of measurements were defined to be taken, up to six channels at a time, digitized and transferred from the HET system's Data General Computer to the online user disk pack of the PDP-11/24. This plan called for an assortment of single axis and triaxial accelerometers to be located

1) On the floor in the HET Lab room.

2) At the laser coolant system inlet and at the end of the coolant line after transition to (turbulent) steady-state flow.

3) At the four corners and 1/2, 1/4 points along the length, width of the top of the HET optical table.

4) On the acoustic/wind screen--i.e., at the corners and at the 1/2, 1/4 points along its length and width edges.

5) On the laser support bench--i.e., at the corners and 1/2 points along its length edges.

In addition, two acoustic microphones, if available, were to be placed at varying loctions both inside and outside of the wind screen

enclosure. The HET system itself was also to be used to obtain simultaneous, two-point relative-phase modulation measurements and corresponding time (not simultaneous), HET maps with RMS error results.

After canvassing AFWL re-availability of triaxial and single axis accelerometers, the author found three ENDEVCO 2262C-25 piezoresistive accelerometers plus corresponding amplifier/signal condition cards (AFWL/NTDE) and three Geo Space HS-10-1 seismometers (AFWL/NTDE) (Refs. 15,16). In addition, two PCB 303A02 piezotronics accelerometers and amplifier/signal conditioning systems were borrowed from the M. E. Department at Texas Tech (Ref. 17). AFWL/ARLB which performs a vibration measurement and analysis service function was also contacted relevant the use of accelerometer sensors, modal analysis equipment, acoustic microphones and, if available, personnel to help with the measurement and analysis program. Unfortunately, ARLB personnel and equipment were not available within the time frame of the author's research project.

Velocity and acceleration measurements have been made using the three Geo Space seismometers (direct measurement of velocity) and the two PCB accelerometers. The seismometers have been the sensors most extensively used to

(a) Measure acceleration, velocity, and motion levels at three locations across the mid-length plane on the optical table and at three locations near the 10-inch beam expander and mirror components (east end of optical table in HET lab).

(b) Measure acceleration, velocity and motion levels at three locations on the floor under the optical table and at a location south of and away from (4 ft) the optical table.

In addition, the seismometers were also used to survey the acceleration, etc. levels in three other ARAA labs. Here measurements were made at locations under and on the optical tables in the (1) Phasar lab, (2) Deformable Mirror Lab (D. Heimlich) and (3) Adaptive Optics Lab (D. DePatie). It should be noted that the optical table in the Phasar Lab also has an attached (somewhat shorter vertically) wind screen enclosure. The deformable mirror and adaptive optics labs also each have a 6+ ft long trench running lengthwise (east-west) along their floors. The optical tables in these three labs are somewhat longer (15 to 18 ft) than the nominal 12 ft table long HET Lab table.

Vertical and horizontal, two-point relative phase modulation measurements and RMS wavefront surface error measurements were made using HET system detector signal conditioning A/D conversion hardware and its Data General signal processing software. It should be noted that the vertical and horizontal, two-point relative phase measurements correspond to pitch and yaw vibratory motion modulations, respectively, of the reference and object beams in the HET system.

III.3 DATA PROCESSING/ANALYSIS

One- (for time sample analysis) and two- (for wavefront surface analyses) dimensional FFT(FFT^{-1})s were put on the ARAA PDP-11/24 computer system by the author. Consideration was also given to anti-aliasing and window smoothing filters tuned to the frequency range and the dynamic bit range of the sampled data (Refs. 18-20). However, these FFT programs, etc. were not used to develop an integrated data

analysis (power spectra, coherence, etc.) software package to process the accelerometer, etc. sensor and HET system-time sample results discussed in this report. This was because the Data General Eclipse S/200 software for A/D conversion and transfer to the PDP-11/24 was not available until late in the author's research project work. The "quick look" results discussed in this report were based on (a) Nicolet 660A digital oscilloscope RMS spectral analysis, differentiation and integration processing and plotting system and (b) HET system relative phase measurement, FFT processing and plotting capability using Data General Eclipse S/200 computer system. The latter was developed during Summer 1983 by Miss Geraldine Cordova under the direction of Captain Joseph T. Evans. It was checked out and then available for use by the last week in July.

III.4 IDENTIFICATION OF VIBRATION NOISE SOURCES IMPACTING SYSTEM PERFORMANCE

III.4.1 Measurement Results

As indicated previously a survey/comparison was made of the velocity and acceleration levels in the ARAA-HET, PHASAR, Deformable Mirror and Adaptive Optics Labs. RMS floor velocity, etc. level results in these four labs indicated that the Deformable Mirror Lab is quietest at 20 to 50 Hz closely followed by the Adaptive Optics Lab. The PHASAR and HET Labs seemed to be somewhat noisier (nominal 5 to 10 dB) in terms of peak velocity and acceleration levels in this 20 to 50 Hz frequency range with the HET Lab the noisier of these two. This

may be correlated with, i.e., due to the fact that the HET Lab is near the center of the first floor of AFWL Building 413 which can be a vibration antinode. Building 413 is an office building which was not structurally and acoustically designed for precision optical measurements.

Comparison of the difference between the RMS acceleration, etc. spectra levels for the table and the floor for each of the four above-mentioned labs indicates nominal 25 to 35 dB differences (level reductions) for the Deformable Mirror and Adaptive Optics Labs and 20 to 25 dB level reductions for the HET (nominal 20 dB) and PHASAR (nominal 25 dB) Labs in the 25 to 45 Hz frequency range. The relevance of the peak levels observed at nominal 30, 40 Hz frequencies will be discussed later. Here it should be noted that the optical tables in the HET and PHASAR Labs both have attached acoustic/wind screen enclosures while the optical tables in the Deformable Mirror and Adaptive Optics Labs do not.

Figures 3a and 3b give representative plots of the RMS spectral decompositions of the velocity levels under and on the HET system optical table with the attached wind screen enclosure. These results show very significant peaks at nominal frequencies of 30 and 40 Hz. Examination of these figures indicates that the floor minus table differences in the 30, 40 Hz levels are 19.8, 21.4 dB, respectively, or a nominal 20 dB. These reductions are somewhat less than the vibration isolation to be expected at the NRC indicated 12 dB/octave fall off above 1 Hz with the six NRC vibration isolation mount support legs. Since comparable NRC vibration isolation legs seem to perform well in corresponding leg configurations in the Deformable Mirror and Adaptive

FIG. 3a. RMS SPECTRAL DECOMPOSITION OF VELOCITY LEVEL UNDER HET OPTICAL TABLE

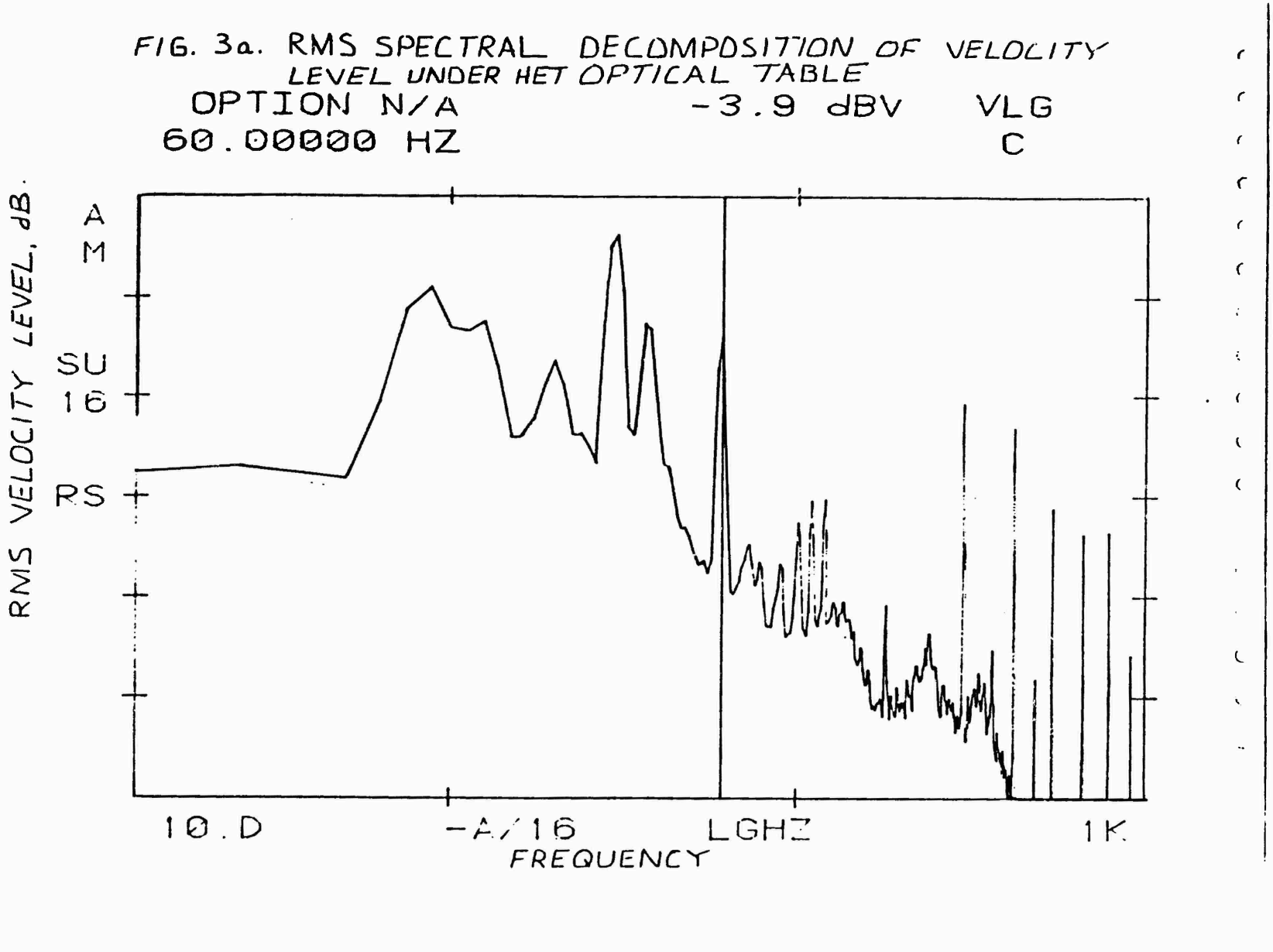
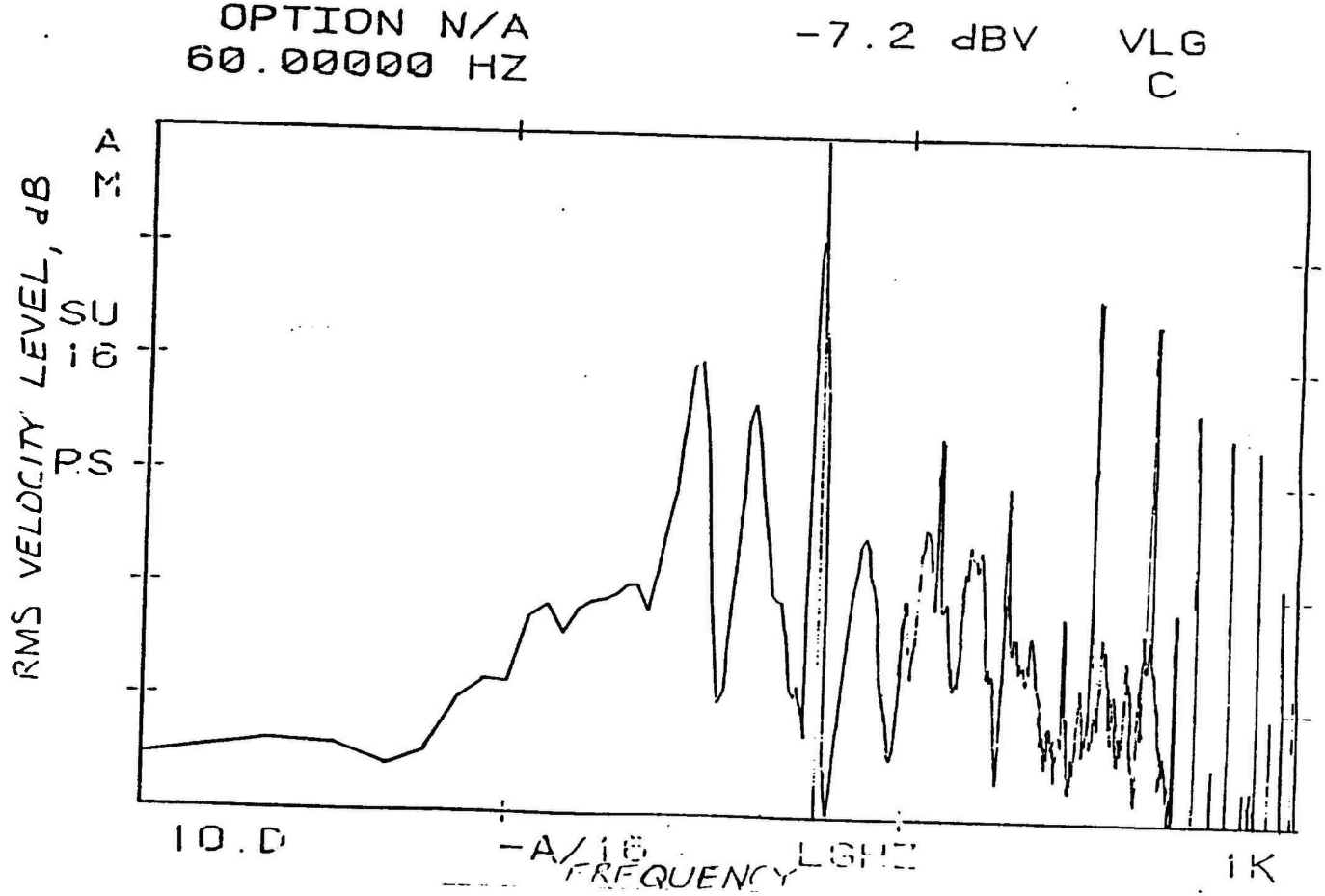


FIG. 3b RMS SPECTRAL DECOMPOSITION OF VELOCITY LEVEL ON HET OPTICAL TABLE



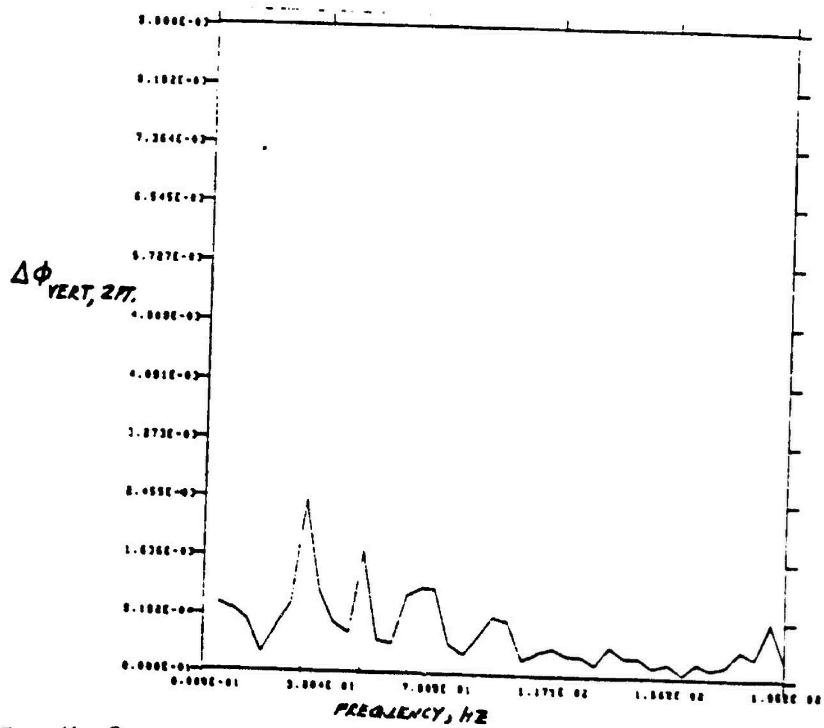


FIG. 4a SPECTRAL DECOMPOSITION OF VERTICAL, 2 POINT RELATIVE PHASE MEASUREMENT FROM HET SYSTEM

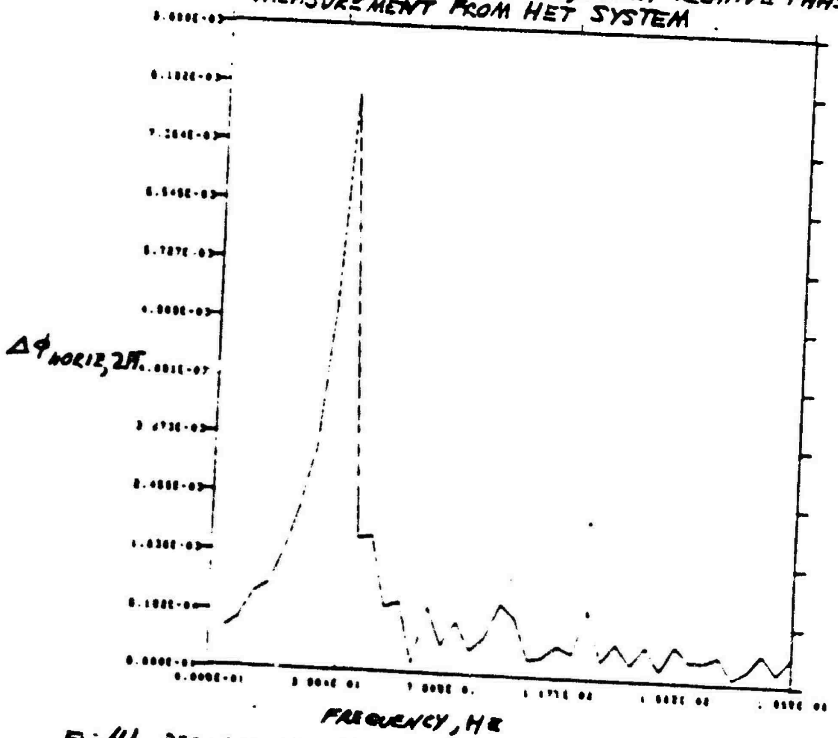


FIG 4b DECOMPOSITION OF HORIZONTAL, 2 POINT RELATIVE MEASUREMENT FROM HET SYSTEM

Optics Labs, these results indicate that additional vibration noise sources as, e.g., acoustic are exciting the system via the acoustic/wind screen enclosure.

Figures 4a and 4b respectively plot spectral decompositions of the vertical and horizontal, two-point relative phase measurements which were made with the HET system. It should be noted that an independent sensor (HET system) using a different physical process--as compared to the seismometers--was used to obtain these results. Figure 4a shows a relatively large RMS motion/relative phase modulation peak at a nominal 30 Hz. This is for the two vertical points configuration in the HET detector array and corresponds therefore to a pitch vibratory modulation of the HET reference and object optical beams. Figure 4b similarly gives a large RMS motion/relative phase modulation peak at a frequency of 42 to 43 Hz. In this case two horizontal points were sampled in the HET detector array. Thus it corresponds to a yaw vibratory modulation of the reference and object optical beams in the HET system.

III.4.2 Simple Analysis of Screen Enclosure, Laser Table Motions WRT Optical Table

Simple spring-mass analyses were made of the translational and yaw motions of

(a) Top member of the acoustic/wind screen with fixed end attachment to the top of the optical table.

(b) Bottom member/ledge of laser bench with fixed end attachment to the bottom of the optical table.

Figure 2 depicts the location of the wind screen and laser bench structures. More detailed figures are included in Appendix A.

Nominal mass, stiffness values were used in the analyses of these motions with respect to the optical table. Here it is assumed that since the rigid body motion natural frequencies of the optical table (as indicated by experimental results and by analysis and time domain computer model results) are in the low frequency range 0 to 5 Hz while the wind screen and laser bench resonant frequencies are in the range 25 to 55 Hz. That is, they are somewhat separated in frequency and thus the wind screen and laser bench can be considered (each separately) as moving with respect to a fixed table. Thus uncoupled, single-degree-of-freedom representations can be used for each of the x,y translation and yaw (wrt z axis) motions. Details of the analysis calculations are given in Appendix A.

Table II gives the natural resonant frequency results, 29.03 Hz and 40.81 Hz calculated for the translational and yaw motions, respectively, of the top member of the wind screen. Table III summarizes corresponding translational and yaw resonant frequency results, 48.09 Hz and 39.88 Hz, respectively, for the laser support bench.

These resonant frequency analysis results taken together indicate that there are translational motion resonant frequencies at nominal values of 29 and 48 Hz. In addition, there are two closely placed yaw motion resonant frequencies at nominal values of 40 and 41 Hz. These results are in close agreement with the significant peaks observed in the RMS seismometer velocity level spectra at nominal frequency ranges of

TABLE II Translational Mode and Yaw Mode Natural Resonant Frequency Results for Motion of Top Member of Acoustic/Wind Screen WRT Optical Table

	$f_{NAT\ RES}$ (Hz)
x_{TRANS} (along table length)	29.03
y_{TRANS} (along table width)	29.03
yaw (angular motion WRT vertical z axis)	40.81

TABLE III Translational Mode and Yaw Mode Natural Resonant Frequency Results for Motion of Bottom Ledge of Laser Support Bench WRT Optical Table

	$f_{NAT\ RES}$ (Hz)
x_{TRANS} (along table length)	48.09
y_{TRANS} (along table width)	48.09
yaw (angular motion WRT vertical z axis)	39.88

29-31 Hz, 40-42 Hz and on occasion 50-52 Hz with both the wind screen and the laser bench structures in place. These results are also in agreement with observed peaks in the independent spectral decompositions of the vertical and horizontal, two-point relative phase measurements made with the HET system (with both the wind screen and the laser bench attached to the optical table).

Specifically, the 29 Hz and 48 Hz natural resonant frequencies are for relative (x,y) translational motions which correspond to roll and pitch optical beam modulations. They correspond to the 32 to 33 Hz and 53 Hz spectral peaks found in the vertical, two-point relative phase measurements directly giving the pitch motion component of the modulation of the HET system optical beams. Similarly, the calculated 40 Hz and 41 Hz yaw natural resonance frequencies correspond to the 43 Hz spectral peak found in the horizontal, two-point relative phase measurements. These measurements directly give the yaw motion component of the modulation of the optical beams in the HET system.

In summary, these results indicate that both the acoustic/wind screen and the laser support bench are important local "sources" of mechanical vibration in the HET system. Together they can cause nominal 25 to 55 Hz frequency modulations of the phase difference between the reference and the object beams. Thus they are apparently significant local sources or causes of the consistently large $\approx \lambda/35$-RMS wavefront surface deviation errors measured with the HET system.

III.5 DESIGN MODIFICATION STUDY AND RESULTS

It was decided to remove the attached acoustic/wind screen--a non-trivial, though relatively easy task--and investigate the impact of this change upon measured acceleration, velocity, etc. levels and upon HET system wavefront sensor performance. The laser support bench, etc. was not also removed because its removal is relatively difficult and would require long term separate mounting modifications to suitably reconfigure the laser to optically drive the HET system. It should be noted that the ideal modification plan here would be to remove only one of these two structures then, replacing that one, only remove the other, and then remove both.

Removal of the acoustic/wind screen was found to give significant reduction of the nominal 29-31 Hz (10 to 15 dB reduction) and 40-43 Hz (15 to 20 dB reduction) vibration noise components for seismometer locations on top of the optical table. However, when the

(a) Differential hologram optical fringes were visually observed for the 10 in. beam wavefronts in the HET system,

and

(b) Corresponding relative phase between the reference and object optical beams was visually observed on the system oscilloscope,

it was seen that removal of the wind screen gave large-scale ($> \lambda/4$ and hence quadrant errors), random oscillations in the low frequency range 0.0 to 2.5 Hz.

Subsequent investigations showed that these oscillations were due to random air movements and turbulence as, e.g., generated by

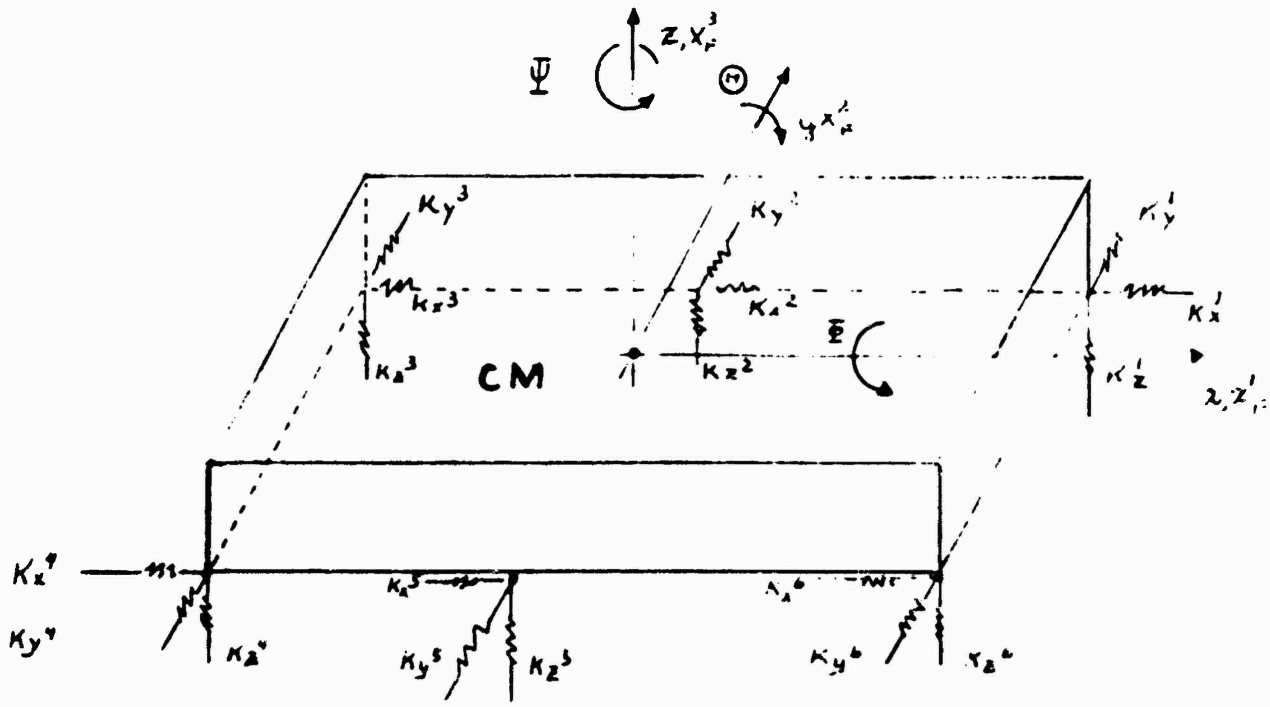


FIG. 5

IDEALIZED SIX DEGREE OF FREEDOM REPRESENTATION FOR RIGID BODY MOTION ANALYSIS OF HET OPTICAL TABLE SYSTEM

oscilloscope and computer, etc. equipment fans and blowers in the HET Lab. These air movements, etc. severely degraded HET system wavefront sensor performance without some form of a wind screen surrounding the optical table and system components.

A system of floor to ceiling curtains with a rudimentary ceiling of plastic film sheeting which completely surrounds, but is not physically attached to, the HET optical table was used as a near term solution to reduce the effects of the random air movements and turbulence. It was found that 10 in. beam RMS wavefront surface deviation errors can be reduced to $\lambda/60$ with the curtains and ceiling in place.

IV. SUPPORTING STUDIES

IV.1 UNDAMPED NATURAL RESONANT FREQUENCY ANALYSIS OF RIGID BODY MOTION OF OPTICAL TABLE

An undamped natural resonant frequency analysis was made of the rigid body motion (RBM) of the HET optical table (Refs. 21-25). Figure 5 depicts the idealized six degree-of-freedom representation used in the analysis. The six degrees of freedom are:

x or x_T^1 = Translational motion along the table length.

y or x_T^2 = Translational motion transverse to the table length, i.e., along table width.

z or x_T^3 = Translational motion in the vertical direction.

ϕ = Roll angular motion about the x axis.

θ = Pitch angular motion about the y axis.

Ψ = Yaw angular motion about the z axis.

TABLE IV Summary of Undamped Natural Resonant Frequencies of HET Optical Table with Six Legs

	$f_{NAT\ RES}$ (Hz)
x_{TRANS} Mode 1 θ_{PITCH}	0.27
x_{TRANS} Mode 2 θ_{PITCH}	2.34
y_{TRANS} Mode 1 ϕ_{ROLL}	0.66
y_{TRANS} Mode 2 ϕ_{ROLL}	2.61
z_{TRANS}	1.12
ψ_{YAW}	2.49

Note that the six legs are assumed to be placed one at each corner and one at the midpoint of each nominal 12 ft length-side.

Here x,y,z (or x_F^1, x_F^2, x_F^3) form a right-handed, fixed inertial coordinate system. Positive x,y,z are in the directions indicated in Figure 5. Positive ϕ, θ, Ψ are defined as positive if they are CCW with respect to the positive x,y,z axes, respectively. Figure 5 also shows the six vibration isolation mount support legs which are represented as three-dimensional linear springs. In this analysis the motions are assumed to be small such that the resulting dynamic system is linear. It is further assumed that the (x,θ) motions and the (y,ϕ) motions are coupled each pair together, while the z and the Ψ motions are each uncoupled

Table IV summarizes the results of this undamped natural resonant frequency analysis. These RBM frequencies range from 0.27 Hz to 2.61 Hz. They are in basic agreement with relative/load peaks which have been observed in the measured acceleration, velocity and motion level spectra for non-zero initial conditions in each of the six degrees of freedom. Appendix B gives details of the analysis. This includes a summary of the equations used to calculate the resonant frequencies given in Table IV.

IV.2 12 STATE VARIABLE, TIME-DOMAIN ANALYSIS OF RIGID-BODY-MOTION OF OPTICAL TABLE, ETC. DESIGNS

A 12 state-variable, nonlinear, time-domain model analysis was made of the rigid-body-motion (RBM) of the HET optical table, etc. system design. The dynamic systems model considers x_F^i, i = 1,2,3--

translational motion of the center of mass and Φ, Θ, Ψ--roll, pitch and yaw degrees of freedom. The dynamic systems model can handle an arbitrary number (> 4) of vibration isolation mount support legs arbitrarily placed under the table. Figure 5, as discussed previously, gives an idealized representation of the system under consideration.

The RBM equations of motion can be defined in state vector equation form as follows*:

$$\frac{d\bar{z}}{dt} = A\bar{z} + B\bar{\bar{F}} \tag{1}$$

in which,

$$\bar{z} = \begin{bmatrix} \bar{x}_F \\ \bar{v}_B \\ \bar{\omega}_B \\ \bar{\epsilon} \end{bmatrix}, \text{ the 12 component state vector} \tag{2}$$

$$\bar{\bar{F}} = \begin{bmatrix} \bar{F} \\ \bar{M} \end{bmatrix}, \text{ the 6 component generalized force vector} \tag{3}$$

$$A = \begin{bmatrix} 0 & T_{FB} & 0 & 0 \\ 0 & \Omega & 0 & 0 \\ 0 & 0 & (I^{-1}\Omega I) & 0 \\ 0 & 0 & \tilde{T}_{FB} & 0 \end{bmatrix}, \text{ 12 x 12 nonlinear system matrix} \tag{4}$$

* See Refs. 26,27 for background development re Eqs. (1-11).

$$B = \begin{bmatrix} 0 & 0 \\ \left(\frac{1}{m}U\right) & 0 \\ 0 & I^{-1} \\ 0 & 0 \end{bmatrix}$$, 12 x 6 compatibility matrix relating generalized forces to time rate of change of state vector components. (5)

$$T_{FB} = \begin{bmatrix} (c\theta c\Psi) & ([s\phi s\theta c\Psi] - [s\Psi c\phi]) & ([c\Psi s\theta c\phi] + [s\Psi s\phi]) \\ (c\theta s\Psi) & ([s\phi s\theta s\Psi] + [c\Psi c\phi]) & ([s\theta c\phi s\Psi] - [c\Psi s\phi]) \\ (-s\theta) & (c\theta s\phi) & (c\theta c\phi) \end{bmatrix} \quad (6)$$

in which, c( ) = cos ( ), s( ) = sin ( ),

$$T_{BF} = (T_{FB})^T \quad (7)$$

Note that

$$\bar{x}_F = T_{FB}\bar{x}_B \qquad \text{Transformation relations between rotating}$$
$$\bar{x}_B = T_{BF}\bar{x}_F \qquad \text{body and fixed inertial vector components.} \quad (8)$$

$$\dot{T}_{FB} = \begin{bmatrix} 1 & (s\phi tn\theta) & (c\phi tn\theta) \\ 0 & (c\phi) & (-s\phi) \\ 0 & (s\phi sc\theta) & (c\phi sc\theta) \end{bmatrix} \quad (9)$$

in which, tn( ) = tan ( ), sc( ) = sec ( )

Note that

$$\frac{d\bar{\Theta}}{dt} = \tilde{T}_{FB}\,\bar{\omega}_B \text{ or } \frac{d}{dt}\begin{bmatrix}\Phi\\ \Theta\\ \Psi\end{bmatrix} = \tilde{T}_{FB}\begin{bmatrix}P\\ Q\\ R\end{bmatrix} \tag{10}$$

in which,

$(\omega_B^1 = P)$, $(\omega_B^2 = Q)$, $(\omega_B^3 = R)$ = so-called "roll, pitch and yaw rates," - wrt body axes.

$$\Omega = \begin{bmatrix}0 & R & -Q\\ -R & 0 & P\\ Q & -P & 0\end{bmatrix} \tag{11}$$

$$\bar{F} = \sum_{r=1}^{r=r_{max}} \bar{F}_{(r)} = T_{BF} \sum_{r=1}^{r=r_{max}} F^{i_F}_{VASL(r)}\,\bar{u}_{i_B};\quad 1 \le i_F, i_B \le 3 \tag{12}$$

$1 \le r \le r_{max}$, the number of vibration absorber support legs.

$\bar{u}_{i_F}$, $\bar{u}_{i_B}$ are unit vectors in frame, body coordinate systems.

$$F^{i_F}_{VASL(r)} = K^{i_F}_{(r)j_F}\left(x^{j_F}_{(r)} - x^{j_F}_{FLOOR(r)}(t)\right) \tag{13}$$

$$1 \le i_F, j_F \le 3$$

$$1 \le r \le r_{max}$$

$$\bar{M} = \sum_{r=1}^{r=r_{max}} \bar{R}_{(r)} \times \bar{F}_{(r)} \qquad (14)$$

This dynamic systems model was implemented in the Fortran IV computer software package HETRBM09, which consists of a driver plus 21 subroutines and 2 function subprograms. This software has been thoroughly checked out. Complete listings of HETRBM09 and of an example input data set file HETRBMIN are given in Appendix C.

HETRBM09 runs have been made which show the coupling of the

(1) $x_F^1 = x$ translation and θ pitch

(2) $y = x_F^2$ translation and ϕ roll

degrees of freedom. The $z = x_F^3$ vertical translation and Ψ yaw degrees of freedom are each essentially uncoupled from all others. The undamped resonant frequency behavior observed with HETRBM09 for non-zero initial condition runs with zero motion input at the floor are in good agreement with the results discussed in Section IV.1.

It should be noted that time did not permit including the effects of:

(a) Damping of the vibration isolation mount support legs.

(b) Mass and inertia effects of additional optical components and structures (as, e.g., wind screen and laser support bench structures) in HETRBM09. Modifications to include these effects can easily be made at a later date. HETRBM09 can also be used to investigate optimum leg placement, etc. designs--i.e., those placement designs which give the best rigid body motion behavior.

IV.3 FINITE ELEMENT METHOD ANALYSIS OF OPTICAL TABLE, ETC. DESIGNS

Here it was desired to use ANSYS or SAPV finite element computer programs to analyze the flexural, torsional, etc. elastic system response of the optical table, the acoustic/wind screen enclosure, the laser support table, and the large beam expander lens and mirror optical components. Time limitations (including the two-week downtime caused by relocation of the CRAY/CDC/IBM large-scale computer system) did not permit the effective use of these finite element method software packages.

However, a normalized, approximate resonant frequency analysis was made of the flexural and torsional vibration modes of the optical table. Here the free/pinned boundary, pinned intermediate support results given in Blevins (Ref. 28) for the resonant frequencies of the multispan elastic beam configuration were used. Tables V, VI and VII present these undamped natural resonant frequency results which have been normalized to a first free-free unloaded, single span flexural frequency of 100 Hz and a corresponding torsional frequency of 170 Hz. These values were given by Newport Research Corporation (Ref. 14) as reasonable approximations for a nominal 12 ft x 4 ft x 1.5 ft HET system optical table.

These results indicate that significant changes in the resonant frequencies and hence mode shapes excited at a given excitation frequency can be effected by a judicious or more optimal placement of the vibration isolation mount support legs. Indeed, Whaley and Pearson

TABLE V Resonant or Natural Frequencies for Optical Table Modeled as a Free-Free Beam with Pinned Intermediate Supports

L = 12.0 ft

i	f_i(Hz) Single Span No Support	f_i(Hz) Triple Span 2 Supports (4 Legs)	f_i(Hz) Quadruple Span 3 Supports (6 Legs)
1	100.0	80.2	162.2
2	275.6	109.3	176.5
3	540.8	515.6	833.0
4	893.7	734.5	1103.4
5	1334.6	891.3	1408.6
6	1863.8	1809.5	1588.5

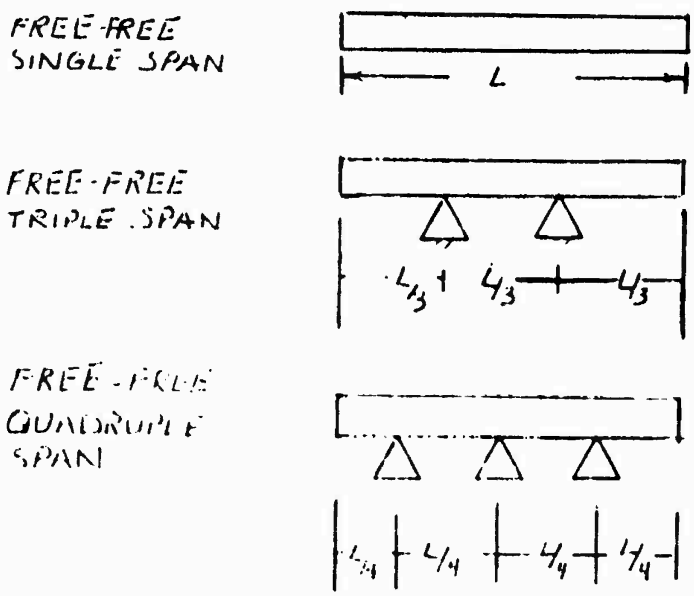
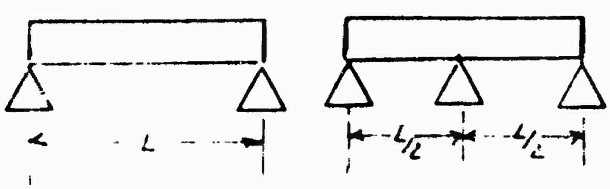


TABLE VI Resonant or Natural Frequencies for Optical Table Modeled as a Pinned-Pinned Beam with Pinned Intermediate Supports

L = 12 ft.

i	f_i(Hz) Single Span 2 Supports (4 Legs)	f_i(Hz) Double Span 3 Supports (6 Legs)
1	44.1	176.5
2	176.5	275.7
3	397.0	705.8
4	706.2	893.2
5	1103.1	1587.8
6	1588.2	1863.8

PINNED PINNED SINGLE SPAN



PINNED-PINNED DOUBLE SPAN

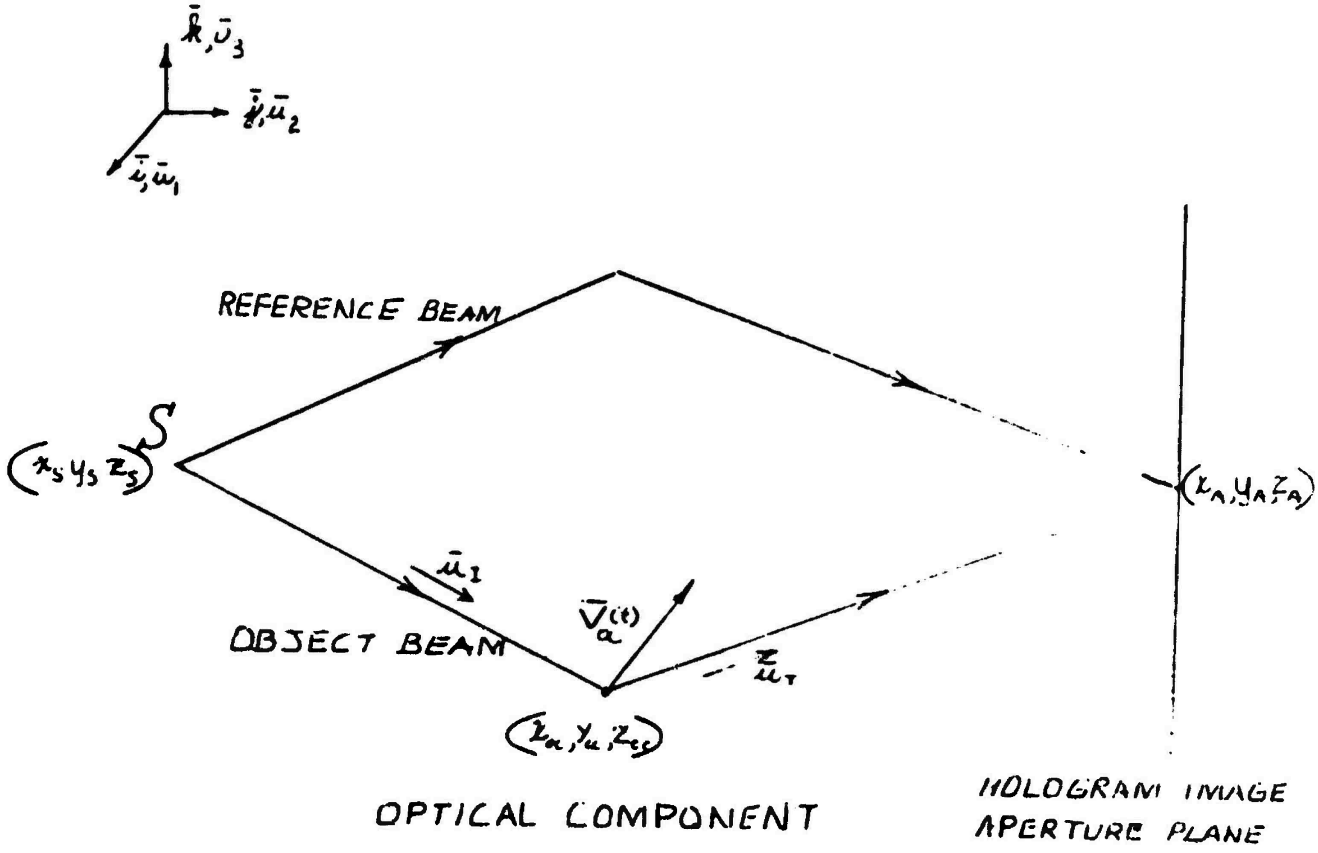


FIG. 6

REPRESENTATIVE LINE DRAWING SHOWING VIBRATORY MOTION GEOMETRY FOR AN OPTICAL COMPONENT IN OBJECT BEAM PATH. NOTE THAT $\bar{u}_2, \bar{u}_1$ ARE THREE-DIMENSIONAL UNIT VECTORS FOR INCIDENT TRANSMITTED OPTICAL BEAMS AT OPTICAL COMPONENT

$$\text{Min } [\bar{V}_{\text{OPT. COMP.}} \cdot (\bar{u}_T - \bar{u}_I)]^2 \qquad (18)$$

Over

Leg Placement Designs

Table Physical Parameters

Table Placement, etc. of Optical Components

IV.4 VIBRATORY/ACOUSTIC MOTION-IMPACT ON FRINGE MODULATION

A review was made of the survey work by Smith (Ref. 30) relevant the impact of motion upon holograms and the fringe modulation function. Smith's consideration is based upon the temporal filtering of holograms formulation introduced by Goodman (Ref. 31). Powell and Stetson also considered the related problem of interferometric vibration analysis and derived equivalent results from a somewhat more complicated point of view (Ref. 32). The related areas of FM modulation (straight line FM slide, constant frequency sinusoidal modulation, etc.) as considered by Nitta (Ref. 33) and the impact of noise on phase-lock-loop, heterodyne system performance as examined by Gardner (Ref. 34), Lindsey (Ref. 35) and Van Trees (Ref. 36) were also reviewed.

The analysis for sinusoidally vibrating motion can be summarized as follows. Consider the reference and object beam geometry as depicted in Fig. 6. Here the optical component in the object beam path is assumed to have vector sinusoidal motion as per

Velocity

$$\bar{V}_{oc} = \bar{V}_{oc}(x_{oc}, y_{oc}, z_{oc}) \cos \omega_T t \qquad (19a)$$

Displacement

$$\Delta\bar{R}_{oc} = \left(\frac{\bar{V}_{oc}}{\omega_T}\right) = \frac{V_{oc}(x_{oc},y_{oc},z_{oc})}{\omega_T}\,\bar{u}_{oc}\,\sin\omega_T t \qquad (19b)$$

in which,

(x_{oc}, y_{oc}, z_{oc}) = Three-dimensional location of "optical center" of optical component under consideration.

$V_{oc}(x_{oc}, y_{oc}, z_{oc})$ = V_{oc} = Magnitude of velocity of optical component. $V_{oc} = [\bar{V}_{oc} \cdot \bar{V}_{oc}]^{1/2}$

$\bar{u}_{oc}$ = Unit vector defining direction of vibratory (i.e., sinusoidal) motion of optical component. $\bar{u}_{oc} = \left(\frac{\bar{V}_{oc}}{V_{oc}}\right)$.

ω_T = Temporal angular velocity of vibratory motion at frequency f. $\omega_T = 2\pi f$.

It should be noted that the optical component could be any lens, mirror, beam splitter, etc. component in the object or the reference optical beam paths or it could even be the wavefront hologram aperture-electronic detector or film recording medium, etc.

The vibratory motion can also arise acoustically in the air through which the optical beam travels. It is assumed that the vibratory motion is small enough such that $\bar{u}_{oc} \cdot \bar{u}_I$ and $\bar{u}_{oc} \cdot \bar{u}_T$ are constant with time. That is, that the changes in the angles between

the direction of the vibratory motion and the directions of the incident and transmitted beams are very small and can be neglected with respect to unity or

$$\underset{\text{over } t}{\text{Max}}\ \bar{u}_{oc} \cdot \bar{u}_I - \underset{\text{over } t}{\text{Min}}\ \bar{u}_{oc} \cdot \bar{u}_I \ll 1 \tag{20a}$$

$$\underset{\text{over } t}{\text{Max}}\ \bar{u}_{oc} \cdot \bar{u}_T - \underset{\text{over } t}{\text{Min}}\ \bar{u}_{oc} \cdot \bar{u}_T \ll 1 \tag{20b}$$

The exposure at point (x_A, y_A, z_A) of the electronic detector or the film recording medium in the hologram aperture plane is

$$E(x_A,y_A,z_A) = \int_{-T/2}^{+T/2} H(x_A,y_A,z_A;t)^* H(x_A,y_A,z_A;t)\,dt \tag{21}$$

in which,

$H(x_A,y_A,z_A;t)^*H(x_A,y_A,z_A;t)$ = the irradiance distribution at the detector/recording medium location.

T = Total sample or exposure time

$(\ )^*$ = Complex conjugate operator

Since,

$$H(x_A,y_A,z_A;t) = R(x_A,y_A,z_A) + O(x_A,y_A,z_A;t), \tag{22}$$

Eq. (21) can be rewritten as

$$E(x_A,y_A,z_A) = E_{REF} + E_{OBJ} + E_{CONJ} + E_{PRIM} \tag{23a}$$

in which,

E_{REF}, E_{OBJ} = Exposures resulting from reference, object complex light waves alone, respectively.

$$E_{CONJ} = E_{PRIM}^{*} \tag{23b}$$

$$E_{PRIM} = R^{*}(x_A, y_A, z_A) \int_{-T/2}^{+T/2} O(x_A, y_A, z_A; t)dt \tag{23c}$$

which results in the primary image.

The primary image term E_{PRIM} is thus the irradiance image term of interest in the following analysis.

Defining

$$\text{Rect}(t/T) = \begin{cases} 1, & |t| < T/2 \\ 0, & |t| > T/2 \end{cases} \tag{24a}$$

and writing the Fourier transform of $O(x_A, y_A, z_A; t)$ as

$$\tilde{O}(x_A, y_A, z_A; \nu) = \int_{-\infty}^{+\infty} O(x_A, y_A, z_A; t)dt \tag{24b}$$

allows Eq. (23c) to be written as

$$E_{PRIM} = TR(x_A, y_A, z_A)^{*} \int_{-\infty}^{\infty} \text{sinc}(\Pi\nu T)\tilde{O}(\tilde{x}_A, y_A, z_A; \nu)d\nu \tag{25}$$

in which

$T\,\text{sinc}(\Pi\nu T) = \dfrac{\sin(\Pi\nu T)}{(\Pi\nu)}$ = Fourier transform of Rect (t/T)

It should be noted that Eq. (25) can be interpreted as a linear filter with transfer function $\text{sinc}(\Pi\nu T)$.

Using the sinusoidal time dependence given in Eq. 19 allows the time dependent phase difference between reference and the object beams to be written as

$$\phi_{REF} - \phi_{OBJ} = \phi_o + \frac{V_{oc}}{\omega_T} \bar{u}_{oc} \cdot (\bar{u}_T - \bar{u}_I) \sin \omega t \tag{26}$$

Thus

$$\begin{aligned} O(x_A, y_A, z_A; t) &= O_o(x_A, y_A, z_A) e^{i\phi(t)} \\ &= O_o(x_A, y_A, z_A) e^{i\phi_o \; i a_{MOD} \sin \omega_T t} \end{aligned} \tag{27}$$

in which

$$a_{MOD} = \frac{V_{oc}(x_{oc}, y_{oc}, z_{oc})}{\omega_T} [\bar{u}_{oc} \cdot (\bar{u}_T - \bar{u}_I)] \tag{28}$$

$k = 2\Pi/\lambda$ = Spatial wave number/frequency.

Substituting Eqs. (27,28) in Eqs. (24b,25) yields

$$E_{PRIM} = TR(x_A^*, y_A, z_A) \; O_o(x_A, y_A, z_A) e^{i\phi_o} \sum_{n=-\infty}^{n=+\infty} J_n\left(\frac{V_{oc}}{\omega_T} [\bar{u}_{oc} \cdot (\bar{u}_T - \bar{u}_I)]\right) \cdot \operatorname{sinc}\left(\frac{n\omega_T T}{2}\right) \tag{29}$$

In Eq. (29), J_n ( ) = nth order Bessel function of the first kind of argument ( ).

Because $\bar{u}_T = \bar{u}_T(x_A, y_A, z_A)$ in Eq. (29), the electronically detected or film recorded fringe modulation

$$M = \left(\frac{2 E_{PRIM} E_{CONJ}}{E_{REF} + E_{OBJ}}\right) \tag{30}$$

is not constant over the hologram image in the aperture plane. This can therefore result in degraded optical fringe patterns as a function of aperture size. Equivalently there is a higher phase detection noise floor (less signal to noise ratio - SNR) which can result in significant wavefront RMS errors (say > $\lambda/35$) as measured with the HET system.

A software package TFVMHF02 has been developed to evaluate the fringe modulation (Eqs. 29,30) as a function of HET sampling time, translational or angular sinusoidal vibration amplitude as a function of frequency, and planar direction of motion with respect to the optical beam. As indicated on the previous page, aperture size effects can also be investigated. TFVMHF02 has been written, data sets defined and was being implemented/checked out on ARAA's PDP-11/24 computer system at AFWL at the end of the author's SFRP work period there.

TFVMHF02 should allow the investigation of meaningful amplitude levels for both mechanical and acoustic vibration source mechanisms. Here curves relating HET sampling time and vibration amplitude as a function of frequency can be validated via measurement experience with the HET system and additional accelerometer, etc. sensors. These validated results can be used as, e.g., to develop design curves specifying maximum structural dynamics-acceleration or g levels. This entails a mapping from acceleration levels in the structural dynamics/mechanical vibrations area to vibratory motion and hence relative phase optical beam modulations per se. The amplitude of these

modulations can be set at values which give acceptable 10 in. (i.e., large diameter) beam, wavefront sensor performance of the HET (or a comparable wavefront sensor) system. If necessary, corresponding acoustic sound pressure level design specification curves can also be developed.

V. CONCLUSIONS

As discussed in Section III, an investigation has been made of mechanical vibration/acoustic sources of the consistently large ($\lambda/35$) RMS deviation errors in 10 in. beam wavefront surfaces measured with the AFWL/ARAA heterodyne holographic interferometer, wavefront sensor (HET) system. Spectral decompositions of vibration level measurements made with seismometer sensors and independent spectral decomposition measurements made with the HET system of pitch, yaw modulations of HET (reference and object) optical beams — showed significant peaks at nominal frequencies of 30 and 40 Hz. Additional seismometer and HET beam modulation peak responses (albeit at somewhat lower levels) have also been seen at nominal frequencies of 50-52 Hz.

Simple undamped resonant frequency analyses of the relative motions with respect to the optical table of the

(a) Top mounted acoustic/wind screen structure

(b) Bottom mounted laser support bench structure

have been made using nominal mass, stiffness values. These analyses gave translational (x,y) and yaw angular motion resonant frequencies corresponding to (roll, pitch) and yaw optical beam modulations of (29.03, 40.81) Hz for the wind screen and (48.09, 39.88) Hz for the laser support table. These results plus the attachment of the wind screen to the top of the optical table indicate that the wind screen is a significant local source (or antenna/amplifier as opposed to absorber) of mechanical/acoustic vibratory motion.

Removal of the acoustic/wind screen enclosure resulted in reduction of the nominal 30 Hz pitch and 40 Hz yaw vibration level and

optical beam modulation components. However, large amplitude, low frequency random motion induced modulations caused quadrant errors ($>\lambda/4$) in (a) the two-point relative phase observed as a function of time on the HET system oscilloscope and (b) the wavefront sensor surface error results indicated in using the HET system detection plus A/D conversion hardware and Zernicke polynomial surface fitting software. These large amplitude, low frequency (0.0 to 2.5 Hz) random modulations were found to be caused by low level, random air movement and turbulence. The use of a system of floor to ceiling curtains and plastic sheeting ceiling as a near term/interim fix to this problem was found to improve HET system performance to the extent that 10 in. beam RMS wavefront surface errors of $<\lambda/60$ could be obtained.

Four supporting studies were also performed as indicated in Section IV. These were:

(1) Resonant frequency analysis for linearized rigid body motion of HET optical table system.

(2) 12 state variable, time-domain analysis of the nonlinear rigid body motion of HET optical table, etc. system designs.

(3) Finite element method, etc. analysis of flexible motion response of optical table system designs.

(4) Impact of mechanical vibration/acoustic motion on relative phase modulation of HET system optical beams and hence on fringe modulation.

Limited work in supporting studies (1), (2) above gave good rigid body motion resonant frequency and time-domain predictions as compared to a limited amount of sensor time function measurements. These

good comparison results indicate that the 12 state-variable, nonlinear time-domain model and the linearized resonant frequency model, when extended to include such effects as vibration isolation support leg damping and realistic nonsymmetric mass distributions can be used to investigate the number, placement of the support legs to minimize the effect on the rigid body motion of the HET optical table system.

The limited results and formulations presented in supporting studies (3), (4) indicate that computer model analytical calculation results (and corresponding modal analysis measurements) can be used to directly investigate the impact of HET system vibratory motion upon its relative phase modulation and aperture-plane fringe modulation response behavior. Therefore, these computer model analytical, etc. studies can be used to optimize or at least improve upon the design of HET and similar optical, large-beam (as, e.g., > 10 in.) wavefront sensor systems.

VI. RECOMMENDATIONS

Based upon the results of the research project reported herein the following recommendations are made:

(1) That ARAA construct and use an unattached "rigid" member wind screen enclosure (4 sides and top) for the HET optical table and components mounted to its top surface. This enclosure can be easily made from a 1 in. steel channel framework plus sliding mylar sheeting sides (≈ 1/4 in. thick) and a fixed mylar top or ceiling member. This structure should be separately supported on its own set of vibration isolation-rubber or similar material pads. It should minimize the effect of random air movements, etc. on HET wavefront/sensor performance.

(2) That usage be investigated of additional local wind screen enclosures at the location of the large-beam expander and mirror, etc. optical components as well as at the table location where the laser beam enters the HET optical system.

(3) Consideration should also be given to the design and use of off-the-shelf components to provide a positive pressure, clean room system environment in the interior of the large wind screen enclosure surrounding the HET optical table.

(4) ARAA should investigate employing separate mounting of the driving laser and the use of fibre-optic or comparable means for light input to the HET system. Alternatively, placement of the laser on top of the optical table should be considered.

(5) An investigation should be made of the optimization/improvement of water input for laser cooling to reduce low frequency

vibration inputs to HET system.

(6) ARAA should consider the design, construction and use of a 1 or 2 mass plus damper, spring elements tuned vibration absorber (yielding 10 to 30+ dB notch at specified design frequency) to reduce the still high 30 Hz, etc. line vertical motion component observed at the top of the HET optical table.

(7) That ARAA use ARLB personnel, equipment to perform modal analysis and acoustic sensor measurements. These measurements can be used to define present optical component vibration mode sensitivities and to suggest changes/perturbations to the HET system resulting in improved wavefront sensor performance. This modal analysis and acoustic sensor work can also be used to head-off potential future wavefront sensor problems relevant the calibration, testing and use of flexible mirror and subaperture-to-full wavefront-reconstruction components.

(8) That the finite element and vibratory/acoustic motion analyses discussed in Sections IV.3, IV.4 be used to (a) Examine in detail the impact of sampling time, vibration amplitude as a function of frequency, direction of vibratory motion and aperture size-upon the relative phase modulation of HET optical beams and upon the HET hologram fringe modulation function. (b) Optimize or improve HET system designs as per selection of the number and placement of the vibration isolation support legs, the table physical parameters and the table placement, etc. of HET optical components.

(9) That calibrated mechanical vibration and acoustic input sources be used to validate analytical predictions of the impact of vibratory motion (Recommendation (8) and Section IV.4) on the relative phase modulation between the reference, object optical beams and upon the fringe modulation/RMS wavefront sensor error performance obtained with the HET system.

REFERENCES

1. Baldwin, George C., An Introduction to Nonlinear Optics, Plenum, N. Y., 1969.

2. Born, Max and Wolf, Emil, Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light, Sixth (Corrected) Edition, Pergamon Press, N. Y., 1980.

3. Hecht, Eugene and Zajac, Alfred, Optics, Addison Wesley, Reading, Mass., 1974. ISBN 0-201-02835-2

4. Malacara, Daniel (Editor), Optical Shop Testing, John Wiley, N. Y., 1978.

5. Wyant, James C., Optical Testing, Optical Sciences Center, University of Arizona, Tucson, Arizona, June 1976.

6. Verdeyen, Joseph T., Laser Electronics, Prentice-Hall, Englewood Cliffs, N. J., 1981.

7. Calendar, Al, "Optics Overview/Summary for Heterodyne Holographic Interferometer Wavefront Sensor System," Unpublished Notes, AFWL/ARAA. Spring 1983.

8. Heterodyne Set-Up Drawing, AFWL/ARAA, Spring 1983.

9. Calendar, Al, "Heterodyne Interferometer-Data Collection Electronics System," Unpublished Notes, AFWL/ARAA, Spring 1983.

10. Gardner, Floyd M., Phaselock Techniques, Second Edition, Wiley-Interscience, N. Y., 1979, Chapters 1-4,9.

11. Perina, Jan, Coherence of Light, Van Nostrand Reinhold, N. Y., 1971.

12. Steel, W. H., Interferometry, Cambridge University Press, England, 1967.

13. "Newport Corporation 1983-1984 Catalog," Newport Research Corporation, Fountain Valley, Calif., pp. 3-29.

14. Telephone conversations with Dennis Siegel, John Turechek, Newport Research Corporation — Relevant (a) Vibration Isolation Mount Support Legs; (b) Rigid Body and Flexural and Torsional Modes of Vibration of HET System Optical Table; (c) Design of Attached Acoustic/Wind Screen, June, July 1983.

15. "HS-10-1 Seismometer: Operational and Frequency Calibration Procedures," Version 1, Revision A, Geo Space Corporation, Houston, Tx., 19 January 1976.

16. "Instruction Manual for ENDEVCO Piezoresistive Accelerometers," Endevco Dynamic Instruments Division, Division of Becton, Dickinson and Company, Copyright 1968 ENDEVCO, Revised January 1972.

17. "PCB Piezotronics Catalog 375," PCB Piezotronics, Inc., Depew, N. Y., Summer 1983.

18. Bendat, Julius S. and Piersol, Allan G., Random Data: Analysis and Measurement Procedures, John Wiley-Interscience, N. Y., 1971.

19. Bendat, Julius S. and Piersol, Allan G., Engineering Applications of Correlation and Spectral Analysis, John Wiley-Interscience, N. Y., 1980.

20. Enochson, Loren D. and Otnes, Robert K., Programming and Analysis for Digital Time Series Data, Shock and Vibration Information Center, Naval Research Laboratory, United States Department of Defense, Washington, D. C., 1968.

21. Meirovitch, Leonard, Analytical Methods in Vibrations, MacMillan, N. Y., 1967, Chapters 7,9,10,11.

22. Meirovitch, Leonard, Elements of Vibration Analysis, McGraw-Hill, N. Y., 1975, Chapters 7,9,11.

23. Thomson, William T., Theory of Vibrations with Applications, Prentice-Hall, Englewood Cliffs, N. J., 1972, Chapters 5-8,10,11.

24. Timoshenko, S., Young, D. H., and Weaver, W., Jr., Vibration Problems in Engineering, Fourth Edition, John Wiley, N. Y., 1974.

25. Warburton, G. B., The Dynamical Behavior of Structures, Second Edition, Pergamon Press, N. Y., 1976.

26. Lipscombe, J. M., "Aerospace Systems," Chapter 5 of Modeling of Dynamical Systems, Vol. 1 (Edited by H. Nicholson), Peter Peregrinus Ltd. on Behalf of the Institution of Electrical Engineers, London and N. Y., 1980.

27. McRuer, D., Ashkenas, I., and Graham, D., Aircraft Dynamics and Automatic Control, Princeton University Press, Princeton, N. J., 1973.

28. Blevins, Robert D., Formulas for Natural Frequency and Mode Shape, Van Nostrand Reinhold, N. Y., 1979, Chapters 5,6,8,11.

29. Whaley, P. W. and Pearson, J, "Computer Aided Design of Passive Vibration Isolators for Airborne Optical Systems," Shock and Vibration Bulletin, No. 49, Pt. 2, September 1979, pp. 81-86.

30. Smith, Howard M., Principles of Holography, J. Wiley-Interscience, N. Y., 1969, Chapters 6,8

31. Goodman, J. W., "Temporal Filtering Properties of Holograms," Applied Optics, Volume 6, 857 (1967).

32. Powell, R. L. and Stetson, K. A., "Interferometric Vibration Analysis by Wavefront Reconstruction," Journal of the Optical Society of America, Volume 55, 1593 (1965).

33. Nitta, G. H., "Frequency-Modulated Signals and Their Power Spectra," Tech Memo No. TM-74-208, Naval Ordinance Laboratory-Corona, Corona, Calif., February 1964.

34. Gardner, Floyd M., Phaselock Techniques, Second Edition, John Wiley-Interscience, N. Y., 1979, Chapters 1-4,9,10.

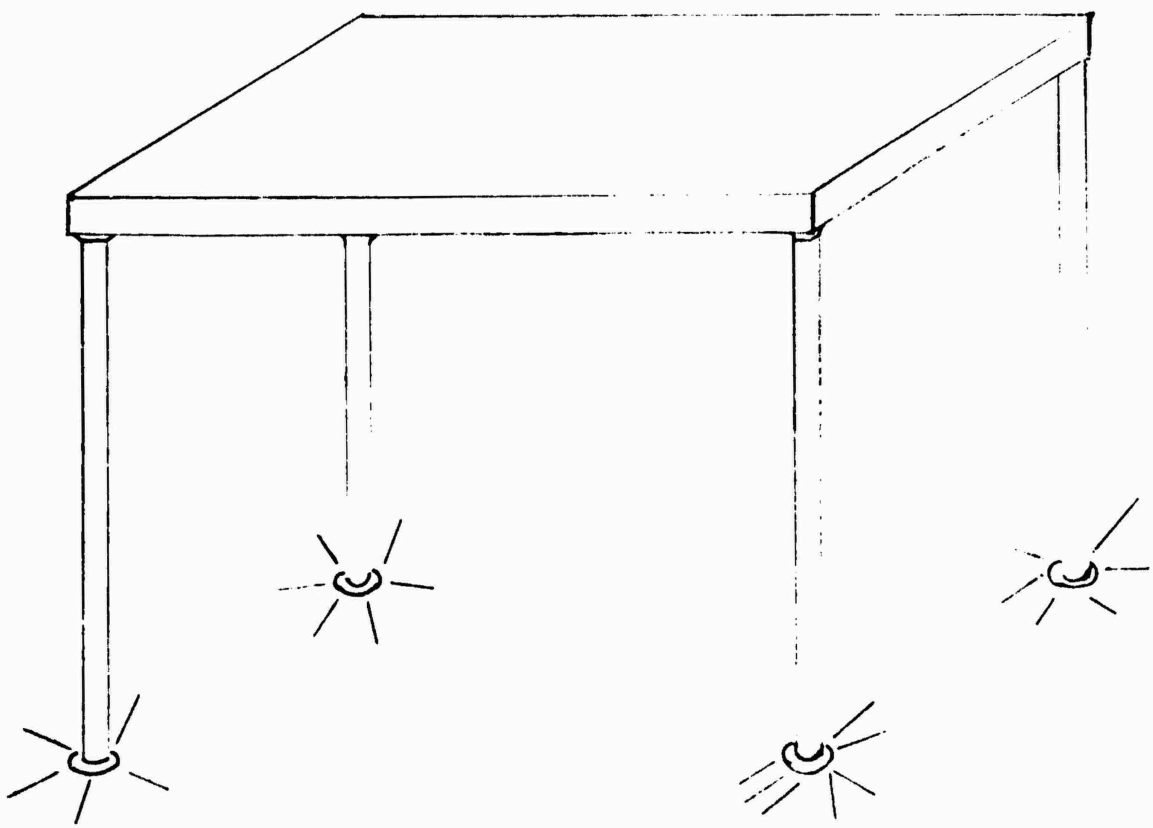
35. Lindsey, W. C., Synchronization Systems in Communications and Control, Prentice-Hall, Englewood Cliffs, N. J., 1972, Chapter 9.

36. Van Trees, Harry L., Detection Estimation and Modulation Theory, Part II: Nonlinear Modulation Theory, John Wiley, N. Y., 1971, Chapters 3,4.

APPENDIX A

Resonant Frequencies for x,y Translational and Yaw Rotational Motions of (1) Acoustic/Wind Screen and (2) Laser Support Bench Structures with Respect to HET Optical Table

FIG. A.1
ACOUSTIC/WIND SCREEN SUPPORT STRUCTURE



4 STEEL SUPPORT TUBES
50 in IN LENGTH
2.375 in OD

TOP IS NOMINAL (12 FT LONG, 4 FT WIDE, 2 9/16 in THICK)
ALUMINUM HONEYCOMB

(1) Acoustic/Wind Screen Structure Attached to Top of HET Optical Table

(a) x,y Translational Motion Undamped Natural Resonant Frequencies (correspond to pitch, roll vibratory modulation of HET optical beams).

$$K_{FLEX\ EQUIV} = 4\left(\frac{24EI}{\ell^3}\right) \tag{1}$$

$$f_{n_{(x,y)}} = \frac{1}{2\Pi}\left(\frac{K_{FLEX\ EQUIV}}{M}\right)^{1/2} \tag{2}$$

(b) Yaw Motion Undamped Natural Resonant Frequency (corresponds to yaw vibratory modulation of HET optical beams)

$$f_{n_{(yaw)}} = \frac{1}{2\Pi}\left(\frac{K_{FLEX\ EQUIV}}{M}\right)^{1/2}\left(\frac{3\ell}{(\ell_{x_{ROOF}}^2 + \ell_{y_{ROOF}}^2)^{1/2}}\right)^{1/2} \tag{3a}$$

$$= f_{n_{(x,y)}}\left(\frac{3\ell}{(\ell_{x_{ROOF}}^2 + \ell_{y_{ROOF}}^2)^{1/2}}\right)^{1/2} \tag{3b}$$

In Eqs. (1-3),

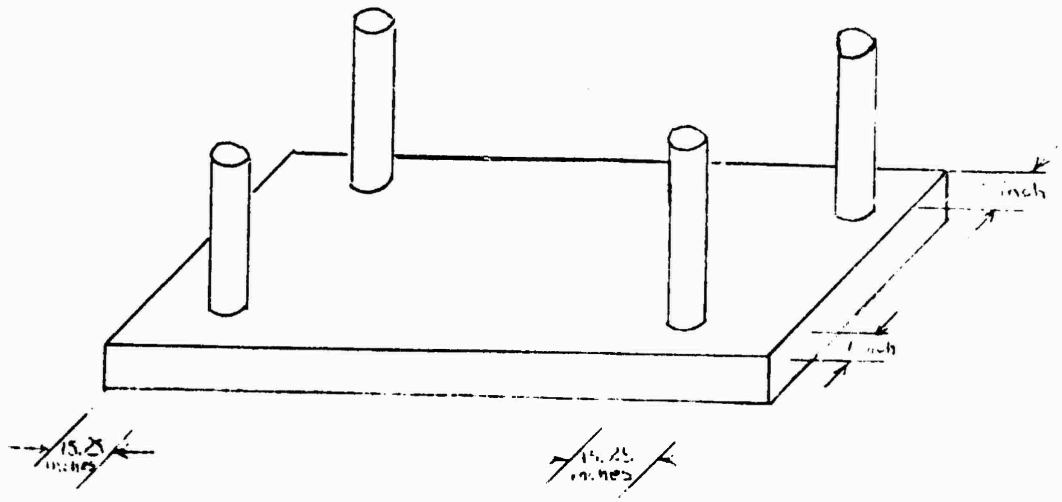
$$M = 150\ lb_m$$

$$E = 30 \times 10^6\ \frac{lb_f}{in^2}$$

$$I_{xx} = I_{yy} = \frac{\Pi(R_o^4 - R_i^4)}{4} = \frac{\Pi(D_o^4 - D_i^4)}{64}$$

FIG. A.2

LASER BENCH AND SUPPORT STRUCTURE



4 STEEL SUPPORT TUBES
14 in IN LENGTH
1.5 in OD

BOTTOM LEDGE IS NOMINAL
67.875 in LONG
14.9375 in WIDE
2.0 in THICK
CAST STEEL CONSTRUCTION

$(R_i, R_o) = (1.0625, 1.1875)$ in.

$\ell = 50$ in.

$(\ell_{x_{ROOF}}, \ell_{y_{ROOF}}) = (72,24)$ in. = one-half length, width of top member of acoustic/wind screen

Substituting into Eqs. (1-3) =>

$$f_{n_{(x,y)}} = 29.03 \text{ Hz} \tag{4a}$$

$$f_{n_{(yaw)}} = 40.81 \text{ Hz} \tag{4b}$$

(2) Laser Support Bench Structure Attached to Bottom of HET Optical Table

(a) x,y Translational Motion Undamped Natural Resonant Frequencies (correspond to pitch, roll vibratory modulation of HET optical beams)

$$K_{FLEX\ EQUIV} = 4\left(\frac{24EI}{\ell^3}\right) \tag{5}$$

$$f_{n_{(x,y)}} = \frac{1}{2\pi}\left(\frac{K_{FLEX\ EQUIV}}{M}\right)^{1/2} \tag{6}$$

(b) Yaw Motion Undamped Natural Resonant Frequency (corresponds to yaw vibratory modulation of HET optical beams)

$$f_{n_{(yaw)}} = \frac{1}{2\pi}\left(\frac{K_{FLEX\ EQUIV}}{M}\right)^{1/2}\left(\frac{3(\bar{\ell}_x^{-2} + \bar{\ell}_y^{-2})^{1/2}\ell}{(\ell_x^2 + \ell_y^2)}\right)^{1/2} \tag{7a}$$

$$= f_{n_{(x,y)}} \left(\frac{3(\bar{\ell}_x^2 + \bar{\ell}_y^2)^{1/2} \ell}{\ell_x^2 + \ell_y^2} \right)^{1/2} \qquad (7b)$$

In Eqs. (5-7),

$M = 571.2\ lb_m$

$E = 30 \times 10^6 \frac{lb_f}{in^2}$

$I_{xx} = I_{yy} = \frac{\pi}{4}\left(R_o^4 - R_i^4\right) = \frac{\pi(D_o^4 - D_i^4)}{64}$

$(R_i, R_o) = (0.625, 0.750)$ in.

$\ell = 14.0$ in.

$(\ell_x, \ell_y) = (33.94, 7.47)$ in. = One-half length, width of bottom ledge member of laser bench.

$(\bar{\ell}_x, \bar{\ell}_y) = (18.69, 6.47)$ in. = x,y components of distance vector from points of attachment of four support members to center of bottom ledge member of laser support bench.

Substituting into Eqs. (5-7) =>

$f_{n_{(x,y)}} = 48.09$ Hz

$f_{n_{(yaw)}} = 39.88$ Hz

APPENDIX B

Resonant Frequencies for Six-Degree-of-Freedom Rigid Body Motion of HET, Etc. Optical Table System

(A) x - Translation, Θ - Pitch Motion Analysis

$$\begin{bmatrix} \omega_{(x,\Theta)_1} \\ \omega_{(x,\Theta)_2} \end{bmatrix} = \left(\frac{\alpha_{11} + \alpha_{22}}{2}\right) \pm \left[\left(\frac{\alpha_{11} + \alpha_{22}}{2}\right)^2 - \alpha_{21}\alpha_{12}\right]^{1/2} \quad (1)$$

in which,

$$\alpha_{11} = \frac{K_{EQ(x)}}{M} \quad (2a)$$

$$\alpha_{12} = \left(\frac{K_{EQ(x)}\ell_z}{M}\right) \quad (2b)$$

$$\alpha_{21} = \left(\frac{K_{EQ(x)}\ell_z}{I_{yy}}\right) \quad (2c)$$

$$\alpha_{22} = \left(\frac{K_{EQ(x)}\ell_z^2 + K_{EQ(z)}\ell_x^2}{I_{yy}}\right) \quad (2d)$$

Here

$$K_{EQ(x)} = 6 \text{ Legs } \left(150 \frac{lb_f}{in}/\text{Leg}\right) = 900 \; lb_f/in$$

$$K_{EQ(z)} = 4 \text{ Legs } \left(64 \frac{lb_f}{in}/\text{Leg}\right) = 256 \; lb_f/in \; *$$

$$I_{yy} = \frac{1}{12} M(\ell_x^2 + \ell_z^2)4$$

*Note that only 4 as opposed to 6 legs supply pitch moment via vertical stiffness force at each leg. This assumes that two middle legs are in vertical symmetry plane of optical table.

(ℓ_x, ℓ_z) = (72.44,9) in. = One-half length, height of HET optical table.

The undamped natural resonant frequencies in Hz follow as

$$\begin{bmatrix} f_{n(x,\Theta)_1} \\ f_{n(x,\Theta)_2} \end{bmatrix} = \frac{1}{2\Pi} \begin{bmatrix} \omega_{n(x,\Theta)_1} \\ \omega_{n(x,\Theta)_2} \end{bmatrix} \tag{3}$$

Substituting into Eqs. (1-3) gives

$$\begin{bmatrix} f_{n(x,\Theta)_1} \\ f_{n(x,\Theta)_2} \end{bmatrix} = \begin{bmatrix} 0.27 \\ 2.34 \end{bmatrix} \text{ Hz} \tag{4}$$

(B) <u>y - Translation, ϕ - Roll Motion Analysis</u>

$$\begin{bmatrix} \omega_{(y,\phi)_1} \\ \omega_{(y,\phi)_2} \end{bmatrix} = \left(\frac{\alpha_{11} + \alpha_{22}}{2}\right) \pm \left[\left(\frac{\alpha_{11} + \alpha_{22}}{2}\right)^2 - \alpha_{21}\alpha_{12}\right]^{1/2} \tag{5}$$

in which,

$$\alpha_{11} = \frac{K_{EQ(y)}}{M} \tag{6a}$$

$$\alpha_{12} = \left(\frac{K_{EQ(y)}\ell_z}{M}\right) \tag{6b}$$

$$\alpha_{21} = \left(\frac{K_{EQ(y)}\ell_z}{I_{xx}}\right) \tag{6c}$$

$$\alpha_{22} = \left(\frac{K_{EQ(y)}\ell_z^2 + K_{EQ(z)}\ell_y^2}{I_{xx}}\right) \tag{6d}$$

and

$$K_{EQ(y)} = 6 \text{ Legs} \times \left(150 \frac{lb_f}{in}/\text{Leg}\right) = 900\ lb_f/in.$$

$$K_{EQ(z)} = 6 \text{ Legs} \times \left(64 \frac{lb_f}{in}/\text{Leg}\right) = 384\ lb_f/in.$$

$$M = 3000\ lb_m$$

$$I_{xx} = \frac{1}{12} M\ [\ell_y^2 + \ell_z^2]4$$

$(\ell_y, \ell_z) = (25.00, 9.00)$ in. = One-half width, height of HET optical table.

The undamped natural resonant frequencies in Hz can be written as

$$\begin{bmatrix} f_{n(y,\phi)_1} \\ f_{n(y,\phi)_2} \end{bmatrix} = \frac{1}{2\pi} \begin{bmatrix} \omega_{n(y,\phi)_1} \\ \omega_{n(y,\phi)_2} \end{bmatrix} \tag{7}$$

Substituting into Eqs. (5-7) yields

$$\begin{bmatrix} f_{n(y,\phi)_1} \\ f_{n(y,\phi)_2} \end{bmatrix} = \begin{bmatrix} 0.66 \\ 2.61 \end{bmatrix} \text{Hz} \tag{8}$$

(C) z - Translation/Vertical Motion Analysis

$$f_{n(z)} = \frac{1}{2\Pi}\,\omega_{n(z)} = \frac{1}{2\Pi}\left(\frac{K_{EQ(z)}}{M}\right)^{1/2} \tag{9}$$

in which,

$$K_{EQ(z)} = 6\ \text{Legs}\left(64\ \frac{lb_f}{in}/\text{Leg}\right) = 384\ \frac{lb_f}{in}$$

$$M = 3000\ lb_m$$

Substituting into Eq. (9) gives

$$f_{n(z)} = 1.12\ \text{Hz} \tag{10}$$

(D) Ψ - Yaw Motion Analysis

$$\omega_{n(\Psi)} = \left(\frac{4K_{(x,y)}(\ell_x^2 + \ell_y^2) + 2K_{(x,y)}\ell_y^2}{\frac{1}{12}M(4)(\ell_x^2 + \ell_y^2)}\right)^{1/2} \tag{11a}$$

$$= \left[\left(\frac{12K_{(x,y)}}{M}\right)\left(1 + \frac{\ell_y^2}{2(\ell_x^2 + \ell_y^2)}\right)\right]^{1/2} \tag{11b}$$

Here

$$K_{(x,y)} = 150\ \frac{lb_f}{in}$$

$$M = 3000\ lb_m$$

$(\ell_x, \ell_y) = (72.44, 25.00)$ in. = One-half length, width of HET optical table.

The undamped natural resonant frequency in Hz can be writen as

$$f_{n(\Psi)} = \frac{1}{2\Pi}\,\omega_{n(\Psi)} \tag{12}$$

Substituting into Eqs. (11,12) yields

$$f_{n(\Psi)} = 2.49 \text{ Hz} \tag{13}$$

(E) Summary of Resonant Frequency Results

Therefore, the undamped natural resonant frequencies for the rigid body motion of the HET optical table can be summarized as

$$\begin{bmatrix} f_{n(x,\theta)_1} \\ f_{n(x,\theta)_2} \\ f_{n(y,\phi)_1} \\ f_{n(y,\phi)_2} \\ f_{n(z)} \\ f_{n(\Psi)} \end{bmatrix} = \begin{bmatrix} 0.27 \\ 2.34 \\ 0.66 \\ 2.61 \\ 1.12 \\ 2.49 \end{bmatrix} \text{ Hz} \tag{14}$$

APPENDIX C

HETRBM09: Fortran IV Software Package for Time-Domain Modeling, Analysis of Rigid Body Motion of HET System Optical Table Designs

```
C     COMMENT: HETRBM09.FTN;1
C     DRIVER PROGRAM HETRBM - DRIVER PROGRAM FOR RIGID BODY
C     MODE - VIBRATION MODEL OF MAIN/SALIENT MECHANICAL/
C     STRUCTURAL ELEMENTS OF THE HETERODYNING-HOLOGRAPHIC-
C     INTERFEROMETER SYSTEM DESIGN.
C     ************************************************************
      EXTERNAL HTRBVM
      COMMON /LUNIO1/LUNI,LUNO,LUNO1
      COMMON /INITC1/XTOTRA(10,3)
      COMMON /MAIN1/NDIM
      COMMON /XPRINT/XPRDUM(50)
      COMMON XFREQ,XOTBL1,XOTBL2,XOTBL3,XMASST,RHOOTB,
     * XKSACT(10,3),XRSACT(10,3),XXPLRA(10,3),XXPLRP(10,3),
     * NSVACT,NSVDIR,INASSO,IGCOP
C     ************************************************************
      DIMENSION APDSEP(50)
      DIMENSION XTRIPB(3,3),XDUM1(3),XDUM2(3)
      REAL T,Y(12),TOUT,RELERR,ABSERR,XFREQ
      REAL TFINAL,TPRINT,WORK(75)
      INTEGER IWORK(5),IFLAG,NEQN
C     ************************************************************
  895 FORMAT(1H0)
  896 FORMAT(1H1)
  897 FORMAT(/)
  898 FORMAT(///)
  899 FORMAT(8E13.5)
  900 FORMAT(8E10.5)
  901 FORMAT(16I5)
  902 FORMAT(8E16.8)
C     ************************************************************
      LUNI = 1
      LUNO = 2
      LUNO1 = 3
      OPEN(UNIT = LUNI,TYPE = 'OLD',NAME = 'HETRBMIN.DAT',
     * ERR = 99)
C     ************************************************************
      NDIM1 = 10
      NDIM2 = 3
      NDIM3 = 1025
      NDIM = NDIM2
      NDIR = 3
      ZER = 0.0
C     ************************************************************
      CALL SDAPRT(APDSEP,1,20,1)
C     DEFINE INITIAL CONDITIONS FOR STATE VARIABLES.
      READ (LUNI,901)NSVEQN
      READ (LUNI,900)XCM10,XCM20,XCM30
      READ (LUNI,900)VCM10,VCM20,VCM30
      READ (LUNI,900)OMG10,OMG20,OMG30
      READ (LUNI,900)XPHI0,XTHT0,XPSI0
      READ (LUNI,900)XXCM10,XXCM20,XXCM30
```

```
      READ (LUNI,900) XXPHI0,XXTHT0,XXPSI0
C     INPUT/DEFINE OPTICAL TABLE SIZE AND MASS PROPERTIES.
      READ (LUNI,900) XOTBL1,XOTBL2,XOTBL3
      READ (LUNI,901) IMASSO,IGCOPT
      READ (LUNI,900) XMASST,RHOOTB
C     INPUT/DEFINE LINEAR SPRING CONSTANTS, DISTANCES FROM CENTER
C     REFERENCE POINT FOR 6 NRC VIBRATION ISOLATION MOUNT LEGS.
      READ (LUNI,901) NSVACT,NSVDIR
      CALL RDPRNT(XKSACT,NSVACT,NSVDIR,NDIM1,1)
      CALL RDPRNT(XRSACT,NSVACT,NSVDIR,NDIM1,1)
      CALL RDPRNT(XXFLRA,NSVACT,NSVDIR,NDIM1,1)
      CALL RDPRNT(XXFLRP,NSVACT,NSVDIR,NDIM1,1)
C     NSVACT = NUMBER OF VIBRATION ISOLATION SUPPORT LEGS.
C     NSVDIR = NUMBER OF DIRECTIONS/DOF FOR TRANSLATIONAL
C              SPRING CONSTANTS FOR VIBRATION ISOLATION LEGS.
C     ******************************************************************
      READ (LUNI,900) XFREQ
C     XFREQ = EXCITATION FREQUENCY AT THE FLOOR.
C     ******************************************************************
      WRITE (LUNO,895)
      CALL RDAPRT(APDSEP,1,20,2)
      WRITE (LUNO,898)
      WRITE (LUNO,901) NSVEQN
      WRITE (LUNO,898)
      WRITE (LUNO,899) XCM10,XCM20,XCM30
      WRITE (LUNO,897)
      WRITE (LUNO,899) YCM10,YCM20,YCM30
      WRITE (LUNO,897)
      WRITE (LUNO,899) OMG10,OMG20,OMG30
      WRITE (LUNO,897)
      WRITE (LUNO,899) XPHI0,XTHT0,XPSI0
      WRITE (LUNO,898)
      CALL RDAPRT(APDSEP,1,20,2)
      WRITE (LUNO,898)
      WRITE (LUNO,899)  XXCM10,XXCM20,XXCM30
      WRITE (LUNO,897)
      WRITE (LUNO,899)  XXPHI0,XXTHT0,XXPSI0
      WRITE (LUNO,898)
      CALL RDAPRT(APDSEP,1,20,2)
      WRITE (LUNO,897)
      WRITE (LUNO,899) XOTBL1,XOTBL2,XOTBL3
      WRITE (LUNO,897)
      WRITE (LUNO,901) IMASSO,IGCOPT
      WRITE (LUNO,897)
      WRITE (LUNO,899) XMASST,RHOOTB
      WRITE (LUNO,898)
      CALL RDAPRT(APDSEP,1,20,2)
      WRITE (LUNO,897)
      WRITE (LUNO,901) NSVACT,NSVDIR
      WRITE (LUNO,897)
      CALL RDPRNT(XKSACT,NSVACT,NSVDIR,NDIM1,2)
      WRITE (LUNO,897)
      CALL RDPRNT(XRSACT,NSVACT,NSVDIR,NDIM1,2)
      WRITE (LUNO,898)
```

```
      CALL RDPRNT(XXFLRA,NSVACT,NSVDIR,NDIM1,2)
      WRITE (LUNO,897)
      CALL RDPRNT(XXFLRP,NSVACT,NSVDIR,NDIM1,2)
      WRITE (LUNO,898)
      CALL RDAPRT(APDSEP,1,20,2)
      WRITE (LUNO,899) XFREQ
      WRITE (LUNO,898)
      WRITE (LUNO,895)
C     ****************************************************************
C     SET INITIAL CONDITIONS.
      Y(1) = XCM10 + XXCM10
      Y(2) = XCM20 + XXCM20
      Y(3) = XCM30 + XXCM30
      Y(4) = VCM10
      Y(5) = VCM20
      Y(6) = VCM30
      Y(7) = OMG10
      Y(8) = OMG20
      Y(9) = OMG30
      Y(10) = XPHI0 + XXPHI0
      Y(11) = XTHT0 + XXTHT0
      Y(12) = XPSI0 + XXPSI0
C     DEFINE INITIAL OR SET POINT DISPLACEMENTS FOR OPTICAL
C     TABLE AT VIBRATION ISOLATION MOUNT LOCATIONS.
      CALL TRXFXB(XTRXFB,XPHI0,XTHT0,XPSI0,NDIR,NDIM2)
      CALL MTRINT(XTOTRA,10,3,NDIM1,ZER)
      DO 1I1=1,NSVACT
      DO 2I2=1,NSVDIR
      XDUM1(I2) = XRSACT(I1,I2)
    2 CONTINUE
      CALL MATPRD(XTRXFB,XDUM1,XDUM2,3,3,NDIM2)
      XTOTRA(I1,1) = XCM10 + XDUM1(1)
      XTOTRA(I1,2) = XCM20 + XDUM1(2)
      XTOTRA(I1,3) = XCM30 + XDUM1(3)
    1 CONTINUE
C     ****************************************************************
C     DEFINE NUMERICAL INTEGRATION, ETC. PARAMETERS.
      RELERR = 0.5E-03
      ABSERR = 0.0
      TFINAL = 20.0
      TPRINT = 0.1
C     INITIALIZE INDEPENDENT VARIABLE TIME.
      T = 0.0
      TOUT = T
      IFLAG = 1
      NEQN = NSVEQN
C     ****************************************************************
      IPRINT = 1
   10 CALL RKF45 (NTRSVS,NEQN,Y,T,TOUT,RELERR,ABSERR,IFLAG,
     * WORK,IWORK)
      WRITE (LUNO,899) T,(Y(I),I=1,NEQN)
      IPRINT = IPRINT + 1
       GO TO (80,20,30,40,50,60,70,80), IFLAG
   20 TOUT = T + TPRINT
```

```
      IF (T .LT. TFINAL) GO TO 10
      GO TO 98
   30 WRITE(LUNO,31) RELERR,ABSERR
      GO TO 10
   40 WRITE(LUNO,41)
      GO TO 10
   50 ABSERR = 1.0E-9
      WRITE(LUNO,31) RELERR,ABSERR
      GO TO 10
   60 RELERR = 10.0*RELERR
      WRITE(LUNO,31) RELERR,ABSERR
      IFLAG = 2
      GO TO 10
   70 WRITE(LUNO,71)
      IFLAG = 2
      GO TO 10
   80 WRITE(LUNO,81)
   98 CONTINUE
C
      CLOSE(UNIT = LUNI,DISP = 'KEEP',ERR = 99)
   99 CONTINUE
      STOP
C
   11 FORMAT(F5.1, 2F15.9)
   31 FORMAT(17H TOLERANCES RESET, 2E12.3)
   41 FORMAT(11H MANY STEPS)
   71 FORMAT(12H MUCH OUTPUT)
   81 FORMAT(14H IMPROPER CALL)
      END
```

```
      SUBROUTINE HTRBVH(T,Y,YP)
      REAL T,Y(12),YP(12)
      REAL HSXTOT,HSYTOT,HSZTOT
      COMMON XFREQ,XOTBL1,XOTBL2,XOTBL3,XMASST,RHOOTB,
     * XKSACT(10,3),XRSACT(10,3),XXFLRA(10,3),XXFLRP(10,3),
     * NSVACT,NSVDIR,IMASSO,IGCOPT
      COMMON /LUNIO1/LUNI,LUNO,LUNO1
      COMMON /INITC1/XTOTRA(10,3)
      COMMON /MAIN1/NDIM
      COMMON /XPRINT/XPRDUM(50)
C     ******************************************************************
      DIMENSION XIXIXJ(3,3),XTRXFB(3,3),XTRXBF(3,3),ATRXFB(3,3),
     * XDUM(3),XXDUM(3),XXXDUM(3),OMGHTR(3,3),XBRIIJ(3,3)
      DIMENSION FVACTV(10,3),XVACTV(10,3)
C     ******************************************************************
      NDIM1 = 10
      NDIM2 = 3
      NDIR = 3
      NDIM = NDIM2
      NEQN = 12
      ACCGRV = 32.174
C     ******************************************************************
      CALL IXIJRP(XOTBL1,XOTBL2,XOTBL3,RHOOTB,XMASST,XIXIXJ,NDIM2,
     * IMASSO,IGCOPT)
      CALL TRXBXF(XTRXBF,Y(10),Y(11),Y(12),NDIR,NDIM2)
      CALL TRXFXB(XTRXFB,Y(10),Y(11),Y(12),NDIR,NDIM2)
      CALL ANGVTR(ATRXFB,Y(10),Y(11),Y(12),NDIR,NDIM2)
      XDUM(1) = Y(4)
      XDUM(2) = Y(5)
      XDUM(3) = Y(6)
      CALL MATPRD(XTRXFB,XDUM,XXDUM,3,3,NDIM2)
      YP(1) = XXDUM(1)
      YP(2) = XXDUM(2)
      YP(3) = XXDUM(3)
C     ******************************************************************
C     DEFINITION OF FSX(IB)TOT.
      IWTOPT = 1
      IWTOPT = 2
      CNSWTO = 1.0
      IF(IWTOPT.EQ.2)CNSWTO = 0.0
      XDUM(1) = 0.0
      XDUM(2) = 0.0
      XDUM(3) = -XMASST*ACCGRV*CNSWTO
      CALL MATPRD(XTRXBF,XDUM,XXDUM,3,3,NDIM2)
      FSXTOT = XXDUM(1)
      FSYTOT = XXDUM(2)
      FSZTOT = XXDUM(3)
      DO 1I1=1,NSVACT
      DO 2I2=1,NSVDIR
      XDUM(I2) = XJPIRP(XXFLRA(I1,I2),XXFLRP(I1,I2),XFREQ,T)
      XXDUM(I2) = XRSACT(I1,I2)
    2 CONTINUE
```

```
      CALL MATPRD(XTRXFB,XXDUM,XXXDUM,3,3,NDIM2)
      DO 3I3=1,NSVDIR
      XVACTV(I1,I3) = (Y(I3) + XXXDUM(I3) - XTOTRA(I1,I3)) -
     * XDUM(I3)
      FVACTV(I1,I3) = -XKSACT(I1,I3)*XVACTV(I1,I3)
    3 CONTINUE
      DO 4I4=1,NSVDIR
      XDUM(I4) = FVACTV(I1,I4)
    4 CONTINUE
      CALL MATPRD(XTRXBF,XDUM,XXDUM,3,3,NDIM2)
      DO 5I5=1,NSVDIR
      FVACTV(I1,I5) = XXDUM(I5)
    5 CONTINUE
      FSXTOT = FSXTOT + FVACTV(I1,1)
      FSYTOT = FSYTOT + FVACTV(I1,2)
      FSZTOT = FSZTOT + FVACTV(I1,3)
    1 CONTINUE
C     ***************************************************************
C     DEFINE OMGMTR 'SPINOR' MATRIX IN TERMS OF (Y(7),Y(8),Y(9))
C           = (R,P,Q) = 'ROLL, PITCH, YAW RATES'.
      CALL OMGDEF(OMGMTR,Y(7),Y(8),Y(9),NDIM2)
      XDUM(1) = Y(4)
      XDUM(2) = Y(5)
      XDUM(6) = Y(6)
      CALL MATPRD(OMGMTR,XDUM,XXDUM,3,3,NDIM2)
      YP(4) = XXDUM(1) + FSXTOT/XMASST
      YP(5) = XXDUM(2) + FSYTOT/XMASST
      YP(6) = XXDUM(3) + FSZTOT/XMASST
      CALL GMINV(3,3,XIXIXJ,XBRIIJ,MTRANK,0)
      IF(MTRANK-3)11,12,11
   11 CONTINUE
      WRITE (LUNO,903)MTRANK
  903 FORMAT(5X,'STOP NO. 11 IN MTRBVM, MTRANK = ',I5)
      STOP
   12 CONTINUE
      XDUM(1) = Y(7)
      XDUM(2) = Y(8)
      XDUM(3) = Y(9)
      CALL MATPRD(XIXIXJ,XDUM,XXDUM,3,3,NDIM2)
      CALL MATPRD(OMGMTR,XXDUM,XDUM,3,3,NDIM2)
C     ***************************************************************
C     DEFINITION OF MSX(IB)TOT = TOTAL 'ROLL, PITCH, YAW' -
C     (R,P,Q) MOMENTS.
      MSXTOT = XDUM(1)
      MSYTOT = XDUM(2)
      MSZTOT = XDUM(3)
      DO 6I6=1,NSVACT
      DO 7I7=1,NSVDIR
      XDUM(I7) = XMSACT(I6,I7)
      XXDUM(I7) = FVACTV(I6,I7)
    7 CONTINUE
      CALL VARAXB(XDUM,XXDUM,XXXDUM,NSVDIR)
      MSXTOT = MSXTOT + XXXDUM(1)
      MSYTOT = MSYTOT + XXXDUM(2)
```

```
      MSZTOT = MSZTOT + XXDUM(3)
    6 CONTINUE
      XDUM(1) = MSXTOT
      XDUM(2) = MSYTOT
      XDUM(3) = MSZTOT
      CALL MATPRD(XBRIIJ,XDUM,XXDUM,3,3,NDIM2)
      YP(7) = + XXDUM(1)
      YP(8) = + XXDUM(2)
      YP(9) = + XXDUM(3)
C     ************************************************************************
      XDUM(1) = Y(7)
      XDUM(2) = Y(8)
      XDUM(3) = Y(9)
      CALL MATPRD(ATRXPB,XDUM,XXDUM,3,3,NDIM2)
      YP(10) = XXDUM(1)
      YP(11) = XXDUM(2)
      YP(12) = XXDUM(3)
      DO 9I9=1,NEQN
      XPRDUM(I9) = YP(I9)
    9 CONTINUE
      RETURN
      END
```

```
      SUBROUTINE RKF45(F,NEQN,Y,T,TOUT,RELERR,ABSERR,IFLAG,WORK,
     1            IWORK)
C
C     FEHLBERG FOURTH-FIFTH ORDER RUNGE-KUTTA METHOD
C
C       WRITTEN BY H.A.WATTS AND L.F.SHAMPINE
C                  SANDIA LABORATORIES
C                 ALBUQUERQUE,NEW MEXICO
C
C     RKF45 IS PRIMARILY DESIGNED TO SOLVE NON-STIFF AND MILDLY
C     STIFF DIFFERENTIAL EQUATIONS WHEN DERIVATIVE EVALUATIONS
C     ARE INEXPENSIVE. RKF45 SHOULD GENERALLY NOT BE USED IS
C     DEMANDING HIGH ACCURACY.
C
C     ABSTRACT
C
C     SUBROUTINE RKS45 INTEGRATES A SYSTEM OF NEQN FIRST ORDER
C     ORDINARY DIFFERENTIAL EQUATIONS OF THE FORM
C           DY(1)/DT = F(T,Y(1),Y(2),.....,Y(NEQN))
C            WHERE THE Y(1) ARE GIVEN AT T.
C     TYPICALLY THE SUBROUTINE IS USED TO INTEGRATE FROM T TO
C     TOUT BUT IT CAN BE USED AS A ONE-STEP INTEGRATOR TO
C     ADVANCE THE SOLUTION A SINGLE STEP IN THE DIRECTION OF TOUT.
C     ON RETURN THE PARAMETERS IN THE CALL LIST ARE SET FOR
C     CONTINUING THE INTEGRATION.  THE USER HAS ONLY TO CALL THE
C     RKFS AGAIN (AND PERHAPS DEFINE A NEW VALUE FOR TOUT).
C     ACTUALLY, RKFS IS AN INTEGRATING ROUTINE WHICH CALLS
C     SUBROUTINE RKFS FOR THE SOLUTION.  RKFS IN TURN CALLS THE
C     SUBROUTINE FEHL WHICH COMPUTES AN APPROXIMATE SOLUTION OVER
C     ONE STEP.
C     RKF45 USES THE RUNGE-KUTTA-FEHLBERG (4,5) METHOD DESCRIBED IN
C     THE REFERENCE.
C     E.FEHLBERG , LOW-ORDER CLASSICAL RUNGE-KUTTA FORMULAS WITH
C     STEPSIZE CONTROL , NASA TR R-315
C
C     THE PERFORMANCE OF RKFS IS ILLUSTRATED IN THE REFERENCE
C     L.F.SHAMPINE,H.A.WATTS,S.DAVENPORT, SOLVING NON-STIFF ORDINARY
C           DIFFERENTIAL EQUATIONS-THE STATE OF THE ART.
C           SANDIA LABORATORIES REPORT SAND75-0182.
C           TO APPEAR IN SIAM REVIEW.
C
C
C     THE PARMETERS REPRESENT-
C       F -- SUBROUTINE F(T,Y,YP) TO EVALUATE DERIVATIVES
C       YP(1)=DY(1)/DT
C     NEQN - NUMBER OF EQUATIONS TO BE INTEGRATED
C     Y(*) - SOLUTION VECTOR AT T
C     T - INDEPENDENT VARIABLE
C     TOUT - OUTPUT AT WHICH SOLUTION IS DESIRED
C     RELERR,ABSERR - RELATIVE AND ABSOLUTE ERROR TOLERANCES FOR
C     LOCAL ERROR TEST.  AT EACH STEP THE CODE REQUIRES THAT
```

```
C         ABS(LOCAL ERROR) .LE. RELERR*ABS(Y) + ABSERR
C       FOR EACH COMPONENT OF THE LOCAL ERROR AND SOLUTION VECTORS
C       IFLAG - INDICATOR FOR STATUS OF INTEGRATION
C       WORK(*) - ARRAY TO HOLD INFORMATION INTERNAL TO RKF45 WHICH
C       IS NECESSARY FOR SUBSEQUENT CALLS. MUST BE DIMENSIONED AT
C       LEAST 3+6*NEQN
C       IWORK(*) -- INTEGRATE ARRAY USED TO HOLD INFORMATION INTERNAL
C        RKF45 WHICH IS NECESSARY FOR SUBSEQUENT CALLS. MUST BE
C       DIMENSIONED AT LEAST 5
C
C
C       FIRST CALL TO RKF45
C
C       THE USER MUST PROVIDE STORAGE IN HIS CALLING PROGRAM FOR THE
C       ARRAY IN THE CALL LIST.    Y(NEQN) . WORK(3+6*NEQN). IWORK(5).
C       DECLARE F IN AN EXTERNAL STATEMENT, SUPPLY SUBROUTINE
C        F(T,Y,YP)
C       AND INITIALIZE THE FOLLOWING PARAMETERS-
C
C        NEQN-NUMBERS OF EQUATIONS TO BE INTEGRATED.  (NEQN .GE. 1)
C       Y(*) - VECTOR OF INITIAL EQUATIONS
C       T - STARTING POINT OF INTEGRATION, MUST BE A VARIABLE
C       TOUT - OUTPUT POINT AT WHICH SOLUTIONIS DESIRED.
C        T = TOUT IS ALLOWED ON THE FIRST CALL ONLY, IN WHICH CASE
C       RKF45 RETURNS WITH IFLAG=2 IF CONTINUATION IS POSSIBLE.
C       RELERR,ABSERR - RELATIVE
C        WHICH MUST BE NON-NEGATIVE. RELERR MUST BE A VARIABLE WHILE
C        ABSERR MAY BE A CONSTANT. THE CODE SHOULD NORMALLY NOT BE
C        USED WITH RELATIVE ERROR CONTROL SMALLER THAN ABOUT 1.E-8.
C        TO AVOID LIMITING PRECISION DIFFICULTIES THE CODE REQUIRES
C        RELERR TO BE LARGER THAN AN INTERNALLY COMPUTED RELATIVE
C        ERROR PARAMETER WHICH IS MACHINE DEPENDENT, IN PARTICULAR,
C        PURE ABSOLUTE ERROR IS NOT PERMITTED. IF A SMALLER THAN
C        ALLOWABLE VALUE OF RELERR IS ATTEMPTED, RKF45 INCREASES
C       RELERR APPROPRIATELY AND RETURNS CONTROL TO THE USER BEFORE
C        CONTINUING THE INTEGRATION.
C       IFLAG - +1,-1 INDICATOR TO INITIALIZE THE CODE FOR EACH NEW
C        PROBLEM. NORMAL INPUT IS +1. THE USER SHOULD SET IFLAG=-1
C        ONLY WHEN ONE-STEP INTEGRATOR CONTROL IS ESSENTIAL IN THIS
C        CASE
C       RKF45 ATTEMPTS TO ADVANCE THE SOLUTION A SINGLE STEP IN THE
C       DIRECTION OF TOUT EACH TIME IT IS CALLED. SINCE THIS MODE
C       OF OPERATION RESULTS IN EXTRA COMPUTING OVERHEAD.  IT SHOULD
C       BE AVOIDED UNLESS NEEDED.
C
C
C       OUTPUT FROM RKF45
C
C       Y(*) - SOLUTION AT T
C       T - LAST POINT REACHED IN INTEGRATION.
C       IFLAG = 2 - INTEGRATION REACHED TOUT. INDICATES SUCCESSFUL
C       RETURN AND IS THE NORMAL MODE FOR CONTINUING INTEGRATION.
C           =2 - A SINGLE SUCCESSFUL STEP IN THE DIRECTION OF TOUT
C                    HAS BEEN TAKEN. NORMAL MODE FOR CONTINUING
```

```
C                 INTEGRATION ONE STEP AT A TIME.
C            = 3 - INTEGRATION WAS NOT COMPLETED BECAUSE RELATIVE
C                 ERROR TOLERANCE WAS TOO SMALL. RELERR HAS BEEN
C              APPROPRIATELY FOR CONTINUING.
C            = 4 - INTEGRATION WAS NOT COMPLETED BECAUSE MORE THAN
C                 3000 DERIVATE EVALUATIONS WERE NEEDED.  THIS IS
C                 APPROXIMATELY 500 STEPS.
C            = 5 - INTEGRATION WAS NOT COMPLETED BECAUSE SOLUTION
C                 VANISHED MAKING A PURE RELATIVE ERROR TEST
C                  IMPOSSIBLE.
C                 MUST USE NON-ZERO ABSERR TO CONTINUE USING THE
C                 ONE-STEP INTEGRATION MODE FOR ONE-STEP IS A GOOD
C                 WAY TO PROCEED.
C            = 6 - INTEGRATION WAS NOT COMPLETED BECAUSE REQUESTED
C                 ACCURACY COULD NOT BE ACHIEVED USING SMALLEST
C                 ALLOWABLE STEPSIZE. USER MUST INCREASE THE ERROR
C                 TOLERANCE BEFORE CONTINUED INTEGRATION CAN BE
C                 ATTEMPTED.
C            = 7 - IT IS LIKELY THAT RKF45 IS INEFFICIENT FOR
C                 SOLVING THIS PROBLEM. TOO MUCH OUTPUT IS
C                 RESTRICTING
C                 THE NATURAL STEPSIZE CHOICE. USE THE ONE-STEP
C                 INTEGRATOR MODE.
C            = 8 - INVALID INPUT PARAMETERS, THIS INDICATOR OCCURS
C                 IF ANY OF THE FOLLOWING IS SATISFIED - NEQN.LE.0
C                        T= TOUT AND IFLAG.NE.+1 OR -1
C                        RELERR OR ABSERR.LT. 0.
C                        IFLAG.EQ.0 OR .LT. -2 OR .GT. 8
C      WORK(*),IWORK(*)--INFORMATION WHICH IS USUALLY OF NO INTEREST
C               TO THE USER BUT NECESSARY FOR SUBSEQUENT CALLS.
C               WORK(1),....,WORK(NEQN) CONTAIN THE FIRST DERIVATES
C               OF THE SOLUTION VECTOR Y AT T. WORK(NEQN+1) CONTAINS
C               THE STEPSIZE H TO BE ATTEMPTED ON THE NEXT STEP.
C               IWORK(1) CONTAINS THE DERIVATE EVALUATION COUNTER.
C
C
C      SUBSEQUENT CALLS TO RKF45
C
C       SUBROUTINE RKF45 RETURNS WITH ALL INFORMATION NEEDED TO
C       CONTINUE THE INTEGRATION. IF THE INTEGRATION REACHED TOUT,
C       THE USER NEED ONLY DEFINE A NEW TOUT AND CALL RKF45 AGAIN.
C       IN THE ONE-STEP INTEGRATOR MODE (IFLAG=-2) THE USER MUST
C       KEEP IN MIND THAT EACH STEP TAKEN IS IN THE DIRECTION
C       OF THE CURRENT TOUT. UPON REACHING TOUT (INDICATED BY
C       CHANGING IFLAG TO 2). THE USER MUST THEN DEFINE A NEW TOUT
C       AND RESET IFLAG TO -2 TO CONTINUE IN THE ONE-STEP
C       INTEGRATOR MODE.
C
C       IF THE INTEGRATION WAS NOT COMPLETED BUT THE USER STILL
C       WANTS TO CONTINUE (IFLAG=3,4 CASES), WE JUST CALLS RKF45
C       AGAIN. WITH IFLAG=3 THE RELERR PARAMETER HAS BEEN ADJUSTED
C       APPROPRIATELY FOR CONTINUING THE INTEGRATION. IN THE CASE
C       OF IFLAG=4 THE FUNCTION COUNTER WILL BE RESET TO 0 AND
C       ANOTHER 3000 FUNCTION EVALUATIONS ARE ALLOWED.
```

```
C
C       HOWEVER, IN THE CASE IFLAG=5, THE USER MUST FIRST ALTER THE
C        ERROR
C        CRITERION TO USE A POSITIVE VALUE OF ABSERR BEFORE
C        INTEGRATION
C        CAN PROCEED. IF HE DOES NOT, EXECUTION IS TERMINATED.
C
C       ALSO, IN THE CASE IFLAG=6, IT IS NECESSARY FOR THE USER TO
C        RESET IFLAG TO 2 (OR -2 WHEN THE ONE-STEP INTEGRATIPON MODE
C        IS BEING USED) AS WELL AS INCREASING EITHER ABSERR,RELERR OR
C        BOTH BEFORE THE INTEGRATION CAN BE CONTINUED. IF THIS IS NOT
C        DONE, EXECUTION WILL BE TERMINATED. THE OCCURANCE OF IFLAG=6
C        INDICATES A TROUBLE SPOT (SOLUTION IS CHANGING RAPIDLY,
C        SINGULARITY MAY BE PRESENT) AND IT OFTEN IS INAVISABLE TO
C        CONTINUE.
C
C       IFLAG=7 IS ENCOUNTERED, THE USER SHOULD USE THE ONE-STEP
C        INTEGRATION MODE WITH THE STEPSIZE DETERMINED BY THE CODE
C        OR CONSIDER SWITCHING TO THE ADAMS CODE DE/STEP.INTRP. IF
C        THE USER INSISTS UPON CONTINUING THE INTEGRATION WITH RKF45
C        HE MUST RESET IFLAG TO 2 BEFORE CALLING RKF45 AGAIN.
C        OTHERWISE, EXECUTION WILL BE TERMINATED.
C
C       IFLAG=8 IS OBTAINED, INTEGRATION CAN NOT BE CONTINUED UNLESS
C        THE INVALID INPUT PARAMETERS ARE CORRECTED.
C
C       IT SHOULD BE NOTED THAT THE ARRAYS WORK, IWORK CONTAIN
C        INFORMATION REQUIRED FOR SUBSEQUENT,ACCORDINGLY, WORK AND
C        IWORK SHOULD BE ALTERED.
C
C
      INTEGER NEQN,IFLAG,IWORK(5)
      REAL Y(NEQN),T,TOUT,RELERR,ABSERR,WORK(1)
C      IF COMPILER CHECKS SUBSCRIPTS, CHANGE WORK(1) TO WORK
C      (3+6*NEQN)

C
       EXTERNAL F
C
       INTEGER K1,K2,K3,K4,K5,K6,K1M
C
C
C      COMPUTE INDICATES FOR THE SPLITTING OF THE WORK ARRAY
C
       K1M=NEQN+1
       K1=K1M+1
       K2=K1+NEQN
       K3=K2+NEQN
       K4=K3+NEQN
       K5=K4+NEQN
       K6=K5+NEQN
C
C      THIS INTERFACING ROUTINE MERELY RELIEVES THE USER OF A
C      LONG CALLING LIST VIA THE SPLITTING APART OF TWO WORKING
```

```
C       STORAGE ARRAYS. IF THIS IS NOT COMPATIBLE WITH THE USER
C       COMPILER, HE MUST USE RKF45 DIRECTLY.
C
        CALL RKFS(F,NEQN,Y,T,TOUT,RELERR,ABSERR,IFLAG,WORK(1),
     1       WORK(K1M),WORK(K1),WORK(K2),WORK(K3),WORK(K4),WORK(K5),
     2       WORK(K6),WORK(K6+1),IWORK(1),IWORK(2),IWORK(3),
     3    IWORK(4),IWORK(5))
C
        RETURN
        END
```

```
      SUBROUTINE RKFS(F,NEQN,Y,T,TOUT,RELERR,ABSERR,IFLAG,YP,
     1            H,F1,F2,F3,F4,F5,SAVRE,SAVAE,NFE,KOP,INIT,JFLAG,
     2            KLAG)
C
C     FEHLBERG FOURTH-FIFTH ORDER RUNGE-KUTTA METHOD
C
C
C     RKFS INTEGRATES A SYSTEM OF FIRST ORDER ORDINARY DIFFERENTIAL
C     EQUATIONS AS DESCRIBED IN THE COMMENTS FOR RKF45.
C     THE ARRAYS YP,F1,F2,F3,F4,AND F5 (OF DIMENSION AT LEAST NEQN)
C     AND THE VARIABLES H,SAVRE,SAVAE,NFE,KOP,INIT,JFLAG,AND KFLAG
C     ARE USED INTERNALLY BY THE CODE AND APPEAR IN THE CALL LIST TO
C     ELIMINATE LOCAL RETENTION OF VARIABLES BETWEEN CALLS.
C     ACCORDINGLY, THEY SHOULD NOT BE ALTERED. ITEMS OF POSSIBLE
C     INTEREST ARE YP-DERIVATIVE OF SOLUTION VECTOR AT T
C           H-AN APPROPRIATE STEPSIZE TO BE USED FOR THE NEXT STEP
C      NFE-COUNTER ON THE NUMBER OF DERIVATIVE FUNCTION EVALUATIONS
C
      LOGICAL HFAILD,OUTPUT
C
      INTEGER NEQN,IFLAG,NFE,KOP,INIT,JFLAG,KFLAG
     REAL Y(NEQN),T,TOUT,RELERR,ABSERR,H,YP(NEQN),
     1  F1(NEQN),F2(NEQN),F3(NEQN),F4(NEQN),F5(NEQN),
     2   SAVRE,SAVAE
C
      EXTERNAL F
C
      REAL A,AE,DT,EE,EEOET,ESTTOL,ET,HMIN,REMIN,RER,S,SCALE,TOL,
     1  TOLN,U26,EPSP1,EPS,YPK
C
      INTEGER K,MAXNFE,HFLAG
C
C     REMIN IS THE MINIMUM ACCEPTABLE VALUE OF RELERR. ATTEMPTS
C     TO OBTAIN HIGHER ACCURACY WITH THIS SUBROUTINE ARE USUALLY
C     VERY EXPENSIVE AND OFTEN UNSUCCESSFUL.
C
      DATA REMIN/1.E-12/
C
C
C     THE EXPENSE IS CONTROLLED BY RESTRICTING THE NUMBER OF
C     FUNCTION EVALUATIONS TO BE APPROXIMATELY MAXNFE. AS SET
C     THIS CORRESPONDS TO ABOUT 500 STEPS.
C
      DATA MAXNFE/3000/
C
C
C     CHECK OUTPUT PARAMETERS
C
C
      IF (NEQN.LT.1) GO TO 10
      IF ((RELERR.LT.0.0) .OR. (ABSERR.LT.0.0)) GO TO 10
```

```
      MFLAG=IABS(IFLAG)
      IF ((MFLAG .EQ. 0) .OR. (MFLAG .GT. 8)) GO TO 10
      IF (MFLAG .NE. 1) GO TO 20
C
C     FIRST CALL, COMPUTE MACHINE EPSILON
C
      EPS = 1.0
    5 EPS = EPS/2.0
      EPSP1 = EPS + 1.0
      IF (EPSP1 .GT. 1.0) GO TO 5
      U26 = 26.0*EPS
      GO TO 50
C
C     INVALID INPUT
   10 IFLAG=8
      RETURN
C
C     CHECK CONTINUATION POSSIBILITIES
C
   20 IF ((T .EQ. TOUT) .AND. (KFLAG .NE. 3)) GO TO 10
      IF (MFLAG .NE. 2) GO TO 25
C
C     IFLAG = +2 OR -2
      IF ((KFLAG .EQ. 3) .OR. (INIT .EQ. 0)) GO TO 45
      IF (KFLAG .EQ. 4) GO TO 40
      IF ((KFLAG .EQ.    .AND. (ABSERR .EQ. 0.0)) GO TO 30
      IF ((KFLAG .EQ. 5) .AND. (RELERR .LE. SAVRE) .AND.
     1    (ABSERR .LE. SAVAE)) GO TO 30
      GO TO 50
C
C     IFLAG = 3,4,5,6,7 OR 8
   25 IF (IFLAG .EQ. 3) GO TO 45
      IF (IFLAG .EQ. 4) GO TO 40
      IF ((IFLAG .EQ. 5) .AND. (ABSERR .GT. 0.0)) GO TO 45
C
C
C     INTEGRATION CANNOT BE CONTINUED SINCE USER DID NOT
C      RESPOND TO THE INSTRUCTIONS PERTAINING TO IFLAG=5,6,7 OR 8
   30 STOP
C
C     RESET FUNCTION EVALUATION COUNTER
   40 NFE=0
      IF (MFLAG .EQ. 2) GO TO 50
C
C     RESET FLAG VALUE FROM PREVIOUS CALL
   45 IFLAG=JFLAG
      IF (KFLAG .EQ. 3) MFLAG=IABS(IFLAG)
C
C     SAVE INPUT IFLAG AND SET CONTINUATION FLAG FOR SUBSEQUENT
C     INPUT CHECKING
   50 JFLAG=IFLAG
      KFLAG=0
C
C     SAVE RELERR AND ABSERR FINPUT ON SUBSEQUENT
```

```
      SAVRE=RELERR
      SAVAE=ABSERR
C
C     RESTRICT RELATIVE ERROR TOLERANCE TO BE AT LEAST AS LARGE
C     AS 2*EPS+REMIN TO AVOID LIMITING PRECISION DIFFICULTIES
C     ARISING FROM IMPOSSIBLE ACCURACY REQUESTS
C
      RER=2.0*EPS+REMIN
      IF (RELERR .GE. RER) GO TO 55
C
C     RELATIVE ERROR TOLERANCE TOO SMALL
      RELERR=RER
      IFLAG=3
      KFLAG=3
      RETURN
C
   55 DT=TOUT-T
C
      IF (NFLAG .EQ. 1) GO TO 60
      IF (INIT .EQ. 0) GO TO 65
      GO TO 80
C
C     INITIALATION ---
C         SET INITIALIZATION COMPLETION INDICATOR,INIT
C
C         EVALUATE INITIAL DERIVATIVES
C         SET COUNTER FOR FUNCTION EVALUATIONS,NFE
C         ESTIMATE STARTING STEPSIZE
C
   60 INIT=0
      KOP=0
C
      A=T
      CALL F(A,Y,YP)
      NFE=1
      IF (T .NE. TOUT) GO TO 65
      IFLAG=2
      RETURN
C
C
   65 INIT=1
      H=ABS(DT)
      TOLN=0
      DO 70 K=1,NEQN
      TOL=RELERR*ABS(Y(K))+ABSERR
      IF (TOL .LE. 0) GO TO 70
      TOLN=TOL
      YPK=ABS(YP(K))
      IF (YPK*H**5 .GT. TOL) H=(TOL/YPK)**0.2
   70 CONTINUE
      IF (TOLN .LE. 0.0) H=0.0
      H=AMAX1(H,.026*AMAX1(ABS(T),ABS(DT)))
      JFLAG=ISIGN(2,IFLAG)
C
```

```
C
C       SET STEPSIZE FOR INTEGRATION IN THE DIRECTION FROM T TO
C        TOUT
C
   80 H=SIGN(H,DT)
C
C       TEST TO SEE IF RKF45 IS BEING SEVERELY IMPACTED BY TOO MANY
C       OUTPUT POINTS
C
        IF (ABS(H) .GE. 2.0*ABS(DT)) KOP=KOP+1
        IF (KOP .NE. 100) GO TO 85
C
C       UNNECESSARY FREQUENCY OUTPUT
        KOP=0
        IFLAG=7
        RETURN
C
   85 IF (ABS(DT) .GT. U26*ABS(T)) GO TO 95
C
C       IF TOO CLOSE TO OUTPUT POINT, EXTRAPOLATE AND RETURN
C
        DO 90 K=1,NEQN
   90 Y(K)=Y(K)+DT*YP(K)
        A=TOUT
        CALL F(A,Y,YP)
        NFE=NFE+1
        GO TO 300
C
C
C       INITIALIZE OUTPUTPOINT INDICATOR
C
   95 OUTPUT=.FALSE.
C
C       TO AVOID PREMATURE UNDERFLOW IN THE ERROR TOLERANCE
C        FUNCTION SCALE THE ERROR TOLERANCES
C
        SCALE=2.0/RELERR
        AE=SCALE*ABSERR
C
C
C       STEP BY STEP INTEGRATION
C
  100 HFAILD=.FALSE.
C
C       SET SMALLEST ALLOWABLE STEPSIZE
C
       HMIN = U26*ABS(T)
C
C       ADJUST STEPSIZE IF NECESSARY TO HIT THE OUTPUT POINT
C       LOOK AHEAD TWO STEPS TO AVOID DRASTIC CHANGES IN THE
C        STEPSIZE AND THUS LESSEN THE IMPACT OF OUTPUT POINTS
C        ON THE CODE.
C
        DT=TOUT-T
```

```
      IF (ABS(DT) .GE. 2.0*ABS(H)) GO TO 200
      IF (ABS(DT) .GT. ABS(H)) GO TO 150

C     THE NEXT SUCCESSFUL STEP WILL COMPLETE THE INTEGRATION
C      TO THE OUTPUT POINT
C
      OUTPUT= .TRUE.
      H=DT
      GO TO 200
C
  150 H=0.5*DT
C
C
C
C     CORE INTEGRATOR FOR TAKING SINGLE STEP
C
C     THE TOLERANCES HAVE BEEN SCALED TO AVOID PREMATURE UNDERFLOW
C      IN COMPUTING THE ERROR TOLERANCE FUNCTION ET,
C     TO AVOID PROBLEMS WITH ZERO CROSSING, RELATIVE ERROR IS
C      MEASURED USING THE AVERAGE OF THE MAGNITUDES OF THE SOLUTION
C      AT THE BEGINING AND END OF A STEP.
C     THE ERROR ESTIMATE FORMULA HAS BEEN GROUPED TO CONTROL LOSS OF
C      SIGNIFICANCE.
C     TO DISTINGUISH THE VARIOUS ARGUMENTS, H IS NOT PERMITTED
C      TO BECOME SMALLER THAN 26 UNITS OF ROUNDOFF IN T.
C     PRACTICAL LIMITS ON THE CHANGE IN THE STEPSIZ TO
C      SMOOTH THE STEPSIZE SELECTION PROCESS AND TO AVOID EXCESSIVE
C      CHATTERING ON PROBLEMS HAVING DISCONTINUTIES.
C     TO PREVENT UNNECESSARY FAILURES,THE CODE USES 9/10 THE
C      STEPSIZE IT ESTIMATES WILL SUCCEED.
C     AFTER A STEP FAILURE, THE STEPSIZE IS NOT ALLOWED TO INCREASE
C      FOR THE NEXT ATTEMPTED STEP. THIS MAKES THE CODE MORE
C      EFFICIENT ON PROBLEMS HAVING DISCONTINUITIES AND MORE
C      EFFECTIVE IN GENERAL SINCE LOCAL EXTRAPOLATION IS BEING USED
C      AND EXTRA CAUTION SEEMS WARRANTED.
C
C
C     TEST NUMBER OF DERIVATIVE FUNCTION EVALUATIONS.  IF OKAY, TRY
C      TO ADVANCE THE INTEGRATION FROM T TO T+H
C
  200 IF (NFE .LE. MAXNFE) GO TO 220
C
C     TOO MUCH WORK
      IFLAG=4
      KFLAG=4
      RETURN
C
C     ADVANCE AN APPROXIMATE SOLUTION OVER ONE STEP OF LENGTH H
C
  220 CALL FEHL(F,NEQN,Y,T,H,YP,F1,F2,F3,F4,F5,F1)
      NFE=NFE+5
C
C     COMPUTE AND TEST ALLOWABLE TOLERANCES VERSUS LOCAL ERROR
C      ESTIMATES AND REMOVE SCALING OF TOLERANCES. NOTE THAT
```

```
C       RELATIVE ERROR IS MEASURED WITH RESPECT TO THE MAGNITUDES OF
C       THE MAGNITUDES OF THE SOLUTION AT THE BEGINING AND END OF
C        THE STEP.
C
      EEOET=0.0
      DO 250 K=1,NEQN
      ET=ABS(Y(K))+ABS(F1(K))+AE
      IF (ET .GT. 0.0) GO TO 240
C
C     INAPPROPIATE ERROR TOLERANCE
      IFLAG=5
      RETURN
C
  240 EE=ABS((-2090.0*YP(K)+(21970.0*F3(K)-15048.0*F4(K)))+
     1                   (22528.0*F2(K)-27360.0*F5(K)))
  250 EEOET=AMAX1(EEOET,EE/ET)
C
      ESTTOL= ABS(H)*EEOET*SCALE/752400.0
C
      IF (ESTTOL .LE. 1.0) GO TO 260
C
C
C     UNSUCCESSFUL STEP
C                  REDUCE THE STEPSIZE, TRY AGAIN
C                  THE DECREASE IS LIMITED TO A FACTOR OF 1/10
C
      HFAILD= .TRUE.
      OUTPUT= .FALSE.
      S=0.1
      IF (ESTTOL .LT. 59049.0) S=0.9/ESTTOL**0.2
      H=S*H
      IF (ABS(H) .GT. HMIN) GO TO 200
C
C     REQUESTED ERROR UNATTAINABLE AT SMALLEST ALLOWABLE STEPSIZE
      IFLAG=6
      KFLAG=6
      RETURN
C
C
C
C     SUCCESSFUL STEP
C                STORE SOLUTION AT T+H
C                AND EVALUATE DERIVATIVES THERE
C
  260 T=T+H
      DO 270 K=1,NEQN
  270 Y(K)=F1(K)
      A=T
      CALL F(A,Y,YP)
      NFE=NFE+1
C
C
C
C     CHOOSE NEXT STEPSIZE
```

```
C       THE INCREASE IS LIMITED TO A FACTOR OF 5
C       IF STEP FAILURE HAS JUST OCCURRED, NEXT STEPSIZE IS NOT
C                       ALLOWED TO INCREASE
C
        S=5.0
        IF (ESTTOL .GT. 1.889568E-4) S=0.9/ESTTOL**0.2
        IF (HFAILD) S=AMIN1(S,1.0)
        H=SIGN(AMAX1(S*ABS(H),HMIN),H)
C
C       END OF CORE INTEGRATOR
C
C
C       SHOULD WE TAKE ANOTHER STEP
C
        IF (OUTPUT) GO TO 300
        IF (IFLAG .GT. 0) GO TO 100
C
C
C       INTEGRATION SUCCESSFULLY COMPLETED
C
C       ONE-STEP MODE
        IFLAG=-2
        RETURN
C
C       INTERNAL MODE
  300 T=TOUT
        IFLAG=2
        RETURN
C
        END
```

```
      SUBROUTINE FEHL(F,NEQN,Y,T,H,YP,F1,F2,F3,F4,F5,S)
C
C     FEHLBERG FOURTH-FIFTH ORDER RUNGE-KUTTA METHOD
C
C     FEHL INTEGRATES A SYSTEM OF NEQN FIRST ORDER ORDINARY
C      DIFFERENTIAL EQUATIONS OF THE FORM
C           DY(1)/DT=F(T,Y(1),-,Y(NEQN))
C      WHERE THE INITIAL VALUESY(1) AND THE INITIAL DERIVATIVES
C       YP(1) ARE SPECIFIED AT THE STARTING POINT T. FEHL ADVANCES
C       THE SOLUTION OVER THE FIXED STEP H AND RETURNS THE FIFTH
C       ORDER (SIXTH ORDER ACCURATE LOCALLY) SOLUTION APPROXIMATION
C       AT T+H IN ARRAY S(1).
C     F1,...,F5 ARE ARRAYS OF DIMENSION NEQN WHICH ARE NEEDED FOR
C        INTERNAL STORAGE.
C     THE FORMULAS HAVE BEEN GROUPED TO CONTROL LOSS OF
C       SIGNIFICANCE.  FEHL SHOULD BE CALLED WITH AN H NOT SMALLER
C       THAN 10F ROUNDOFF IN T SO THAT THE VARIOUS
C       INDEPENDENT ARGUMENTS CAN BE DISTING.
C
C
      INTEGER NEQN
      REAL Y(NEQN),T,H,YP(NEQN),F1(NEQN),F2(NEQN),F3(NEQN),
     1   F4(NEQN),F5(NEQN),S(NEQN)
C
      REAL CH
      INTEGER K
C
      CH=H/4.0
      DO 221 K=1,NEQN
  221 F5(K)=Y(K)+CH*YP(K)
      CALL F(T+CH,F5,F1)
C
      CH=3.0*H/32.0
      DO 222 K=1,NEQN
  222  F5(K)=Y(K)+CH*(YP(K)+3.0*F1(K))
      CALL F(T+3.0*H/8.0,F5,F2)
C
      CH=H/2197.0
      DO 223 K=1,NEQN
  223  F5(K)=Y(K)+CH*(1932.0*YP(K)+(7296.0*F2(K)-7200.0*F1(K)))
      CALL F(T+12.0*H/13.0,F5,F3)
C
      CH=H/4104.0
      DO 224 K=1,NEQN
  224  F5(K)=Y(K)+CH*((8341.0*YP(K)-845.0*F3(K))+
     1                (29440.0*F2(K)-32832.0*F1(K)))
      CALL F(T+H,F5,F4)
C
      CH=H/20520.0
      DO 225 K=1,NEQN
  225  F1(K)=Y(K)+CH*((-6080.0*YP(K)+(9295.0*F3(K)-
```

```
     1       5643.0*P4(K)))+(41040.0*P1(K)-28352.0*P2(K)))
        CALL F(T+H/2.0,P1,P5)
C
C       COMPUTE APPROXIMATE SOLUTION AT T+H
C
        CH=H/7618050.0
        DO 230 K=1,NEQN
  230    S(K)=Y(K)+CH*((902880.0*YP(K)+(3855735.0*P3(K)-
     1       1371249.0*P4(K)))+(3953664.0*P2(K)+277020.0*P5(K)))
C
        RETURN
        END
```

```
      SUBROUTINE RDAPRT(IPLT,N1,N2,KEY)
      COMMON /LUNIO1/LUNI,LUNO,LUNO1
      DIMENSION IPLT(N2)
  903 FORMAT(20A4)
      GO TO (1,2,3),KEY
    1 CONTINUE
      READ (LUNI,903) (IPLT(I),I=N1,N2)
      GO TO 4
    2 WRITE (LUNO,903) (IPLT(I),I=N1,N2)
      GO TO 4
    3 WRITE (LUNO1,903) (IPLT(I),I=N1,N2)
    4 CONTINUE
      RETURN
      END



      SUBROUTINE RDPRNT(A,MMAX,NMAX,NDIM,KEY2)
      COMMON /LUNIO1/LUNI,LUNO,LUNO1
      DIMENSION A(NDIM,1)
  898 FORMAT(5E16.8)
  899 FORMAT(8E13.5)
  900 FORMAT(8E10.5)
      DO 1I=1,MMAX
      GO TO (2,3,5),KEY2
    2 CONTINUE
      READ (LUNI,900) (A(I,J),J=1,NMAX)
      GO TO 4
    3 CONTINUE
      WRITE (LUNO,899) (A(I,J),J=1,NMAX)
      GO TO 4
    5 CONTINUE
      WRITE (LUNO1,898) (A(I,J),J=1,NMAX)
    4 CONTINUE
    1 CONTINUE
      RETURN
      END
```

```
      SUBROUTINE VCTINT(X,N,XO)
      DIMENSION X(N)
      DO 1I=1,N
      X(I)=0.0+XO
    1 CONTINUE
      RETURN
      END



      SUBROUTINE MTRINT(A,N,M,NDIM,AO)
      DIMENSION A(NDIM,1)
      DO 1I=1,N
      DO 2J=1,M
      A(I,J)=0.0+AO
    2 CONTINUE
    1 CONTINUE
      RETURN
      END



      SUBROUTINE RDPRT1(X,N,KEY2)
      COMMON /LUNIO1/LUNI,LUNO,LUNO1
      DIMENSION X(N)
  897 FORMAT(8E10.5)
  898 FORMAT(5E16.8)
  899 FORMAT(8E13.5)
  900 FORMAT(8E10.3)
      GO TO (1,2,4,5,6),KEY2
    6 CONTINUE
      WRITE (LUNO1,898) (X(I),I=1,N)
      GO TO 3
    5 CONTINUE
      READ (LUNI,898) (X(I),I=1,N)
      GO TO 3
    4 CONTINUE
      WRITE (LUNO1,898) (X(I),I=1,N)
      GO TO 3
    1 CONTINUE
      READ (LUNI,900) (X(I),I=1,N)
      GO TO 3
    2 CONTINUE
      WRITE (LUNO,899) (X(I),I=1,N)
    3 CONTINUE
      RETURN
      END
```

```
      SUBROUTINE GMINV(NR,NC,A,U,MR,MT)
      DIMENSION A(1),U(1),S(30)
      COMMON/MAIN1/NDIM
      COMMON/INOU/KIN,KOUT
      NDIM1=NDIM+1
      TOL=1.E-14
      ADV=1.E-24
      MR=NC
      NRM1=NR-1
      TOL1=0.
      JJ=1
      DO 10 J=1,NC
      S(J)=DOT(NR,A(JJ),A(JJ))
      IF(S(J).GT.TOL1)TOL1=S(J)
   10 JJ=JJ+NDIM
      TOL1=ADV*TOL1
      ADV=TOL1
      JJ=1
      DO 100 J=1,NC
      FAC=S(J)
      JM1=J-1
      JRM=JJ+NRM1
      JCM=JJ+JM1
      DO 20 I=JJ,JCM
   20 U(I)=0.
      U(JCM)=1.0
      IF(J.EQ.1) GO TO 54
      KK=1
      DO 30 K=1,JM1
      IF(S(K).EQ.1.0) GO TO 30
      TEMP=-DOT(NR,A(JJ),A(KK))
      CALL VADD(K,TEMP,U(JJ),U(KK))
   30 KK=KK+NDIM
      DO 50 L=1,2
      KK=1
      DO 50 K=1,JM1
      IF(S(K).EQ.0.) GO TO 50
      TEMP=-DOT(NR,A(JJ),A(KK))
      CALL VADD(NR,TEMP,A(JJ),A(KK))
      CALL VADD(K,TEMP,U(JJ),U(KK))
   50 KK=KK+NDIM
      TOL1=TOL*FAC+ADV
      FAC=DOT(NR,A(JJ),A(JJ))
   54 IF(FAC.GT.TOL1) GO TO 70
      DO 55 I=JJ,JRM
   55 A(I)=0.
      S(J)=0.
      KK=1
      IF(S(K).EQ.0.) GO TO 65
      DO 65 K=1,JM1
      TEMP=-DOT(K,U(KK),U(JJ))
      CALL VADD(NR,TEMP,A(JJ),A(KK))
```

```
 65 KK=KK+NDIM
    FAC=DOT(J,U(JJ),U(JJ))
    NR=NR-1
    GO TO 75
 70 CONTINUE
    S(J)=1.0
    KK=1
    DO 72 K=1,JM1
    IF(S(K).EQ.1.) GO TO 72
    TEMP=-DOT(NR,A(JJ),A(KK))
    CALL VADD(K,TEMP,U(JJ),U(KK))
 72 KK=KK+NDIM
 75 FAC=1./SQRT(FAC)
    DO 80 I=JJ,JRM
 80 A(I)=A(I)*FAC
    DO 85 I=JJ,JCM
 85 U(I)=U(I)*FAC
100 JJ=JJ+NDIM
    IF(NR.EQ.NR.OR.NR.EQ.NC) GO TO 120
    IF(NT.NE.0)WRITE(KOUT,110)NR,NC,NR
110 FORMAT(I3,1HX,I2,8H M: RANK,I2)
120 NEND=NC*NDIM
    JJ=1
    DO 135 J=1,NC
    DO 125 I=1,NR
    II=I-J
    S(I)=0.
    DO 125 KK=JJ,NEND,NDIM
125 S(I)=S(I)+A(II+KK)*U(KK)
    II=J
    DO 130 I=1,NR
    U(II)=S(I)
130 II=II+NDIM
135 JJ=JJ+NDIM1
    RETURN
    END
```

```
      FUNCTION DOT(NR,A,B)
      DIMENSION A(1),B(1)
      EPSTOL = 0.5E-08
      DOT=0.
      DO 1 I=1,NR
      IF(ABS(A(I)).LE.EPSTOL)GO TO 1
      IF(ABS(B(I)).LE.EPSTOL)GO TO 1
      DOT = DOT + A(I)*B(I)
    1 CONTINUE
      RETURN
      END


      SUBROUTINE VADD(N,C1,A,B)
      DIMENSION A(1),B(1)
      DO 1 I=1,N
    1 A(I)=A(I)+C1*B(I)
      RETURN
      END
```

```
      SUBROUTINE TRIBXF(TR,PHI,THETA,PSI,N,NDIM1)
C     TRANSFORMATION FROM XFRAME(INERT) TO XBODY
C     COORDINATE SYSTEM COMPONENTS.
      REAL*8 CPH,SPH,CTH,STH,CPS,SPS
      DIMENSION TR(NDIM1,1)
      CPH = COS(PHI)
      SPH = SIN(PHI)
      CTH = COS(THETA)
      STH = SIN(THETA)
      CPS = COS(PSI)
      SPS = SIN(PSI)
C     XBODY(I) = TR(I,J)*XFRAME(J)
      TR(1,1) = CTH*CPS
      TR(1,2) = CTH*SPS
      TR(1,3) = -STH
      TR(2,1) = -CPH*SPS + SPH*STH*CPS
      TR(2,2) = CPH*CPS + SPH*STH*SPS
      TR(2,3) = SPH*CTH
      TR(3,1) = SPH*SPS + CPH*STH*CPS
      TR(3,2) = -SPH*CPS + CPH*STH*SPS
      TR(3,3) = CPH*CTH
      IPRTOP = 2
      IPRTOP = 1
      IF(IPRTOP.EQ.2)CALL RDPRNT(TR,N,N,NDIM1,2)
      RETURN
      END
```

```
      SUBROUTINE TRIFIB(TR,PHI,THETA,PSI,N,NDIM1)
C     TRANSFORMATION FROM XBODY TO XFRAME(INERT)
C     COORDINATE SYSTEM COMPONENTS.
      REAL*8 CPH,SPH,CTH,STH,CPS,SPS
      DIMENSION TR(NDIM1,1)
      CPH = COS(PHI)
      SPH = SIN(PHI)
      CTH = COS(THETA)
      STH = SIN(THETA)
      CPS = COS(PSI)
      SPS = SIN(PSI)
C     XFRAME(I) = TR(I,J)*XBODY(J)
      TR(1,1) = CTH*CPS
      TR(1,2) = -CPH*SPS + SPH*STH*CPS
      TR(1,3) = SPH*SPS + CPH*STH*CPS
      TR(2,1) = CTH*SPS
      TR(2,2) = CPH*CPS + SPH*STH*SPS
      TR(2,3) = -SPH*CPS + CPH*STH*SPS
      TR(3,1) = -STH
      TR(3,2) = SPH*CTH
      TR(3,3) = CPH*CTH
      IPRTOP = 2
      IPRTOP = 1
      IF(IPRTOP.EQ.2)CALL RDPRNT(TR,N,N,NDIM1,2)
      RETURN
      END
```

```
      SUBROUTINE ANGVTR(TR,PHI,THETA,PSI,N,NDIM1)
C     TRANSFORMATION FROM XBODY TO XFRAME(INERT)
C     COORDINATE SYSTEM COMPONENTS.
      REAL*8 CPH,SPH,CTH,STH,CPS,SPS,TNTH,SCTH
      DIMENSION TR(NDIM1,1)
      CPH = COS(PHI)
      SPH = SIN(PHI)
      CTH = COS(THETA)
      STH = SIN(THETA)
      CPS = COS(PSI)
      SPS = SIN(PSI)
      TNTH = STH/CTH
      SCTH = 1.0/CTH
C     ANGVFR(I) = TR(I,J)*ANGVBD(J)
      TR(1,1) = 1.0
      TR(1,2) = SPH*TNTH
      TR(1,3) = CPH*TNTH
      TR(2,1) = 0.0
      TR(2,2) = CPH
      TR(2,3) = -SPH
      TR(3,1) = 0.0
      TR(3,2) = SPH*SCTH
      TR(3,3) = CPH*SCTH
      IPRTOP = 2
      IPRTOP = 1
      IF(IPRTOP.EQ.2)CALL RDPRNT(TR,N,N,NDIM1,2)
      RETURN
      END
```

```
      SUBROUTINE IXIJTR(XIXIJ,XL1,XL2,XL3,XMB,NDIM1,IGCOPT,ITROPT)
C     SRSP TO DEFINE SHIFTED INERTIA MATRIX/2ND ORDER TENSOR.
C     I.E., AXES ARE SHIFTED (SAY FROM CM TO ARBITRARY POINT O).
C     NOTE THAT CONST1 = 1 FOR XCM TO POINT XO TRANSFORMATION.
C                CONST1 = 2 FOR XO TO POINT XCM TRANSFORMATION.
      DIMENSION XIXIJ(NDIM1,1)
      GCONST = 32.174
      CONST1 = 1.0
      IF(ITROPT.EQ.2) CONST1 = -1.0
      XMASS = XMB
      IF(IGCOPT.EQ.2)XMASS=XMASS/GCONST
      IGCOPT = 1
      XL1SQ = XL1*XL1
      XL2SQ = XL2*XL2
      XL3SQ = XL3*XL3
      XIXIJ(1,1) = XIXIJ(1,1) - CONST1*XMASS*(XL2SQ + XL3SQ)
      XIXIJ(1,2) = XIXIJ(1,2) + CONST1*XMASS*(XL1*XL2)
      XIXIJ(1,3) = XIXIJ(1,3) + CONST1*XMASS*(XL1*XL3)
      XIXIJ(2,1) = XIXIJ(2,1) + CONST1*XMASS*(XL2*XL3)
      XIXIJ(2,2) = XIXIJ(2,2) - CONST1*XMASS*(XL1SQ + XL3SQ)
      XIXIJ(2,3) = XIXIJ(2,3) + CONST1*XMASS*(XL2*XL3)
      XIXIJ(3,1) = XIXIJ(3,1) + CONST1*XMASS*(XL3*XL1)
      XIXIJ(3,2) = XIXIJ(3,2) + CONST1*XMASS*(XL3*XL2)
      XIXIJ(3,3) = XIXIJ(3,3) - CONST1*XMASS*(XL1SQ + XL2SQ)
      IPRTOP = 2
      IPRTOP = 1
      IF(IPRTOP.EQ.2) CALL RDPRNT(XIXIJ,3,3,NDIM1,2)
      RETURN
      END
```

```
      SUBROUTINE IXIJRP(XL1,XL2,XL3,RHOB,XMB,XIXIJ,NDIM1,
     * IMASSO,IGCOPT)
      DIMENSION XIXIJ(NDIM1,1)
C     XL1 = RECTANGULAR PARALLELEPIPED LENGTH IN X1 OR U1
C           BODY ROLL AXIS DIRECTION.
C     XL2 = RECTANGULAR PARALLELEPIPED LENGTH IN X2 OR U2
C           BODY PITCH AXIS DIRECTION.
C     XL3 = RECTANGULAR PARALLELEPIPED LENGTH IN X3 OR U3
C           BODY YAWAXIS DIRECTION.
C     IMASSO,IGCOPT = MASS CALCULATION, GC UNIT CONVERSION
C           FLAGS, RESPECTIVELY.
C     RHOB = MASS DENSITY OF BODY.
C     XMB = MASS OF THE BODY IF KNOWN A PRIORI.
C     NDIM1 = ROW DIMENSION ALLOCATION FOR INERTIA MATRIX
C           (2ND ORDER TENSOR).
      ZER = 0.0
      UNI = 1.0
      GCONST = 32.174
      IF(IMASSO-1) 1,1,2
    1 CONTINUE
      XMASS = RHOB*XL1*XL2*XL3
      GO TO 3
    2 CONTINUE
      XMASS = XMB
    3 CONTINUE
      IF(IGCOPT.EQ.2)XMASS = XMASS/GCONST
      IGCOPT = 1
      CALL MTRINT(XIXIJ,3,3,NDIM1,ZER)
      XFACT1 = (1.0/12.0)*XMASS
      XL1SQ = XL1*XL1
      XL2SQ = XL2*XL2
      XL3SQ = XL3*XL3
      XIXIJ(1,1) = XFACT1*(XL2SQ + XL3SQ)
      XIXIJ(2,2) = XFACT1*(XL1SQ + XL3SQ)
      XIXIJ(3,3) = XFACT1*(XL1SQ + XL2SQ)
      IPRTOP = 2
      IPRTOP = 1
      IF(IPRTOP.EQ.2)CALL RDPRNT(XIXIJ,3,3,NDIM1,2)
      XMB = XMASS
      RETURN
      END
```

```
      SUBROUTINE OMGDEF(OMGMTR,R,P,Q,NDIM1)
C     'ROLL, PITCH, AND YAW RATE' - 'SPINOR' MATRIX IN
C     BODY COORDINATE REFERENCE SYSTEM.
      DIMENSION OMGMTR(NDIM1,1)
      ZER = 0.0
      CALL MTRINT(OMGMTR,3,3,NDIM1,ZER)
      OMGMTR(1,2) = R
      OMGMTR(1,3) = -Q
      OMGMTR(2,3) = P
      OMGMTR(2,1) = -OMGMTR(1,2)
      OMGMTR(3,1) = -OMGMTR(1,3)
      OMGMTR(3,2) = -OMGMTR(2,3)
      RETURN
      END


      SUBROUTINE VECAXB(A1,A2,A3,B1,B2,B3,C1,C2,C3)
      C1 = A2*B3 - A3*B2
      C2 = A3*B1 - A1*B3
      C3 = A1*B2 - A2*B1
      RETURN
      END



      SUBROUTINE VARAXB(A,B,C,NDIM)
      DIMENSION A(NDIM),B(NDIM),C(NDIM)
      C(1) = A(2)*B(3) - A(3)*B(2)
      C(2) = A(3)*B(1) - A(1)*B(3)
      C(3) = A(1)*B(2) - A(2)*B(1)
      RETURN
      END



      SUBROUTINE EQVMTR(A,B,N1,N2,NDIMA,NDIMB)
      DIMENSION A(NDIMA,1),B(NDIMB,1)
      DO 1I1=1,N1
      DO 2I2=1,N2
      B(I1,I2) = A(I1,I2)
    2 CONTINUE
    1 CONTINUE
      RETURN
      END
```

```
      SUBROUTINE MATPRD(A,X,XK,N1,N2,NDIM1)
      DIMENSION A(NDIM1,1),X(N2),XK(N1)
      EPSTOL = 1.0E-08
      DO 1I1=1,N1
      SUM = 0.0
      DO 2I2=1,N2
      IF(ABS(A(I1,I2)) .LE. EPSTOL)GO TO 2
      IF(ABS(X(I2)) .LE. EPSTOL)GO TO 2
      SUM = SUM + A(I1,I2)*X(I2)
    2 CONTINUE
      XK(I1) = SUM
    1 CONTINUE
      RETURN
      END



      FUNCTION XJFIRF(AMPFLR,PHSFLR,XFREQ,T)
C     DEFINITION OF X(JF) FLOOR MOTION AS FUNCTION OF TIME
C     AT (IR)TH VIBRATION ISOLATION SUPPORT LEG LOCATION.
      PI = 3.1415926535
      PIT2 = 2.0*PI
      EPSTOL = 0.5E-06
      ZER = 0.0
      DEGRAD = PI/180.0
      PPHSFL = DEGRAD*PHSFLR
      OMEGA = PIT2*XFREQ
      AAMPLR = AMPFLR
      CSOMPH = COS(OMEGA*T - PPHSFL)
      IF(ABS(AAMPLR) - EPSTOL) 1,1,2
    2 CONTINUE
      IF(ABS(CSOMPH) - EPSTOL) 1,1,3
    1 CONTINUE
      XJFIRF = ZER
      GO TO 4
    3 CONTINUE
      XJFIRF = AAMPLR*CSOMPH
    4 CONTINUE
      RETURN
      END
```

HETRBMIN

```
**************************************************************
  12
    0.000       0.000       0.000
    0.000       0.000       0.000
    0.000       0.000   1.000E-03
    0.000       0.000       0.000
    0.000       0.000       0.000
    0.000       0.000       0.000
   12.073       4.167       1.500
   2    2    1
   3000.       0.000
   6    3
   1500.       1500.       768.0
   1500.       1500.       768.0
   1500.       1500.       768.0
   1500.       1500.       768.0
   1500.       1500.       768.0
   1500.       1500.       768.0
   6.037       2.084      -0.750
   0.000       2.084      -0.750
  -6.037       2.084      -0.750
  -6.037      -2.084      -0.750
   0.000      -2.084      -0.750
   6.037      -2.084      -0.750
   0.000       0.000       0.000
   0.000       0.000       0.000
   0.000       0.000       0.000
   0.000       0.000       0.000
   0.000       0.000       0.000
   0.000       0.000       0.000
   0.000       0.000       0.000
   0.000       0.000       0.000
   0.000       0.000       0.000
   0.000       0.000       0.000
   0.000       0.000       0.000
   0.000       0.000       0.000
   1.000
```

APPENDIX D

TFVMHF02: Fortran IV Software Package for Temporal Filtering Analysis of the Effect of Vibratory/Acoustic Motion on Fringe Modulation

```
C     COMMENT: TFVMHF02.FTN;1
C     DRIVER PROGRAM TFVMHF - DRIVER FOR TEMPORAL FILTERING
C     OF HOLOGRAMS RELEVANT THE IMPACT OF VIBRATORY MOTION
C     ON DIFFERENTIAL HOLOGRAM FRINGE MODULATION FUNCTION.
C     ****************************************************************
      IMPLICIT REAL*8 (A-H,O-Z)
      COMMON /LUNIO1/LUNI,LUNO,LUNO1
C     ****************************************************************
      DIMENSION APDSEP(50)
      DIMENSION XJNKA(500),XYNKA(500)
      DIMENSION F1(15),F2(15),DELF12(15)
      DIMENSION VAMPEX(50),TSAMPH(15),XFREQ(50),ALPHAE(15),
     * BETAEX(15)
      DIMENSION XFRMDF(50),MXTMDF(50)
C     ****************************************************************
  895 FORMAT(1H0)
  896 FORMAT(1H1)
  897 FORMAT(/)
  898 FORMAT(///)
  899 FORMAT(8E13.5)
  900 FORMAT(8E10.5)
  901 FORMAT(16I5)
  902 FORMAT(8E16.8)
C     ****************************************************************
      LUNI = 1
      LUNO = 4
      LUNO1 = 5
      OPEN(UNIT = LUNI, TYPE = 'OLD', NAME = 'TFVMHFIS.DAT',
     * ERR = 99)
C     ****************************************************************
      NDIM1 = 500
      NDIM2 = 50
      NDIM3 = 15
      NDIMPR = 50
      EPSTOL = 0.5E-15
      UNI = 1.0
      ZER = 0.0
      PI = 3.1415926535
      PIT2 = 2.0*PI
      PIS2 = PI/2.0
      DEGRAD = PI/180.0
      RADDEG = 180.0/PI
C     ****************************************************************
      CALL RDAPRT(APDSEP,1,20,1)
      READ (LUNI,901) NDESMX,MXVAMP,IF1F2D,MXALPH,MXBETA,MXTSMP
      READ (LUNI,900) CLFV,HETFRQ,HTLMDA
      CALL RDPRT1(VAMPEX,MXVAMP,1)
      CALL RDPRT1(F1,IF1F2D,1)
      CALL RDPRT1(F2,IF1F2D,1)
      CALL RDPRT1(DELF12,IF1F2D,1)
      CALL RDPRT1(ALPHAE,MXALPH,1)
```

```
      CALL RDPRT1(BETAEX,MXBETA,1)
      CALL RDPRT1(TSAMPH,MXTSMP,1)
C     ****************************************************************
      WRITE (LUNO,896)
      CALL RDAPRT(APDSEP,1,20,2)
      WRITE (LUNO,898)
      WRITE (LUNO,901) NBESMX,MXVAMP,IF1F2D,MXALPH,MXBETA,MXTSMP
      WRITE (LUNO,898)
      WRITE (LUNO,899) CLPV,RETFRQ,RTLMDA
      WRITE (LUNO,898)
      CALL RDAPRT(APDSEP,1,20,2)
      WRITE (LUNO,897)
      CALL RDPRT1(VAMPEX,MXVAMP,2)
      WRITE (LUNO,898)
      CALL RDAPRT(APDSEP,1,20,2)
      WRITE (LUNO,897)
      CALL RDPRT1(F1,IF1F2D,2)
      WRITE (LUNO,897)
      CALL RDPRT1(F2,IF1F2D,2)
      WRITE (LUNO,897)
      CALL RDPRT1(DELF12,IF1F2D,2)
      WRITE (LUNO,898)
      CALL RDAPRT(APDSEP,1,20,2)
      WRITE (LUNO,897)
      CALL RDPRT1(ALPHAE,MXALPH,2)
      WRITE (LUNO,897)
      CALL RDPRT1(BETAEX,MXBETA,2)
      WRITE (LUNO,898)
      CALL RDAPRT(APDSEP,1,20,2)
      WRITE (LUNO,897)
      CALL RDPRT1(TSAMPH,MXTSMP,2)
      WRITE (LUNO,898)
      CALL RDAPRT(APDSEP,1,20,2)
      CALL RDAPRT(APDSEP,1,20,2)
      WRITE (LUNO,898)
C     ****************************************************************
      CALL ALVCPD(ALPHAE,MXALPH,DEGRAD,1)
      CALL ALVCPD(BETAEX,MXBETA,DEGRAD,1)
      DO 1I1=1,MXVAMP
      VVAMEX = VAMPEX(I1)
      DO 2I2=1,IF1F2D
      FF1 = F1(I2)
      FF2 = F2(I2)
      DDLF12 = DELF12(I2)
      CALL XVECCL(XFREQ,JMXFRQ,NDIMFR,FF1,FF2,DDLF12,1)
      DO 3I3=1,MXALPH
      AALPHA = ALPHAE(I3)
      DO 4I4=1,MXBETA
      BBETAE = BETAEX(I4)
      DO 5I5=1,MXTSMP
      TTSMPH = TSAMPH(I5)
      DO 6I6=1,JMXFRQ
      XXFREQ = XFREQ(I6)
      XOMEGA = PIT2*XXFREQ
```

```
      XK = PIT2/HTLNDA
      XKA = XK*VVAMEX
      XKARG = XKA*(DCOS(AALPHA) + DCOS(BBETAE))
      LIM = NBESMX
      CALL SPBESS(XKARG,ZER,LIM,EPSTOL,XJNKA,XYNKA,DUMMY,
     * IP,NDIM1)
      NNBSMX = LIM
      SNCARG = OMEGA*TTSMPH/2.0
      EPSVNM = 1.0
      CCSWTA = 1.0
      SSNWTA = 0.0
      CSWTAR = DCOS(SNCARG)
      SNWTAR = DSIN(SNCARG)
      ICNTER = ICNTER + 1
      SUM = 0.0
      DO 7I7=1,NNBSMX
      XI7 = I7
      IF(I7 - 1) 10,10,11
   10 CONTINUE
      SNCFCN = UNI
      GO TO 12
   11 CONTINUE
      EPSVNM = 2.0
      IF(DABS(SNCARG) - EPSTOL)10,10,13
   13 CONTINUE
      SNCFCN = SSNWTA/(XI7*SNCARG)
   12 CONTINUE
      SUM = SUM + EPSVNM*XJNKA(I7)*SNCFCN
      CALL CPLXQR(CCSWTA,SSNWTA,CSWTAR,SNWTAR,XDUM1,XDUM2,1)
      CCSWTA = XDUM1
      SSNWTA = XDUM2
    7 CONTINUE
      XFRMDF(I6) = SUM
      NXTMDF(I6) = NNBSMX
    6 CONTINUE
      CALL RDPRT1(XFRMDF,JMXFRQ,2)
      WRITE (LUNO,898)
      CALL RDIPRT(NXTMDF,1,JMXFRQ,2)
      WRITE (LUNO,898)
    5 CONTINUE
    4 CONTINUE
    3 CONTINUE
    2 CONTINUE
    1 CONTINUE
      CLOSE (UNIT = LUNI,DISP = 'KEEP',ERR = 99)
   99 CONTINUE
      STOP
      END
```

```
      SUBROUTINE SPBESS(X,UI,NTH,EP,BBJ,BBY,CN,IP,NDIMJN)
      DIMENSION BBJ(NDIMJN),BBY(NDIMJN)
      IMPLICIT DOUBLE PRECISION (A -H, O - Z)
      NTHMAX = 399
      IF(X - 1.0E-06)101,101,10
  101 CONTINUE
      WRITE (LUNO,102)
  102 FORMAT(/20X,18HSTOP 101 IN BESSEL   /)
      STOP
   10 CONTINUE
      IF(NTH - NTHMAX)12,12,121
  121 CONTINUE
      WRITE (LUNO,122)
  122 FORMAT(/20X,18HSTOP 121 IN BESSEL       /)
      STOP
  141 CONTINUE
      WRITE (LUNO,142)
  142 FORMAT(/20X,18HSTOP 141 IN BESSEL       /)
      STOP
   12 CONTINUE
      IF(UI*(1.0-UI))141,210,210
  210 CONTINUE
      IP = 1
C
C     ESTIMATE THE BESSEL FUNCTION OF MAXIMUM ORDER.
C
      NU = 0
      XOT = X/2.0
      U = UI
      XTU = XOT**UI
      GAM = GAMMA(UI + 1.0)
      TI = XTU/GAM
      PI = 3.1415926535
      TOX = 1.0/XOT
      W = TOX/PI
   15 CONTINUE
      TH = TI
      NU = NU + 1
      U = U + 1.0
      TI = XOT*TH/U
      IF(TI - EP)151,15,15
  151 CONTINUE
      NU = NU - 2
      U = U - 2.0
C
C     IF NU .LT. NTH THEN SET NTH = NU AND IP = 2
C
      IF(NU - NTH)161,16,16
  161 CONTINUE
      IP = 2
      NTH = NU
```

```
   16 CONTINUE
      BJP = TM
C
C     COMPUTE J(NU+1)/J(NU) BY CONTINUED FRACTION
C
      U = U + 1.0
      NU = NU + 1
      IF(U - X)201,201,20
  201 CONTINUE
      PRINT 202,U,X
  202 FORMAT(/10X,11HMAX ORDER = 1PE14.7,10X,10HARGUMENT = ,
     * 1PE14.7,/,10X,40HCONTINUED FRACTION EVALUATION IS SUSPECT /)
   20 CONTINUE
      AKMT = 0.0
      AKMO = 1.0
      BKMT = 1.0
      BKMO = TOX*U
      RK = AKMO/BKMO
      RM = 0.0
      SA = -1.0
      UK = U
      DP = 1.0E-08
   21 CONTINUE
      IF(DABS(1.0 - RM/RK) - DP)22,211,211
  211 CONTINUE
      RM = RK
      UK = UK + 1.0
      SB = UK*TOX
      AK = SB*AKMO + SA*AKMT
      BK = SB*BKMO + SA*BKMT
      RK = AK/BK
      AKMT = AKMO
      AKMO = AK
      BKMT = BKMO
      BKMO = BK
      GO TO 21
   22 CONTINUE
      BJ = BJP/RK
C
C     IF X .LT. 8 COMPUTE NORMALIZATION CONSTANT, Y(UI), AND
C     Y(UI + 1) BY SERIES (PATH 80). IF X.GE. 8 USE PHASE
C     AMPLITUDE METHOD.
C
      IF(X - 8.0)80,221,221
  221 CONTINUE
      NU = NU - 1
      U = U - 1.0
      GO TO (30,40),IP
   30 CONTINUE
      NN = NU - NTM
      DO 26I=1,NN
      BJM = U*TOX*BJ - BJP
      NU = NU - 1
      U = U - 1.0
```

```
      BJP = BJ
      BJ = BJM
   26 CONTINUE
   40 CONTINUE
      BBJ(NU + 1) = BJ
      DO 35I=1,NTH
      BJM = U*TOI*BJ - BJP
      BBJ(NU) = BJM
      BJP = BJ
      BJ = BJM
      NU = NU - 1
      U = U - 1.0
   35 CONTINUE
      AL1 = UI*UI - 0.25
      A2 = 0.5*AL1
      AL2 = AL1*AL1
      A4 = AL2/24.0 - AL1/4.0
      AL3 = AL2*AL1
      A6 = (1.0/80.0)*AL3 - (7.0/20.0)*AL2 + 0.75*AL1
      AL4 = AL3*AL1
      A8 = (1.0/1792.0)*AL4 - (95.0/224.0)*AL3 + (807.0/224.0)*
     * AL2 - (315.0/56.0)*AL1
      AL5 = AL4*AL1
      A10 = (7./2304.)*AL5 - (35./72.)*AL4 + (1975./192.)*AL3 -
     * (58.)*AL2 + (315./4.)*AL1
      TS = 1.0/X
      T2 = TS*TS
      T4 = T2*T2
      T6 = T4*T2
      T8 = T6*T2
      T10 = T8*T2
      TPHI = 1.0 + A2*T2 + A4*T4 + A6*T6 + A8*T8 + A10*T10
      PHI = TPHI/TS - (UI + 0.5)*(PI/2.0)
      A2 = 0.25*AL1
      A4 = (5./32.)*AL2 - (3./8.)*AL1
      A6 = (15./128.)*AL3 - (37./32.)*AL2 + (15./8.)*AL1
      A8 = (195./2048.)*AL4 - (611./256.)*AL3 + (1821./128.)*AL2  -
     * (315./16.)*AL1
      A10 = (663./8192.)*AL5 - (4199./1024.)*AL4 + (29811./512.)*AL3 -
     * (2223./8.)*AL2 + (2835./8.)*AL1
      BNU = 1.0 + A2*T2 + A4*T4 + A6*T6 + A8*T8 + A10*T10
      SQW = SQRT(W)
      ANU = SQW*BNU
      CS = DCOS(PHI)
      SS = DSIN(PHI)
      FJ = ANU*CS
      FY = ANU*SS
      CN = FJ/BBJ(1)
      FJP = CN*BBJ(2)
      AB = DABS(FJ)
C
C     IF J(UI) IS NOT CLOSE TO ZERO, CALCULATE Y(UI +1) FROM THE
C     WRONSKIAN CONDITION. IF IT IS CALCULATE THE DERIVATIVE OF
C     Y(UI).
```

```
C
      IF(AB - 0.001) 36,361,361
  361 CONTINUE
      FYP = (FY*FJP - W)/FJ
      GO TO 50
   36 CONTINUE
      A2 = (5./2.)*A2
      A4 = (9./2.)*A4
      A6 = (13./2.)*A6
      A8 = (17./2.)*A8
      A10 = (21./2.)*A10
      ANP = 0.50 + A2*T2 + A4*T4 + A6*T6 + A8*T8 + A10*T10
      ANP = -TS*SQW*ANP
      PHIP = W/(ANU*ANU)
      FYP = ANP*SS + PHIP*FJ
      FYP = UI*TS*FY - FYP
      FJP = ANP*CS - PHIP*FY
      FJP = UI*TS*FJ - FJP
      CN = FJP/BBJ(2)
      GO TO 50
C
C     SERIES CALCULATION OF NORMALIZATION CONSTANT, Y(UI), AND
C     Y(UI + 1). START THE SERIES SUMMATION AT THE LARGEST EVEN
C     INTEGER. IF NU IS ODD SET JP = 1, IF NU IS EVEN SET
C     JP = 2.
C
   80 CONTINUE
      KU = NU/2
      ANOT = GAM/XTU
      AN = (UI + 2.)*ANOT
      TOP = 2.0/PI
      BN = 0.0
      IF(UI .GT. 0.99995)GO TO 81
      BN = TOP*(UI + 2.)*ANOT*ANOT/(1. - UI)
   81 CONTINUE
      TUI = UI*UI
      IF(KU - 2) 84,841,841
  841 CONTINUE
      DO 83K=2,KU
      KM = K - 1
      ANM = AN
      BNM = BN
C     AK = (UI + 2*K)*(UI + KM)/(K*(UI + 2*KM))
      TWSK = 2*K
      XKM = KM
      XK = K
      TWSKM = 2*KM
      RK = (UI +TWSK)*(UI +XKM)/(XK*(UI +TWSKM))
      AN = RK*ANM
C     BN = -RK*(TUI + KM)*BNM/(K - UI)
      BN = -RK*(TUI + XKM)*BNM/(XK - UI)
   83 CONTINUE
   84 CONTINUE
      CN = 0.0
```

```
      FY = 0.0
      FYP = 0.0
      BJD = 0.0
      BND = 0.0
      JP = 1
      IF(2*KU - NU)851,85,851
  851 CONTINUE
      JP = 2
   85 CONTINUE
      NU = NU - 1
      U = U - 1.
      N = KU
      GO TO (86,95),IP
   86 CONTINUE
      NN = NU - NTM
      DO 87I=1,NN
      BJM = U*TOI*BJ - BJP
      GO TO (88,89),JP
   88 CONTINUE
      ANP = AN
      BNP = BN
      NP = N
      CN = CN + ANP*BJP
      FY = FY + BNP*BJP
      FYP = FYP + 0.50*(BND-BNP)*BJD
      BJD = BJ
      BND = BNP
      N = N - 1
C     RN = NP*(UI + 2*N)/((UI + 2*NP)*(UI + N))
      XN = N
      XNP = NP
      RN = XNP*(UI + 2*XN)/((UI + 2*XNP)*(UI + XN))
      AN = RN*ANP
C     BN = -BNP*(NP - UI)*RN/(TUI + N)
      BN = -BNP*(XNP - UI)*RN/(TUI + XN)
      JP = 2
      GO TO 90
   89 CONTINUE
      JP = 1
   90 CONTINUE
      NU = NU - 1
      U = U - 1.
      BJP = BJ
      BJ = BJM
   87 CONTINUE
   95 CONTINUE
      BBJ(NU + 1) = BJ
      DO 105I=1,NTM
      BJM = TOI*U*BJ - BJP
      BBJ(NU) = BJM
      GO TO (97,98),JP
   97 CONTINUE
      ANP = AN
      BNP = BN
```

```
      NP = N
      CN = CN + ANP*BJP
      FY = FY + BNP*BJP
      FYP = FYP + 0.50*(BND-BNP)*BJD
      BND = BNP
      BJD = BJ
      N = N - 1
C     RN = NP*(UI + 2*N)/((UI + 2*NP)*(UI + N))
      XN = N
      XNP = NP
      RN = XNP*(UI + 2*XN)/((UI + 2*XNP)*(UI + XN))
      AN = RN*ANP
C     BN = -BNP*(NP - UI)*RN/(TUI + N)
      BN = -BNP*(XNP - UI)*RN/(TUI + XN)
      JP = 2
      GO TO 99
   98 CONTINUE
      JP = 1
   99 CONTINUE
      NU = NU -1
      U = U - 1.
      BJP = BJ
      BJ = BJN
  105 CONTINUE
      AN = ANOT
      IF(UI - 0.00005)701,701,107
  701 CONTINUE
      BN = TOP*(0.5772156649 + ALOG(XOT))
      GO TO 108
  107 CONTINUE
      BN = 1./TANF(UI*PI) - ANOT*ANOT/(PI*UI)
  108 CONTINUE
      CN = CN + AN*BBJ(1)
      FYP = -W*AN*AN*BBJ(1) + (BN - 0.50*BNP)*BBJ(2) + 3.*UI*FY/X
     * - FYP
      FY = FY + BN*BBJ(1)
      CN = 1.0/CN
      FY = CN*FY
      FYP = CN*FYP
      FJ = CN*BBJ(1)
      FJP = CN*BBJ(2)
      IF(DABS(FJ) - 0.001)50,501,501
  501 CONTINUE
      FYP = (FY*FJP - W)/FJ
C
C     NORMALIZE THE J(NU) AND CALCULATE THE Y(NU).
C
   50 CONTINUE
      BBJ(1) = FJ
      BBJ(2) = FJP
      BBY(1) = FY
      BBY(2) = FYP
      U = UI + 1.
      NT = NTH - 1
```

```
      DO 51K=1,NT
      BBY(K+2) = TOX*U*BBY(K+1) - BBY(K)
      BBJ(K+2) = CN*BBJ(K+2)
      U = U + 1.
   51 CONTINUE
      RETURN
      END



      FUNCTION GAMMA(X)
      IMPLICIT DOUBLE PRECISION (A-H,O-Z)
      IF(DABS(X) - 1.0E-06)100,100,10
  100 CONTINUE
      WRITE (LUNO,101)
  101 FORMAT(/20X,17HSTOP 101 IN GAMMA  /)
      STOP
   10 CONTINUE
      SUM = X + X*X*(0.5772156649 + X*(-0.6558780715 + X*(
     * -0.0420026350 + X*(0.1665386114 + X*(-0.0421977346 +
     * X*(-0.0096219715 +X*(0.0072189432 + X*(-0.0011651676 +
     * X*(-0.0002152417 +X*(0.0001280503 +X*(-0.0000201348 +
     * X*(-0.0000012505 + X*(0.0000011330 +X*(-0.0000002056 +
     * X*(0.0000000061 + X*(0.0000000050 + X*(-0.0000000012 +
     * X*0.0000000001)))))))))))))))
      GAMMA = 1.0/SUM
      RETURN
      END
```

```
      SUBROUTINE RDAPRT(IPLT,N1,N2,KEY)
      COMMON /LUNIO1/LUNI,LUNO,LUNO1
      DIMENSION IPLT(N2)
  903 FORMAT(20A4)
      GO TO (1,2,3),KEY
    1 CONTINUE
      READ (LUNI,903) (IPIT(I),I=N1,N2)
      GO TO 4
    2 WRITE (LUNO,903) (IPLT(I),I=N1,N2)
      GO TO 4
    3 WRITE (LUNO1,903) (IPLT(I),I=N1,N2)
    4 CONTINUE
      RETURN
      END



      SUBROUTINE RDPRNT(A,NMAX,MMAX,NDIM,KEY2)
      IMPLICIT REAL*8 (A-H,O-Z)
      COMMON /LUNIO1/LUNI,LUNO,LUNO1
      DIMENSION A(NDIM,1)
  898 FORMAT(5E16.8)
  899 FORMAT(8E13.5)
  900 FORMAT(8E10.5)
      DO 1I=1,MMAX
      GO TO (2,3,5),KEY2
    2 CONTINUE
      READ (LUNI,900) (A(I,J),J=1,MMAX)
      GO TO 4
    3 CONTINUE
      WRITE (LUNO,899) (A(I,J),J=1,MMAX)
      GO TO 4
    5 CONTINUE
      WRITE (LUNO1,898) (A(I,J),J=1,MMAX)
    4 CONTINUE
    1 CONTINUE
      RETURN
      END
```

```
      SUBROUTINE ALVCPD(X,N,AL,IOPT)
      IMPLICIT REAL*8 (A-H,O-Z)
      DIMENSION X(N)
      DO 3I=1,N
      GO TO (1,2),IOPT
    1 CONTINUE
      X(I)=X(I)*AL
      GO TO 3
    2 CONTINUE
      X(I)=X(I)/AL
    3 CONTINUE
      RETURN
      END



      SUBROUTINE RDIPRT(J,IMIN,IMAX,IGTOPT)
      COMMON /LUNIO1/LUNI,LUNO,LUNO1
      DIMENSION J(IMAX)
  901 FORMAT(16I5)
      GO TO (1,2,4),IGTOPT
    4 CONTINUE
      WRITE (LUNO1,901) (J(I),I=IMIN,IMAX)
      GO TO 3
    1 CONTINUE
      READ (LUNI,901) (J(I),I=IMIN,IMAX)
      GO TO 3
    2 CONTINUE
      WRITE (LUNO,901) (J(I),I=IMIN,IMAX)
    3 CONTINUE
      RETURN
      END
```

```
      SUBROUTINE VCTINT(X,N,XO)
      IMPLICIT REAL*8 (A-H,O-Z)
      DIMENSION X(N)
      DO 1I=1,N
      X(I)=0.0+XO
    1 CONTINUE
      RETURN
      END


      SUBROUTINE MTRINT(A,N,M,NDIM,AO)
      IMPLICIT REAL*8 (A-H,O-Z)
      DIMENSION A(NDIM,1)
      DO 1I=1,N
      DO 2J=1,M
      A(I,J)=0.0+AO
    2 CONTINUE
    1 CONTINUE
      RETURN
      END
```

```
      SUBROUTINE XVECCL(X,NP1,NDIM,X1,X2,DELX,IOPT)
      COMMON /LUNIO1/LUNI,LUNO,LUNO1
      IMPLICIT REAL*8 (A-H,O-Z)
      DIMENSION X(NDIM)
  897 FORMAT(/)
  898 FORMAT(///)
  899 FORMAT(5E16.8)
  901 FORMAT(16I5)
      WRITE (LUNO,899)X1,X2,DELX
      GO TO (2,3,2),IOPT
    2 CONTINUE
      NOX1X2=(X2-X1)/DELX+1.01
      WRITE (LUNO,901)NOX1X2,NP1
      NOX1X2=IABS(NOX1X2)
      NP1=NOX1X2
      WRITE (LUNO,901)NOX1X2,NP1
      GO TO (4,3,3),IOPT
    4 CONTINUE
      XCONST=1.0
      IF(X1.GT.X2)XCONST=-XCONST
      DO 1I=1,NOX1X2
      X12IM1=I-1
      X(I)=X1+X12IM1*DELX*XCONST
    1 CONTINUE
      GO TO 5
    3 CONTINUE
      CALL RDPRT1(X,NP1,1)
    5 CONTINUE
      WRITE (LUNO,897)
      CALL RDPRT1(X,NP1,2)
      PRINT 898
      RETURN
      END
```

```
      SUBROUTINE RDPRT1(X,N,KEY2)
      IMPLICIT REAL*8 (A-H,O-Z)
      COMMON /LUNIO1/LUNI,LUNO,LUNO1
      DIMENSION X(N)
  897 FORMAT(8E10.5)
  898 FORMAT(5E16.8)
  899 FORMAT(8E13.5)
  900 FORMAT(8E10.3)
      GO TO (1,2,4,5,6),KEY2
    6 CONTINUE
      WRITE (LUNO1,898) (X(I),I=1,N)
      GO TO 3
    5 CONTINUE
      READ (LUNI,898) (X(I),I=1,N)
      GO TO 3
    4 CONTINUE
      WRITE (LUNO1,898) (X(I),I=1,N)
      GO TO 3
    1 CONTINUE
      READ (LUNI,900) (X(I),I=1,N)
      GO TO 3
    2 CONTINUE
      WRITE (LUNO,899) (X(I),I=1,N)
    3 CONTINUE
      RETURN
      END
```

```
      SUBROUTINE CPLXQR(Y1R,Y1I,Y2R,Y2I,Y3R,Y3I,NOPT)
      IMPLICIT REAL*8 (A-H,O-Z)
      REAL*8 X1R,X1I,X2R,X2I,X3R,X3I
      X1R = Y1R
      X1I = Y1I
      X2R = Y2R
      X2I = Y2I
      X3R = 0.0
      X3I = 0.0
      GO TO (1,2), NOPT
    1 CONTINUE
      X3R = X1R*X2R - X1I*X2I
      X3I = X1R*X2I + X1I*X2R
      GO TO 3
    2 CONTINUE
      XDEN = X2R*X2R + X2I*X2I
      X3R = X1R*X2R + X1I*X2I
      X3I = X1I*X2R - X1R*X2I
      X3R = X3R/XDEN
      X3I = X3I/XDEN
    3 CONTINUE
      Y3R = X3R
      Y3I = X3I
      RETURN
      END



      FUNCTION TANF(X)
      IMPLICIT REAL*8 (A-H,O-Z)
      SINX = DSIN(X)
      COSX = DCOS(X)
      IF(DABS(COSX) - EPSTOL) 1,1,2
    1 CONTINUE
      COSX = EPSTOL
    2 CONTINUE
      TANF = SINX/COSX
      RETURN
      END
```

1983 USAF-SCEE SUMMER FACULTY RESEARCH PROGRAM

SPONSORED BY THE

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

CONDUCTED BY THE

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

USE OF COLLOIDAL GAS APHRONS FOR TREATING HAZARDOUS WASTES

Prepared by:	Donald L. Michelsen Assistant Professor
	Wade L. Auten Graduate Student
Department and University:	Department of Chemical Engineering Virginia Tech Blacksburg, VA 24061
Research Location:	Environics Laboratory Engineering and Services Center Tyndall AFB, FL 32403
USAF Research:	Captain Terry Stoddart
Date:	August 31, 1983
Contract No.	F 49620-82-C-0035

USE OF COLLOIDAL GAS APHRONS (CGA'S) FOR TREATING HAZARDOUS WASTES

by
Donald L. Michelsen
Wade L. Auten

Abstract

Colloidal Gas Aphrons, a microdispersion of stable small air bubbles (15 to 50 mμ) in air, have proven to be effective in separation processes. The purpose of this laboratory study was to explore their effectiveness in treating hazardous wastes by 1) evaluating their flow characteristi s through soil matrices (for removal or scouring of pollutants and for in-situ biodegradation); 2) quantifying their adhesion and retention in various saturated matrices (impoundments) for subsequent in-situ biodegradation; and 3) determining the feasibility of an extraction/flotation process for removing ortho dichlorobenzene (dissolved hazardous organics) from contaminated water using CGA's for flotation.

Extensive results on flow of CGA through a coarse golf course sand documented the importance of surfactant, type and concentration, and pretreatment on CGA flow characteristics.

The adhesion and retention of CGA's in a flooded impoundment were very promising. Water was displaced with 0.56 to 1.25 air/water volume ratio injected in the saturated sand and 75% retention of these bubbles one month later.

Finally, while testing was limited, CGA's provided encouraging phase separation (clarification) in an extraction/flotation process for removing soluble ortho dichlorobenzene from water using a good decane solvent (high partition coefficient).

Acknowledgement

The authors would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for the opportunity and financial support to conduct research with the Environic Laboratory of the Engineering Services Center located at Tyndall AFB, Florida. They would like to praise the entire Environics Division for general support and for making the experience meaningful. The cooperation level was very high indeed.

From a personal standpoint, the authors would like to thank Captain Terry Stoddart and Captain Randy Gross for their helpful guidance and comments based on considerable laboratory and field experience. Their interaction was important. Finally, without the assistance, cooperation and jesting of Sergeants Brown, Manikas, Niemeier, and Storch, the work would not have been completed and certainly not as pleasantly.

I. INTRODUCTION

Colloidal Gas Aphrons can be made by mixing a small amount of a surfactant into water (nominally 0.25 to 1.0 g. per l.) and agitating in a very high speed mixer. The resulting cream can contain as much as 70% small air bubbles dispersed in water. In a CGA, the gas or bubble is encapsulated in a thin "soap" film that is so tenacious, the bubbles do not coalesce, even when pressed together. Hence, the bubbles remain very small (on the order of 25 μm) and present an enormous surface area. CGA's must be clearly distinguished from the so-called "bubbles" produced by dissolved air precipitation, sparging, or electrolysis which are perhaps 100 to 1,000 times larger, rise to the surface rapidly and then coalesce quickly. CGA's were first investigated by Sebba (1971, 1975) and named microfoams because they were thought to be foams made with small bubbles. Subsequently, closer investigation (Sebba, 1982) showed these were really true bubbles encapsulated in a thin film of water considerably thicker than a monolayer.

Because CGA's are so stable and small, they remain suspended in solution and can flow through channels such as exist in a sand bed. Larger, unstable "bubbles" would be filtered out in such a situation. Another important characteristic is that CGA's can be pumped by conventional methods (e.g., centrifugal pumps) without deterioration. Finally, techniques for improved flotation separation have evolved with the development of these new types of bubble systems. A number

of these applications have been discussed in an article by Sebba and Barnett (1981).

The key properties of CGA's are as follows:

1.) CGA's are relatively stable to temperatures close to boiling point.

2.) Viscosity and flow characteristics of CGA's in water are similar to water phase and, therefore CGA's can be pumped easily.

3.) Small size, $10^{-14}m^3$ in volume, when 25 μ diameter, determines colloidal properties.

4.) Have a durable elastic encapsulating shell akin to the soap film surrounding an ordinary soap bubble.

5.) Coalescence does not occur when bubbles collide, though they may cluster then drift away.

6.) Smaller bubbles are under higher pressure than larger bubbles. Thus, smaller ones disappear and larger ones grow thus enhancing creaming at the surface.

7.) Water surface adjacent to the bubble will have a monolayer of surfactant oriented hydrophobic groups facing the bubble.

8.) CGA's can be partially coated with selective solvents for extraction, adsoprtion.

Because of the unique characteristics and potential as a processing technique, this work with the Environics Laboratory during the Summer, 1983 has been directed toward the feasibility of using

colloidal gas aphrons for resolving several hazardous waste problems.

II. OBJECTIVES OF THE RESEARCH EFFORT

Three different situations were targeted for the use of CGA's in treating hazardous waste and work was completed to determine the technical feasibility of each. The principle objective of the summer was to study the flow characteristics of CGA's through one dimensional consolidated soil matrices. The long range purpose of this effort was to determine the possibility of delivering CGA's (air and water) to and through various soil matrices in order to flush (scour) hazardous materials from these soil matrices and in addition to enhance in situ biodegradation of organic contaminants. The technique could have application for treating release of jet fuels, trichloroethylene, and gasoline which have migrated into the subsoils. Limited earlier studies had demonstrated the ability of CGA's to scour out gasoline in one-dimensional soil column tests, to penetrate a soil matrix (with microorganisms and nutrients) and to be retained in the soil matrix for a long period of time.

The objective of the second portion of this research effort was to produce CGA's from various surfactant and surfactant concentrations, and quantify their adhesion (removal) and retention in various saturated soils and selected materials. Limited tests completed at Virginia Tech had shown earlier that CGA's are retained in an unconsolidated soil matrix covered with water such as an impoundment. The development of these techniques for injecting air

(or oxygen) into unconsolidated soils (impoundments) would provide a means to alter impoundment sediments and/or swamps from anaerobic to aerobic conditions. These aerobic conditions would encourage localized in situ biodegradation of hazardous organics with or without the simultaneous injection of aerobic microorganisms or nutrients. Another possibility would be to use CGA's to deliver air (oxygen), microorganisms and nutrients into a french drain/biodegradation ditch filled with a porous matrix in order to biodegrade (treat), in situ, a slow flowing contaminated aquifer.

Finally, CGA's have displayed very effective removal of suspended fines, and emulsified organics from waste water because of their high surface area and excellent (slow rising) flotation characteristics. Limited work had also been documented on the removal of dissolved organics (phenol) from waste waters using solvent coated CGA's. The objective of the third portion of this research project was to evaluate the removal of low concentrations of dissolved ortho dichlorobenzene (5 to 10 mg/l) using small quantities of selective low solubility organics. The first step was to screen several solvents for their ability to extract ortho dichlorobenzene from water. Subsequently, the technique was to disperse and/or emulsify the solvent with a brief, intensive agitation and extract the ortho dichlorobenzene from water.

Then, several CGA's were tested for their ability to float (separate) the fine oil droplet (extract). Minimal solvent usage and losses are critical, particularly if treating large quantities of

contaminated ground water. The use of an extraction/flotation technique has application for removing dissolved hazardous contaminants from ground water perhaps after air stripping to remove the lower boiling fractions. The extraction/flotation treatment might prove to be an attractive alternate for activated carbon adsorption.

III. THE FLOW OF COLLOIDAL GAS APHRONS (CGA'S) THROUGH A MEDIUM SAND CONSOLIDATED MATRIX

The primary effort for the summer was to determine the possibility of delivering air (or oxygen) to and through a medium golf course sand (14 to 115 Tyler mesh) containing small amounts of fine clay (silt). The hydraulic conductivity was measured between 120 to 160 cm/hr for each test. The first author of this report concentrated his first 4 weeks on setting up test facilities, debugging the system, conducting 6 preliminary tests, and developing a work plan for Wade Auten, the second author, to carry out beginning in late June 1983.

Figure 1 shows a flow diagram of the final test apparatus used for conducting the one-dimension soil pack column tests. The system had the capability of generating CGA's from a wide range of surfactant solution, and delivering them at constant pressure with recycle to a one dimensional soil pack for testing. A back pressure controller assured constant upstream pressure with the recycle returning to the CGA generator.

A laboratory CGA stability test of potential surfactants was completed first to screen surfactants for column testing. These results are shown on Figure 2. The steeper the slope, the more stable

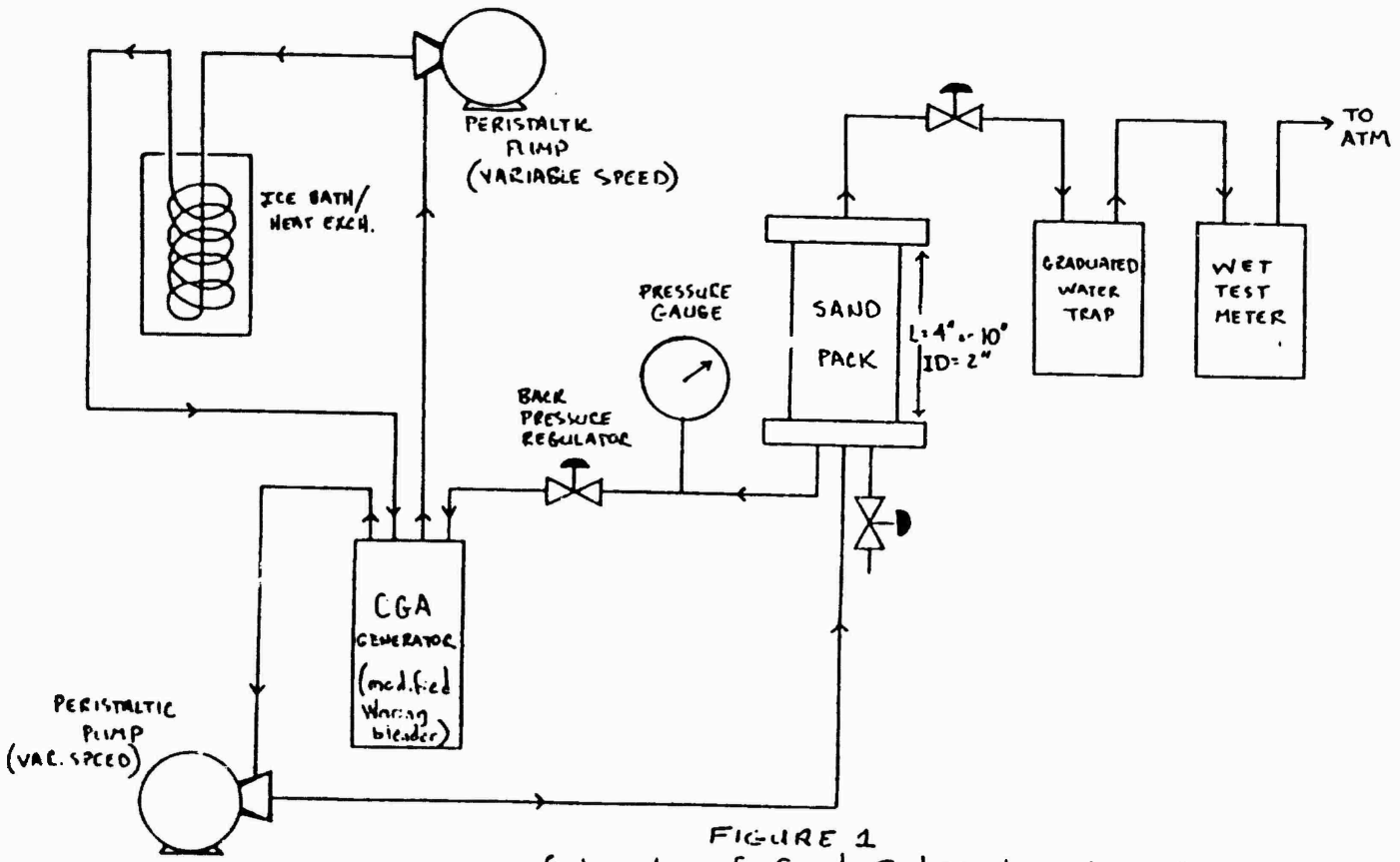


FIGURE 1
Schematic of Sand Bed Testing Apparatus

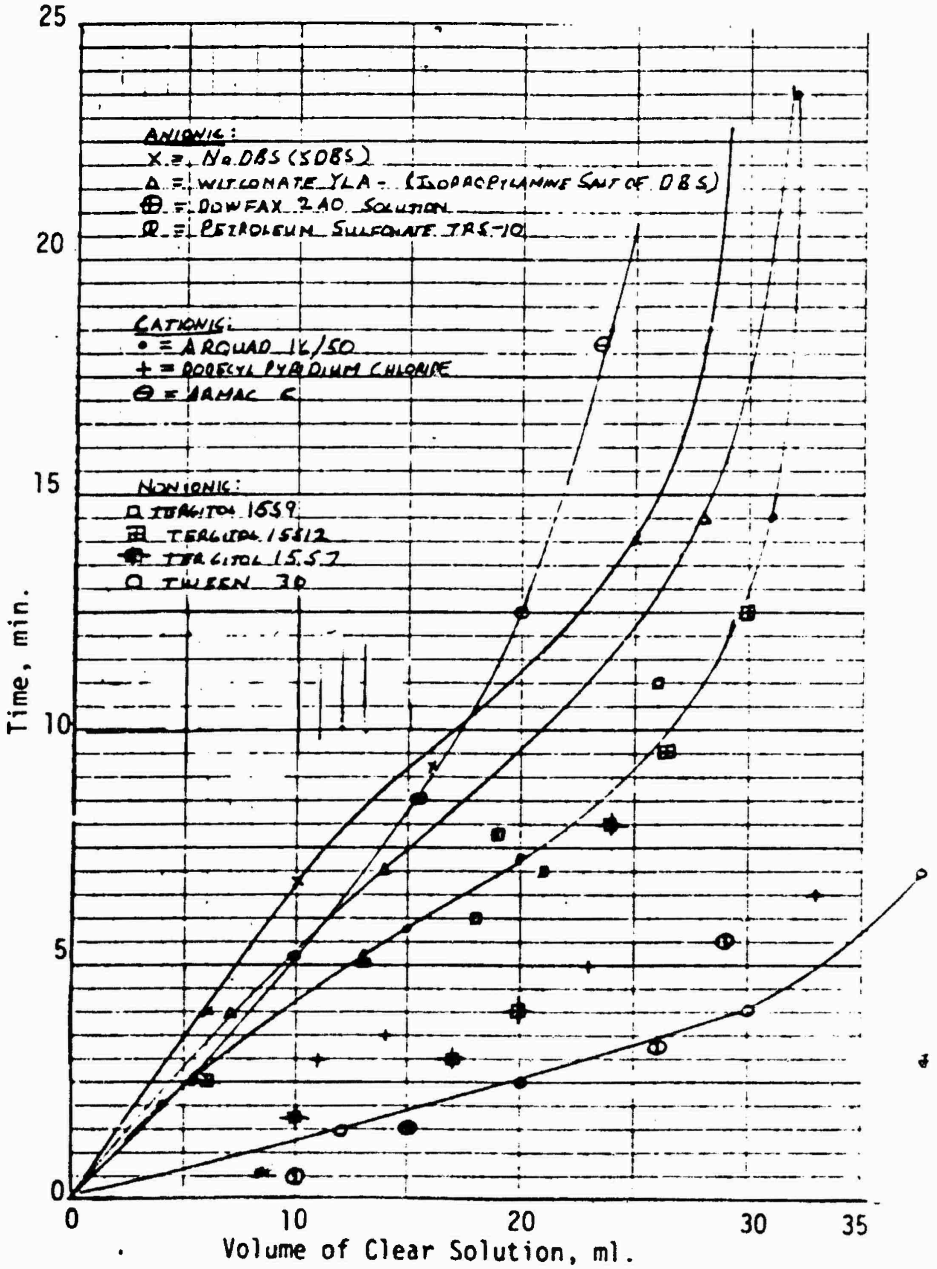


Figure 2

Stability of CGA's in 100 Ml.

Settling Test

the CGA and according more likely to remain coherent in flowing through a soil matrix, assuming no soil or contaminatn interference. On the other hand, if the goal is only to deliver CGA's to or just into a soil surface before breakdown, then an unstable CGA may be desired. Surfactants for testing in the one dimensional soil pack tests were chosen based on these laboratory studies.

Having screened several surfactants, the major effort then was to conduct a series of tests using the sand bed testing apparatus. The experimental procedure used for the column tests was as follows:

1. Size locally-obtained quartz sand to 14- x 115-mesh.
2. Clean sand by repeated rinsing.
3. Carefully pack column to insure:
    a) complete water saturation, and
    b) most stable packing density.
4. Flush with water until:
    a) effluent water is completely clear, and
    b) effluent flow rate is constant for a given pressure drop.
5. Determine hydraulic conductivity (a measure of permeability).
6. Prepare surfactant solution(s) to desired concentrations.
7. Pretreat column if desired by flushing concentrated surfactant solution (at least two pore volumes) through column.

8. Begin CGA generation.
9. Recycle CGA's through and ice bath to control temperature.
10. Add CGA's to sand pack at a controlled pressure.
11. Monitor one or more of the following variables:
    a) time,
    b) movement of CGA interface through column,
    c) pressure drop across the column (usually independent),
    d) total effluent flow rate,
    e) liquid effluent flow rate,
    f) first appearance of true CGA's in effluent,
    g) CGA generation quality.
12. At end of test, reflush column with water until pressure drop stabilizes.
13. Repeat hydraulic conductivity measurement (optional).

Based on the initial studies there is resistance to CGA flow through consolidated soil systems compared to either air or water. Several studies were completed on the movement of CGA's at varying pressure drops through a sand-packed glass column. Figure 1 snows the apparent relationship between CGA movement and pressure drop when CGA's made from 0.5 g/l surfactant solutions are pumped through a packed 4-inch column. From Figure 3 it can be seen that the anionic surfactant produced a CGA which displaced the greatest volume of water from the saturated sand pack. Conversely, CGA's made from a cationic

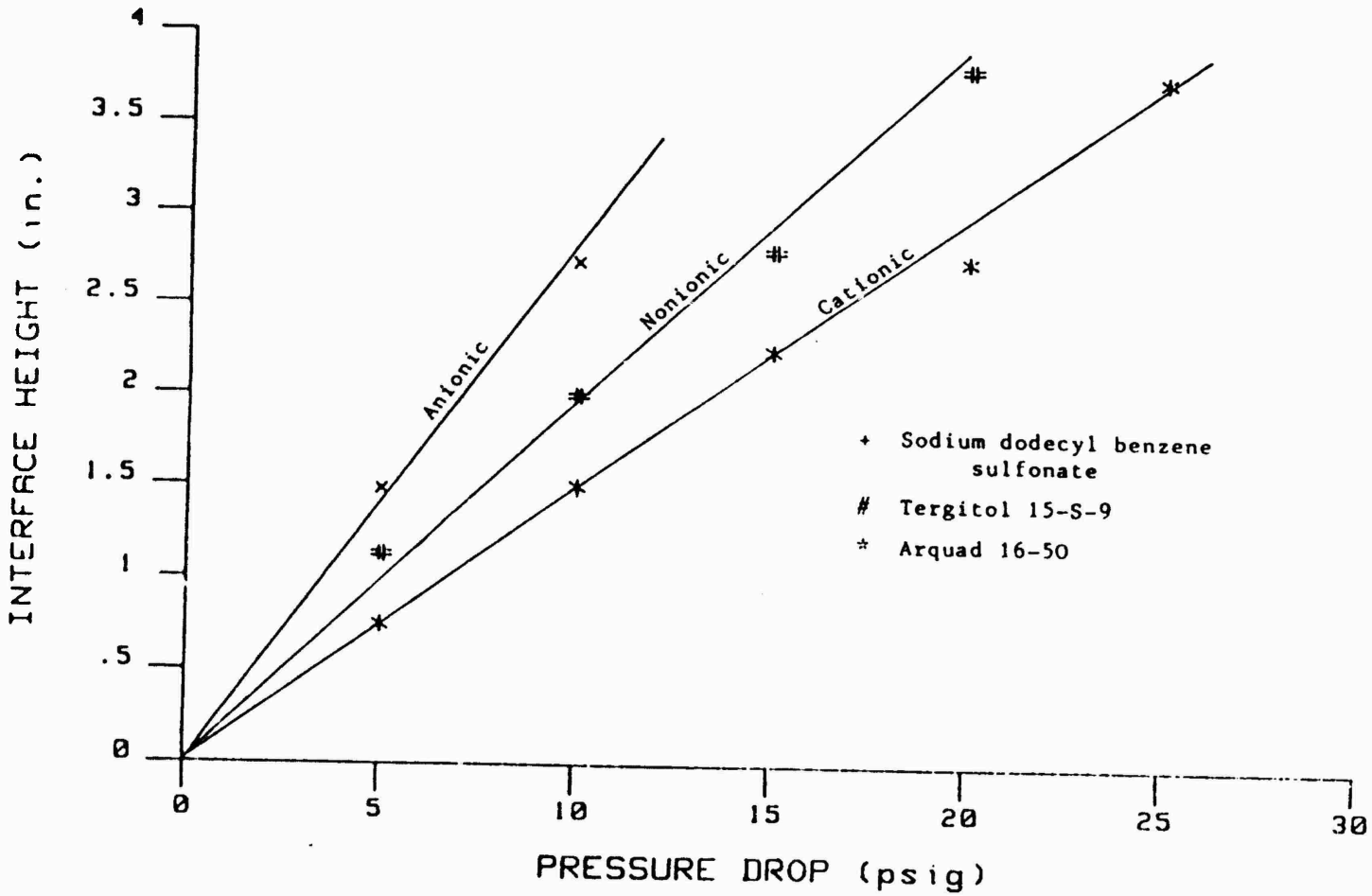


Figure 3: Water displacement by CGA's from a 4-inch packed column as a function of pressure. No column pretreatment.

surfactant displaced the least volume of water at a given pressure while the nonionic surfactant gave intermediate results. This trend is not surprising if the net surface change on quartz sand is negative, as expected. In such a case, the number of negatively-charged surface sites is greater than the number of positively-charged sites and so cationic surfactants would be preferentially adsorbed. Cationic CGA's would then be destablized more quickly by the bed than either nonionic or anionic CGA's.

Based on these results, the resistance to CGA flow through a sand pack was thought to be primarily a function of the degree of loading of the bed with CGA's. It is known that, under some circumstances, CGA's can be approximated as rigid spheres. Because the internal pressure in a bubble is inversely proportional to the bubble's radius, micron-size bubbles like CGA's can have significant internal pressures relative to their environment. If CGA's remain stable after entry into a soil system, their effect on the pack should be similar to that observed if solids, an order of magnitude smaller in size than the sand particles, were loaded into the interstitial regions of the bed. In such an event, the permeability would be observed to decrease linearly with increased loading. On a macroscopic scale, the increased loading would also increase the pressure drop across the pack required for additional loading. This simple physical model was originally thought to account for the observed linearity in the data in Figure 3. But if so, then the relationship between CGA movement and pressure drop should be independent of surfactant type and a

function only of the CGA size.

A more accurate model of CGA flow behavior through a soil matrix requires consideration of surface effects. Figure 3 clearly indicates some degree of selectivity in the propagation of CGA's with different surface properties through a sand pack. In reality, CGA's moving through the pore volume of a sand bed are not only preferentially adsorbed onto the matrix but are also destabilized as a result of surfactant stripping. This was shown in a test conducted in a 10-inch column, Figure 4. During this test, an attempt was made to flow CGA's through a medium to medium-coarse sand matrix. Preliminary results (Figure 3) had shown that significant pressure drops would be required to get CGA movement over even short distances. However, by pretreating with surfactants, we were able to show that even at pressure drops as low as 2.5 psi essentially all of the water initially saturating a sand pack could be displaced with CGA's. The sand pack was pretreated with a moderately concentrated surfactant solution (10.0 g/l of anionic sodium dodecyl benzene sulfonate, SDBS) which minimized subsequent surfactant stripping on addition of CGA's. Essentially, the stability of the CGA's (made from 1.0 g/l SDBS) was enhanced and they were capable of traversing the sand pack intact.

Presumably, pretreatment of sand pack with a concentrated surfactant of the same ionic character as the CGA's to be added subsequently just served to accelerate the time required to reach steady-state flow through the bed. In this light, the points in Figure 3 may be viewed as representing pseudo-steady state conditions.

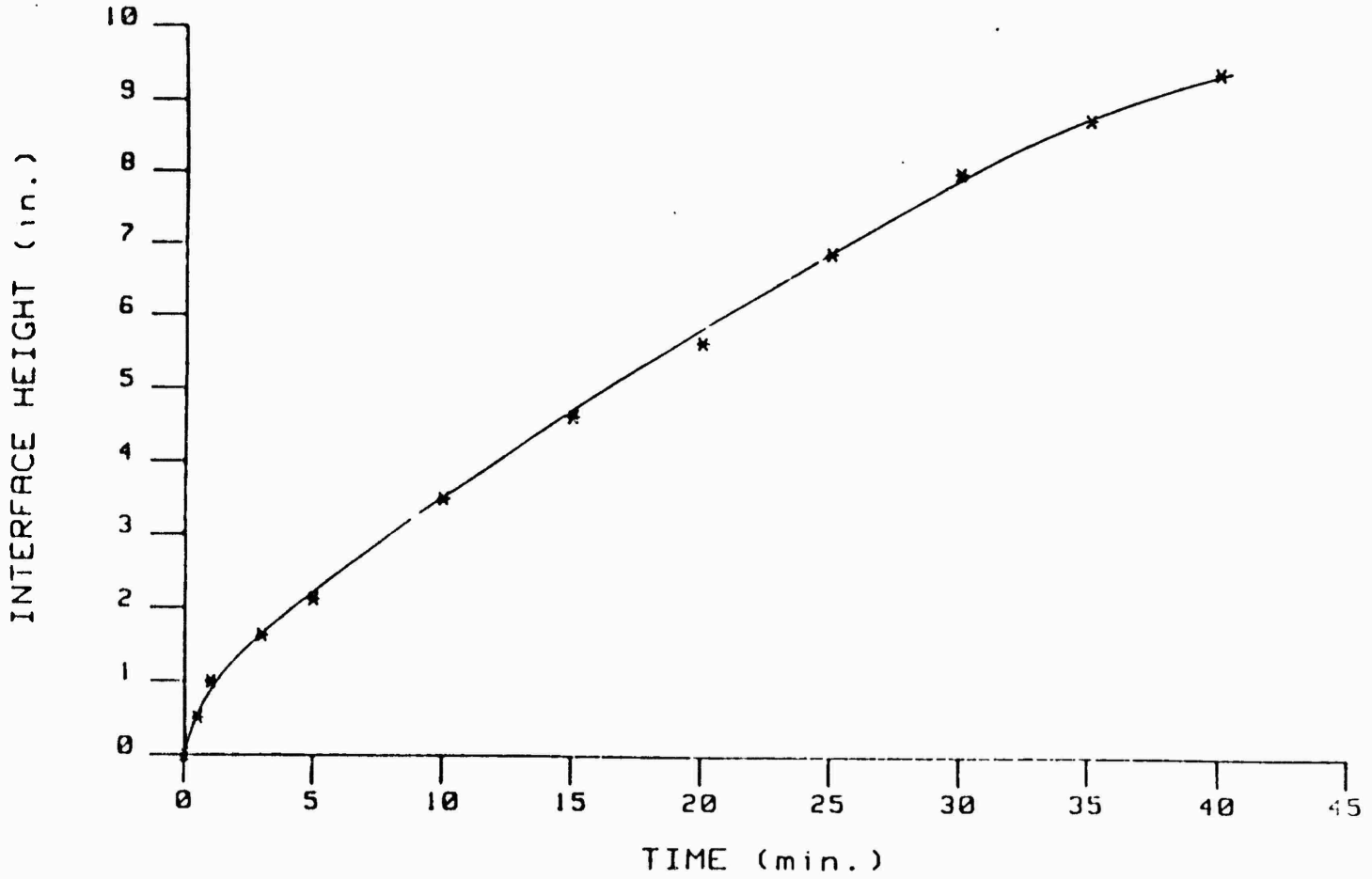
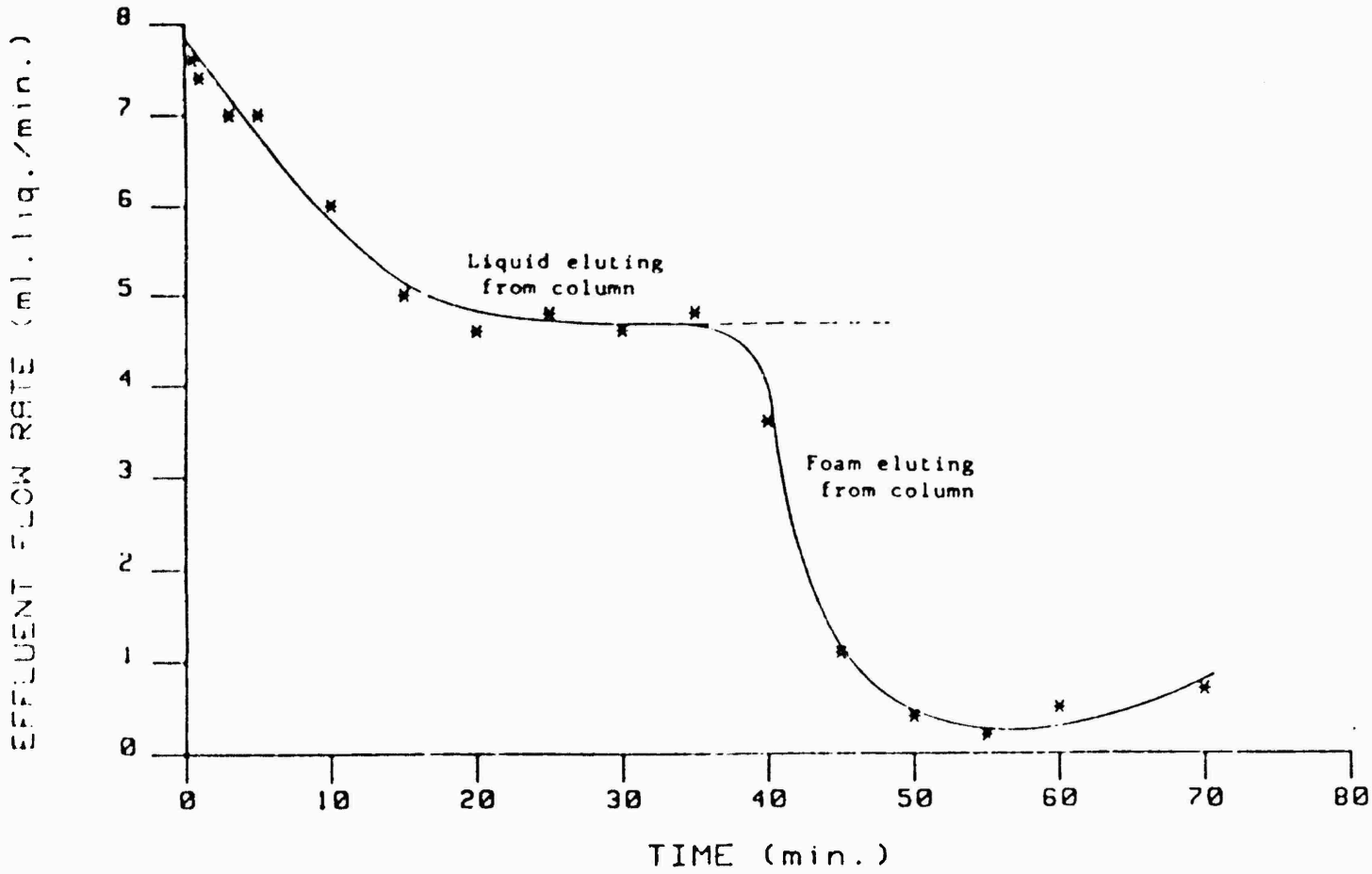


Figure 4: Water displacement by CGA's from a 10-inch column as a function of time at a constant pressure of 2.5 psig. (CGA's made with 1.0 g/l SDBS; pretreatment with 600 ml of 10.0 g/l SDBS).

So for a given pressure drop, CGA movement through a bed is limited by the rate of adsorption of surfactant onto the matrix. Only after a region of the bed is stabilized by sufficient surfactant adsorption can the CGA front move further into the pack.

Figure 5 shows effluent flow rate data for the same test which yielded Figure 4. Until about 40 minutes the CGA front was actively displacing water initially, saturating the sand pack and no effluent liquid flow rate was greater than 4 ml/min. At about 20 minutes into the test, the effluent flow rate seemed to statilize at about 4.6 ml/min. Presumably, this rate would have been maintained for an even longer period given a longer column. But at about 40 minutes (ref. Figure 4) the CGA front had essentially displaced all of the initial water in the column. At this time, as noted in Figure 5, the first appearance of air, as foam, occurred at the cell outlet. The net liquid flow rate dropped substantially even though the combined water and air flow rate was observed to increase to about 8 ml/min.

Figure 6 shows the effect of pressure drop on effluent flow rate for a system in which a surfactant-poor CGA made with a 0.05 g/l SDBS solution (anionic) was pumped through a 4-inch packed column. In this particular test, the sand pack was not pretreated with any surfactant. Yet, CGA's comparable to freshly generated CGA's in quality were eluted at pressures about 15 psi. and the effluent flow rate climbed dramatically as the effluent drop approached 30 psi, the maximum operating pressure for the apparatus. Obviously, at pressures above 15 psi internal resistances to flow had been negated. If these



91-89

Figure 5: Effluent flow rate from a 10-inch packed column at a constant pressure of 2.5 psig. (CGA's made with 1.0 g/l SDBS; pretreatment with 600 ml of 10.0 g/l SDBS).

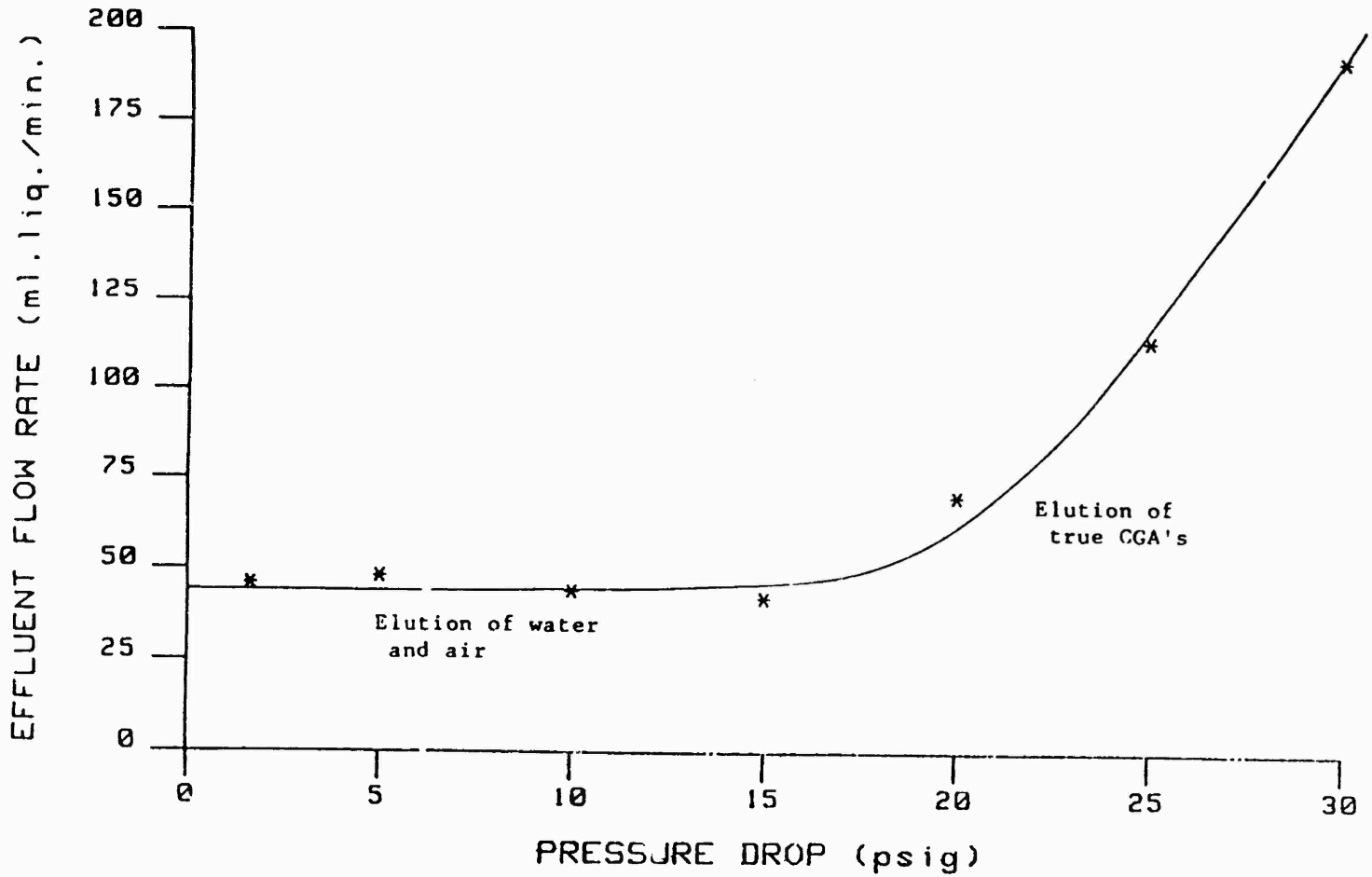


Figure 6: Effluent (liquid) flow rate as a function of pressure. (Pressure increased at regular 5-minute intervals; CGA's made with 0.05 g/l SDBS; no column pretreatment).

resistances are primarily due to CGA adsorption, as believed, then this pressure represents the minimum force needed to prevent adsorption of CGA's onto the sand matrix. At pressures below this critical value, effluent flow is essentially independent of pressure and contains no visible sign of CGA's.

Flow in the high pressure regime of Figure 6 is observed to increase at a dramatic rate at these high pressures. Flow behavior seems practically independent of the matrix. Because CGA diameters are at least an order of magnitude smaller than the sand comprising the matrix, their flow at these pressures should closely approximate those of a pure liquid, assuming negligible surface effects under these conditions. So, to have similar flow behavior at lower pressures where surface effects may be significant, effort must be made to condition the system so that these effects are minimized. As mentioned previously, one approach to achieving this end is pretreatment of the sand matrix so that surfactant adsorption from the CGA's is minimized. In Table I, data are compiled showing the relative effects of surfactant pretreatment on CGA movement through a packed column using CGA's made from an anionic, nonionic, and cationic surfactant. In each case, a packed 4-inch column was pretreated with 300 ml of 10.0 g/l of surfactant of the same type as the CGA's to be added. In each case, surfactant pretreatment is observed to decrease the pressure drop needed to achieve complete displacement of the water initially saturating the sand pack.

TABLE I

EFFECT OF SURFACTANT PRETREATMENT ON CGA

PROPAGATION THROUGH A SAND PACK (UPFLOW)

A. ANIONIC SAA, SDBS - 0.5 g/l CGA

	CGA ADVANCE (HEIGHT OF INTERFACE/COLUMN LENGTH)	
ΔP (psig)	NO PRETREATMENT	W/PRETREATMENT
5	0.38	complete
10	0.69	-
15	complete	-
30	foam	CGA

B. NONIONIC SAA, TERGITOL 15-S-9 - 0.5 g/l CGA

	CGA ADVANCE (HEIGHT OF INTERFACE/COLUMN HEIGHT)	
ΔP (psig)	NO PRETREATMENT	W/PRETREATMENT
5	0.28	0.63
10	0.50	0.78
15	0.70	complete
1[illegible]	-	CGA
20	0.95	
25	complete (foam)	

C. CATIONIC SAA, ARQUAD 16-50 - 0.5 g/l CGA w/no pretreatment
0.1 g/l CGA w/pretreatment

	CGA ADVANCE (HEIGHT OF INTERFACE/COLUMN HEIGHT)	
ΔP (psig)	NO PRETREATMENT	W/PRETREATMENT
5	0.19	0.69
10	0.38	complete (foam)
15	0.56	-
20	0.69	-
25	0.94	-
30	complete (foam)	still foam

Figure 7 shows elution volume as a function of time for a flow system in which a 0.08 g/l SDBS solution was used to make the CGA. Plotted are data taken at three relatively low pressures so that complete displacement of the liquid initially occupying the sand pack never occurred. The data show the predicted increase in elution rate, or volume, with increasing pressure which is constant once the CGA front has been established. Of more significance is the apparent linearity which exists between elution volume (rate) with time at a given pressure. The nonlinear regions reflected at the beginning of each pressure change just reflects the step-wise movement of the CGA interface through the column and the subsequent water displacement.

Figure 8 shows that by reducing the concentration of anionic surfactant used for making CGA's, the total elution volume increases markedly. At lower concentration, the CGA's destablize faster because of the low surfactant concentration used in making the CGA's and more rapid adsorption of the surfactant upon entering the sand pack. As indicated on Figure 9, after displacing the interstitial water (water initially in sand pack) the bulk flow is mostly air. Surprisingly, in testing with the nonionic Tergitol 15-S-9 surfactant, Figure 10, the elutant flow rate was considerably lower (1/20) the flow of elutant when using a CGA produced from SDBS. This suggests retained stability of the CGA's made from this nonionic detergent. In addition, the air flow measured as a percentage of total effluent volume was also lower, see Figure 11. Notice, also that the leveling off of air ratio was not as prominent for the nonionic. There is order difference in terms

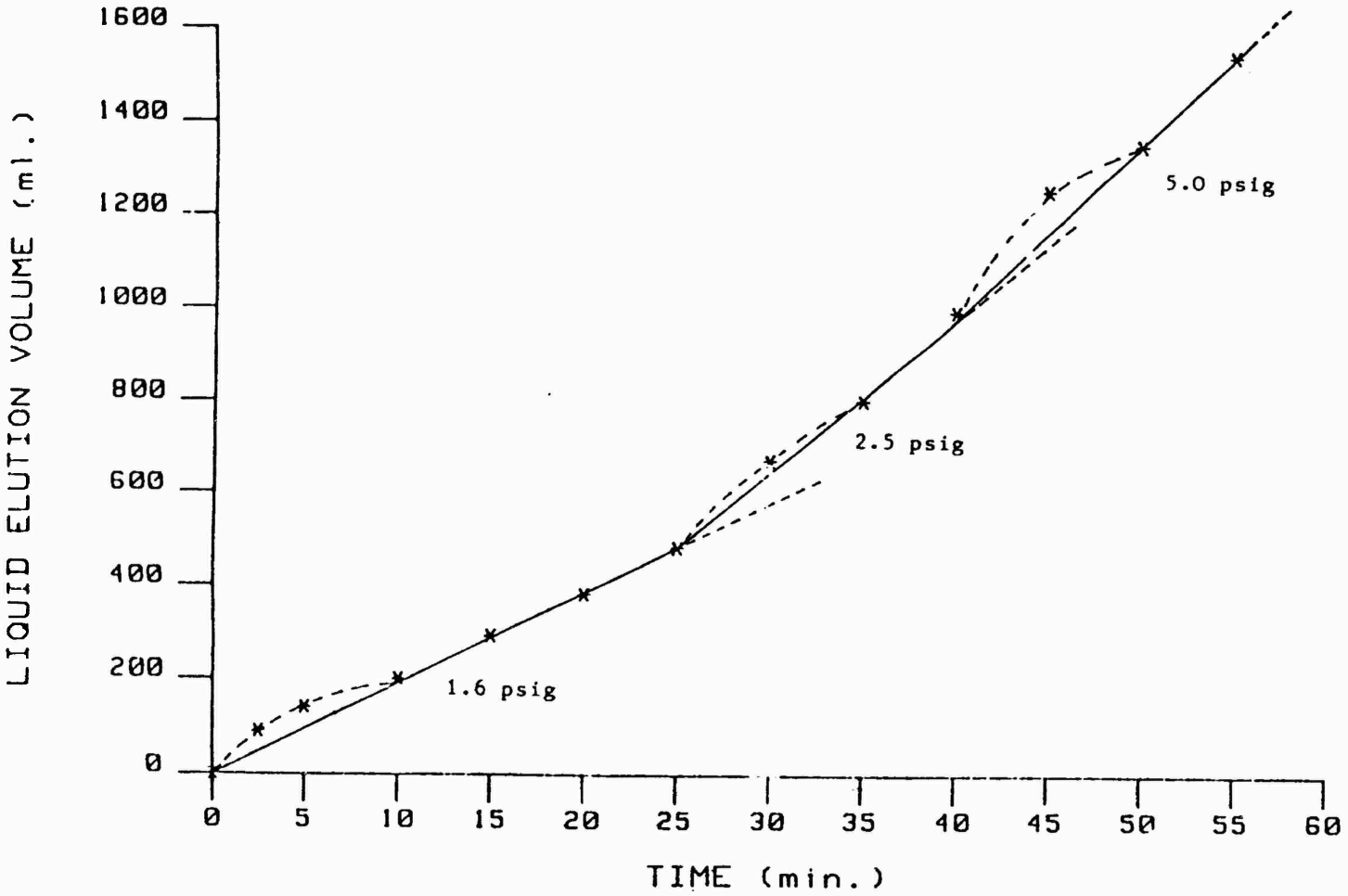


Figure 7: Elution volume (liquid) as a function of time at 3 different pressure drops. (CGA's made with 0.08 g/l SDBS; no column pretreatment).

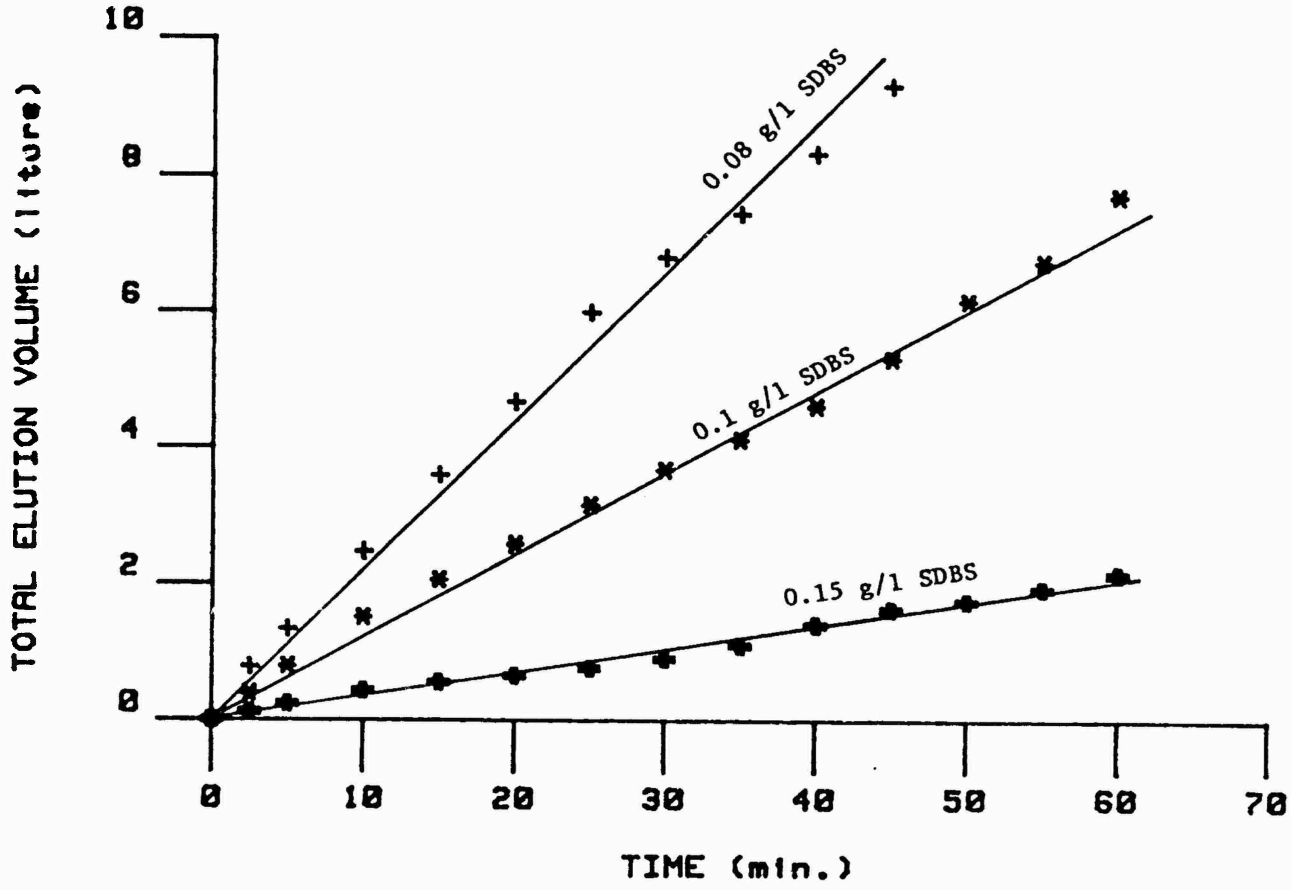


Figure 8: Flow of an anionic CGA (sodium dodecyl benzene sulfonate) through a 4-inch packed column as a function of surfactant concentration. (No column pretreatment.)

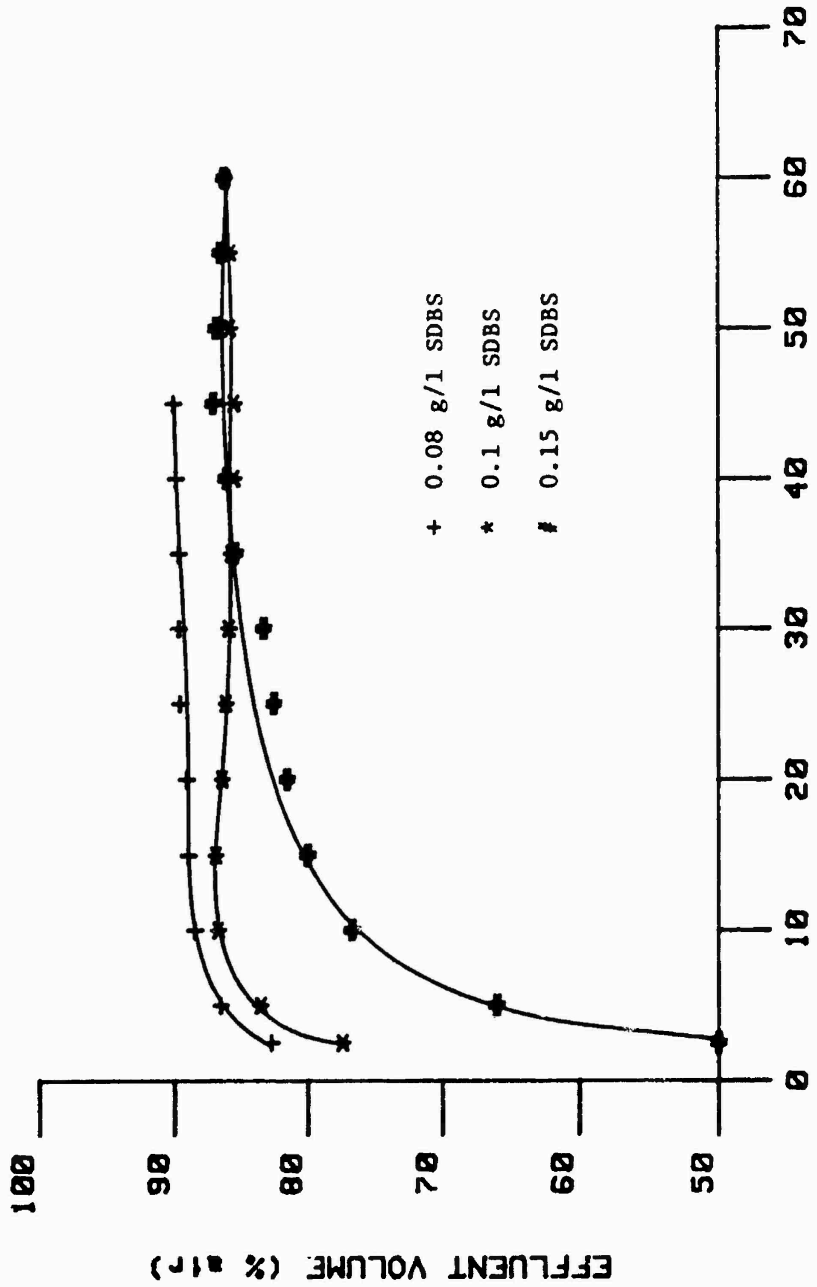


Figure 9: Air flow measured as a percentage of total effluent volume for a system in which CGA's are made with different concentrations of an anionic surfactant (sodium dodecyl benzene sulfonate). (No column pretreatment).

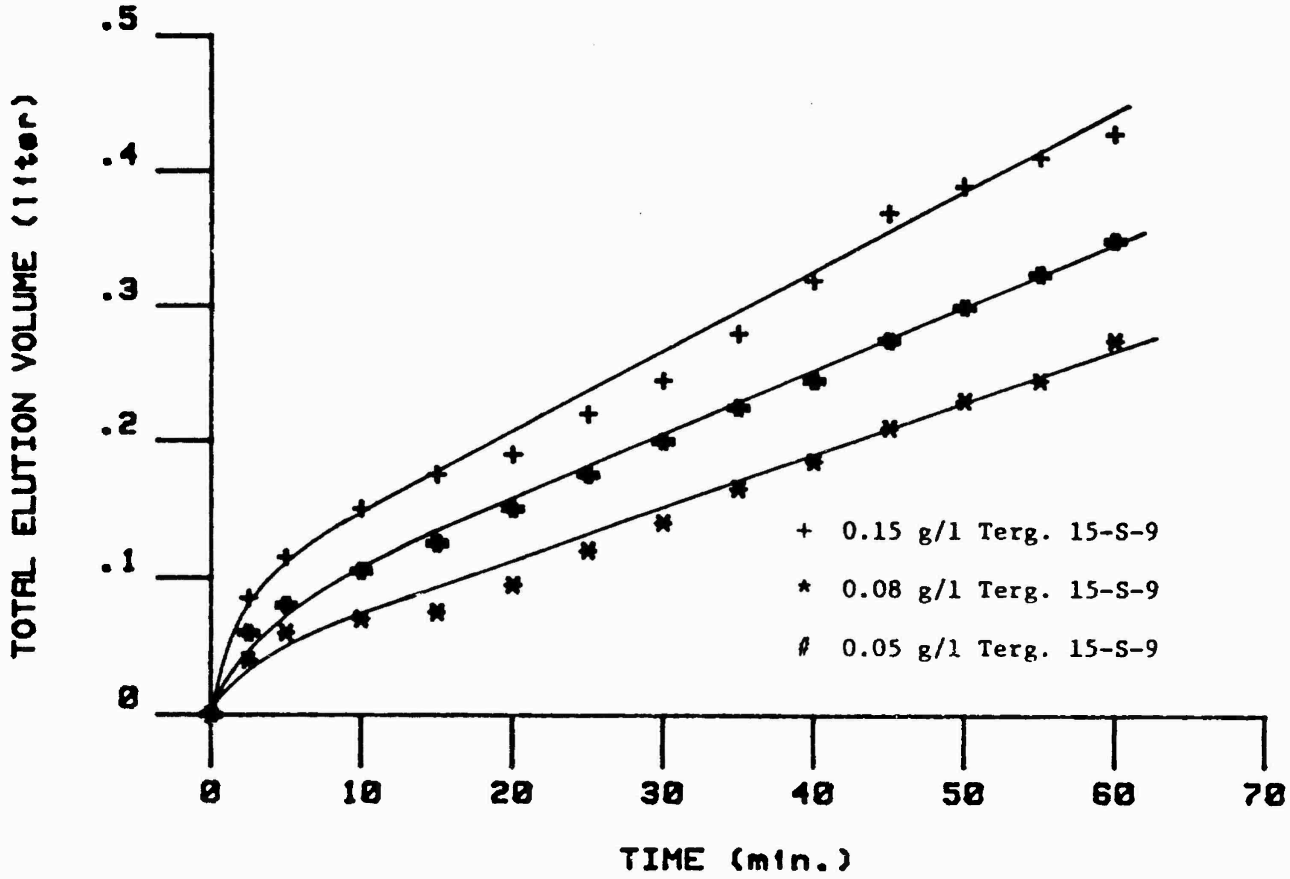


Figure 10 : Flow of a nonionic CGA (Tergitol 15-S-9) through a 4-inch packed column as a function of surfactant concentration. (No column pretreatment.)

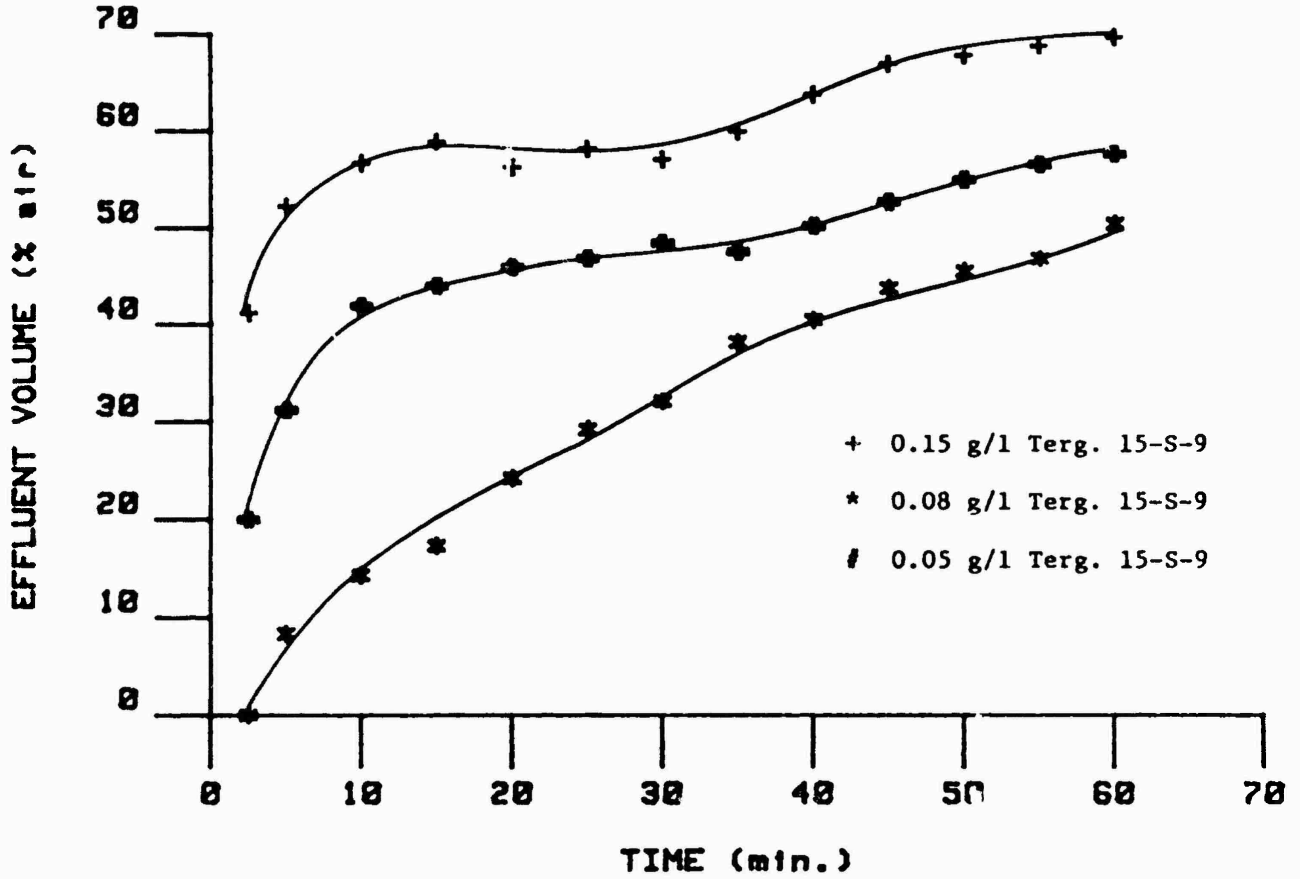


Figure 11: Air flow measured as a percentage of total effluent volume for a system in which CGA's were made with different concentrations of a nonionic surfactant (Tergitol 15-S-9). (No column pretreatment.)

of the effect on concentration on flow in Figure 10 and 11. The differences are slight and could be a function of soil packing or error in blending the surfactant concentration.

Based on experimental work performed to date, it is believed that surface effects are the primary contributors in controlling flow of CGA's through a consolidated soil (sand) matrix. By sufficient manipulation, such as pretreatment, these surface effects can be minimized and CGA's can be made to flow significant distances through a medium to medium-coarse matrix at reduced pressures.

Following is a list of some of the proposed potential advantages of CGA stripping of actual soil contaminants over surfactant floods or other presently applicable stripping techniques.

Based on these tests and earlier studies, potential advantages of CGA's for contaminant removal from soil systems are as follows:

1. CGA's have been shown to displace essentially all of the freely entrained water initially saturating a sand pack.
2. CGA's displace more than twice as much of the water in a sand pack than can be achieved with a simple air purge.
3. CGA's offer potential for added mechanical scouring action as a result of their tortuous trek through a soil's interstitial pore volume.
4. Potential exists for adding flocculants or coagulants (pretreatment) to enhance CGA stripping efficiency.
5. Compared to surfactant floods, CGA stripping offers the potential for reduced surfactant usage.

IV. ADHESION AND RETENTION OF CGA'S IN UNCONSOLIDATED SAND AND OTHER NATURAL AND BIOLOGICALLY ACTIVE MATERIALS

Procedure:

The adhesion and retention laboratory studies were conducted in a 500 ml wide mouth Erlenmeyer Flasks. The procedure was to add 400 g. of "as is" golf course sand (or material being tested) to a series of these flasks. Then water was added to completely saturate the unconsolidated matrix and raise the level in the flask to the 500 ml. mark. This weight albeit not that precise was recorded, and the water level was then lowered to 400 ml by decanting off a portion of the water covering the unconsolidated matrix.

A commercial blender was used to generate CGA's using several typical surfactants at various concentrations. The CGA's generation procedure involved adding about 250 ml of surfactant solution to the blender, operating the blender at low speed 17,000 rpm for 1/2 minute, and then at high speed, 21,000 for 1 minute. After switching the blender back to low agitation, a FMI positive displacement pump was used to deliver the CGA's to the Erlenmeyer Flask until the level rose again to 500 ml mark. A three prong fork (rake) made using 1/16" stainless steel tubing was used to inject the aphrons into the submerged matrix. Injection took 2 to 4 minutes. The rake was moved through the soil while injection was occurring. In several runs, a 100 ml graduate flask was filled with CGA's and allowed to destabilize in order to determine the quality of CGA's delivered to the flask (% air in the CGA mixture).

After careful adjustment of the water level to 500 ml, the flask

was reweighed and the difference in weight from original flask before sparging represents the cc of air (aphrons) adhering to or retained within the matrix. Knowing the cc of CGA's retained and the quality of CGA's produced and injected, the fraction (or percentage) of CGA bubbles retained can be determined from the following relationship.

$$\% \text{ CGA Retention} = \left\{ \frac{\text{cc of Hold-Up Gas}}{\text{cc of Hold-Up Liquid} \left(\frac{\text{CGA Quality}}{1-\text{CGA Quality}} \right)} \right\} 100$$

The Erlenmeyer flasks were weighed, allowed to stand uncovered, and reweighed intermittently after bringing water level back to the 500 ml mark. Some surface evaporation occurred.

Results and Discussion of Results:

A series of runs were made using the golf course sand (as is) and injecting CGA's generated from sodium dodecyl benzene sulfonate (SDBS) and also Tergitol 15-S-12. These results are shown on Figures 12 and 13. The amount of air adhesion (pick-up) and also its retention compared favorably with the total water volume in the sand measured between 80 and 90 cc. Some expansion of the bed also occurs but if 50 cc is picked up, the air to water volume ratio is somewhere between 56 to 1.25 (50/(90-50) to 50/90). The volume of air to water is very high.

Similar tests with direct room air injection into the bed using the same three prong probe yielded retentions of 5 to 6 cc of air at a maximum. While some retention was certainly observed, the weight

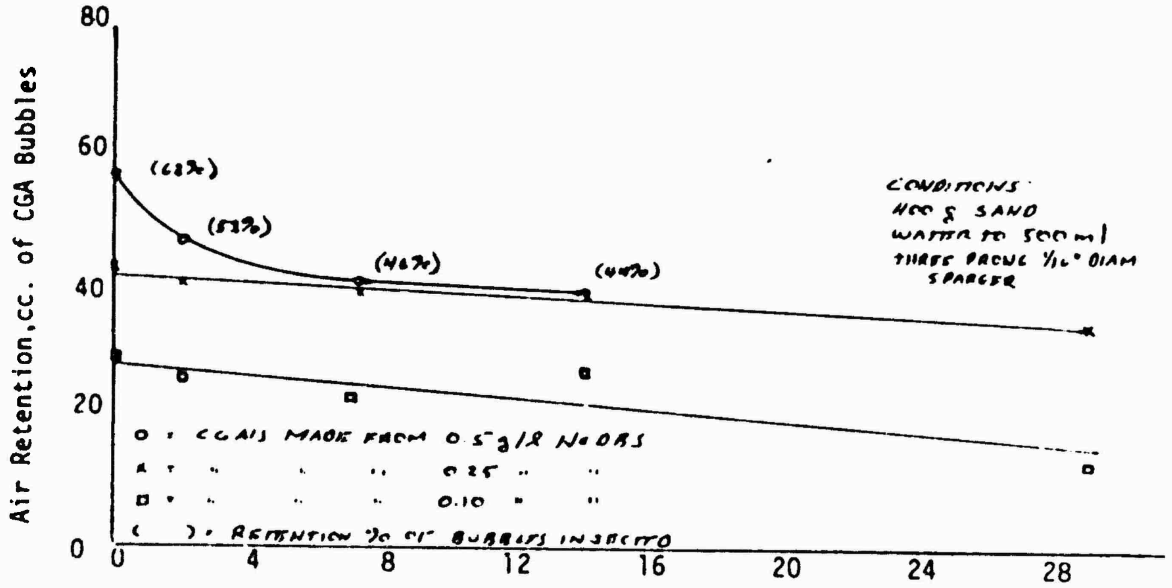


Figure 12

Injection and Retention of Colloidal Gas Aphrons in Laboratory Sparging Test Using Sodium Dodecyl Benezene Sulfonate

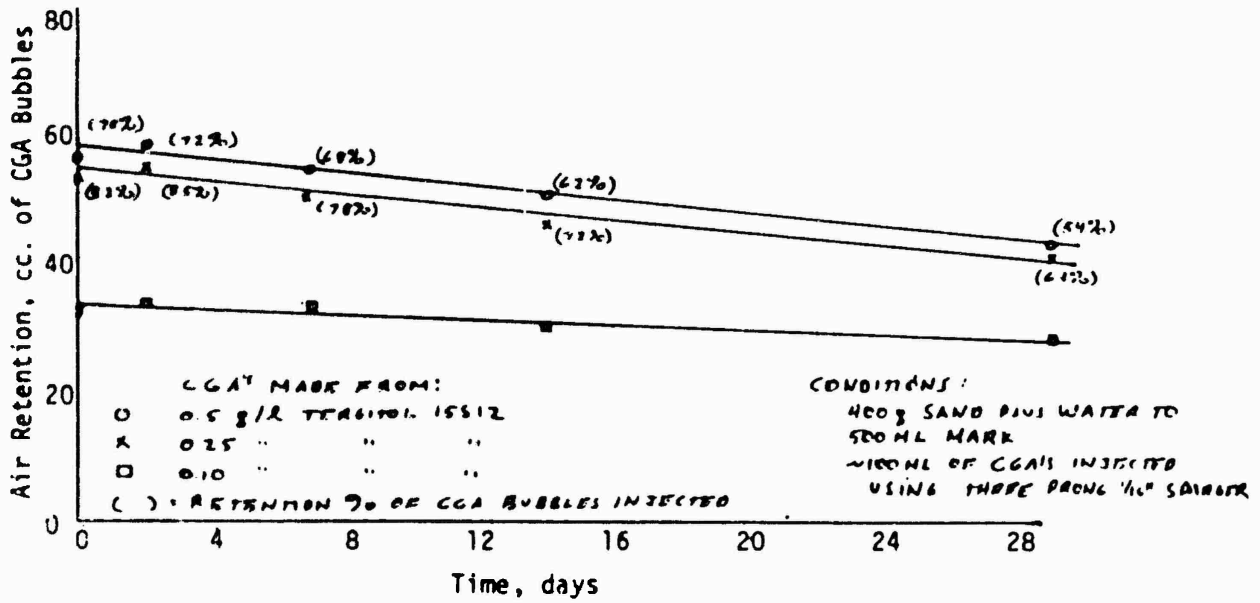


Figure 13
Injection and Retention of Colloidal Gas Aphrons
in Laboratory Sparging Tests Using Tergitol 15S12
(Nonionic) Surfactant

accuracy is limited by the ability to refill the flask consistently to 500 ml. In any case, the level of retention was small.

Another series of tests were conducted to determine the retention of CGA's on a pea gravel and oyster shells. Again some retention was observed and agitation by a spatula demonstrated this. However, the cc of retention was within the precision of refilling the flask to the 500 ml mark. In general, the oyster shells appeared to be more effective at CGA retention. With both materials, the CGA's made with a cationic surfactant, 1 g/l resulted in considerable flotation of the soil fines. The positively charged CGA's apparently bonded or complexed with the negatively charged silt mixed in with these materials.

Samples of both the pea gravel and ground oyster shells were fixed in a plastic mesh bag and placed in the trickling filter of the Tyndall AFB waste treatment plant. An active biological culture developed on the surface. After ten days, the samples were retrieved, hastily weighed and CGA retention studies completed. Considerable flotation of the microorganism occured and 5 to 25 ml of CGA's were retained. The foam froth in the neck above the 500 ml mark made precise measurement of the weight impossible after adding the CGA's.

A final series of tests were completed on standard EPA soil, Clarkburg Borough Soil Pit, Typic Hapludults (U.S.D.A. 1975). This soil is being used for in situ treatment studies being conducted by JRB Associates for U.S. EPA, Edison, N.J. Laboratory (Tony Tafuri, Contract Officer). The soil contains low organic carbon, and 8% by

weight clay sized particles although the fine are mostly quartz. When 100 cc of a 69% CGA quality mixture made from 0.5 g/l solution of SDBS was injected into 400 g (dry weight) of the saturated soil covered with water between 34.3 to 46.7 cc of CGA air bubbles were retained. Similiarly, when 100 cc of a 68% CGA quality mixture made from a 0.5 g/l solution of Tergitol 15-S-12 was injected between 18.6 to 23.2 cc of the air bubbles were retained. The pick-up with the SDBS was better, but only 18.6 to 38% of the CGA bubbles injected adhered to the soil. A week later, the CGA bubbles were still very dispersed with little coalescence apparent.

The adhesion and retention was generally not as good as with the golf coarse sand. However, this "standardized" test soil seemed to pack poorly and have a low bulk density. The particle density also seemed low. During the test, the bubbles just expanded the soil matrix and escaped up through the very unconsolidated rather fluidized soil water mixture. The authors question the choice of this soil as "typical" of region 2 Superfund sites.

V. REMOVAL OF SOLUBLE HAZARDOUS ORGANICS USING EXTRACTION/CGA SEPARATION TECHNIQUES

The purpose of this third study was to evaluate the possibility of using an extraction/CGA separation technique (using small quantities of solvents) for the removal of a soluble hazardous organic from water. Ortho dichlorobenzene was selected as a typical hazardous organic chemical with moderate water solubility and toxicity for these tests. It cannot be air stripped, is not easily biodegraded and has

been identified in contaminated water sources.

The first step in this limited study was to study and screen a number of solvents to determine their feasibility as a good solvent for removing ortho dichloro benzene from water. Eight available solvents were selected to provide a range of chemical properties.

The test procedure used was as follows:

1) 400 ml of 5 to 7.5 ppm ortho dichlorobenzene was placed in ten 500 to 1,000 cc separatory funnels. Saturated ortho dichlorobenzene (solubility 100 to 145 mg/l at 20 and 25°C respectively) was diluted twenty times and served as the common feed.

2) Four (4) ml of the solvent being tested was added to its designated separatory funnel - a total of 8 samples and 2 blanks.

3) A ground stopper top was placed on each funnel and each separatory funnel was shaken by hand vigorously for 5 minutes.

4) Observations were made two hours later and the following day.

5) After draining about 50 ml, the contents of the separatory funnel (180 ml) was placed in a brown bottle with about 30 ml of head space and sealed in a new septum capped clean bottle.

6) After shaking the contents, the vapor liquid mixture was allowed to come to equilibrium. Head space

analyses were completed on the solvent treated samples and on the two blanks (no solvent).

The results of this study are summarized on Table II. The most promising solvents for removing ortho dichlorobenzene (ODCB) appeared to be the decane and hexadecane - the straight chain paraffins. Based on ODCB in the head space, the decane removed an estimated 76% of the ODCB while not being dissolved into the water phase (low solubility). Clearly several of the branched organics extracted ortho dichlorobenzene from water; however, their solubility seems high as indicated by the high solvent content of head space.

The use of head space analyses for both the extractant and raffinate level is subject to considerable error. Aside from assumptions of equilibrium within the bottle and no needle contamination from any liquid on the bottom surface of the septum the sample should be a soluble single phase. Any solvent phase including dissolved ODCE floating on the surface of the raffinate or the two phase skimmed surface layer will give distorted head space analyses.

The analyses of both the extract and raffinate layers samples are best achieved by extraction of the solvent tested and ortho dichlorobenzene into a second solvent followed by centrifugation with careful attention to weights. For screening purposes, the hope was that head space analyses would provide indications of extraction and separation performance. In fact, the visual observations in the separation funnel are probably more important. The selection of the best solvent requires additional study following classical analytical

TABLE II

SOLVENT EFFECTIVENESS FOR REMOVING ORTHO DICHLORO BENZENE FROM WATER

	Solvent	Point	Density	Solubility in Water	Water Layer Head Space Solvent	Water Layer GC Area O-DCB	Appearance of Separatory Funnel	Comments	Estimated* ODCB Removal
1.	4-Tert Buty Toluene	192.8°C	0.857		581,919	0 to 1274	2 hrs - haze, slow settling 24 hrs - crystal clear, good separation	good separation, but moderate solubility, may be expensive	95%
2.	1,2,3-Trimethyl Benzene	176	0.89		970,818	low or masked	2 - haze, slow settling, emulsion oil layer 24 - o'k separation, hazy oil, bubbles in water	high solvent solubility	
3.	Diethyl Benzene	183.7	0.862		694,697	low or masked	2 - clear, organic bubbles adhering to glass 24 - good separation, solvent adhering to glass	high solvent solubility	
4.	m Diisopropyl Benzene	203	0.856		295,294	11,369	2 - haze, slow settling 24 - poor separation, hazy H_2O and oil layers	poor separation	52 %
5.	Hexadecane	287	0.7749	.0009 mg/l	95,676	4,754	2 - quite clear 24 - good emulsified oil layer separation, bubbles adhering to glass	known low solubility	80%
6.	Dibutylphthalate	340	1.0465		119,032	331 to 847	2 - poor, very hazy adhering to glass 24 - poor separation, oil sticking to sep. funnel	poor separation	96 %
7.	n-Decane	173	0.730	.009	21,344 to 134,737	0 to 5,699	2 - good, emulsion in top layer 24 - good separation, both phases clear, some oil adhering to sep. funnel	known low solubility, cheap	76 %
8.	Polyethylene Glycol			High			soluble, one phase only	no separation	
9.	No Solvent					19,478*	single plase		
10.	No Solvent					28,158*	single phase		
	Contaminant: Ortho Dichloro Benzene	179		00 (20°C) 45 (25°C)					

* estimated ODCB Removal = $\left\{\frac{(19,478 + 28,158)}{2} - \text{ODCB Area For Solvent}\right\} \Big/ \left(\frac{19,478 + 28,158}{2}\right)$

techniques as described in Barbarl & King (1982).

However, because of time constraints, it was decided to initiate the extraction/CGA separation process study using decane - a straight chain paraffin, easily biodegraded and relatively cheap. This choice was a good one. Barbarl and King (1982) recently determined equilibrium distribution (partition) coefficients for extraction of 17 different priority pollutants including ortho dichlorobenzene from water into undecane - a reasonable model for kerosene. A partition coefficient of 3030 was measured for ortho dichlorobenzene.

The goal of the next phase of the study was to evaluate several extraction/CGA flotation combinations for the removal of ortho dichlorobenzene from water using decane. The colloid gas aphrons were used to float the decane solvent after extraction of ortho dichlorobenzene was completed.

The first step in the test procedure was to pour 250 cc of a 5 to 7.5 ppm ODCB solution into blender (saturated solution of ODCB diluted twenty times). Small amounts of decane in some cases containing a drop of Tergitol 15-S-7 (oil soluble nonionic surfactant) were added. The air was dispersed (possibly emulsified) using a high speed blender (flat disk agitator operating at 21,000 rpm) to allow extraction of the ODCB to occur. The hazy mixture was then poured into a 1,000 ml flotation separatory funnel. A CGA flotation mixture was then formed in a blender using 21,000 rpm agitation for 30 seconds and 200 ml of the gas/liquid CGA mixture was sparged at 1 cc/sec into the flotation separation funnel using a dip stick. The separation was observed two

minutes after CGA injection was completed and the water (lower) layer drained into a second holding separatory funnel. A sample of the bulk lower water layer ("A" samples) and most of the surface skum ("B" and "C" samples) were taken and placed in a septum capped bottles for subsequent head space analyses. The holding separatory funnel was observed the following day, but time did not allow for sampling and GC analysis. Aside from the ODCB feed used, only the head space from bulk lower water layers, IIIA, VIA and VIIA and the surface foam IIIB, VIB and VIIIB and VIIIC were analyzed. A constant 0.5 ml vapor (head space) sample was injected into a 140°C isothermally operated GC equipped with a SE-30 column and a flame ionization detector. The recorder determined the integrated area. A description of the tests conducted, observations made, and GC analyses completed are shown on Table III.

While only preliminary, a combination of 0.05 cc decane plus a small amount of an oil surfactant (Tergitol 15-2-7) can be effectively dispersed and removed (clarified) from a much larger amount of water containing low concentrations of ODCB, runs VI and VIII. The GC analysis for VI indicated a preferential concentration of the decane in the skim layer 387,366/49,970 = 7.8. Similarly, the foam VIIIC from run VIII (a duplicate of VI) showed high concentration of decane (436,196) and ODCB (1080) comparable to VIB fro decane concentration and favorable to feed 1 x A, 1 x B for ODCB concentration. However, the bulk liquid VIIIA apparently still contained considerable decane (212,956).

TABLE III

TEST RESULTS FOR EXTRACTION/CGA FLOTATION OF ORTHO DICHLORO BENZENE FROM WATER USING TRACE DECANE SOLVENT

	Addition to 250 ml ODCB Solution	200 ml of of CGA's Made From A Solution of	Observation in Flotation Separation Funnel* 2 min.	Observation in Separation Holding Funnel Next Day	Gas Chromatographic Peak Area
I.	0.5 cc Decane	1 g/l. NaDBS	poor separation very hazy	four layers formed - definite oil skim, two oil emulsion in water layers, and a clear bottom water layer	
II.	0.5 cc Decane	1 g/l NaDBS (72% quality)	poor separation but less haze than I	slight haze in water layer, slight oil layer present	
III.		2 g/l Tergitol 15-S-7		clear but slight foam on surface	IIIA Decane 4021 ODCB 259 IIIB Decane 755 ODCB 201
IV.	0.5 Decane	No CGA's Added	poor separation hazy	three layers formed - oil skim, emulsified cloudy water, and clear bottom water layer	
V.	0.5 cc Decane 1 drop Tergitol 15-S-7	1 g/l Tergitol 15-S-9 (62% quality)	poor separation, very hazy like I, more foam than I	hazy water layer densifying up the column plus oil layer with some bubbles	
VI.	0.05 cc Decane 1 drop Tergitol 15-S-7	1 g/l Tergitol 15-S-9	good separation, much better than II, more foam than II	very clear water layer with oil layer (difficult to see)	VIA Decane 49970 ODCB 133 VIB Decane 387366 ODCB masked?
VII.	0.5 cc Decane + 0.05 cc Dodecyl Pyridium Chloride	1 g/l NaDBS	poor, not decanted DPC not soluble in decane	discarded	
VIII.	same as VI		good separation, considerable foam	not quite as clear as VI	VIIIA Decane 212956 ODCB 7 VIIIB Decane 103084 ODCB 107 VIIIC Decane 436198 ODCB 1980
IX.	feed, head space two minutes after CGA sparge				IXA ODCB 1130 IXB ODCB 665

*two minutes after CGA sparge

The results are encouraging in terms of demonstrating the use of small amounts of a solvent (0.05 cc) in a mixer/separator to extract quantities of a toxic, hazardous organic such as ortho dichlorobenzene from 250 cc of contaminated water and particularly the ability of the flotation process using CGA's to rapidly float the decane while leaving a clear water layer in the batch/separator batch tests.

The potential for using straight chain hydrocarbons to extract soluble ortho dichlorobenzene has been documented by Barbari and King, (1982). They experimentally determined the equilibrium distribution (partition) coefficient (weight fraction) of 3030 for the extraction of ortho dichlorobenzene from water using undecane. Undecane was chosen as a reasonable model for kerosene. Furthermore, if results for extraction of chlorobenzene with straight chain hydrocarbons is any indication the equilibrium distribution coefficients for ortho dichlorobenzene will increase as the paraffin chain length decreases, e.g., heptane or hexane, see again Barbari and King (1982). A balance must be made between good ortho dichlorobenzene extraction and the increased solvent solubility (losses) of the shorter chained hydrocarbons.

The key, however, to processing success using small quantities of solvent for extraction is effective removal of the dispersed solvent (extractant). While GC analysis of the head space is indicative of this separation, a better measure is the clarity of the raffinate and extract phases. The unique combination of small amounts of decane plus trace Tergitol 15-S-3 solvent could be effectively clarified from

a large volume of raffinate using a CGA made from Tergitol 15-S-9. These results have been duplicated recently back at the Blacksburg campus of Virginia Tech.

VI. CONCLUSIONS AND RECOMMENDATIONS

Flow of CGA's through a medium sand consolidated matrix.

1. The flow of CGA's and water/air breakdown products is very much a function of surfactant type and concentration level and degree of pretreatment provided. Vertical slice and/or aquarium three dimensional tests would be helpful to determine the effectiveness of CGA's to transport horizontally through a soil matrix for possible scouring or to provide a delivery formulation for in-situ flushing by the CGA liquid or in-situ biodegradation by the CGA entrapped air (or O_2, O_3, etc.)

Adhesion and Retention of CGA's in Unconsolidated Matrices

2. The injection adhesion and retention of sparged CGA's in a simulated impoundment resulted in air to water volume ratios in the flooded sand of .56 to 1.25. After a month with CGA's made from various levels of Tergitol 15-S-9, over 75% of the air remained entrapped in the matrix.

   The feasibility of using these CGA's (perhaps made from O_2) for biodegrading organics in situ should be explored. Limited studies on campus have already proved its effectiveness for biodegrading phenol in situ using pseudomonis pudita mixed with CGA's during injection. There must be a number of situations where in situ treatment

techniques using specific or a broad range of micro-organisms could be used to facilitate biodegradation of organics in impoundments sediments or buried underground below the water table.

3. The use of straight chain paraffin such as decane in an extraction/flotation process for removal of small amounts of ortho dichlorobenzene from water looks promising. With a favorable partition coefficient, initial studies have demonstrated an effective separation (phase clarity) using CGA technology. The recommendation is to initiate a more complete laboratory process development study to evaluate various possible solvents for their extraction capability (Distribution Coefficient) and to evaluate separation characteristics using CGA's as well as oil aphrons for extraction/flotation processing.

VII. BIBLIOGRAPHY

Barbari, T. A. and C. J. King, 1982, "Equilibrium Distribution Coefficients for Extraction of Chlorinated Hydrocarbons and Aromatics from Water into Undecane," Env. Sci. and Tech., Vol. 16, No. 9, pp. 624-627.

Sebba, Felix, 1979, "The Behavior of Minute Oil Droplets Encapsulated in a Water Film." Colloid and Polymers Sci. 257, 392, (1979).

Sebba, Felix, 1982, "Investigation of the Modes of Contaminant Capture in CGA (MG) Foams," Report to O.W.R.T., 14-34-0001-0489, October 1982.

Sebba, Felix, "Preparation and Uses of Polyaphrons." Presented at Colloid and Surface Symposium of ACS, Blacksburg, 1982. Accepted by "Chemistry and Industry."

Sebba, F. and S. M. Barnett, 1981, "Separation Using Colloidal Gas Aphrons," Proc. 2nd Int. Congress of Chem. Eng., IV, pp. 27-31, 1981.

Verschueren, K., 1977, "Handbook of Environmental Data on Organic Chemicals," Van Nostrand Reinhold Co., 114.

U. S. Department of Agriculture, 1975, "Soil Taxonomy: Basic System of Soil Classification for Making and Interpreting Soil Survey," Agriculture Handbook No. 436.

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

TRANSPORT AND ELECTRON PARAMAGNETIC RESONANCE STUDIES OF INFRARED DETECTOR MATERIALS

Prepared by:	George K. Miner, Ph. D.
Academic Rank:	Associate Professor of Physics
Department and University:	Department of Physics University of Dayton
Research Location:	Materials Laboratory Wright-Patterson Air Force Base
USAF Research Contact:	Patrick M. Hemenger, Ph. D.
Date:	August 22, 1983
Contract No:	F-9620-82-C-0035

TRANSPORT AND ELECTRON PARAMAGNETIC RESONANCE STUDIES OF INFRARED DETECTOR MATERIALS

by

George K. Miner, Ph. D.

ABSTRACT

Through interaction with the Materials Laboratory, an experimental apparatus for Hall Effect and resistivity measurements has been established in the Physics Department at the University of Dayton. It operates over a temperature range of 4 to 400K. The system has been tested with silicon and gallium arsenide and compared to a comparable system at the Materials Laboratory. It has proved to be reliable for supplemental measurements for that laboratory. An electron paramagnetic resonance spectrometer has been used to examine similar gallium arsenide materials and comparisons made to earlier silicon work. Preliminary work indicates that with some dopants there are weak but as-yet unidentified signals at 77 and 300K. There remains the possibility of identification of the defects responsible.

VII. ACKNOWLEDGEMENTS

The SCEEE Fellow wishes to acknowledge the interaction with and the helpfulness of his Effort Focal Point, Dr. Patrick Hemenger, and his associate, Dr. William Mitchel, both of AFWAL/MLPO. The careful assistance of graduate student Mark Pippin and undergraduate Christopher Emmerich were invaluable in this project. The useful interaction with colleagues Dr. Joseph Lang, Dr. Perry Yaney, and Dr. Rex Berney is greatly appreciated. In addition the support of two department chairmen during this period, Dr. James Schneider and Dr. Michael O'Hare, is valued.

The Fellow wishes to acknowledge the support of the Air Force Systems Command, the Air Force Office of Scientific Research, and the Air Force Materials Laboratory. Also the assistance of the Southeastern Center for Electrical Engineering Education is acknowledged.

I. INTRODUCTION

An important concern of the Materials Laboratory at Wright-Patterson Air Force Base is the characterization of semiconductors for infrared sensitive detectors. The behavior of these materials depends on very low concentrations of impurities and defects in the semiconductor. By combining Hall Effect and resistivity measurements as a function of temperature, the electrical transport properties (carrier concentration, Hall mobility, and conductivity) can be determined. From the temperature dependence of the carrier concentration, the identities and concentrations of the impurities can be inferred. More complex defects may be analyzed using other techniques, for example, electron paramagnetic resonance. The identities of these defects often emerge as a result of annealling studies.

In the past at the Materials Laboratory major attention has been focused on silicon. A more recent thrust includes increased interest in gallium arsenide. The SCEEE Fellow involved in this study has experience in transport studies, radiation damage, and electron paramagnetic resonance. Due to the availability of an experimental apparatus for the Hall Effect at his university and the general interest at the Materials Laboratory for such measurements, an upgrade of the university capability for supplemental measurements was very appropriate. The experience of the Fellow in electron paramagnetic resonance and the availability of a university spectrometer coupled with the fact that such a facility is not available at the Materials Laboratory, made pursuit of this method for the identification of defects very desirable. The Fellow's previous experience with electron paramagnetic resonance in silicon was also attractive. His radiation damage experience could be helpful in future annealling experiments. He was involved in this type of work earlier with silicon.

II. OBJECTIVES

The following is the research goals and objectives statement submitted to SCEEE on April 26, 1983:

"During the summer 1983 appointment I will work with Dr. Patrick Hemenger of the Materials Laboratory at Wright-Patterson Air Force Base. The work of the group with which he is associated is concerned with the characterization of semiconductor materials for infrared detectors. Past work has been largely involved with silicon doped with various acceptors and donors. A newer thrust involves a greater interest in gallium arsenide. Due to my experience in electron paramagnetic resonance and transport studies, the following possibilities will be explored.

"Electron paramagnetic resonance studies will be attempted in the gallium arsenide materials. Comparisons will be made to measurements in silicon. In this work and a related search of the literature, an attempt will be made to determine if this technique would be useful to the Materials Laboratory in its characterization of gallium arsenide.

"Transport studies are regularly conducted at the Materials Laboratory. An independent experimental facility for Hall Effect and resistivity measurements exists in the Physics Department at the University of Dayton. Measurements on common samples will be conducted in both laboratories and comparisons made. It is expected that both silicon and gallium arsenide will be involved. The degree of agreement will be examined. This work should lead to the possibility of supplemental measurements for the Materials Laboratory in the university laboratory.

"The result of the ten-week summer appointment should lead to a relationship with the Materials Laboratory and an involvement in its semiconductor materials characterization program."

III. EXPERIMENTAL FACILITIES

A) Transport Studies Apparatus

The techniques used in the transport studies are standard techniques as described in Putley.[1] Improvements have been described by the Materials Laboratory[2] and elsewhere.[3] For the transport studies apparatus the sample configuration is a six point probe. The carefully controlled current enters one end electrode and exits the other. A measurement of a voltage between two side-by-side electrode points yields the resistivity. By measuring the perpendicular voltage between two opposite electrode points with an appropriately applied magnetic field, the Hall coefficient is determined. These measurements lead to the carrier concentration, the mobility and dopant energy levels.

During the summer project the existing apparatus at the University of Dayton was upgraded to approach the quality of the comparable facility at the Materials Laboratory. The existing five-point sample configuration was replaced with a six-point one for use with a Hall bar. A new high resistance Deutsch connector was installed as well as a second thermocouple. Improvements were made in magnetic field calibration and temperature control. The early measurements were taken over a temperature range of 77 to 300K, that is, from liquid nitrogen temperature up to room temperature. By using liquid helium the range was extended down to 4K and with the use of a heating coil up to 400K. The temperaure range is now the same as is used by the standard Hall Effect apparatus at the Materials Laboratory.

B) Electron Paramagnetic Resonance Spectrometer

The electron paramagnetic resonance equipment that was used in the study is a modified form of the Varian

Associates Model 4500 series EPR Spectrometer. This x-band spectrometer operates at about 9.5 Gigahertz, or about 9.2 Gigahertz with a quartz dewar inserted. The external magnetic field for the Zeeman splitting is provided by a Varian 12-inch regulated and water-cooled magnet with field scanning provisions. The microwaves are produced by a klystron, are distributed by a microwave bridge, and are detected by a crystal detector. The original hybrid tee bridge has been replaced by a more efficient three-port-circulator arrangement. The klystron is stabilized with automatic frequency control. The spectrometer operates in the absorption mode.

This spectrometer has a 100-kilohertz crystal-controlled oscillator which generates the field modulation, and a high gain amplifier and phase detector for the sensing of the EPR signal. The spectrometer can detect $10^{12} \pm 25\%$ spins. Due to the weak intensity of the EPR lines observed in the materials of concern, a signal averaging technique was employed using a Tracor Northern TN-1710. Typically 25 sweeps produced a useful spectrum. Also a test was conducted of the three available cavities and the most sensitive was selected for these measurements because of the very weak signals involved.

For magnetic field calibration and intensity calibration, a manganese marker was permanently installed in the EPR cavity. The marker employed was forsterite, a material characterized[4] earlier by the SCEEE Fellow. Its location in the cavity was such that its phase was opposite to that of the sample, making differentiation easy. The marker has a six-line hyperfine spectrum which spans some 435 gauss centered near $g = 2$. These lines were used for relative g-value (spectroscopic splitting factor) measurements and intensity measurements. Both g-value and intensity calibrations were made using a standard reference of 0.00033% pitch in KCl.

C) Annealling Equipment

The silicon heat treatment was done in a Blue M Stabil-Glow tube furnace with temperature control. The samples were on a quartz boat in a quartz tube and were bathed in an argon atmosphere during heating and cooling. The samples were sandwiched between ultra-high purity silicon wafers during treatment. The furnace was brought to temperature before the boat was inserted. At the end of the anneal time, the boat was quickly pulled to the end of the quartz tube and kept in the argon atmosphere until its return to room temperature. All samples were given an RCA wash cycle before each anneal and were measured by EPR just afterwards.

When isochronal anneals were done, the time at temperature was one-half hour and the temperatures range in 100 degrees steps from 250°C to 750°C or 1050°C. The isothermal anneals were performed at 550°C for total times ranging from one-half hour to 16 hours.

IV. MEASUREMENTS AND RESULTS

A) Transport Studies

With the improved transport studies apparatus, several measurements were made. The first series involved silicon with indium (Si:In) and gallium (Si:Ga) dopants. Both materials had been well characterized at the Materials Laboratory.[5] The results of initial measurements are shown for Si:In in Figure 1 and for Si:Ga in Figure 2. In these figures the data points are displayed as plus signs and the solid line is a fit via the charge balance equation. The goodness-of-fit is described by a chi-square. Both fits are good as displayed by the chi-square and the results agree well with the previous Materials Laboratory measurements.

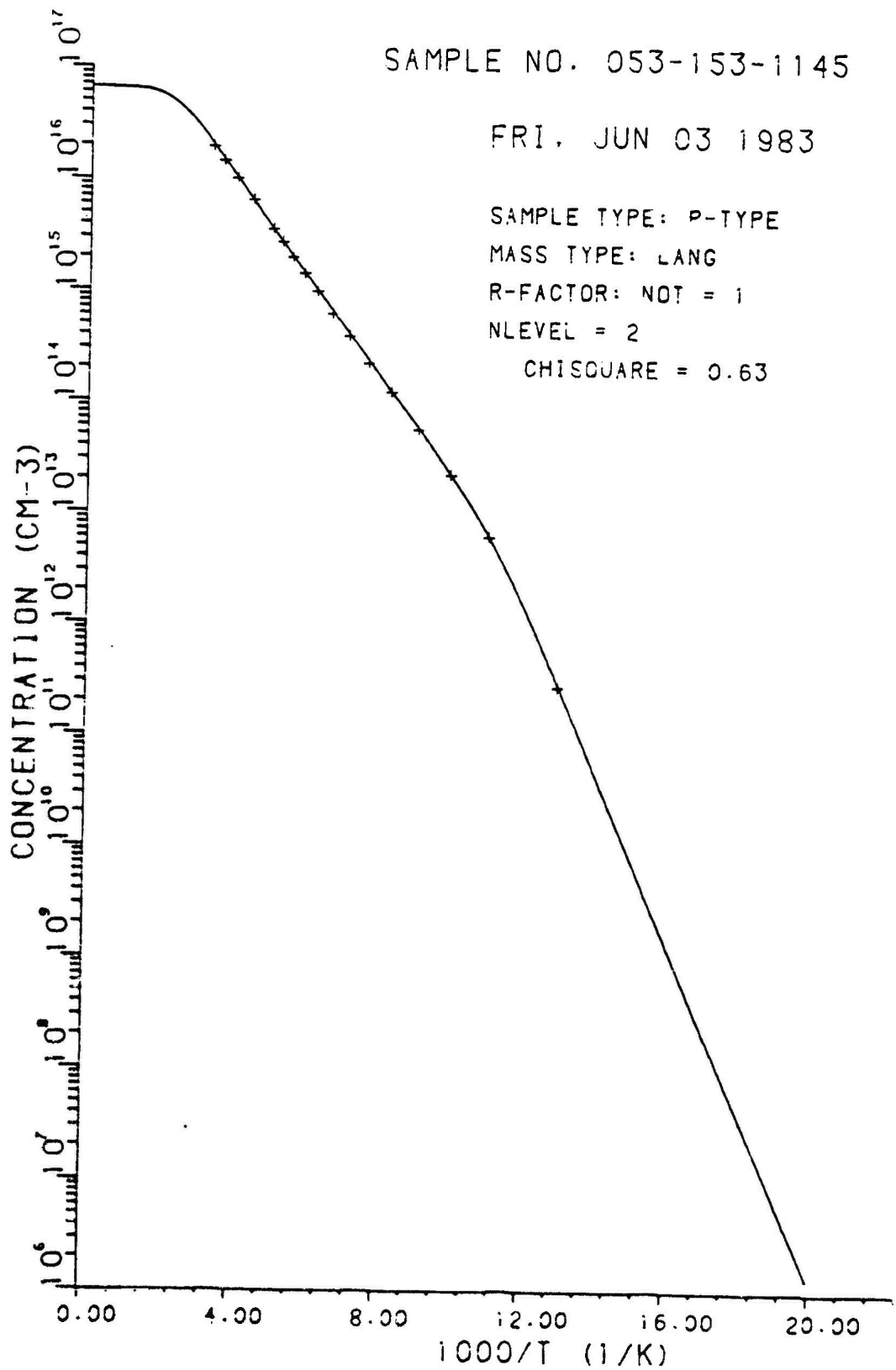


Figure 1. Silicon with Indium Dopant

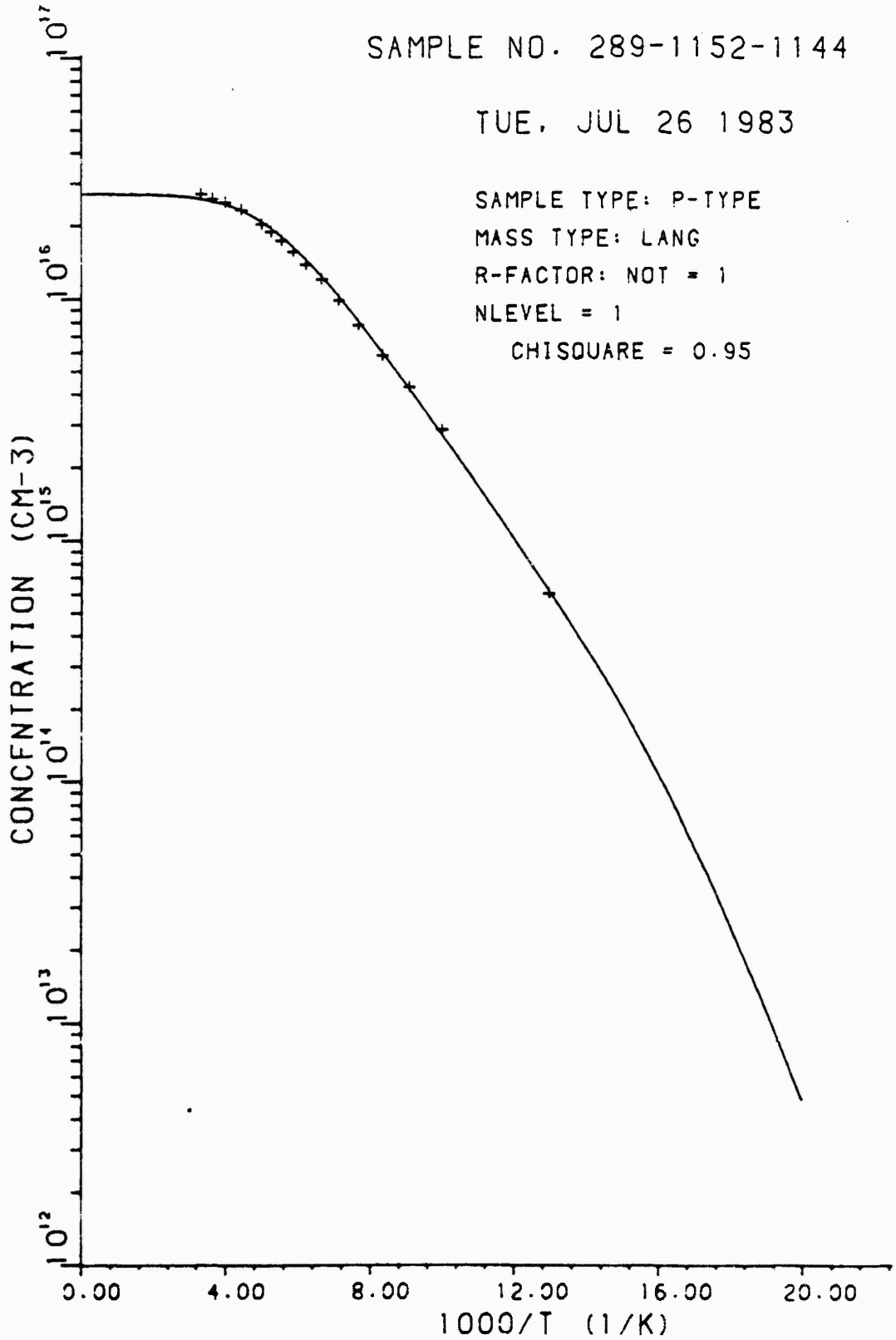


Figure 2. Silicon with Gallium Dopant

These good fits and the agreement with the values of parameters measured at the Materials Laboratory provide a validation of the new experimental arrangement.

Both of the above data sets are for a temperature range of 77 - 300K. When the range was extended to 35 - 393K, the chi-square increased noticeably. This less desirable fit was traced to the low temperature data, where temperature control is much more critical. In the low temperature range (4 - 77K) better temperature control is required. This fit is displayed in Figure 3 for the Si:Ga sample.

After this equipment test, some new samples of gallium arsenide doped with indium (GaAs:In) and antimony (GaAs:Sb) were examined. The data is displayed for GaAs:In in Figure 4 and for GaAs:Sb in Figure 5. Both fits are very good over the range reported. Both materials have a very high resistivity and therefore cannot be measured at the lower temperatures, a difficulty shared by the comparable apparatus at the Materials Laboratory. The higher temperature limit is imposed by the 400K temperature design of the apparatus, also the limit at the Materials Laboratory. New values for the activation energies have been determined however.

In order to analyze the data, it is necessay to know the Hall scattering factor, the r-factor. It is the ratio of the Hall mobility to the conductivity mobility. The Hall mobility is determined by the experiment, but the conductivity mobility is required in the data analysis. The Hall scattering factor as a function of temperature has been calculated at the Materials Laboratory,[6] and that result has been used in this analysis. Also refined values for the effective mass computed at the Materials Laboratory[7] were used in these calculations. In addition, it is assumed here that the g for each level is four consistent with the best current information at the Materials Laboratory.

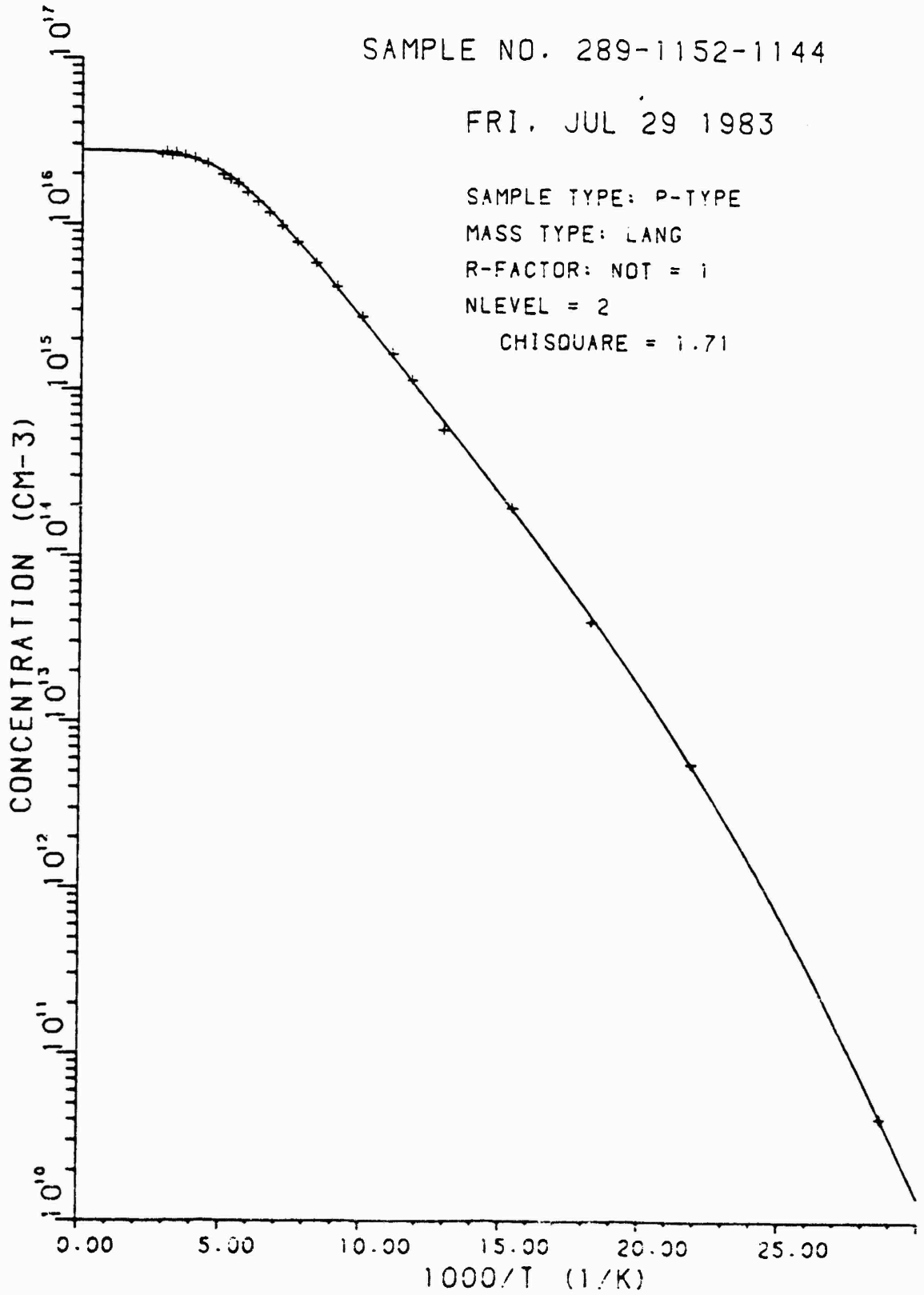


Figure 3. Silicon with Gallium, Large Temperature Range

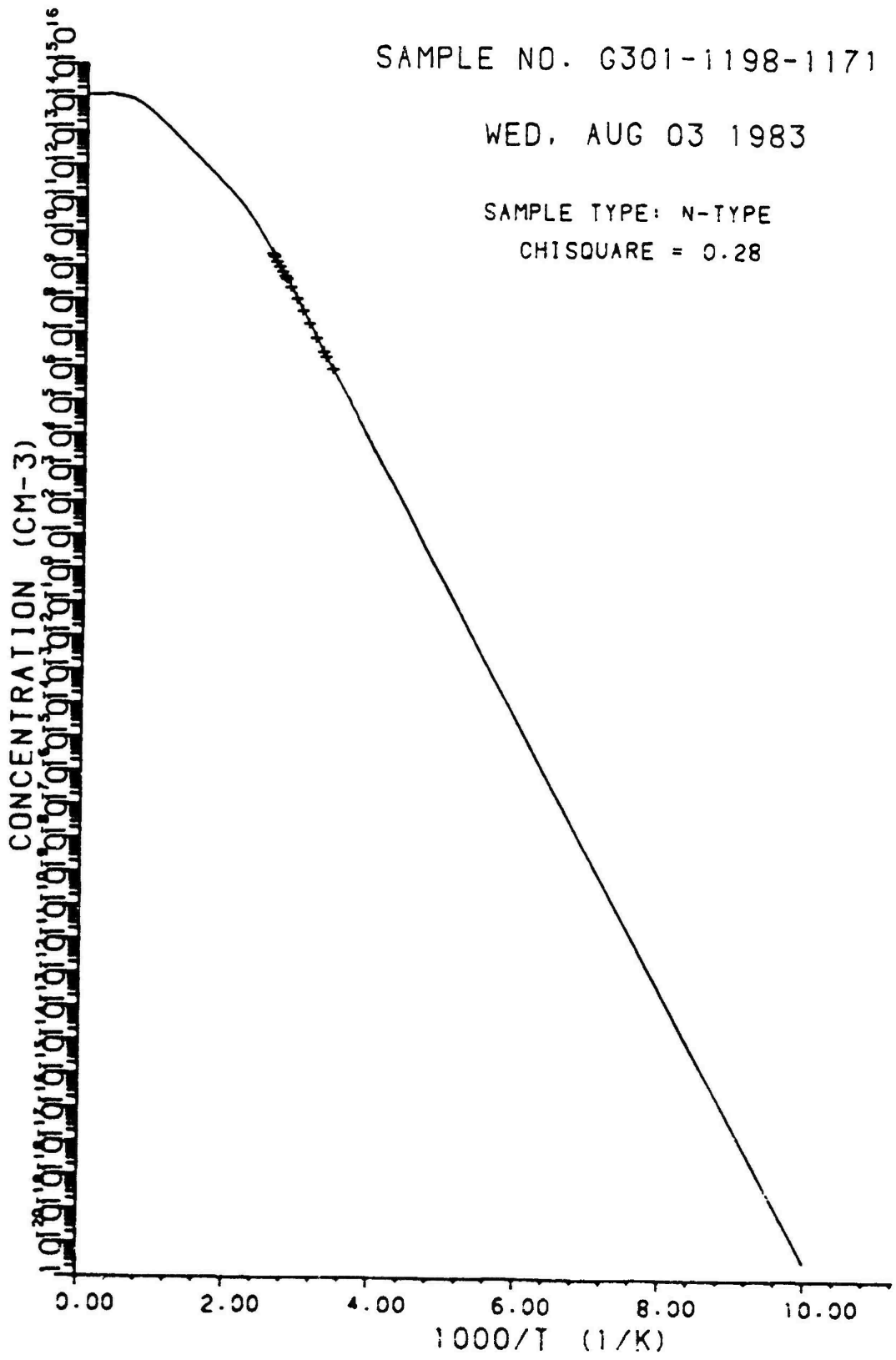


Figure 4. Gallium Arsenide with Indium Dopant

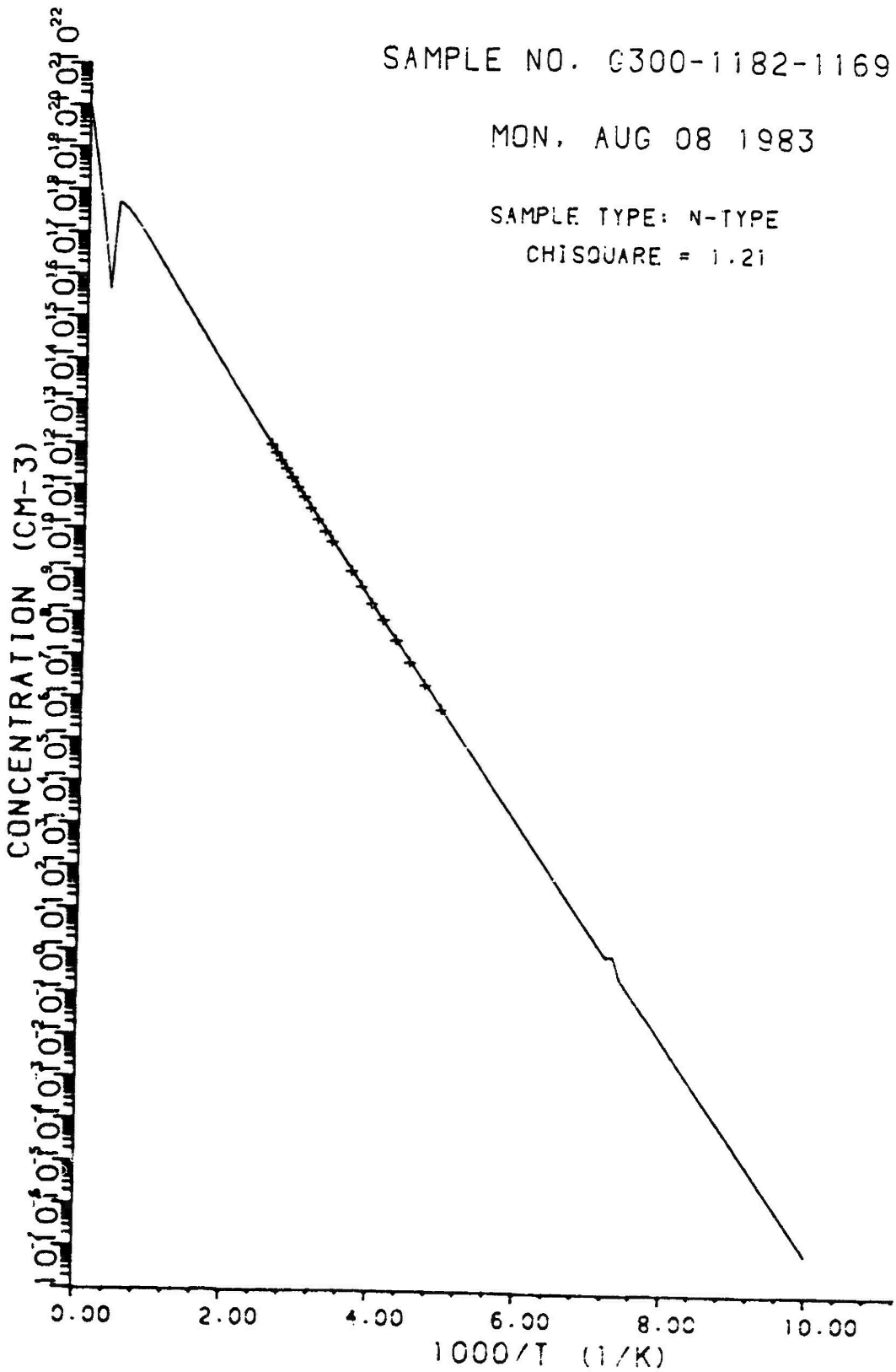


Figure 5. Gallium Arsenide with Antimony Dopant

The parameters derived from the best fit experimental runs are summarized in the following table.

MATERIALS LAB SAMPLE NUMBER	TEMP(K) RANGE	NO. OF LEVELS	N_D (cm^{-3})	ENERGY-1 $(N_A)_1$	ENERGY-2 $(N_A)_2$	X^2
Si:In	77-300	2		0.180eV	0.141eV	0.63
0053-0153-1145			9.50E12	5.65E16	8.40E15	
Si:Ga	77-300	1		0.068eV	-	0.95
0289-1152-1144			1.69E14	2.72E16	-	
Si:Ga	35-353	2		0.069eV	0.021eV	1.71
0289-1152-1144			1.74E15	2.74E16	1.73E15	
GaAs:In	292-387	1		0.778eV	-	0.28
G301-1198-1171			4.56E11	1.17E15	-	
GaAs:Sb	203-393	1		0.484eV	-	1.21
G300-1182-1169			9.63E18	1.04E21	-	

In the table N_D is the number of donors and X^2 is the chi-square of the fit. ENERGY-1 and $(N_A)_1$ are the activation energy and number of acceptors of type 1. If a two-level fit was successful similar parameters are given for the second acceptor.

For the silicon samples the energies are generally close to accepted values,[8] with perhaps the indium doped sample a little high. With a poorer chi-square they are in excellent agreement. The second level detected in gallium can also be improved with a poorer chi-square. For the gallium arsenide samples there are no accepted values. However, the results obtained are in good agreement with preliminary data from the Materials Laboratory.

B) Electron Paramagnetic Resonance Experiments

Initial efforts in electron paramagnetic resonance centered on improving the sensitivity and in calibrating the magnetic field and intensity. Then a series of preliminary experiments were conducted on gallium arsenide, with guidance from previous work on silicon.

A literature search yielded the most information on gallium arsenide doped with chromium,[9] GaAs:Cr. As a result this material was investigated first, both at room temperature (300K) and at liquid nitrogen temperature (77K). The 77K measurements were done with nitrogen gas passed through a heat exchanger immersed in liquid nitrogen, and therefore the temperature was just above 77K. For the GaAs:Cr sample, no reliable signal was detected at 300K and only a very weak line at 77K. The work was then directed toward two other dopants of interest because of their semi-insulating possibilities.[10] [Many references other than those listed herein were found; a review[11] includes many of them.]

In samples of GaAs:In at 300K and at 77K a line was detected. An example is shown in Figure 6A. As the perpendicular to the sample wafer was rotated relative to the magnetic field of the spectrometer, the magnetic field location of the line was observed to move and its width was found to change. When samples of GaAs:Sb were measured, two lines were observed, both at 300K and 77K. When the sample is rotated, one of these lines appears to be fixed at a value of $g = 2.0723$ while the other moves in field. This pattern is shown in Figure 6B. This data suggests a complicated rotation pattern about this axis; the pattern may become simplier about another axis. Also computer modeling of the spectra could aid in isolating the parameters of individual lines.

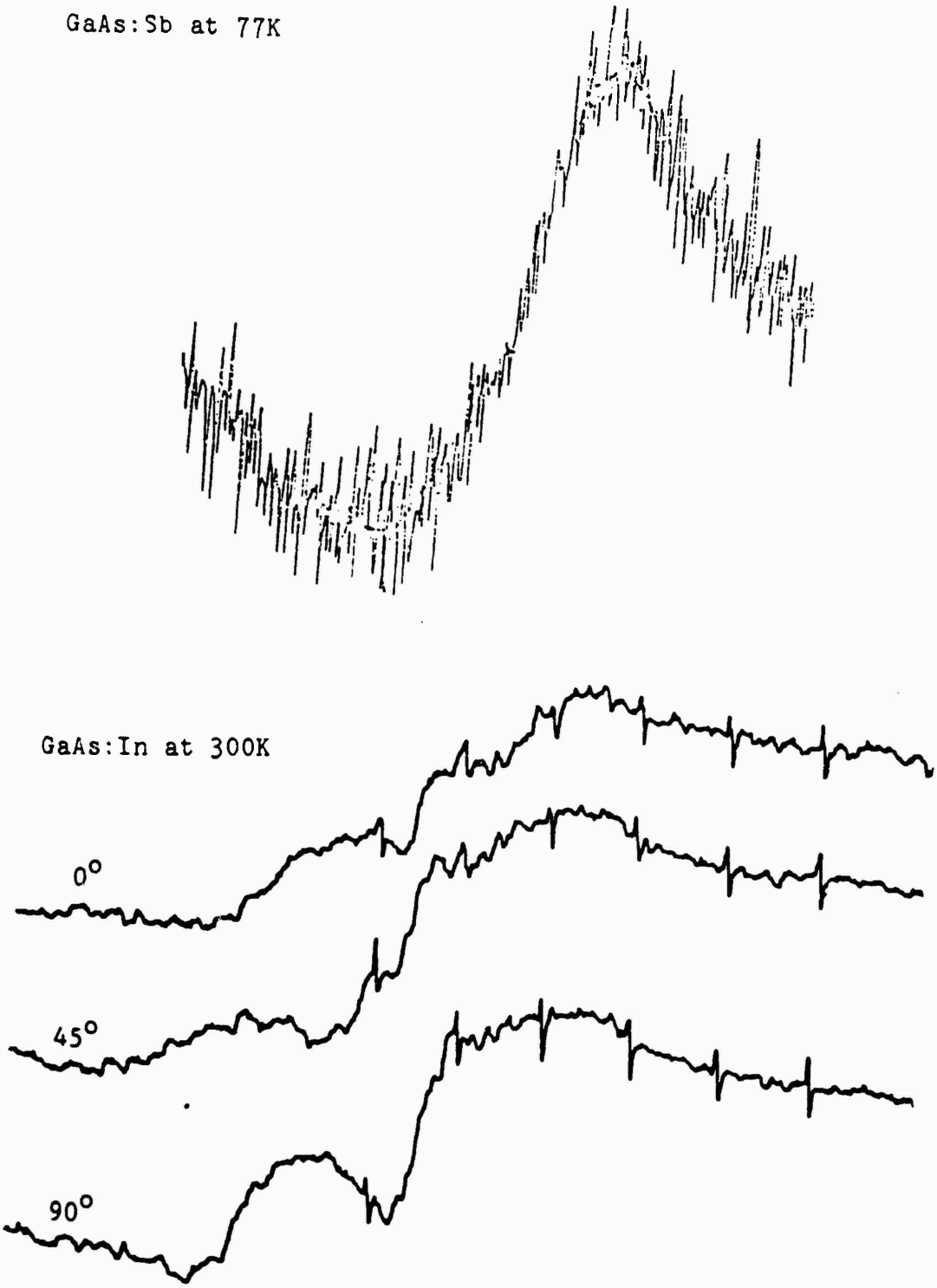


Figure 6. EPR Traces from GaAs

The data from these GaAs measurements is summarized in the following table.

DOPANT	TEMPERATURE (K)	g-VALUE/LINEWIDTH(Gauss) 0°	45°	90°
Cr	77	2.0579/24g		
In	300	2.0058/23g		2.2435/165g
In	77	2.0662/197g	2.1044/316g	2.0784/237g
Sb	300	2.0844/87g		
Sb	77	2.0004/211g	2.0704/230g	2.1044/342g

In the table the temperatures are nominal figures. The g-values are accurate only to about 0.1% and the line-widths are measured between the two points of inflection. Ordinarily the defects are identified by g-values and relaxations during annealling studies. No literature g-value information has been located for these two dopants. The limited time available for this project prevented any serious annealling experiments, however some annealling has been done and more studies are recommended for the future. It appears that such work could shed light on the defects involved.

V. CONCLUSIONS

The ten-week effort of this summer project leads to the following conclusions:

1) A reasonable Hall Effect/resistivity apparatus has been established in the Physics Department at the University of Dayton and results of measurements on silicon and gallium arsenide are in good agreement with comparable measurements

at the Materials Laboratory.

2) Measurements conducted during this project have yielded values for impurity energy levels previously unreported for indium or antimony in gallium arsenide.

3) In gallium arsenide doped with indium or antimony, the electron paramagnetic resonance spectrometer in the Physics Department at the University of Dayton can detect weak signals at room temperature and at liquid nitrogen temperature. Further measurements and annealling studies could yield useful information on the defects responsible.

VI. RECOMMENDATIONS

The ten-week summer research project resulted in the following recommendations:

1) Supplemental measurements for the Materials Laboratory should be made on the University of Dayton Hall apparatus.

2) The Hall apparatus should be enhanced by improving the temperature control at the low temperature range, 4 - 77K.

3) Additional electron paramagnetic resonance measurements should be made on the University of Dayton spectrometer in order to learn more about the nature of the defects involved. Computer simulation of spectra should be pursued to isolate adjacent spectral lines.

4) For electron paramagnetic resonance measurements, the samples should be carefully x-ray oriented due to the complicated rotation patterns observed.

5) Annealling studies should be conducted to aid in the identification of the semiconductor defects.

6) The possibility of electron paramagnetic resonance measurements at liquid helium temperaure should be explored.

VIII. REFERENCES

1 Putley, E. H., The Hall Effect and Semi-Conductor Physics, New York, NY, Dover Publications, 1960.

2 Hemenger, Patrick M., "Measurement of High Resistivity Semiconductors Using the van der Pauw Method," Rev. Sci. Instrum. 44, 698 (1973).

3 Look, D. C., and J. W. Farmer, "Automated, High Resistivity Hall Effect and Photoelectronic Apparatus," J. Phys. E: Sci. Instrum. 14, 472 (1981).

4 Miner, G. K., T. P. Graham, and G. T. Johnston, "The Preparation and Calibration of a Convenient EPR Field Marker and Intensity Reference," Rev. Sci. Instrum. 43, 1297 (1972).

5 Hemenger, P. M., Private Communication.

6 Lang, J. E., Private Communication.

7 Lang, Joseph E., Frank L. Madarasz, and Patrick M. Hemenger, "Temperature Dependent Density of States Effective Mass in Nonparabolic p-Type Silicon," J. Appl. Phys. 54, 3612 (1983).

8 Lang, J. E., Private Communication.

9 Kaufmann, U., and J. Schneider, "Chromium as a Hole Trap in GaP and GaAs," Appl. Phys. Lett. 36, 747 (1980).

10 Goltzene, A., B. Meyer, and C. Schwab, "Electron Paramagnetic Resonance of Extended Defects in Semi-insulating GaAs," J. Appl. Phys. 53, 4541 (1982).

11 Blakemore, J. S., "Semiconducting and Other Major Properties of GaAs," J. Appl. Phys. 53, R123 (1982).

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

Synergetic Maneuvers

Prepared by:	Dr. Don Mittleman
Academic Rank:	Professor
Department and University:	Mathematics Oberlin College
Research Location:	Air Force Wright Aeronautical Laboratories, Flight Dynamics Laboratory, High Speed Aero Performance Branch.
USAF Research:	David T. Johnson
Date:	September 1, 1983
Contract No:	F49620-82-C-0035

Synergetic Maneuvers

by

Don Mittleman

Abstract

A procedure that may allow the specification of a synergetic maneuver from one Keplerian orbit to another, lying in a different plane and with different pedigree, depends on knowing the equations of motion of the vehicle on the transfer orbit. When the transfer orbit is also Keplerian, solutions to this problem for different impulse requirements have been given. When the transfer orbit is not Keplerian but subject to drag, a solution is known for an optimal, two-impulse transfer trajectory subject to a very special form of the drag. To attempt to apply the technique to the case of interest, hypersonic flight through a resisting atmosphere, requires the reformulation of the differential equations. Initially, these equations are cast in a form suitable for numerical integration on a high-speed computer. The desired form needs to be more amenable for theoretical analysis. This reformulation is done in this report. Suggestions for continued research in this area are offered.

Acknowledgement

The author wishes to express his thanks and appreciation to the Air Force Systems Command, the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for offering him this opportunity to spend a productive and interesting summer at the Flight Dynamics Laboratory of the Air Force Wright Aerodynamical Laboratories, Wright-Patterson AFB, Ohio. To the Flight Dynamics Laboratory, and in particular to the High Speed Aero Performance Branch, he wishes to acknowledge their hospitability and interest in his work.

He would like also to extend his thanks to David T. Johnson for suggesting this area of research and for his continuing support of this effort. Our conversations were very helpful.

Finally, thanks are due to Donald J. Jezewski of NASA's Johnson Space Center, Houston, Texas, for meaningful discussions during the course of this work.

I. Introduction

The problem addressed in this report is the formulation of the equations of motion governing the orbit-to-orbit transfer of a vehicle. Only the aerodynamic forces and the equations incorporating these are considered; the inclusion of aerodynamic heating as a constraint is deferred to a later report.

Orbit-to-orbit transfers have been considered by many authors. References are not intended to be inclusive of the literature; rather, they are directed more as indicators of the thinking that underlies the material discussed herein. The problems that have been successfully solved include: a) transfer between circular orbits lying in the same plane and requiring minimum energy:[1] b) transfer between Keplerian orbits using two fixed impulses[2]; c) transfer between Keplerian orbits when the transfer orbit contains a velocity-dependent term and requiring minimum energy[3]. D.J. Jezewski,[4] using an impulse optimization method and the results of an atmospheric guidance logic, studied the problem of going from a specified initial Keplerian orbit to a transfer orbit consisting of three parts to a fixed Keplerian orbit. The three parts of the transfer orbit are an initial phase which is Keplerian, an intermediate phase which is motion through an atmosphere, and a third phase which is again Keplerian. The intermediate phase, the atmospheric portion of the solution, is discussed in Hill[5].

With the exception of this intermediate phase, which is a numerical solution of a system of differential equations requiring

approximately 10 minutes on a high-speed computer (Univac 1110 series), all of the problems mentioned above are solved in closed form and in terms of known functions. The purpose of this report is to attempt to approximate this numerical computation by a closed form solution in terms of known functions.

The proposed method of solution is based on the following observations. Historically, the solution of the differential equations of motion for an orbital trajectory is an initial value problem. Given the equations of motion and the initial conditions, one studies the motion, the position and velocity, along the trajectory, generally as a function of time or in terms of some other variable closely related to time. That is not the case for orbit-to-orbit transfer. Here, one starts the vehicle at some point, to be determined as a condition of solution of the problem, on an initial Keplerian orbit, with an initial velocity, also to be determined as a condition of the solution. Having started the vehicle with these initial conditions, the equations of motion of the vehicle through the resisting medium are operative until the vehicle reaches some point, again possibly to be determined as a condition of the solution of the problem, on a prescribed final orbit, with a final velocity, again possibly to be determined as a condition of solution of the problem. For a problem such as this, it is clear that only the initial and final positions and velocities of the vehicle are of interest (subject, of course, to the physical constraints on the vehicle during the transition along the transfer trajectory). For these reasons, when we try to find a solution both of whose endpoints may be prescribed, we refer to the problem as a boundary value problem.

If one could write the solution to the initial value problem in closed form and in terms of known functions, (something we have not, in general, learned to do) then, at least conceptually, one should be able to specify the conditions for the boundary value problem and determine the values of the constants of integration in terms of the boundary conditions. The algebraic complications are formidable and recourse is made to an alternate approach. The alternative method depends on deriving the integrals of motion. Thus, Jezewski & Mittleman[3] depends in part on Jezewski & Mittleman[6].

II. Objectives

The main objective of this project is to recast the equations of motion for hypersonic flight mechanics, generally given in the form of a system of differential equations suitable for solution by numerical methods on a high speed computer, into a form better suited for analytic study. In particular, the formulation chosen for this analytic study had been used successfully in studying orbit-to-orbit transfer in the Keplerian case and when the transfer orbit is modified by a specialized drag. This relationship is established and the several terms used in one formulation are identified and related to their correspondents in the other.

III. The Equation of Motion in a Resisting Medium

The development of the equations of motion in this section is taken from Vinh, Busemann, Culp[7], specifically chapters 2 and 13. For completeness only, the relevent equations are repeated here. The reader is referred to their text for a full description. A dual number system is used in this section for the equations given. When

written in parenthesis, such as (2-27), the number is the same as in their text; when written in brackets, such as [3-4], the number will be the one used in this report. Thus, [3-4] will be the fourth equation in section III.

Consider a spherical rotating planet and let $OX_1Y_1Z_1$ be an inertial coordinate system with O at the sphere's center and with the OX_1Y_1 plane coincident with the equitorial plane. Let OXYZ be a coordinate system fixed in the planet and rotating with it with constant angular velocity . At time t = 0, OX_1 and OX coincide. Newton's equation

$$m \frac{d\vec{V}}{dt} = \vec{F} \qquad (2\text{-}3) \qquad [3\text{-}1]$$

valid in the inertial system, becomes in the rotating system

$$m \frac{d\vec{V}}{dt} = \vec{F} - 2m\vec{\omega} \times \vec{V} - m\vec{\omega} \times (\vec{\omega} \times \vec{r}) \qquad (2\text{-}13) \qquad [3\text{-}2]$$

where now $\vec{V}$ is the velocity in the rotating coordinate system. Using spherical coordinates fixed in the planet, the position vector $\vec{r}$ is given by its magnitude r, its longitude θ (measured from the OX axis positively eastward), and its latitude φ (measured from the equatorial plane along a meridian, positive northward).

A third coordinate system Oxyz is introduced. The Ox axis is along the vector $\vec{r}$, the y-axis, orthogonal to Ox, is in the equitorial plane and positive in the direction of motion, and Oz is chosen to form a right-handed system.

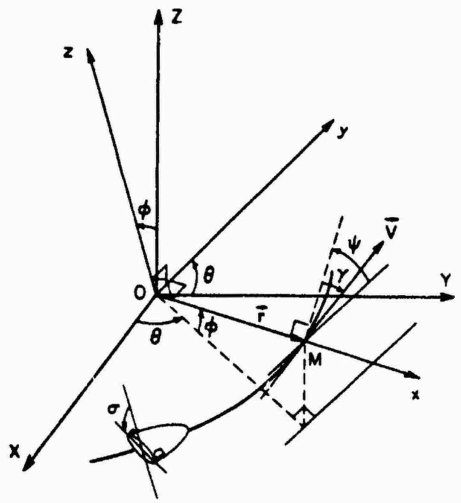


Fig. 2-3. Coordinate systems.

If this coordinate system is translated so that its origin coincides with the end-point of the position vector $\vec{r}$, then Ox is in the direction of the outward drawn normal to the sphere with center O and radius r, and the Oyz plane, called the local horizontal plane, is tangent to the sphere at that point. The Oy direction is tangent to the circle of latitude and we shall assume positively directed eastward and Oz is tangent to the meridian circle and positively directed northward.

The velocity vector $\vec{V}$ lies in the local horizonal plane only for "level" flight. Otherwise, $\vec{V}$ makes an angle γ, the flight path angle, with this plane; $\gamma > 0$ implies the vehicle is climbing. The angle ψ, called the heading, is measured in the tangent plane from the Oy direction (i.e. from the direction of the circle of latitude) and is positive when the vehicle is heading north of east.

Let $\vec{i}, \vec{j}, \vec{k}$ be unit vectors in the direction Ox, Oy, Oz respectively. Then we may write

$$\vec{r} = r\vec{i} \qquad (2\text{-}14) \qquad [3\text{-}3]$$

$$\vec{V} = (V\sin\gamma)\vec{i} + (V\cos\gamma\cos\psi)\vec{j} + (V\cos\gamma\sin\psi)\vec{k} \qquad (2\text{-}15) \qquad [3\text{-}4]$$

$$\vec{\omega} = (\omega\sin\phi)\vec{i} + (\omega\cos\phi)\vec{k} \qquad (2\text{-}16) \qquad [3\text{-}5]$$

From these three equations, we calculate

$$\vec{\omega} \times \vec{V} = -(\omega V \cos\gamma \cos\phi \cos\psi)\,\vec{i} + \omega V(\sin\gamma \cos\phi - \cos\gamma \sin\phi \sin\psi)\,\vec{j} + \omega V \cos\gamma \sin\phi \cos\psi\,\vec{k} \qquad (2\text{-}17) \qquad [3\text{-}6]$$

$$\vec{\omega} \times (\vec{\omega} \times \vec{r}) = -\omega^2 r \cos^2\phi\,\vec{i} + \omega^2 r \sin\phi \cos\phi\,\vec{k} \qquad (2\text{-}18) \qquad [3\text{-}7]$$

values needed in eq. (2-13), [3-2].

The gravitational component of the force F is

$$m\vec{g} = -mg(r)\,\vec{i} \qquad (2\text{-}19) \qquad [3\text{-}8]$$

A drag force $\vec{D}$, oppositively directed to the velocity vector $\vec{V}$, and a lift force $\vec{L}$, orthogonal to $\vec{D}$, comprise the aerodynamic forces $\vec{A}$. We assume a thrust vector $\vec{T}$ lying in the $(\vec{D},\vec{L})$ plane and making an angle ϵ with the velocity vector $\vec{V}$. Then $T\cos\epsilon$ is the component of $\vec{T}$ along the velocity $\vec{V}$ and $T\sin\epsilon$ is the component of $\vec{T}$ along $\vec{L}$. For convenience, the aerodynamic and propulsive forces are combined, and we define

$$F_T = T\cos\epsilon - D \qquad [3\text{-}9.1]$$

$$F_N = T\sin\epsilon + L \qquad (2\text{-}20) \qquad [3\text{-}9.2]$$

Thus F_T is the component of the combined aerodynamic and propulsive forces along $\vec{V}$ and F_N is their component along $\vec{L}$. If F_T and F_N are written in the $(\vec{i},\vec{j},\vec{k})$ system, we have:

$$\vec{F}_T = (F_T \sin\gamma)\,\vec{i} + (F_T \cos\gamma \cos\psi)\,\vec{j} + (F_T \cos\gamma \sin\psi)\,\vec{k} \qquad (2\text{-}21) \qquad [3\text{-}10.1]$$

$$\vec{F}_N = (F_N \cos\sigma \cos\gamma)\,\vec{i} - (F_N \cos\sigma \sin\gamma \cos\psi + F_N \sin\sigma \sin\psi)\,\vec{j} - (F_N \cos\sigma \sin\gamma \sin\psi - F_N \sin\sigma \cos\psi)\,\vec{k} \qquad (2\text{-}24) \qquad [3\text{-}10.2]$$

where σ is the angle between the vector $\vec{L}$ and the $(\vec{r},\vec{V})$ plane, which, in Fig.(2.4), is referred to as the vertical plane.

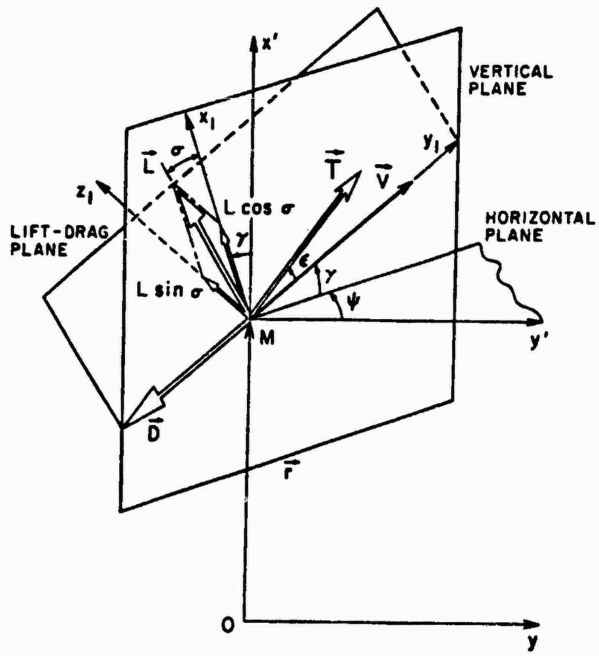


Fig. 2-4. Aerodynamic forces and thrust components.

Equation (2-13), [3-2] was written for the OXYZ coordinate system whereas the development given above was in the Oxyz, or equivalently, the $(\vec{i},\vec{j},\vec{k})$ system. To obtain the equivalent of eq. (2-13), [3-2] in the $(\vec{i},\vec{j},\vec{k})$ system, we need the angular velocity $\vec{\Omega}$ of the Oxyz system with respect to the OXYZ system. This is

$$\vec{\Omega} = (\sin\phi \frac{d\theta}{dt})\vec{i} - (\frac{d\phi}{dt})\vec{j} + (\cos\phi \frac{d\theta}{dt})\vec{k} \qquad (2\text{-}25) \qquad [3\text{-}11]$$

Then

$$\frac{d\vec{i}}{dt} = \vec{\Omega} \times \vec{i} = (\cos\phi \cdot \frac{d\theta}{dt})\vec{j} + (\frac{d\phi}{dt})\vec{k} \qquad [3\text{-}12.1]$$

$$\frac{d\vec{j}}{dt} = \vec{\Omega} \times \vec{j} = -(\cos\phi \frac{d\theta}{dt})\vec{i} + (\sin\phi \frac{d\theta}{dt})\vec{k} \qquad (2\text{-}26) \qquad [3\text{-}12.2]$$

$$\frac{d\vec{k}}{dt} = \vec{\Omega} \times \vec{k} = -(\frac{d\phi}{dt})\vec{i} - (\sin\phi \frac{d\theta}{dt})\vec{j} \qquad [3\text{-}12.3]$$

From eq. (2-14), [3-3] and using eq.(2-16), [3-5] we get

$$\frac{d\vec{r}}{dt} = (\frac{dr}{dt})\vec{i} + (r\cos\phi \frac{d\theta}{dt})\vec{j} + (r\frac{d\phi}{dt})\vec{k} \qquad (2\text{-}27) \qquad [3\text{-}13]$$

Comparing eqs. (2-27), [3-13] and (2-15), [3-4] we get the three kinematic equations.

$$\frac{dr}{dt} = V \sin\gamma \qquad [3\text{-}14.1]$$

$$\frac{d\theta}{dt} = \frac{V \cos\gamma \cos\psi}{r \cos\phi} \qquad (2\text{-}28) \qquad [3\text{-}14.2]$$

$$\frac{d\phi}{dt} = \frac{V \cos\gamma \sin\psi}{r} \qquad [3\text{-}14.3]$$

If we compute $\frac{d\vec{V}}{dt}$ from eq. (2-15), [3-4] and simplify using eqs. (2-26), [3-12] and (2-28), [3-14] we have

$$\begin{aligned}\frac{d\vec{V}}{dt} = &\left[\sin\gamma \frac{dV}{dt} + V\cos\gamma \frac{d\gamma}{dt} - \frac{V^2}{r}\cos^2\gamma\right]\vec{i} \\ &+ \left[\cos\gamma\cos\psi \frac{dV}{dt} - V\sin\gamma\cos\psi \frac{d\gamma}{dt} - V\cos\gamma\sin\psi \frac{d\psi}{dt}\right. \\ &\left. + \frac{V^2}{r}\cos\gamma\cos\psi(\sin\gamma - \cos\gamma\sin\psi\tan\phi)\right]\vec{j} \\ &+ \left[\cos\gamma\sin\psi \frac{dV}{dt} - V\sin\gamma\sin\psi \frac{d\gamma}{dt} + V\cos\gamma\cos\psi \frac{d\psi}{dt}\right. \\ &\left. + \frac{V^2}{r}\cos\gamma(\sin\gamma\sin\psi + \cos\gamma\cos^2\psi\tan\phi)\right]\vec{k}\end{aligned} \qquad [3\text{-}15]$$

(2-29)

Equation (2-29), [3-15] is used to replace $\frac{d\vec{V}}{dt}$ in the left-hand side of eq. (2-13), [3-2]; the right hand side of eq. (2-13) is found from $\vec{\omega}\times\vec{V}$, eq. (2-17), [3-6] from $\vec{\omega}\times(\vec{\omega}\times\vec{r})$ eq. (2-18), [3-7] and F, from eqs. (2-19), [3-8], (2-21), [3-10.1] and (2-24), [3-10.2]. These substitutions result in three equations which are then solved for $\frac{dV}{dt}$, $\frac{d\gamma}{dt}$ and $\frac{d\psi}{dt}$.

$$\frac{dV}{dt} = \frac{1}{m}F_T - g\sin\gamma + \omega^2 r\cos\phi\,(\sin\gamma\cos\phi - \cos\gamma\sin\phi\sin\psi) \qquad [3\text{-}16.1]$$

$$\begin{aligned}V\frac{d\gamma}{dt} = &\frac{1}{m}F_N\cos\sigma - g\cos\gamma + \frac{V^2}{r}\cos\gamma + 2\omega V\cos\phi\cos\psi \\ &+ \omega^2 r\cos\phi\,(\cos\gamma\cos\phi + \sin\gamma\sin\phi\sin\psi)\end{aligned} \qquad [3\text{-}16.2]$$

$$\begin{aligned}V\frac{d\psi}{dt} = &\frac{1}{m}\frac{F_N\sin\sigma}{\cos\gamma} - \frac{V^2}{r}\cos\gamma\cos\psi\tan\phi \\ &+ 2\omega V(\tan\gamma\cos\phi\sin\psi - \sin\phi) - \frac{\omega^2 r}{\cos\gamma}\sin\phi\cos\phi\cos\psi\end{aligned} \qquad [3\text{-}16.3]$$

(2-31)

For a non-rotating planet and stationary atmosphere, $\vec{\omega} = 0$ and eqs. (2-31), [3-16] reduce to

$$\frac{dV}{dt} = \frac{1}{m} F_T - g \sin\gamma \qquad [3\text{-}17.1]$$

$$V \frac{d\gamma}{dt} = \frac{1}{m} F_N \cos\sigma - g\cos\gamma + \frac{V^2}{r}\cos\gamma \qquad (2\text{-}32) \qquad [3\text{-}17.2]$$

$$V \frac{d\psi}{dt} = \frac{1}{m} \frac{F_N \sin\sigma}{\cos\gamma} - \frac{V^2}{r} \cos\gamma \cos\psi \tan\phi \qquad [3\text{-}17.3]$$

For a non-thrusting vehicle, F_T and F_N simplify to -D and L respectively and eqs. (2-32), [3-17] become:

$$\frac{dV}{dt} = -\frac{D}{m} - g \sin\gamma \qquad [3\text{-}18.1]$$

$$V\frac{d\gamma}{dt} = \frac{L\cos\sigma}{m} - g\cos\gamma + \frac{V^2}{r}\cos\gamma \qquad (2\text{-}34) \qquad [3\text{-}18.2]$$

$$V\frac{d\psi}{dt} = \frac{L\sin\sigma}{m\cos\gamma} - \frac{V^2}{r}\cos\gamma\cos\psi\tan\phi \qquad [3\text{-}18.3]$$

The formula for the drag is given by

$$D = -\frac{\rho S C_D V^2}{2}$$

and for the lift

$$L = \frac{\rho S C_L V^2}{2}$$

where ρ is the density of the resisting medium, S the shape factor, and C_D and C_L are the drag and lift coefficients respectively, of the vehicle.

In summary, the equations of interest are:

$$\frac{dr}{dt} = V \sin\gamma \qquad [3\text{-}19.1]$$

$$\frac{d\theta}{dt} = \frac{V \cos\gamma \cos\psi}{r \cos\phi} \qquad [3\text{-}19.2]$$

$$\frac{d\phi}{dt} = \frac{V \cos\gamma \sin\psi}{r} \qquad [3\text{-}19.3]$$

$$\frac{dV}{dt} = -\frac{\rho S C_D V^2}{2m} - g \sin\gamma \qquad [3\text{-}19.4]$$

$$V \frac{d\gamma}{dt} = \frac{\rho S C_L V^2}{2m} \cos\sigma - \left(g - \frac{V^2}{r}\right) \cos\gamma \qquad [3\text{-}19.5]$$

$$V \frac{d\psi}{dt} = \frac{\rho S C_L V^2}{2m \cos\gamma} \sin\sigma - \frac{V^2}{r} \cos\gamma \cos\psi \tan\phi \qquad (13\text{-}1) \qquad [3\text{-}19.6]$$

Again, the reader is reminded that this presentation is taken from Vinh, Busemann, Culp[7], chapter 2 and 13. For further details, the reader is referred to that text.

IV. An Alternate Formulation for the Equations of Motion

In Mittleman & Jezewski[8] an analytic solution to the two-body problem with a specified drag model is obtained. In this section, a more general procedure that includes these results is presented[9].

The general two-body problem may be described by the vector differential equation

$$\ddot{\vec{r}} + \alpha_1 \dot{\vec{r}} + \alpha_2 \vec{r} + \alpha_3 (\vec{r} \times \dot{\vec{r}}) = 0 \qquad [4\text{-}1]$$

where $\vec{r}$ is the radius vector from the center of mass of the attracting body to the particle and the α_i ($i = 1,2,3$) are scalar coefficients dependent upon $\vec{r}$, $\dot{\vec{r}}$, the characteristics of the resisting medium as well as those of the vehicle. Dots ($\dot{}$) denote differentiation with respect to the independent variable, time. In

all that follows, we specifically exclude rectilinear motion, i.e. we assume that $\vec{r}$, $\dot{\vec{r}}$ and rxr are three non-coplanar vectors. Thus, $\vec{r}$, $\dot{\vec{r}}$ and $\vec{r}\times\dot{\vec{r}}$ form a basis for the space. We introduce a (ξ,η,ζ) coordinate system as follows.

Let

$$\xi = \frac{\vec{r}}{r} \quad , \quad r = |\vec{r}| \qquad [4\text{-}2]$$

$$\eta = \frac{r\dot{\vec{r}} - \dot{r}\vec{r}}{h} \quad , \quad h = |\vec{r}\times\dot{\vec{r}}| \qquad [4\text{-}3]$$

From these definitions we conclude that

a) ξ is a unit vector. This is obvious from the definition.

b) ξ and η are orthogonal. We show that the inner (dot) product of ξ and η is zero.

$$\xi\cdot\eta = \left(\frac{\vec{r}}{r}\right)\cdot\left(\frac{r\dot{\vec{r}} - \dot{r}\vec{r}}{h}\right) = \frac{1}{rh}\left(r\,\vec{r}\cdot\dot{\vec{r}} - \dot{r}\,\vec{r}\cdot\vec{r}\right)$$

$$= \frac{1}{rh}\left(r^2\dot{r} - r^2\dot{r}\right) = 0$$

c) η is a unit vector. From

$$\xi\times\eta = \left(\frac{\vec{r}}{r}\right)\times\left(\frac{r\dot{\vec{r}} - \dot{r}\vec{r}}{h}\right) = \frac{\vec{r}\times\dot{\vec{r}}}{h} = \frac{\vec{r}\times\dot{\vec{r}}}{|\vec{r}\times\dot{\vec{r}}|}$$

Thus, $|\xi\times\eta| = 1$. However,

$$|\xi\times\eta| = |\xi||\eta|\sin\left(\frac{\pi}{2}\right) = |\eta| = 1$$

d) ξ , η , and $\zeta = \xi\times\eta$ are unit vectors, forming a right-handed orthonormal coordinate system.

At this point, we introduce a pseudo-time $\tau = \tau(t)$. In terms of τ

$$\vec{r}(\tau) = r(\tau)\,\xi(\tau)$$

$$\dot{\vec{r}}(\tau) = (r'\xi + r\xi')\,\dot{\tau}$$

where the prime (') indicates differentiation with respect to τ.

We show that for the proper choice of $\tau(t)$, $\eta = \xi'$.

From the definition of η, we find

$$\eta = \frac{[r(r\xi' + r'\xi) - r r'\xi]\dot{\tau}}{h} \quad , \quad h = |r\xi \times (r\xi' + r'\xi)\dot{\tau}|$$

$$\eta = \frac{r^2 \xi' \dot{\tau}}{r^2 |\xi'||\dot{\tau}|} = \frac{\xi' \dot{\tau}}{|\xi'||\dot{\tau}|}$$

We choose τ to be *i)* an increasing function of t, to insure that $\dot{\tau} > 0$ and *ii)* we choose τ so that $|\xi'| = 1$. (It is not important, at this point, to know the precise relation between τ and t. That relationship will become apparent during the solution of the problem. In the Keplerian case, τ is the true anomoly; in the case discussed on Mitt-leman & Jezewski[8], it is a central angle; in Vinh, Busemann, Culp[7], it is the parameter s.(Eq. (13-6), pg. 229). The first two cases are obvious from the development described in the references; the third assertion will be proved in this paper.

We now can determine $\ddot{\vec{r}}$ in terms of τ.

$$\ddot{\vec{r}} = (r\xi' + r'\xi)\ddot{\tau} + (r\xi'' + 2r'\xi' + r''\xi)\dot{\tau}^2 \qquad [4\text{-}4]$$

Using these values for $\vec{r}$, $\dot{\vec{r}}$, and $\ddot{\vec{r}}$, eq. [4-1] becomes

$$\xi'' + \nu_1 \xi' + \nu_2 \xi + \nu_3 (\xi \times \xi') = 0 \qquad [4\text{-}5]$$

where

$$\nu_1 = \frac{\ddot{\tau}}{\dot{\tau}^2} + \frac{2r'}{r} + \frac{\alpha_1}{\dot{\tau}}$$

$$\nu_2 = \frac{r'\ddot{\tau}}{r\dot{\tau}^2} + \frac{r''}{r} + \frac{r'\alpha_1}{r\dot{\tau}} + \frac{\alpha_2}{\dot{\tau}^2}$$

$$\nu_3 = \frac{r\alpha_3}{\dot{\tau}}$$

e) We prove: $\nu_1 \equiv 0$ and $\nu_2 \equiv 1$

The dot product of ξ, and eq. [4-5] yields

$$\xi \cdot \xi'' + \nu_2 = 0$$

However, $(\xi \cdot \xi'') = (\xi \cdot \xi')' - (\xi' \cdot \xi') = -1$

Therefore, $\nu_2 \equiv 1$

The dot product of ξ' and eq. [4-5] yields

$$\xi' \cdot \xi'' + \nu_1 = 0$$

However, $(\xi' \cdot \xi'') = \frac{1}{2} (\xi' \cdot \xi')' = 0$

Therefore, $\nu_1 = 0$

With ν_3 as defined above, eq. [4-5] may be written

$$\xi'' + \xi + \nu_3 (\xi \times \xi') = 0 \qquad [4\text{-}6]$$

f) The vector ξ satisfies a linear, third order differential equation.

The eq. [4-6] is obviously non-linear. If $\nu_3 = 0$, then ξ is a solution of $\xi'' + \xi = 0$ and clearly the assertion is true. If $\nu_3 \neq 0$, then, dividing by ν_3 and differentiating, we get

$$\left(\frac{\xi'' + \xi}{\nu_3} \right)' + (\xi \times \xi')' = 0 \qquad [4\text{-}7]$$

However, $(\xi \times \xi')' = (\xi' \times \xi') + (\xi \times \xi'') = (\xi \times \xi'')$

The cross product of ξ and eq. [4-6] is

$$\xi \times \xi'' + \nu_3 [\xi \times (\xi \times \xi')] = 0$$

From the vector identity for the triple cross product

$$\xi \times (\xi \times \xi') = (\xi \cdot \xi') \xi - (\xi \cdot \xi) \xi' = -\xi'$$

Therefore,

$$(\xi \times \xi')' = (\xi \times \xi'') = -\nu_3 [\xi \times (\xi \times \xi')] = \nu_3 \xi'$$

Substituting this into eq. [4-7], we have

$$\left(\frac{\xi'' + \xi}{\nu_3} \right)' + \nu_3 \xi' = 0 \qquad [4\text{-}8]$$

which when expanded becomes

$$\nu_3 \xi''' - \nu_3' \xi'' + \nu_3 (1+\nu_3^2) \xi' - \nu_3 \xi = 0 \qquad [4\text{-}8]$$

g) An observation. That the vector ξ satisfies a <u>linear</u> vector third order differential equation need not be surprising. Since the constants of integration of any linear differential equation enter linearly in the solution, we should expect that

$$\xi = \vec{A}_1 f_1(\tau) + \vec{A}_2 f_2(\tau) + \vec{A}_3 f_3(\tau)$$

where the vectors $\vec{A}_1$, $\vec{A}_2$ $\vec{A}_3$ are three constants and the scalar functions $f_i(\tau)$, $i = 1,2,3$ are linearly independent. Since the vector constants $\vec{A}_i$ are linearly independent, they form a basis for the three-space and thus any vector, and in particular ξ, is expressible as a linear combination of these constants.

Conversely, if $\vec{A}_1$, $\vec{A}_2$, $\vec{A}_3$ are any three linearly independent vectors, then the vector function

$$\xi(\tau) = \vec{A}_1 f_1(\tau) + \vec{A}_2 f_2(\tau) + \vec{A}_3 f_3(\tau)$$

satisfies a linear third order differencial equation, provided, of course, that f_1''', f_2''', f_3''' exist.

V. <u>The Meaning of $\nu_1 = 0$ and $\nu_2 = 1$.</u>

Recall the definition of

$$h = |\vec{r} \times \dot{\vec{r}}| = r^2 \dot{\tau}$$

and differentiate with respect to τ.

$$h' = 2rr'\dot{\tau} + r^2 \frac{\ddot{\tau}}{\dot{\tau}} = r^2 \left[\frac{2r'}{r} + \frac{\ddot{\tau}}{\dot{\tau}^2} \right] \dot{\tau}$$

Substituting from $\nu_1 = 0$, we find

$$h' + r^2 \alpha_1 = 0 \qquad [5\text{-}1]$$

When $\nu_1 = 0$ is combined with $\nu_2 = 1$ that equation becomes

$$\frac{r''}{r} - 2\left(\frac{r'}{r}\right)^2 + \frac{\alpha_2}{\dot{\tau}^2} = 1 \qquad [5-2]$$

If we let $u = \frac{1}{r}$ and $\dot{\tau} = h/r^2$, eq. [5-2] becomes

$$u'' + u = \frac{\alpha_2}{h^2 u^3} \qquad [5-3]$$

The value of ν_3 in eq. [4-6] or eq. [4-8] is

$$\nu_3 = \frac{\alpha_3}{h u^3} \qquad [5-4]$$

The quantity h is, of course, the angular momentum and eq. [5-1] indicates how h changes with τ and the "forces" as given in α_1. If eq. [5-1] is integrable, and it is known to be so in the case of Keplerian motion when $\alpha_1 = 0$ and in the case $\alpha_1 = \alpha_0/r^2$ where α_0 is a constant, then one integral of the motion would be known. When this integral for the angular momentum is substituted into eq. [5-3], that becomes a second order differential equation in u. If $\alpha_2 = \mu u^3$, the classical case of an inverse square gravitational field, eq. [5-3] becomes linear, provided, of course, that h is a function of τ alone. It is in this sense that we say that the system of equations [4-1] is decoupled. Note also that with this decoupling, the out-of-plane motion of the vehicle is now governed solely by eq. [4-6] or eq. [4-8] with ν_3 as given by [5-4].

The equations [5-1], [5-2] and [4-7], with ν_3 as given in [5-4] constitute a system of differential equations that, for specific values of $\alpha_1, \alpha_2, \alpha_3$, define the motion of a particle under the most general conditions. We turn our attention now to determining $\alpha_1, \alpha_2, \alpha_3$ in terms of the aerodynamic and gravitational forces as given in Section III.

VI. Aerodynamic Values for $\alpha_1, \alpha_2, \alpha_3$.

We start with the differential equation

$$\ddot{\vec{r}} + \alpha_1 \dot{\vec{r}} + \alpha_2 \vec{r} + \alpha_3 (\vec{r} \times \dot{\vec{r}}) = 0 \qquad [6\text{-}1]$$

Working in the $\vec{i}$, $\vec{j}$, $\vec{k}$ coordinate system, we recall eqs. (2-14),[3-3], (2-15),[3-4], and (2-29),[3-15].

$$\vec{r} = r\,\vec{i} \qquad (2\text{-}14) \qquad [6\text{-}2]$$

$$\vec{V} = (V\sin\gamma)\,\vec{i} + (V\cos\gamma\cos\psi)\,\vec{j} + (V\cos\gamma\sin\psi)\,\vec{k} \qquad (2\text{-}15) \qquad [6\text{-}3]$$

$$\begin{aligned}\frac{d\vec{V}}{dt} = {} & \left[\sin\gamma\frac{dV}{dt} + V\cos\gamma\frac{d\gamma}{dt} - \frac{V^2}{r}\cos^2\gamma\right]\vec{i} \\ & + \left[\cos\gamma\cos\psi\frac{dV}{dt} - V\sin\gamma\cos\psi\frac{d\gamma}{dt} - V\cos\gamma\sin\psi\frac{d\psi}{dt}\right. \\ & \left. + \frac{V^2}{r}\cos\gamma\cos\psi(\sin\gamma - \cos\gamma\sin\psi\tan\phi)\right]\vec{j} \\ & + \left[\cos\gamma\sin\psi\frac{dV}{dt} - V\sin\gamma\sin\psi\frac{d\gamma}{dt} + V\cos\gamma\cos\psi\frac{d\psi}{dt}\right. \\ & \left. + \frac{V^2}{r}\cos\gamma(\sin\gamma\sin\psi + \cos\gamma\cos^2\psi\tan\phi)\right]\vec{k}\end{aligned} \qquad [6\text{-}4]$$

(2-29)

If the values of $\vec{r}$, $\vec{V} \equiv \dot{\vec{r}}$ and $\frac{d\vec{V}}{dt} = \ddot{\vec{r}}$ as given by these equations are substituted into eq. [6-1], we obtain the three scalar equations

$$\begin{aligned}(V\sin\gamma)\alpha_1 + r\alpha_2 &= P_1 \\ (V\cos\gamma\cos\psi)\alpha_1 - (rV\cos\gamma\sin\psi)\alpha_3 &= P_2 \\ (V\cos\gamma\sin\psi)\alpha_1 + (rV\cos\gamma\cos\psi)\alpha_3 &= P_3\end{aligned} \qquad [6\text{-}5]$$

where

$$P_1 = -\sin\gamma\frac{dV}{dt} - V\cos\gamma\frac{d\gamma}{dt} + \frac{V^2}{r}\cos^2\gamma$$

$$P_2 = -\cos\gamma\cos\psi\frac{dV}{dt} + V\sin\gamma\cos\psi\frac{d\gamma}{dt} + V\cos\gamma\sin\psi\frac{d\psi}{dt} - \frac{V^2}{r}\cos\gamma\cos\psi(\sin\gamma - \cos\gamma\sin\psi\tan\phi)$$

$$P_3 = -\cos\gamma\sin\psi\,\frac{dV}{dt} + V\sin\gamma\sin\psi\,\frac{d\gamma}{dt} - V\cos\gamma\cos\psi\,\frac{d\psi}{dt}$$
$$-\frac{V^2}{r}\cos\gamma\left(\sin\gamma\sin\psi + \cos\gamma\cos^2\psi\tan\phi\right).$$

The second and third of eqs. (6-5) are equivalent to

$$\cos\psi\,(V\cos\gamma\,\alpha_1) - \sin\psi\,(Vr\cos\gamma\,\alpha_3) = P_2$$
$$\sin\psi\,(V\cos\gamma\,\alpha_1) + \cos\psi\,(Vr\cos\gamma\,\alpha_3) = P_3$$

which in turn are easily solvable for $(V\cos\gamma\,\alpha_1)$ and $(Vr\cos\gamma\,\alpha_3)$.

$$V\cos\gamma\,\alpha_1 = P_2\cos\psi + P_3\sin\psi$$
$$Vr\cos\gamma\,\alpha_3 = P_3\cos\psi - P_2\sin\psi$$

If P_2 and P_3 are replaced by their values as given above, we find

$$(V\cos\gamma)\,\alpha_1 = -\cos\gamma\,\frac{dV}{dt} + V\sin\gamma\,\frac{d\gamma}{dt} - \frac{V^2}{r}\sin\gamma\cos\gamma \qquad [6\text{-}6]$$

$$(Vr\cos\gamma)\,\alpha_3 = -V\cos\gamma\,\frac{d\psi}{dt} - \frac{V^2}{r}\cos^2\gamma\cos\psi\tan\phi \qquad [6\text{-}7]$$

The value for α_2 is found from the first of the eqs. [6-5].

$$r\alpha_2 = P_1 - V\sin\gamma\,\alpha_1 = -\frac{V}{\cos\gamma}\,\frac{d\gamma}{dt} + \frac{V^2}{r} \qquad [6\text{-}8]$$

If, in equations [6-6], [6-7] and [6-8], $\frac{dV}{dt}$, $\frac{d\gamma}{dt}$ and $\frac{d\psi}{dt}$ are replaced by their values as given in eqs. (2-32), [3-17], repeated for convenience,

$$\frac{dV}{dt} = \frac{1}{m}F_T - g\sin\gamma$$
$$V\frac{d\gamma}{dt} = \frac{1}{m}F_N\cos\sigma - g\cos\gamma + \frac{V^2}{r}\cos\gamma \qquad (2\text{-}32)$$
$$V\frac{d\psi}{dt} = \frac{1}{m}\frac{F_N\sin\sigma}{\cos\gamma} - \frac{V^2}{r}\cos\gamma\cos\psi\tan\phi$$

we arrive finally at

$$\alpha_1 = \frac{1}{mV}\left[F_N \cos\sigma \tan\gamma - F_T\right] \qquad [6\text{-}9.1]$$

$$\alpha_2 = \frac{q}{r} - \frac{1}{mr} F_N \frac{\cos\sigma}{\cos\gamma} \qquad [6\text{-}9.2]$$

$$\alpha_3 = -\frac{F_N \sin\sigma}{mVr\cos\gamma} \qquad [6\text{-}9.3]$$

Of course, these values for the α_i are predicated on a non-rotating planet with a stationary atmosphere (actually, we assume $\omega = 0$ in eqs.(2-31), [3-16]; for a rotating sphere, $\omega \neq 0$ and we would use the values of $\frac{dV}{dt}$, $\frac{d\gamma}{dt}$ and $\frac{d\psi}{dt}$ as given in eqs. (2-31),[3-16].*

VII. Recommendations

a) The first and most obvious extension of the method presented would be to include the case of a rotating planet and derive the equivalent of eqs. [6-9].

b) The third order linear differential equation [4-8] should be investigated for different V_3. Since the vectors ξ and ξ' that enter into that equation are unit orthogonal vectors and since the solution must also satisfy the non-linear eq. [4-6], necessary conditions will be imposed on the relationship among the constants of integration. It may be possible to determine this relationship for the general case. All of this effort would be directed to understanding the out-of-plane motion of the vehicle.

---

*That derivation will be given in a subsequent report.

c) Physical constraints (e.g. aerodynamic heating) on the motion have been ignored. Eventually, these should be included.

d) It is possible to choose α_1, α_2, and α_3 so that the eqs. [5-1], [5-3], and [5-4] are integrable in an analytic form involving known functions. It would be too much to expect that the cases of interest for the aerodynamicist would be among these. It may be possible however, to bound the cases of interest between integrable cases and thus get some feeling for the nature of the solution.

REFERENCES

1) Lawten, D.F., Optimal Trajectories for Space Navigation, 1963, Butterworth, London

2) Jezewski, D.J. and Mittleman, D., An Analytic Approach to Two-Fixed Impulse Transfers Between Keplerian Orbits. Journal of Guidance, Control, and Dynamics, Vol.5, No.5, Sep-Oct 1982, 458-464

3) Jezewski, D.J. and Mittleman,D., The Optimal, Two-Impulse Transfer Trajectory Affected by Drag, NASA JSC-17775, (81-FM-57) Dec 1981.

4) Jezewski, D.J., Impulsive Optimization and Aerodynamic Braking, NASA JSC-18767 (83-FM-2), Jan 1983.

5) Hill, Oliver, A Proposed Atmospheric Guidance Logic for a Maneuvering Orbital Transfer Vehicle, NASA JSC-18620 (82-FM-58) Oct 1982.

6) Jezewski, D.J. and Mittleman, Don, Integrals of Motion for the Classical Two-Body Problem with Drag; Int. Journal Non-Linear Mechanics, Vol 18., No. 2, 1983, 119-124.

7) Vinh, N.X., Busemann, A., Culp, R.D., Hypersonic and Planetary Entry Flight Mechanics, 1980, The University of Michigan Press, Ann Arbor.

8) Mittleman, D. and Jezewski, D., An Analytic Solution to the Classical Two-Body Problem with Drag, Celestrial Mechanics, Vol 28 (1982) 400-413.

9) Some of the material presented in section IV is from: Mittleman, D. & Jezewski, D., A Class of Problems in Celestial Mechanics Reducible to a Perturbed Harmonic Oscillation, NASA-JSC-IN-(82-FM-55), Sept. 1982

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

Final Report

APPROACHES TO SYNTHESIS OF SOME NOVEL POLYBENZIMIDAZOLE MONOMERS

Prepared by:	Dr. Charles G. Moseley
Academic Rank:	Associate Professor
Department and University:	Department of Chemistry The Ohio State University at Lima
Research Location:	Materials Laboratory, Non-Metallic Materials Division, Polymer Branch, Wright-Patterson Air Force Base, OH
USAF Research:	Dr. Robert Evers
Date:	September 26, 1983
Contract No.:	F49620-82-C-0035

APPROACHES TO SYNTHESIS OF SOME NOVEL POLYBENZIMIDAZOLE MONOMERS

by

Charles G. Moseley

ABSTRACT

Rigid rod aromatic-heterocyclic polymers such as polybenzimidazoles have high strength and high thermooxidative stability and are useable as structural materials for aerospace vehicles. Thus, a literature search was carried out and methods of synthesis proposed for two model compounds and three novel monomers which could possibly be polymerized to useful polybenzimidazoles. The monomers studied are N^1,N^5-diaryl-1,2,4,5-benzenetetramines, N^1,N^4-diphenyl-1,2,4,5-benzenetetramine, and 4-amino-3-anilinobenzoic acid. Experimental work was ther carried out in the laboratory in an attempt to synthesize one of the monomers (N^1,N^4-diphenyl-1,2,4,5-benzenetetramine) by adding ammonia to N,N′-diphenyl-p-benzoquinonediimine. The desired reaction did not occur in six attempts made under various conditions. Recommendations for further research in this area were then made.

Acknowledgement

The author wishes to thank the Air Force Systems Command, the Air Force Office of Scientific Research, and the Southeastern Center for Electrical Engineering Education for allowing him to spend ten very interesting and profitable weeks in the Materials Laboratory at the Air Force Wright Aeronautical Laboratory, Wright Patterson AFB, OH. He wishes to particularly thank the Polymer Branch for its hospitaity and for the use of its facilities.

In addition, the author wishes to thank Dr. Robert Evers for suggesting this area of research and for his valuable assistance in carrying out this project.

I. INTRODUCTION:

The Air Force Wright Aeronautical Laboratories and the Air Force Office of Scientific Research are engaged in a research and development program directed toward the synthesis and processing of ultra-high strength, thermooxidatively stable polymers for use as structural materials in aerospace vehicles. The objective is the attainment of mechanical properties for a structural material comparable with those currently obtained with fiber reinforced composites, but with significantly higher environmental resistance and without the use of a fiber reinforcement. The materials chosen for this effort are the rigid rod aromatic-heterocyclic polymers. Although the physical and chemical properties of these polymers show promise for the achievement of the program objectives, these aromatic-heterocyclic polymers present special processing and fabrication problems because of the all-para, rigid rod character of the molecules. A continuing Materials Laboratory in-house research program as well as related contractual programs in academic and industrial laboratories are addressing various aspects of the processing science and fabrication of these polymers. This research includes synthesis efforts directed toward the preparation of novel rigid rod aromatic-heterocyclic polymers which possess either improved processing characteristics or enhanced mechanical/physical properties.

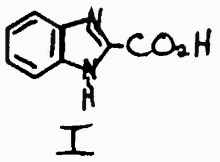
The present project involves the synthesis of several previously unreported novel monomers which are expected to undergo polycyclo-condensation reactions with selected aromatic diacids to yield rigid rod polybenzimidazole polymers which exhibit improved processing characteristics over known rigid-rod polymers.

II. OBJECTIVES:

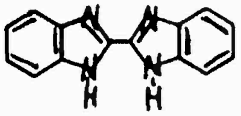
The primary objective of this project was to find efficient methods to synthesize several previously unknown monomers which could possibly be polymerized to novel polybenzimidazoles with high thermooxidative stabilities and application as aerospace structural materials. During the first part of the project a thorough search of the chemical literature was carried out to identify methods which have been used to prepare compounds similar to the desired compounds and which might serve as good routes to these compounds. During the second part of the project laboratory work was carried out in an attempt to synthesize one or more of the desired compounds.

III. LITERATURE SEARCH:

A. Synthesis of 1-H-benzimidazole-2-carboxylic acid (I) and 2,2 -bi-1H-benzimidazole (II).

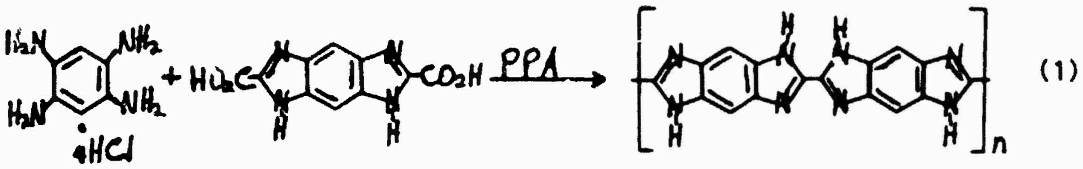


I



II

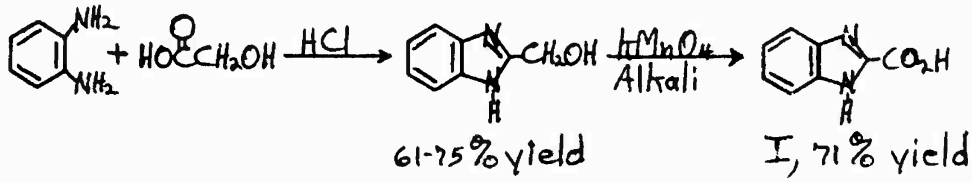
These two compounds were investigated because I is a model for the monomer and II is a model for the polymer produced in the reaction below:



(1)

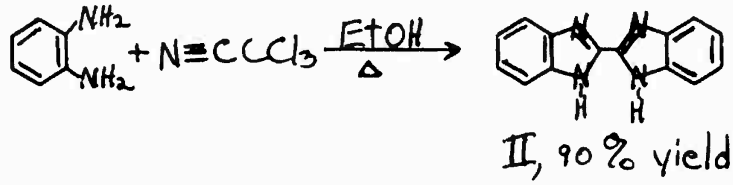
Dr. Jim Cain at this laboratory will investigate and report on this system.

The best synthesis found for I was [1,2,3]:



(2)

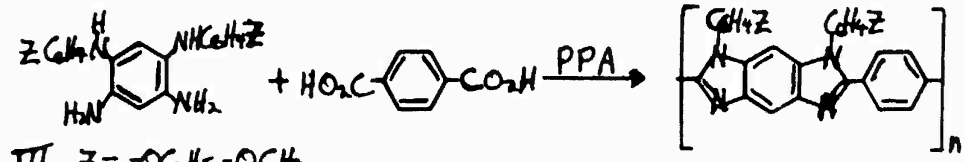
The best synthesis found for II was[4]:



(3)

B. Synthesis of N^1,N^5-diaryl-1,2,4,5-benzenetetramines (III).

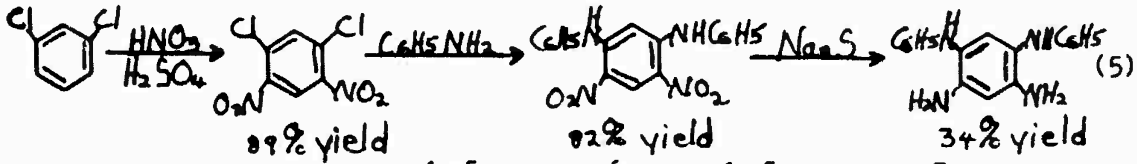
These compounds were investigated for use in the reaction to prepare the phenylated polybenzimidazoles shown below:



(4)

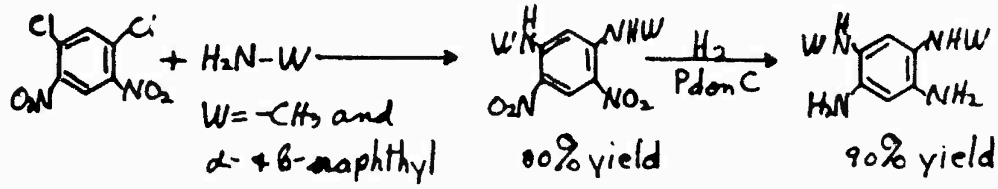
III, Z = $-OC_6H_5$, $-OCH_3$, $-N(CH_3)_2$, $-NO_2$

The parent compound (Z = H) has been prepared[5]:



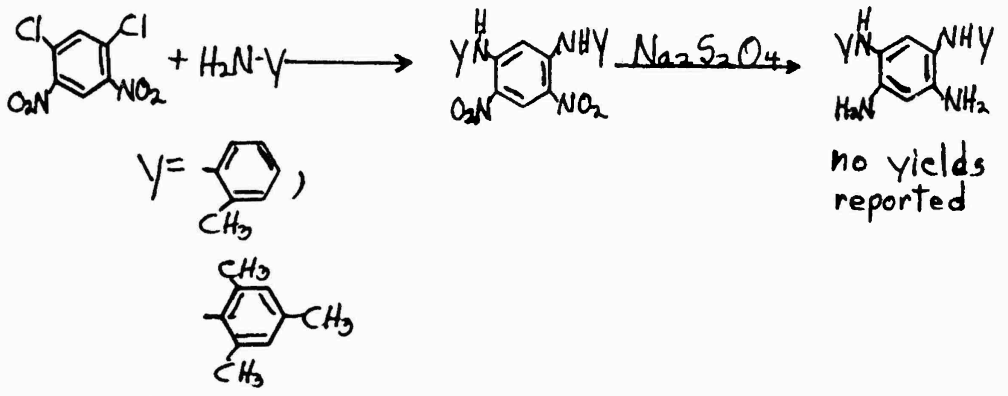
(5)

The corresponding N^1,N^5-dimethyl[6]- and N^1,N^5-dinaphthyl[7]- compounds have been prepared by very similar processes:



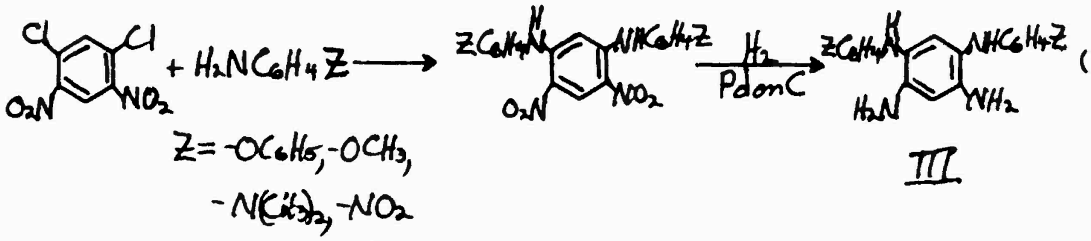
(6)

The closest reference found to the desired compounds was:[8]



(7)

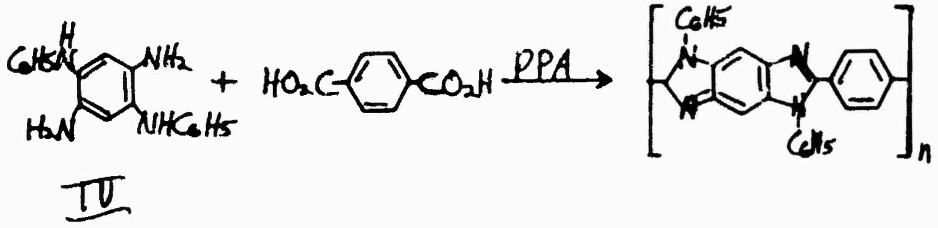
Based on these reactions it is likely that III could be prepared as shown below:



(8)

C. Synthesis of N^1,N^4-diphenyl-1,2,4,5-benzenetetramine (IV)

This compound was investigated for use in the reaction to prepare the phenylated polybenzimidazole shown below:

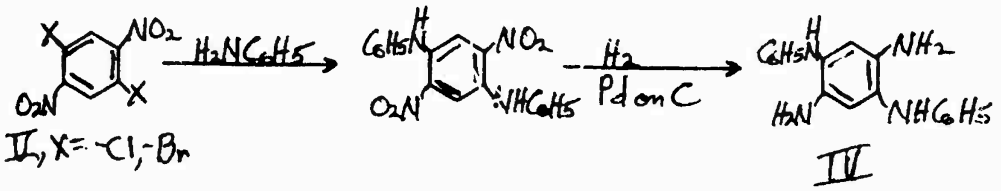


(9)

No references to IV having been prepared were found in the chemical literature. Methods which have been used to make related compounds and

which may possibly be used for synthesis of IV are:

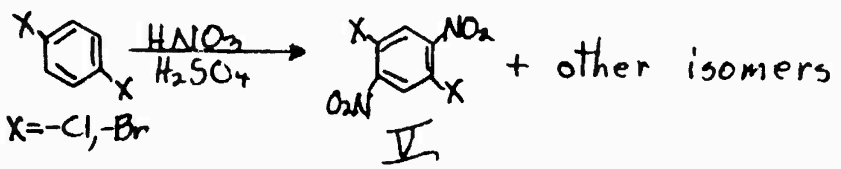
1) Reaction of 1,4-dihalo-2,5-dinitrobenzene (V) with aniline followed by reduction:



(10)

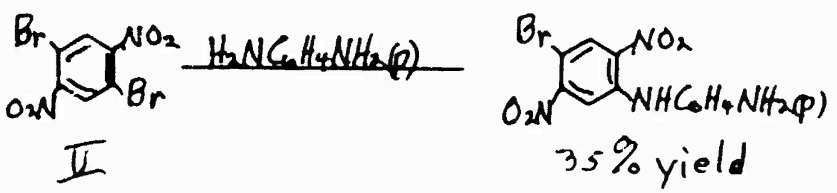
There are at least two potential problems with this approach:

a) No really good synthesis for V has been reported; it has been prepared by the laborious and inefficient separation (by crystallization) of the isomers formed by nitration of p-dihalobenzene[9,10,11]:



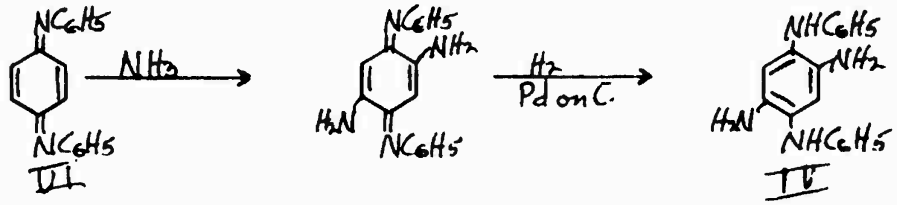
(11)

b) It may be difficult (or impossible) to replace both of the halogen atoms in V by reaction with aniline. In the closest example found to this reaction,[10] the second Br atom would not react:



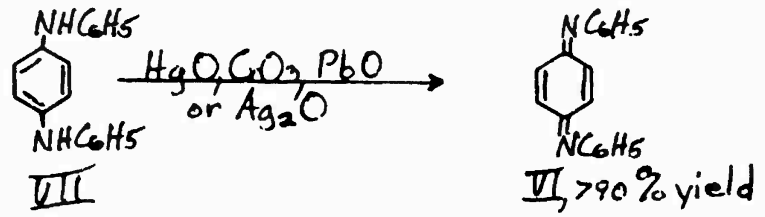
(12)

2) Reaction of N,N -diphenyl-p-benzoquinonediimine (VI) with NH_3 followed by reduction:



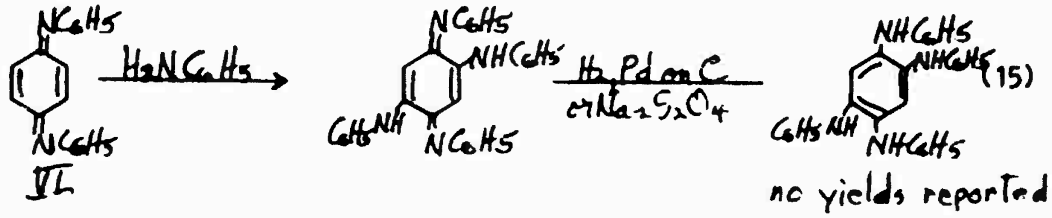
(13)

VI has been prepared by oxidation of N,N'-diphenyl-p-phenylenediimine (VII) which is commercially available[12,13,14]:



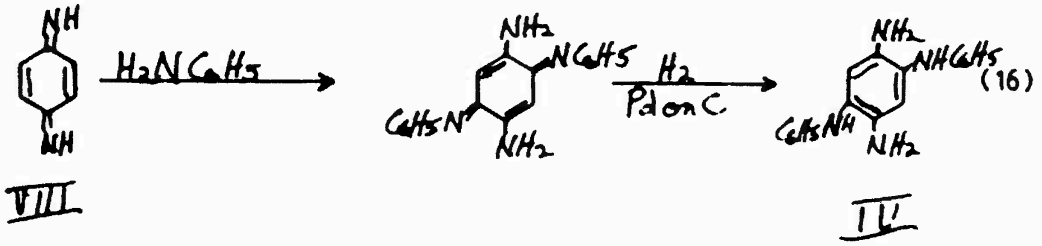
(14)

The closest example found to the desired reaction is[12,14]:



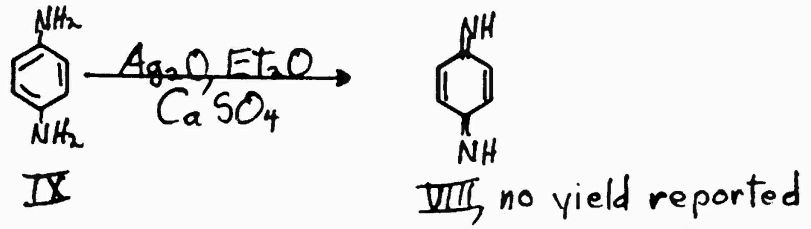
(15)

3) Reaction of p-benzoquinonediimine (VIII) with aniline followed by reduction:



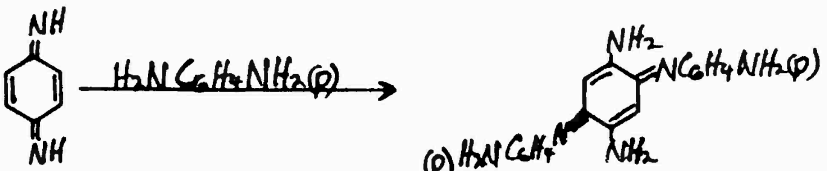
(16)

VIII has been prepared by oxidation of p-phenylenediamine (IX)[15,16] which is commercially available:



IX VIII, no yield reported (17)

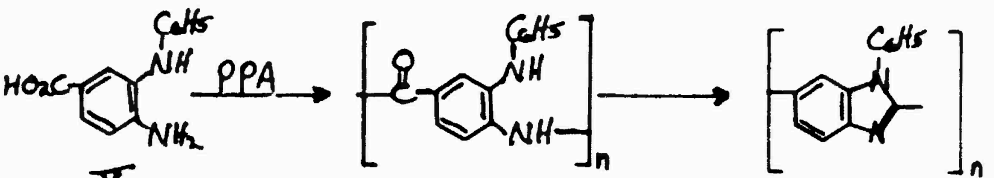
The closest example found to the desired reaction is[17]:



VIII no yield reported (18)

D. Synthesis of 4-amino-3-anilinobenzoic acid (X)

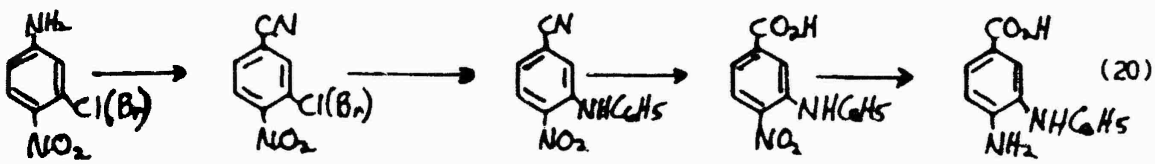
This compound was investigated for use in the reaction to prepare the phenylated polybenzimidazole shown below:



X (19)

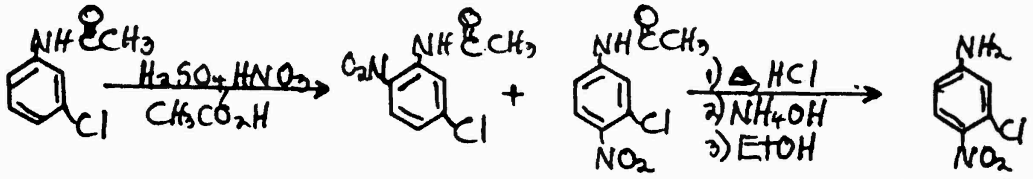
No references to X having been prepared were found in the chemical literature. Methods which have been used to make related compounds and which may possibly be used for synthesis of X are:

1) Conversion of 3-halo-4-nitroaniline (XI) to 3-halo-4-nitrobenzonitrile (XII) to X



XI XII X (20)

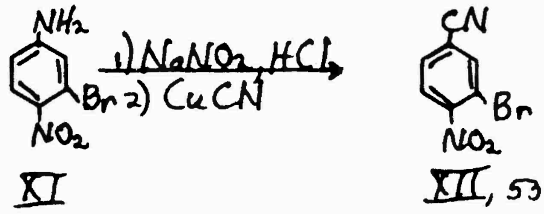
The best synthesis found for XI is[18]:



(21)

XI, 32% yield

XI is readily converted to XII by diazotization[19]:

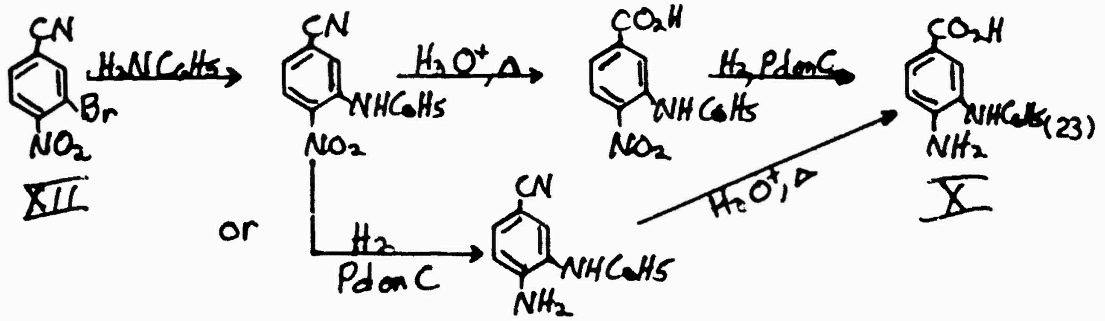


(22)

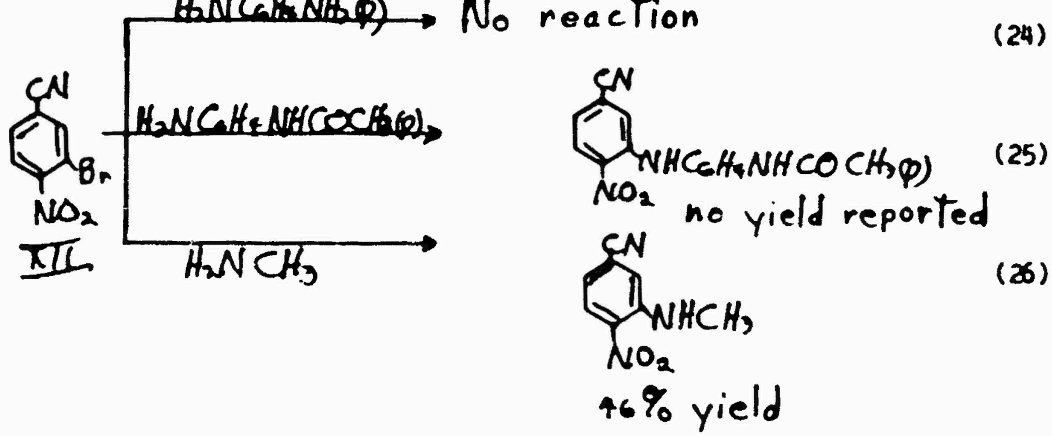
XI

XII, 53% yield

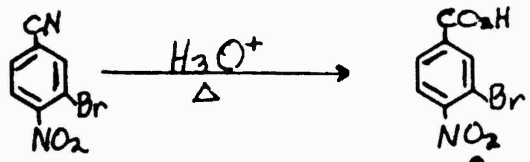
The route from XII to X with the greatest probability of success is:



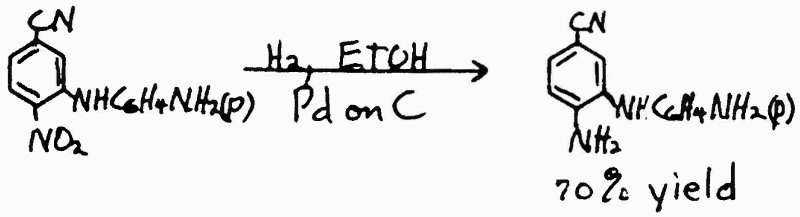
There is a potential problem with this route, however. No reports were found in the chemical literature of XII reacting with aniline, and XII is reported to react with some amines but not with others in the desired manner[19,20]:



Reported reactions on similar compounds suggest that the other steps in equations 21 should go well[19,20]:

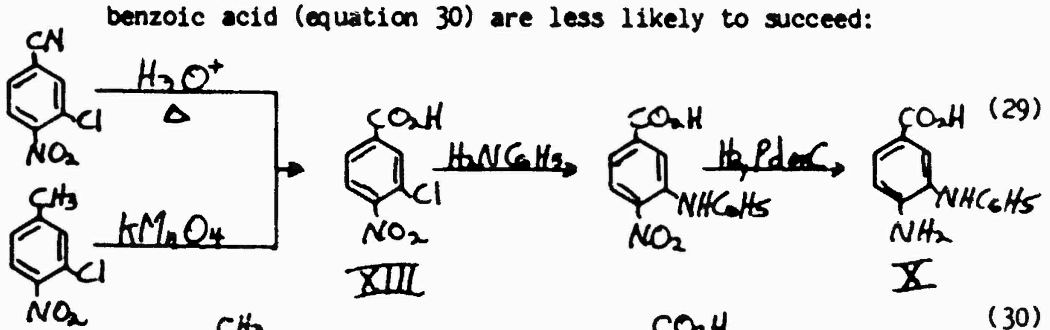


(27)

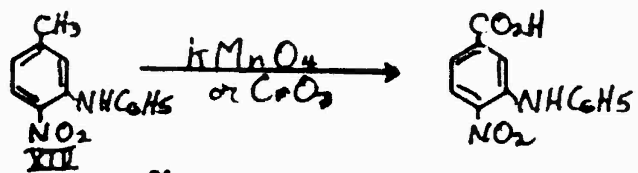


(28)

2) Other aproaches to X which involve either the reaction of 3-halo-4-nitrobenzoic acid (XIII) with aniline (equation 29) or the oxidation of 3-anilino-4-nitrotoluene (XIV) to the corresponding benzoic acid (equation 30) are less likely to succeed:

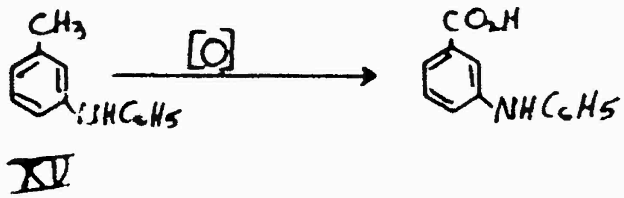


(29)



(30)

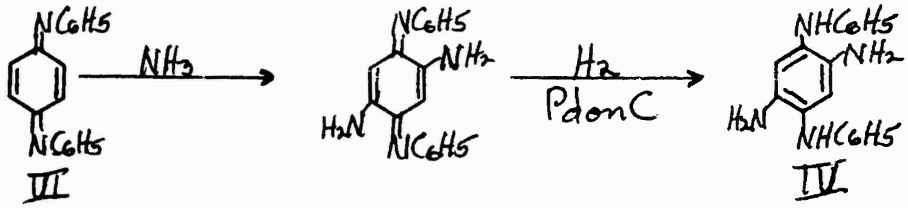
It has been reported[21] that XIII reacts with piperidine at less than one hundredth the rate of XII, and no reports of oxidation of an anilinotoluene (XV) to the corresponding anilinobenzoic acid could be found (equation 31):



(31)

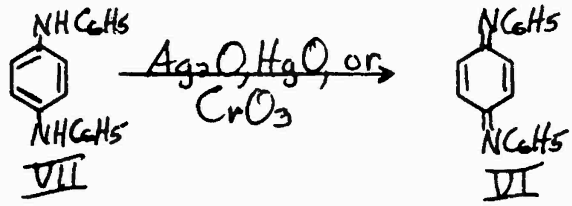
IV. EXPERIMENTAL WORK

All of the experimental work performed in the laboratory on this project was directed toward synthesis of N^1,N^4-diphenyl-1,2,4,5-benzenetetramine (IV) by addition of ammonia to N,N'-diphenyl-p-benzoquinonediimine (VI) followed by reduction:



(13)

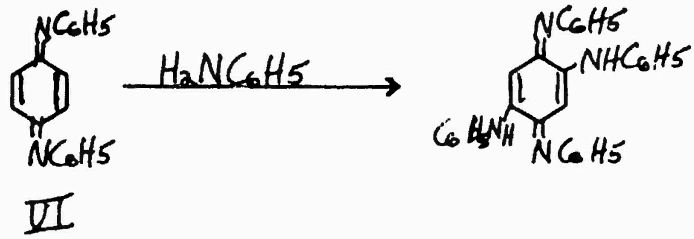
It was first necessary to prepare VI by oxidation of the commercially available N,N'-diphenyl-p-phenylenediamine VII):



(14)

Six different preparations were carried out using three different oxidizing agents (Ag_2O[22], HgO[12], and CrO_3[13]) as reported in the literature. Of the three only CrO_3 gave a good yield (69%) of pure product.

Although aniline added to VI to produce the expected product[12] (equation 32) in a test run, none of the six attempts made to add ammonia to VI under various conditions were successful. In each case either



(32)

unreacted starting material (VI) was recovered or a black very impure material was produced. The results are summarized in Table I.

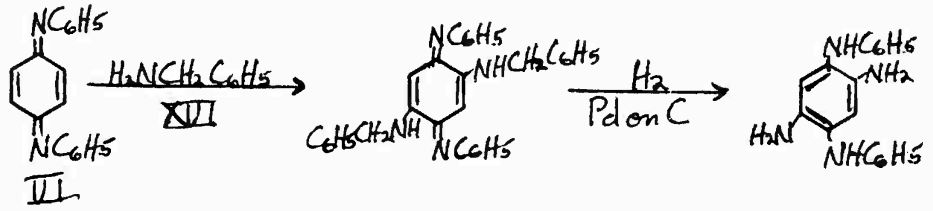
Table I

Results of Adding Ammonia[1] to N,N′-Diphenyl-p-benzoquinonediimine (VI)

Solvent	Catalyst	Temperature	Reaction Time	Results
Acetic Acid	None	0°C	15 min	Impure VI
Acetic Acid[2]	None	100°C	5 min	Black oil
THF	None	65°C	4 hr	VI
THF	p-TSA[3]	25°C	20 1/2 hr	VI
THF	p-TSA[3]	150°C[4]	6 hr	Black solid[5]
THF[6]	Acetic Acid	25°C	21 hr	VI

1. The ammonia was added as 28.1% in NH_4OH in 10-fold excess over the theoretically required amount.
2. The ammonia was added to VI before it was dissolved in acetic acid.
3. p-Toluenesulfonic acid monohydrate.
4. The reaction was carried out in a bomb under 200 psi pressure.
5. The unidentified product melted above 250°C and gave very broad peaks in the IR spectrum.
6. The ammonia was added to VI dissolved in THF before the acetic acid was added.

Finally, two attempts were made to add ammonia bonded to a protecting group (benzylamine, XVI) to VI (equation 33) with the hope that the protecting group could be removed to give the



(33)

desired product (IV). One reaction was carried out in acetic acid solvent and the other in chloroform in the presence of p-toluenesulfonic acid. Although time did not allow an extensive work up or analysis, neither reaction appeared to have produced the desired product.

V. RECOMMENDATIONS

Based upon the results obtained thus far, it does not appear likely that it will be possible to synthesize N^1,N^4-diphenyl-1,2,4,5-benzenetetramine (IV) by simple addition of ammonia to N,N^1-diphenyl-p-benzoquinonediimine (VI) (Equation 13).

It is recommended that the investigation of the reaction of VI with ammonia bonded to a protecting group be continued. Various reaction conditions and other protecting groups should be evaluated.

If this approach is not successful, it is recommended that the reaction of p-benzoquinonediimine (VIII) with aniline (equation 16) be investigated. Finally, if this reaction does not work, the reaction of 1,4-dihalo-2,5-dinitrobenzene (V) with aniline (equation 10) could be investigated.

REFERENCES

1. Lane, T. J. and J. M. Daly, J. Am. Chem. Soc., 81, 2953 (1959).
2. Copeland, R. A. and A. R. Day, J. Am. Chem. Soc., 65, 1072 (1943).
3. Phillips, M. A., J. Chem. Soc., 2393 (1928).
4. Holan, G., E. L. Samuel, B. C. Ennis, and R. W. Hinde, J. Chem. Soc. C, 20 (1967).
5. Vogel, H. and C. S. Marvel, J. Polymer Sci., A1, 1531 (1963).
6. Mitsuhaski, K. and C. S. Marvel, J. Polymer Sci., A3, 1681 (1965).
7. Kehayoglou, A. H. and G. P. Karayannides, J. Macromol. Sci. Chem., A18 (2), 237 (1982).
8. Jois, H. S., A. Kuppusami, and B. L. Manjunath, J. Indian Chem. Soc., 16, 43 (1939).
9. Hammond, G. and F. J. Modic, J. Am. Chem. Soc., 75, 1385 (1953).
10. Sunde, C. J., G. Johnson, and C. F. Kade, J. Org. Chem., 4, 548 (1939).
11. Jackson, C.L. and D. F. Calhane, J. Am. Chem. Soc., 28, 451 (1902).
12. Hughes, G. M. K. and B. C. Saunders, J. Chem. Soc., 3814 (1956).
13. J. Piccard, Chem. Ber., 46, 1835 (1913).
14. Taimr, L. and J. Pospisil, Angew. Makromol. Chem. 9253 (1980).
15. Willstatler, R. and A. Pfannenstiel, Chem. Ber., 37, 4605 (1904).
16. Corbett, J. F., J. Chem. Soc. B, 213 (1969).
17. Corbett, J. F., Proc. Joint Conf. Cosmet. Sci., 159 (1968).
18. Hodgson, H. H., and A. Kershaw, J. Chem. Soc., 2919 (1929).
19. Wibberly, D. G. and F. F. Stephens, J. Chem. Soc., 3336 (1950).
20. Holliman, F. G., B. A. Jeffery, and D. J. H., Brock, Tetrahedron, 19 (11), 1841 (1963).
21. Greizerstein, W., R. A. Bonelli, and J. A. Brieux, J. Am. Chem. Soc., 84, 1026 (1962).
22. Honzel, J. and M. Metalova, Tetrahedron, 25, 3641 (1969).

23. Velluz, G. A. and R. Heymes, Bull. Soc. Chem. Fr., 1012 (1954).

24. El Amin, B., G. M. Anantharamaiah, G. P. Royer, and G. E. Means, J. Org. Chem., 44, 3442 (1979).

25. Green, Theodora W., Protective Groups in Organic Synthesis, New York, John Wiley & Sons, 1981, p. 272.

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

BOUNDARY CORRECTIONS FOR LOW SPEED, SOLID WALL WIND TUNNELS

Prepared by:	Dr. Dale F. Moses
Academic Rank:	Associate Professor
Department:	Aerospace Engineering and Engineering Mechanics Department
University:	San Diego State University
Research Location:	Flight Dynamics Laboratory, Aeromechanics Division, Aerodynamics & Airframe Branch
USAF Research:	Noel H. Allen
Date:	August 18, 1983
Contract No:	F49620-82-C-0035

BOUNDARY CORRECTIONS FOR LOW SPEED, SOLID WALL WIND TUNNELS

by

Dale F. Moses

ABSTRACT

A plan of experiments is discussed for developing a subsonic, three dimensional wall correction code, based on a method due to Sears, into a form which can be efficiently used with solid wall wind tunnels. The inherent accuracy of the code is addressed and ways are identified for reducing calculational errors. The state-of-the-art of transonic, three-dimensional, slotted wall boundary corrections is reviewed and recommendations are made for further research in this area.

ACKNOWLEDGEMENTS

The author would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for providing him with the opportunity to spend a very worthwhile and interesting summer at the Air Force Flight Dynamics Laboratory. He would also like to acknowledge the Aerodynamics and Airframe Branch for its support and assistance.

Finally, he would like to thank Noel M. Allen and Capt Steve Reznick of the Experimental Group for their help in planning the work for this summer.

I. INTRODUCTION

The wind tunnel test section interferes with the flow about a model by introducing perturbations into this flow field. These perturbations can have the effect of 1) changing the streamline pattern (lift and induced drag); 2) altering the free-stream dynamic pressure (blockage); 3) reflecting shock waves in high speed flows, etc. Wall effects on measured model forces and moments can be significant. At low speeds, the effect on angle-of-attack near the stall can be 15% (ref. 1) and for high angle-of-attack ($\alpha > 60°$) effects from model blockage can be pronounced (ref. 8). In attempting to verify small improvements in the efficiency of transport aircraft at transonic speeds, it is found that the wall effects are of the same order of magnitude as the improvements being tested (ref. 9). Finally, the perturbations due to the presence of the walls forces the test engineer to use smaller models which has the disadvantages of limiting the wind tunnel in its Reynolds number range as well as requiring greater precision in the manufacturing of the model.

Traditionally there have been two approaches to correcting for these boundary interferences 1) direct calculation (for example references 2-4) and 2) adjustment of the model flow field (for example references 5-7). Wall interference has been handled exclusively by method (1) above until the advent of the adaptive wall concept in the early 1970's. This concept allows for the elimination of wall interference, within experimental error, but has the disadvantages of requiring considerable additional auxiliary flow control equipment or wall flexing attachments.

A method has been demonstrated (ref. 1) for calculating the wall effect on lift and induced drag for low speed, solid wall wind tunnels which eliminates the disadvantages of the adaptive wall approach but still provides an improved correction over classical methods. This improved method, which stems from the adaptive wall scheme of W.R. Sears (ref. 6), forms the basis for this present SFRP project.

The Sears method (see ref. 1) is based on the fact that if the values of certain chosen flow parameters on a cylindrical surface in the vicinity of the wind tunnel model are consistent with unconfined flow to infinity then the flow field within this surface can be considered to be unconfined. The method employs a computer program written by the author which takes measurements of flow parameters acquired during a wind tunnel test and iterates the values to those consistent with unconfined flow. The wall correction to lift and induced drag is then deduced from the difference in quantities related to the measured and unconfined values of the flow parameters. Additional research is needed, during the SFRP, to: 1) determine the best number and distribution of parameter measurements for practical, production testing and 2) improve the accuracy of the method.

II. OBJECTIVES

The main objective for the effort this summer was to refine the Sears method program to the point where it could be used, with minimum user interaction, as a production wall correction code for low-speed, solid wall wind tunnels. Another objective was also included to explain the difference in force measurements on a model at the same Reynolds number and transonic Mach number but in different test section inserts.

The objectives were, then:

1) Determine the best location and number of wall static pressure measurements which will give wall corrections superior to the present state-of-the-art for low speed, solid wall wind tunnels. This will include determining how many, if any, sidewall pressures need to be taken. Because test sections tend to be comparatively short it is necessary to determine what minimum length is necessary in order to make an accurate wall correction calculation.

2) Investigate the accuracy of the subject method of subsonic wall corrections. The sources of error in the computation of the boundary corrections (exclusive of instrument errors) will be identified and ways of reducing the overall error determined.

3) At some time in the past, tests were performed on a certain model in the Transonic Gasdynamics Facility wind tunnel at the same conditions but using two different test section inserts. The results did not agree. It was agreed to assess the state-of-the-art in three dimensional, slotted wall boundary corrections to see if wall effects could explain the disagreement in the test data.

III. Refinement of the Sears method

This effort required the performance of a series of tests on a generalized research model conducted in the TGF wind tunnel. These tests were to be as follows:

TEST 1

PURPOSE: Assess the effect of tunnel boundary layer on the wall static pressures at locations "A" (see attached sketch).

CONFIGURATION: Subsonic test section. Empty tunnel.

TEST CONDITIONS: Mach numbers .2, .4, .6, .7, .8 (for M > .6 match conditions with LRC-16T tests).

MEASUREMENTS: Perturbation pressure, $p_\infty - p_i$, at 60 locations on test section floor, 60 locations on ceiling and 12 locations on each sidewall. (p_i = wall static pressure at i^{th} locations). Record Reynolds number per foot.

TEST 2a

PURPOSE: Assess the effect of proximity of crescent apparatus on model perturbation pressures produced at wall locations. Determine the necessary upstream and downstream extent of pressure measurements in order to obtain an accurate wall correction. (Here accurate means an improvement over standard methods).

CONFIGURATION: Subsonic test section. Full span (11 in.) model at forward model position and at tunnel center-line.
Roll angle = 0° (wings horizontal).
Angle-of-attack α = 19° for Mach numbers M = .2, .4, .6, .7 and .8.

TEST CONDITIONS: Mach number schedule given above (Reynolds numbers and Mach numbers > .6 on all tests should match LRC-16T data).

MEASUREMENTS: Perturbation pressures at wall taps in Figure 1 and at 12 locations on each sidewall.

TEST 2b

PURPOSE: Obtain equivalent sidewall perturbation pressures for Test 2a. Assess the necessity for sidewall pressures and the minimum numbers and locations needed for accurate wall correction.

CONFIGURATION: Same as Test 2a except model is rolled 90° (wings vertical).

This test is otherwise the same as Test 2a.

TEST 3a

PURPOSE: Make flow parameter measurements necessary for wall correction computation for normal model test position. Assess the effect of proximity of crescent apparatus.

CONFIGURATION: Subsonic test section. Full span (11 in.) model at window center-line position and at tunnel center-line. Roll angle = 0° (wings horizontal). Angle-of-attack and Mach number as follows:

M	α, degrees
.2	-3, 0, +3, 5, 7, 9, 12, 14, 16, 19
.4	0, 5, 19
.6	0, 5, 19
.8	-3, 0, +3, 5, 7, 9, 12, 14, 16, 19

TEST CONDITIONS: Mach number schedule given above (Reynolds number and Mach numbers > .6, on all tests, should match LRC-16T data).

MEASUREMENTS: (1) perturbation pressures, $p_{\infty} - p_i$, at wall taps in figure 1; (2) Schlieren photos of model flow field extending to test section boundaries for each Mach number and angle-of-attack; (3) Perturbation pressures at sidewall tap location.

TEST 3b

PURPOSE: Obtain equivalent sidewall pressures for Test 4a.

CONFIGURATION: Roll angle = 90° (wings vertical).

This test is otherwise identical in all respects with Test 3a.

The first consideration to be addressed using the test results was the minimum number of wall static taps (both number across the test section and down the test section) that are needed in order to still compute a wall correction with any accuracy superior to the conventional method of images (e.g., ref. 4). In order to do this, perturbation pressures would be measured at 60 locations on the floor and ceiling in accordance with Figure 1. The number 60 was arrived at as the largest number that could be handled practically by the tunnel pressure system. Since the test section is square, the model can be rolled 90°, as explained in Test 2b above, and the pressures measured again to get equivalent sidewall values. By selectively deleting pressure values from certain locations the effect of number and distribution on the resultant wall correction calculation can be determined. In this case we can study the optimum longitudinal spacing of the rows of tops and the number (if any) and locations of required sidewall perturbation pressures.

Perturbation pressures were to be measured far upstream to determine their streamwise variation to see if this variation could be adequately represented by a theoretical calculation instead. In this way we could determine how far upstream we need to make the pressure measurements.

Forward and normal model positions were to be tested in order to determine the effect of the angle-of-attack drive mechanism ("crescent") on the wall pressures.

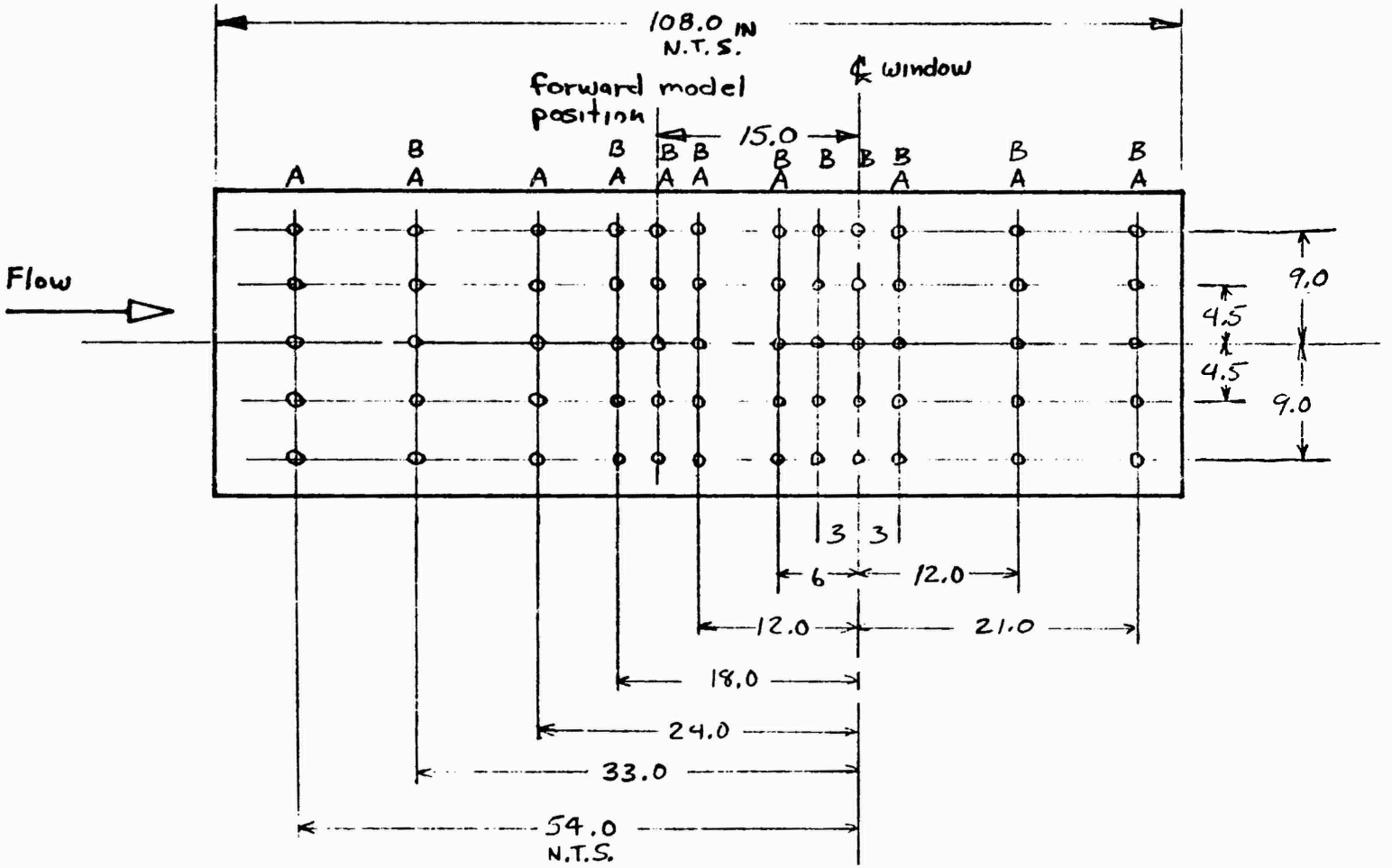


FIGURE 1: LOCATION OF FLOOR & CEILING STATIC TAPS

Tests were to be conducted at higher Mach numbers in order to determine the wall correction accuracy limits due to compressibility effects.

Unfortunately the planned tests were not conducted during the ten week summer period. This was due to delays occasioned by the preceding tunnel user, delays in obtaining and calibrating the necessary low range pressure tranducers, and finally the fact that, during calibration, these transducers were inadvertently over pressured and damaged beyond repair. This happened in the seventh week, leaving insufficient time for replacement.

In place of the transducers, a U-tube micromanometer was connected to sample selected rows of perturbation pressures. Because of the time necessary to make the measurements, the manometer could not replace the transducers for the complete series of tests. The only row of pressures which seemed to be good was the floor center-line longitudinal row. These results are shown in Figure 2. Because of questions about the manometer data the plot is only presented to show that the floor center-line pressures are following the expected trend and that the values had the order of magnitude that was expected.

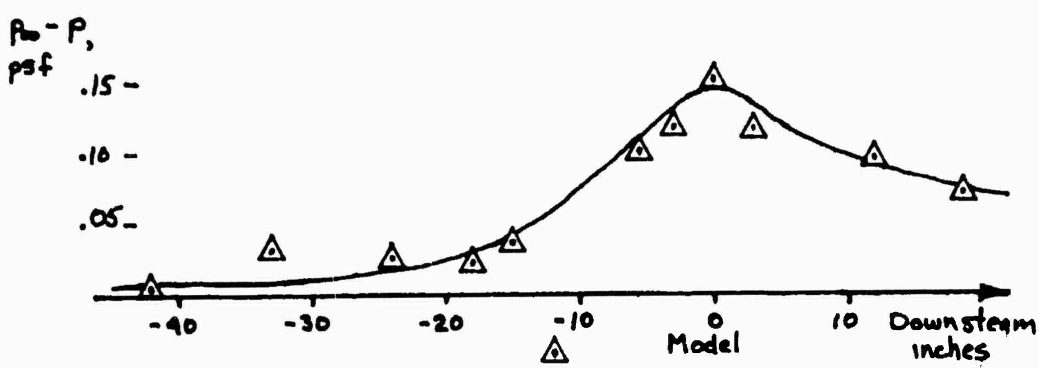


FIGURE 2: LONGITUDINAL CENTER-LINE PERTURBATION PRESSURES ON THE LIFT SIDE OF THE MODEL

IV. Accuracy of the Sears wall correction program

In order to investigate the inherent accuracy of the wall correction method we consider the computational sequence. The perturbation pressures measured on the tunnel boundaries are smoothed by finding the best fit least squares surface to the data. The surface is composed of a linear combination of products functions of the streamwise coordinate alone and functions of the cross-flow coordinate alone. This surface is then integrated from $-\infty$ to a given point i to obtain ϕ_i.

The perturbation pressures (and therefore ϕ_i) are obtained in a cylindrical surface which is coincident with the wind tunnel test section surface. This surface can be thought to divide the space into the near field and far field. To compute the far field flow, the surface is subdivided into i dipole elements, each with a boundary value point at its centroid. The ϕ_i above can then be used to fix the outer flow using the relation

$$A_{ij}\mu_j = \phi_i \qquad (1)$$

where A_{ij} is an influence coefficient matrix. This linear system can then be solved for the dipole strength μ_j. The consistent cross-flow perturbation velocity can now be computed from

$$B_{ij}\mu_j = v_i \qquad (2)$$

v_i is then compared with the measured value, v_m (which in this test is zero) to get an error signal, δv_i where

$$\delta v_i = v_i - v_m \qquad (3)$$

A relaxed value of this error signal is then used to fix the near-field flow using

$$B'_{ij}\delta\mu'_j = k\delta v_i \qquad (4)$$

When this linear system is solved for $\delta'\mu_j$ one can calculate an increment to ϕ_i,

$$A'_{ij}\delta\mu'_j = \delta\phi_i \qquad (5)$$

The first iteration is completed when $\delta\phi_i$ are used as an update to ϕ_i in the far-field. A new error signal (equation (3)) is computed, using the previous v_i in place of v_m. The iterations continue until the error signal becomes as small as desired. Reference 10 shows that if a linear system.

$$a_{ij}x_j = c_i \qquad (6)$$

is solved for particular values of a_{ij} abd c_i which are not the true values then the error in the solution, δx_j, satisfies a similar equation, namely

$$a_{ij}\,\delta x_j = \delta c_i - x_j\delta a_{ij}$$

For our case $\delta a_{ij} = 0$ so we get

$$a_{ij}\delta x_j = \delta c_i$$

which can be rewritten as

$$\delta x_j = a^{-1}{}_{ij}\delta c_i \qquad (7)$$

Exact errors δc_1 are not known so it is convenient to say that to a confidence of 95% $-\varepsilon \leq \delta c_i \leq \varepsilon$. If we apply equation (7) to equation (1) and use, for purposes of discussion, values from ref. 1 we find that if $\varepsilon \approx 10\%$ then $\delta x_j \approx 4\%$ from equation (7). This assures that there is negligible error in the determination of $A^{-1}{}_{ij}$. This is probably not quite true because the conditioning of the matrix A_{ij} is not determined by the IMSL subroutine used to invert it. It is planned to employ instead a subroutine from the LINPACK library which does provide a conditioning figure of merit.

An error bound of $\varepsilon \cong 10\%$, above, is rather conservative and can be estimated closer from its contributing components. The integral for ϕ_i is divided into two integrals; one integrates the perturbation pressures due to a singularly representation of the model from $-\infty$ to the upstream - most row of measurements, and the second integrates the least squares surface fit from this location to the point i. Errors in ϕ_i are therefore due to the adequacy of the singularity representation of the model and to the goodness of fit of the data smoothing operation. The goodness of fit is primarily dependent upon the choices for the functions of x and functions of y that go to make up the smoothed surface (ref. 1).

The next step in the iteration is to determine v_i from equation (2). Using the well known expression from error analysis, we can say that w_{v_i} is the error bound on v_i to, say, 95% confidence, so that

$$(v_i)_{true} = v_i \pm w_{v_i} \tag{8}$$

to 95% confidence (ref. 11). From equation (2) and ref. 11, we can write

$$W^2_{v_i} = (B_{i1}W_{\mu_1})^2 + (B_{i2}W_{\mu_2})^2 + \ldots + (B_{ij}W_{\mu_j})^2 \tag{9}$$

since, in our case, $W_{B_{ij}} = 0$. The uncertainly in v_i can thus be estimated from the matrix B_{ij} and the uncertainty, or error, in the μ's above.

The error in determining $\delta\mu_j$ from equation (4) can now be estimated in a way similar to that of μ_j and likewise the uncertainty in $\delta\phi_i$ in equation (5) follows the procedure used for v_i.

It is therefore seen that in the process of iterating to convergence the initial error in ϕ_i is compounded at each iteration. Since convergence to unconfined flow may take $\sim$ 45 iterations the final error could be significant. To eliminate this source of error in the wall correction calculation a one-step convergence procedure due to ref. 12 will be incorporated. This will also have the effect of halving the computational time.

V. CONVENTIONAL WALL CORRECTIONS FOR THREE-DIMENSIONAL, TRANSONIC WIND TUNNELS

The task of determining if wall corrections could explain the disagreement of data between the subsonic and transonic test sections at the same Mach number depended to some extent on tests that were to be conducted following those of section III above. In investigating the effect of the walls, compressibility should be taken into account. A search of the literature and other sources revealed that presently there is no suitable method for calculating boundary corrections for slotted, transonic wind tunnel. Methods exist, such as ref. 13, which calculates wall interference from flow field measurements but require the computation of the flowfield within the test section using either small disturbance or full potential equations. Such flowfield computation codes are available and could be adapted to this purpose. Specifically, the Sears method discussed above could very easily be used at transonic (compressible) speeds if the subsonic flow field computation subroutine were replaced with one that computed compressible flow fields. This would be the only significant change. Possible compressible codes are treated in references 14-16.

VI. RECOMMENDATIONS

The Sears method program for computing subsonic wall corrections has the potential of providing accurate corrections to lift and induced drag, and the author believes it can be extended to include blockage and pitching moment without much difficulty. The numerical errors will be substantially diminished by the incorporation of the one-step convergence as mentioned previously. It is recommended that wind tunnel tests be performed to develop the proogram into a production code which will improve the accuracy of the data taken at the TGF.

Because wall interference can be significant at transonic speeds and since the principles of the Sears method program applies to this regime as well, it is recommended that the program be adapted to transonic testing by the inclusion of a compressible flow field computation module. The resulting transonic program should then be demonstrated on actual wind tunnel tests using suitable flow parameter measuring instruments.

REFERENCES

1. Moses, Dale F., An Improved Method for Wind-Tunnel Wall-Corrections Deduced by Iterating from Measured Wall Static Pressure, Doctoral Dissertation, University of Arizona, 1980.

2. Lotze, I., "Correction of Downwash in Wind Tunnels of Circular and Elliptic Sections," NACA TM 801, July, 1936.

3. Theodorson, Theodore, "The Theory of Wind Tunnel Wall Interference," NACA Rept 410, 1931.

4. Silverstein, A., and J. A. White, "Wind-tunnel Interference with Particular Reference to Off-center Positions of the Wing and to the Downwash at the Tail," NACA Rept. No. 547, 1935.

5. Ferri, A., and P. Baronti, "A Method for Transonic Wind-tunnel Corrections," AIAA Journal, Vol. 11, No. 1, Jan., 1973.

6. Sears, W. R., R. J. Vidal, J. C. Erickson, Jr., and A. Ritter, "Interference-free Wind-tunnel Flows by Adaptive-wall Technology," AIAA Journal of Aircraft, Vol. 14, No. 11, Nov., 1977, p. 1042.

7. Harney, D. J., "Three-Dimensional Test Experience With a Transonic Adaptive Wall Wind Tunnel," AFWAL-TR-83-3028, Flight Dynamics Lab, Wright-Patterson AFB, March 1983.

8. Unpublished Wind Tunnel Study by General Dynamics, Convair Divisions, San Diego, CA.

9. Personal Communication, Aerodynamics Laboratory, Boeing Commerical Airplane Division, Seattle, WA.

10. Hildebrand, F. B., Introduction to Numerical Analysis, 2nd ED, McGraw-Hill Book Company, 1974.

11. Kline and McClintock, "Uncertainties in Single-Sample Experiments," Mechanical Engineering, 1953.

12. Dowell, E. H., "Control Laws for Adaptive Wind Tunnels," AIAA Journal, Vol 19, N 11, Nov.,1981, pp 1486-1488.

13. Rizk, M. H., Hafex, M., Murman, E. M., and Lovell D., "Transonic Wind Tunnel Wall Interference Corrections for Three-Dimensional Models," AIAA Paper 82-0588, AIAA 12th Aerodynamic Testing Conference Proceedings, March 22-24, 1982, pp 120-130.

14. Boppe, C. W., "Transonic Flow Field Analysis for Wing-Fuselage Configurations," NASA CR-3243, May, 1980.

15. Mercer, J. E., and Murman, E. M., "Application of Transonic Potential Calculations to Aircraft and Wind Tunnel Configurations," AGARD CP-285, May, 1980.

16. Hinson, B. L., and Burdges, K. P., "An Evaluation of Three-Dimensional Transonic Codes Using New Correlation-Taylored Test Data," AIAA Paper 80-0003, Jan., 1980.

16. Hinson, B. L., and Burdges, K. P., "An Evaluation of Three-Dimensional Transonic Codes Using New Correlation-Taylored Test Data," AIAA Paper 80-003, Jan., 1980

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

COMBINED TIME SPACE FILTERING FOR HF ANTENNA ARRAY SYSTEMS

Prepared by:	Randolph L. Moses
Academic Rank:	Instructor
Department and University:	Department of Electrical Engineering Virginia Polytechnic Institute and State University
Research Location:	Rome Air Development Center Communications Division Communications Transmission Branch
USAF Rearch:	Richard N. Smith
Date:	September 1, 1983
Contract No.:	F49620-82-C-0035

COMBINED TIME SPACE FILTERING

FOR HF ANTENNA ARRAY SYSTEMS

by

Randolph L. Moses

ABSTRACT

This report investigates methods for combined time-space adaptive filtering of HF receiver antenna array data. These two-dimensional filters attenuate interference and noise components in the received data to facilitate effective channel equalization processing. The filtering algorithms are derived by first modeling the received data array as a two-dimensional ARMA process. Algorithms for estimating the AR coefficients in this model are then developed. These AR coefficients are used as the weights of the two-dimensional filter. In addition, computationally efficient recursive algorithms that combine the AR coefficient estimation and data filtering operations are derived. Suggestions for performance evaluation of these filters and recommendations for future research in this area are also presented.

ACKNOWLEDGEMENTS

I would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for providing me with the opportunity to spend a very enriching summer at Rome Air Development Center. I also wish to thank those at Rome Air Development Center, and in particular the Communications Division and the Communications Transmission Branch for their hospitality and support throughout this research effort.

I would also like to acknowledge the many useful discussions with John Graniero, Jim Periard and Paul VanEtten and the computer support from Russell Brown.

Finally, I would like to thank Richard Smith for his help in formulating the research problem, and his guidance and support throughout the summer.

I. INTRODUCTION

In the past several years there has been a marked increase in communication systems that operate in the HF range (3-30MHz). This increase has been spawned by advances in the understanding of the HF communication channel and by the development of effective channel equalization processors. Data rates of as high as 9600 bits/sec have been achieved in systems incorporating these channel equalizers.[1] Adaptive receiver antenna arrays are also being used to provide improved performance in spatially disturbed environments.

By combining spread spectrum techniques with adaptive receiver antenna arrays, HF communication systems are potentially able to provide reliable communication links even in electromagnetically unfriendly environments. However, the performance of channel equalizers necessary in such systems is very sensitive to interference signals. Effective operation in this type of environment is contingent upon effective signal prefilters that attenuate these interference signals before channel equalization. Because knowledge of frequency content and source location of many interference signals are not known apriori and may vary with time, these prefilters must be adaptive.

The most common adaptive filtering technique applied to this problem consists of cascading spatial nulling algorithms with temporal spectral whitening filters.[2] However, it appears possible to realize a further improvement in interference rejection by combining the spatial and temporal filtering tasks into a single multidimensial process. Frost [3] developed one such technique that assumes the desired signal direction is known (an often impractical assumption). Monzingo, et.al.[4]

have investigated maximum likelihood algorithms. Also, related algorithms have been developed for the processing of geophysical data [5,6].

This report investigates combined time-space signal processing algorithms for prefiltering of array element data. These algorithms do not assume known signal direction. The procedure taken is to model the received data as a two-dimensional autoregressive moving-average (ARMA) process, then to propose particular algorithms for estimating the coefficients in the ARMA model. The AR coefficients are then used as the weights in a two-dimensional data prefilter. We also present two computationally efficient recursive algorithms that are more amenable to implementation. Finally, we present recommendations for further research on this problem.

II. OBJECTIVES

The primary objective of this research effort was to investigate models and algorithms for combined time-space prefiltering of receiver antenna array data. The objective of the prefiltering is to increase the signal-to-interference ratio (SIR) and signal-to-noise ratio (SNR) at the filter output. The primary application of this research effort is in HF antenna array receivers; however, such filters are needed in other receiver systems as well as for clutter suppression of radar return data. In this report we consider the very practical special case that the antennas are equally spaced and collinear.

Our specific objectives were:

1) To investigate combined time-space data and filter

models that facilitate effective and efficient filtering techniques.

2) To develop both block processing and recursive filtering algorithms that are based on these models.

3) To outline procedures for performance evaluation of these algorithms and to identify important areas for future research.

III. DATA AND FILTER MODELS

In order to propose models for the data and prefilter, it is necessary to first understand the basic receiver configuration. The basic antenna array receiver is shown in Figure 1. After amplification, the M received analog signals are shifted from the carrier frequency to baseband and lowpass filtered, then sampled. The result after N samples is a complex-valued discrete time data array $\{x(m,n)\}$ for $1 \leq m \leq M$ and $1 \leq n \leq N$. This data is then passed through the two-dimensional prefilter to attenuate the noise and interference. The M filter outputs are combined as a weighted sum, and this sum is sent to the channel equalizer and receiver decoding processors.

The received signal at each antenna consists of a desired signal $s(t)$ incident on the array at some angle θ_s, along with any of several interference signals $i_k(t)$, each incident at some angle θ_k (see Figure 1). The incident angles are assumed unknown. In addition there is additive background noise $n(t)$. Because of possible signal and interference source movement, and because of HF channel disturbances, the amplitudes and incident angles of these signal components may vary with time at a rate that is assumed to

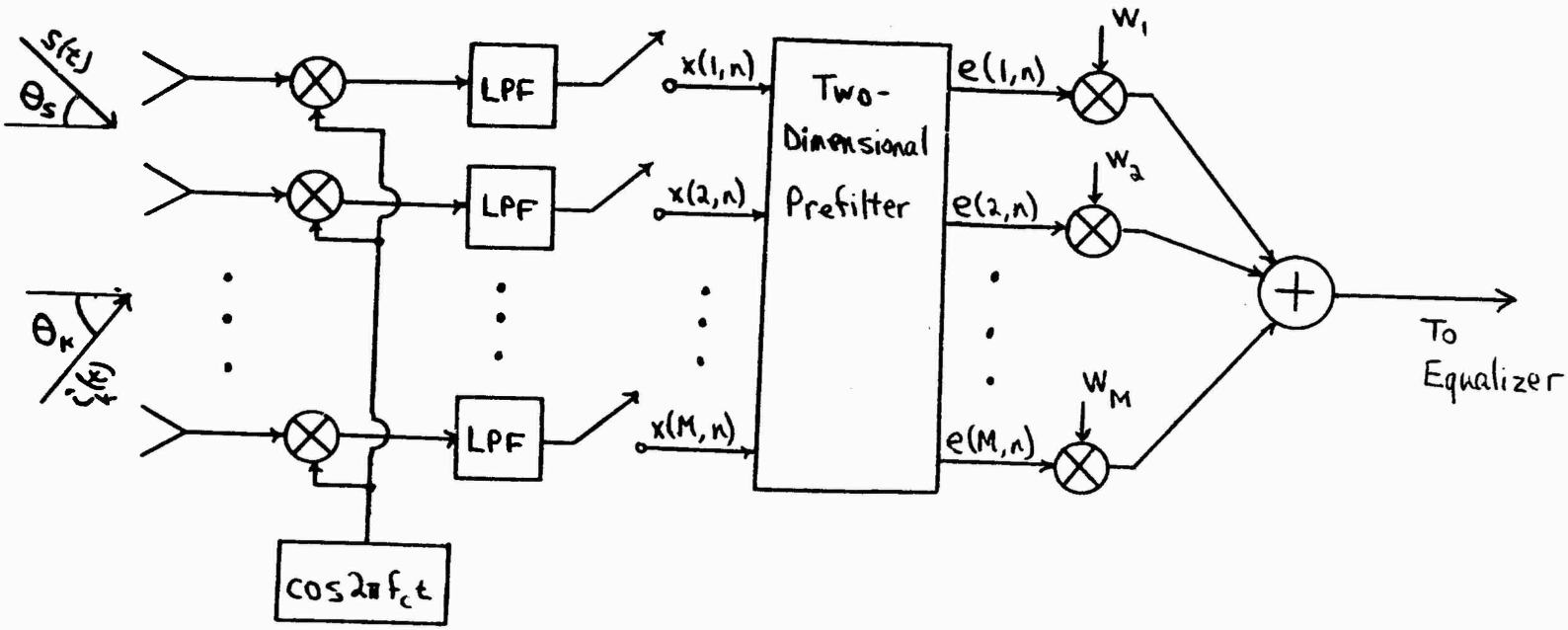


Figure 1. The Basic HF Antenna Array Receiver

be slow compared to the data sampling rate. These signal components are all present in the discrete data array measurements $\{x(m,n)\}$; in addition there is a noise component $w(m,n)$ due to receiver noise, round-off, etc. Thus, we may write

$$x(m,n) = s(m,n) + \sum_{k=1}^{K} i_k(m,n) + w(m,n) + n(m,n) \qquad (1)$$

Due to the dispersive nature of the HF channel, signals arriving from point sources are spread slightly over a small spatial frequency range. We shall assume that this spreading function can be modeled by an ARMA function. An ARMA model is robust since as the spread with approaches zero, the spreading function becomes a limiting ARMA (1,1) function.

The desired signal $s(m,n)$ is a spread spectrum signal, and is well modeled as white noise in the temporal frequency domain. [7] If we assume an ARMA spatial spreading function, then a two dimensional (2-D) ARMA model results. Similarly, each interference signal is often modeled as a complex exponential or as an ARMA process in the temporal domain [7]. Since a complex exponential can be modeled as the limit of an ARMA (1,1) process, by again assuming an ARMA spatial spreading function a 2-D ARMA model is obtained for each interference signal as well. The noise processes can also be modeled as 2-D ARMA processes [6]. Thus we have modeled $x(m,n)$ as the sum of independent ARMA processes, which is itself an ARMA(p_1, q_1, p_2, q_2) process, where p_1 and q_1 are the AR and MA orders in the spatial frequency direction, and p_2 and q_2 are the AR and MA orders in the temporal frequency direction. The corresponding power spectral density function is given by

$$S_x(z_1, z_2) = \left| \frac{B(z_1, z_2)}{A(z_1, z_2)} \right|^2 \qquad (2)$$

where $B(z_1, z_2)$ is a polynomial of degree q_1 in z_1 and q_2 in z_2, and $A(z_1, z_2)$ is a polynomial of degree p_1 in z_1 and p_2 in z_2.

It is apparent that interference signals and some noise appear as areas of high power density in the 2-D frequency plane. An appropriate interference rejection filter, then, is one that spectrally whitens the data, as this flattens such peaks. If the signal power is below the background noise power level (which is often the case in practice), a whitening filter has little effect on the desired signal, so SIR and SNR gains are realized at the filter output. If the signal power is substantially above the noise power level, a whitening filter attempts to attenuate the signal along with the interference. This effect is unavoidable when the signal incident angle is unknown, as the whitening filter cannot distinguish between signal and interference. However, even in this case SIR and SNR gains may be realized at the filter output.

The optimum whitening filter for $\{x(m,n)\}$ has transfer function from each input to each output of the form

$$H(e^{j\omega_1}, e^{j\omega_2}) = \frac{A(e^{j\omega_1}, e^{j\omega_2})}{B(e^{j\omega_1}, e^{j\omega_2})} \qquad (3)$$

where A and B are given in equation (2). Since H contains a denominator term, it is an infinite impulse response (IIR) filter. IIR filters are undesirable because they pose a stability problem, (especially when the filters are time varying, as is the case here), and because the B coefficients are often difficult to estimate accurately without a large amount of data (especially when there are sharp nulls in the spectrum). Because of these problems and because spectral nulls are not formed by interference signals, there seems to be no practical reason to incorporate the coefficients

into the filter. A practical alternative filter is

$$H(e^{j\omega_1}, e^{j\omega_2}) = A(e^{j\omega_1}, e^{j\omega_2}) \tag{4}$$

where A is obtained by either:

1. Modeling the data with the ARMA model in equation (2) and using only the estimated AR coefficients in (4).
2. Modeling the data with an AR model and using the estimated AR coefficients in (4).

The main advantage of the first alternative is that sharp spectral peaks present in the data (say, from narrowband jammers) are more effectively nulled; the main advantage of the second alternative is that some of the MA filtering is performance because the entire spectrum is approximated by the AR model. Thus, the two filter models represent a tradeoff between effectiveness in eliminating narrowband interference and effectiveness of spectrally whitening the data.

IV. TWO-DIMENSIONAL ARMA ALGORITHMS

In this section we propose one class of AR coefficient estimation algorithms derived by considering the observed data as a sample from a two-dimensional stationary ARMA process. The derivation makes use of the fact that the antenna array elements are collinear and equally spaced, which results in in efficient use of the data in generating autocorrelation estimates.

Consider the 2-D semi-casual ARMA recursion

$$\sum_{i=-p_1}^{p_2} \sum_{j=0}^{p_3} a_{ij}\, x(m-i, n-j) = \sum_{i=-q_1}^{q_2} \sum_{j=0}^{q_3} b_{ij}\, v(m-i, n-j), \quad a_{00} = 1 \tag{5}$$

If we follow the earlier suggestion of using only AR coefficients in the filtering operation, the corresponding prefilter outputs

are given by

$$e(m,n) = \sum_{i=-p_1}^{p_2} \sum_{j=0}^{p_3} a_{ij}\, x(m-i, n-j) \quad , \; 1 \le m \le M \tag{6}$$

It can be seen from (6) that p_1 and p_2 must be chosen so that $1 \le m-i \le M$ for $-p_1 \le i \le p_2$. In general this may require a different choice of p_1 and p_2 for each of the M filter outputs. Thus, a different set of AR coefficients is in general necessary for each of the M outputs.

An effective AR coefficient estimation procedure can be derived by appealing to the well-known Yule Walker equations, which are found by multiplying both sides of equation (5) by $x^*(m-k, n-l)$ and then taking the expected value to give

$$r_x(k,l) + \sum_{\substack{i=-p_1 \\ (i,j)\neq(0,0)}}^{p_2} \sum_{j=0}^{p_3} a_{ij}\, r_x(k-i, l-j) = 0 \quad \text{for } l > q_3 \tag{7}$$

By replacing exact autocorrelations in (7) with ones estimated from the given data, we can form a matrix of approximate Yule Walker equations whose solution yields AR coefficient estimates. To this end, define the matrix equation

$$\hat{\underline{r}} + \hat{R}\,\hat{\underline{a}} = \underline{\varepsilon} \tag{8}$$

where each row is an estimate of equation (7) for a particular choice of the pair (k,l), $\hat{R}$ and $\hat{\underline{r}}$ are a matrix and a vector of autocorrelation estimates, and $\hat{\underline{a}}$ is a vector of (unknown) AR coefficient estimates. Because autocorrelation estimates are used in equation (8), the right hand side is not in general equal to zero.

The solution to equation (8) represents the prefilter coefficient determination algorithum. If the number of rows in (8) is equal to the number of AR coefficients then $\underline{\varepsilon}$ can be made

equal to zero and

$$\hat{\underline{a}} = -\hat{R}^{-1}\hat{\underline{r}} \tag{9}$$

However, choosing a greater than minimal number of rows in (6) generally results in improved AR coefficient estimates [5]. In this case the AR coefficient estimate is the least squared error solution

$$\hat{\underline{a}} = -\left[\hat{R}^{\dagger}\hat{R}\right]^{-1}\hat{R}^{\dagger}\hat{\underline{r}} \tag{10}$$

The dagger symbol ($\dagger$) denotes complex-conjugate transposition. It can be seen from equations (9) and (10) that generally improved AR coefficient estimates are obtained at a cost of increase computational burden.

There are several standard autocorrelation estimators that may be used in generating $\hat{R}$ and $\hat{r}$. One choice is the unbiased estimator

$$\hat{r}_x(k,l) = \frac{1}{(M-|k|)(N-|l|)} \sum_i \sum_j x(k+i, l+j)\, x^*(i,j) \tag{11}$$

where

$$i = \begin{cases} 1, 2, \ldots, M-k & , k \geq 0 \\ k+1, k+2, \ldots, M & , k < 0 \end{cases}$$

$$j = \begin{cases} 1, 2, \ldots, N-l & , l \geq 0 \\ l+1, l+2, \ldots, N & , l < 0 \end{cases}$$

It is important to note that $\hat{r}_x(k,l)$ is formed not only by summing over time lags but also by summing over spatial lags. This is a consequence of assuming collinear, equally space antennas. We shall see that this procedure uses the data more efficiently then do procedures developed in the next section.

The values of k and l in each row of equation (6) should generally be chosen small to maximize the number of data lags being summed to form autocorrelation estimates (see (11)). Equally important, however, is that by judicious selection of the

(k, l) pairs, and by properly ordering the rows of (8), Toeplitz structures in the matrix $\hat{R}$ can be obtained. This can significantly reduce the computational burden of inverting $\hat{R}$ or $\hat{R}^{\dagger}\hat{R}$ [11,12].

Up to now we have assumed that the data array $\{x(m,n)\}$ is stationary. In fact, channel fading and movement of the transmitter and interference source locations produce temporal nonstationarities in the data. One way of adapting to the nonstationarities is to use only the last $\tilde{N}$ data samples from each antenna when calculating the AR coefficient estimates. This method actually treats the data as being stationary over the last $\tilde{N}$ samples; thus, the selection of $\tilde{N}$ is predicated on how rapidly the data statistics vary. A second method is to incorporate an exponential forgetting factor λ in the autocorrelation estimates, wherein data lags "k" samples in the past are weighted by λ^k, for some $0<\lambda\leq 1$ [13]. In this way, autocorrelation estimates (and therefore AR coefficient estimates) are primarily generated from recent data. The forgetting factor method is attractive because it is easily implemented in recursive algorithms.

It can be seen that the AR coefficient estimators (9) and (10) presented in this section not only provide flexibility in model order selection, but also make efficient use of the given data in generating autocorrelation estimates. However, in order to obtain the "optimal" AR coefficients at each time n, one must perform a matrix inversion every time a new data point arrives (i.e., at each sample interval). Even though efficient matrix inversion procedures are available due to the special structure of $\hat{R}$, the

computational requirements of this procedure can be very high. One way to decrease this computational burden is to update the AR coefficients at every N sample intervals, and use these coefficients in the prefilter until the next update, however, a more attractive alternative is to develop recursive versions of the algorithms. Such recursive versions determine optimal AR coefficients at every sample interval by updating the previous estimates rather than by solving for them from scratch every time. Unfortunately, no such recursive algorithm is available for the general algorithm. In the next section, though, we present a subclass of this algorithm for which computationally efficient recursive implementations are available.

V. MULTIDIMENSIONAL RECURSIVE ALGORITHMS

In this section we consider a class of AR coefficient estimation algorithms that facilitate efficient recursive implementations. These methods treat the data array as an Mxl vector ARMA process. Both block processing and recursive algorithms are presented for the "prewindow method". They are straightforward extensions to the multidimensional case of algorithms presented in [9], [10], [14]. The multidimensional case for real data is also found in [13]. The reader is referred to these papers for details of the derivations presented below.

In order to develop these multidimensional instrumental variable algorithms, consider the data array $\{x(m,n)\}$ as a set of N Mxl vectors $\{\underline{x}(n)\}$ where

$$\underline{x}(n) = [x(1,n), x(2,n), \ldots, x(M,n)]', \quad 1 \le n \le N \tag{12}$$

The corresponding ARMA recursion is given by

$$\underline{x}(n) + \sum_{i=1}^{p} A_i \underline{x}(n-i) = \sum_{j=0}^{q} B_j \underline{v}(n-j) \tag{13}$$

where $\{\underline{v}(n)\}$ is an Mxl vector white noise process.

Each row of (13) is seen to be a special case of (5), where

$$p_1+p_2=M-1 \quad , \quad q_1=q_2=0 \quad , \quad p_3=p \quad , \quad q_3=q \quad , \quad a_{0j}=0 \text{ for } j\neq 0.$$

Implicitly, then, (13) assumes an AR(M-1) model in the spatial direction.

The block processing algorithms are developed by first writing the prefiltering operation for every time up to N as

$$\begin{bmatrix} \underline{x}'(q+1) \\ \underline{x}'(q+2) \\ \vdots \\ \underline{x}'(N) \end{bmatrix} + \begin{bmatrix} 0 & & \underline{0} & \\ x'(q+1) & & & \\ x'(q+2) & x'(q+1) & \cdots & x'(q+1) \\ \vdots & \ddots & \ddots & \vdots \\ x'(N-1) & x'(N-2) & \cdots & x'(N-p) \end{bmatrix} \begin{bmatrix} A_1' \\ A_2' \\ \vdots \\ A_p' \end{bmatrix} = \begin{bmatrix} \underline{e}(q+1) \\ \underline{e}'(q+2) \\ \vdots \\ \underline{e}'(N) \end{bmatrix} \tag{14}$$

$$\underline{X}_N + X_{N,p} \cdot A_{N,p} = \underline{e}_N$$

Here, $\underline{e}(n)$ is the prefilter output at time N, and $A_{N,p}$ is the (as yet unknown) matrix of AR coefficient estimates.

Also, define

$$Y_{N,t} \triangleq \begin{bmatrix} 0 & 0 & \underline{0} \\ x'(1) & 0 & \\ \vdots & x'(1) & \cdots \quad x'(1) \\ x'(N-q-1) & x'(N-q-2) & \cdots \quad x'(N-q-t) \end{bmatrix} \tag{15}$$

where $t \geq p$. Premultiplying (14) by $Y_{N,t}^{\dagger}$ yields

$$Y_{N,t}^{\dagger} \underline{x}_N + Y_{N,t}^{\dagger} X_{N,p} A_{N,p} = Y_{N,t}^{\dagger} \underline{e}_N \tag{16}$$

which is an approximation to the Yule Walker equations, and similar to equation (8). Since the right hand side of (16) has expected value of zero, an appropriate choice for the AR coefficient matrix is

$$A_{N,p} = -\left[Y_{N,p}^{\dagger} X_{N,p}\right]^{-1} Y_{N,p}^{\dagger} \underline{x}_N \tag{17}$$

when t=p and the least squared error solution

$$A_{N,p} = \left\{\left[Y_{N,t}^{\dagger} X_{m,p}\right]^{\dagger}\left[Y_{N,t}^{\dagger} X_{n,p}\right]\right\}^{-1}\left\{\left[Y_{N,t}^{\dagger} X_{N,p}\right]^{\dagger}\left[Y_{N,t}^{\dagger} \underline{x}_{N}\right]\right\} \qquad (18)$$

when t>p. Equations (17) and (18) are block processing AK coefficient estimation algorithms, and are particular versions of (9) and (10). The desired data prefiltering operation is performed after solving for $A_{N,p}$ by employing the last row of equation (14) to find $\underline{e}(N)$.

As mentioned earlier, it is desired to find computationally efficient recursive algorithms that filter the data "optimally" at every time interval. These take the form of "fast recursive" algorithms. They are actually adpative filtering algorithms that calculate the prefilter output $\underline{e}(N)$ directly. They are fast in the sense that they require only $O(M^3 p)$ computations to implicitly determine $A_{N,p}$ from $A_{N-1,p}$ and to calculate the "optimal" filter output $\underline{e}(N)$. Two algorithms are presented; the first generates $\underline{e}(N)$ using the exact solution to $A_{N,p}$ from (17), and the second generates e(N) using an approximation to $A_{N,p}$ in (18).

When t=p, the exact recursive filter $A_{p,p}$ is found by extending the algorithm in [10] to the multidimensional case. The resultant lattice algorithm is presented below. Here "n" is the (discrete) time, $f^{x}_{m,n}$, $f^{y}_{m,n}$, $b^{x}_{m,n}$ and $b^{y}_{m,n}$ are Mx1 vectors, $\gamma_{m,n}$ is a scalar, $\sigma_{m,n}$, $\tau_{m,n}$, $\mu_{m,n}$ and $\omega_{m,n}$ are MxM matrices, λ is the exponential forgetting factor, and $\underline{y}(n)=\underline{x}(n-q)$.

1) Initial conditions: $\sigma_{m,q}=\tau_{m,q}=\mu_{m,q}=\omega_{m,q}=0$ for m=0,1,...,p-1

2) Set n=q+1

3) $f^x_{0,n}=b^x_{0,n}=\underline{x}(n)$

$f^y_{0,n}=b^y_{0,n}=\underline{y}(n)$

$\gamma_{0,n}=0$

$\mu_{0,n}=\lambda\mu_{0,n-1} + \underline{y}^*(n)\ \underline{x}'(n)$

$\omega_{0,n}=\lambda\omega_{0,n-1} + \underline{y}^*(n)\ \underline{x}'(n)$

4) For m=0 to min(p-1,n-q-1):

$$f^x_{m+1,n} = f^x_{m,n} - (\sigma_{m,n})' (\omega_{m,n-1})'^{-1} b^x_{m,n-1}$$

$$f^y_{m+1,n} = f^y_{m,n} - (\tau_{m,n})^* (\omega_{m,n-1})^{*-1} b^y_{m,n-1}$$

$$b^x_{m+1,n} = b^x_{m,n-1} - (\tau_{m,n})' (\mu_{m,n})'^{-1} f^x_{m,n}$$

$$b^y_{m+1,n} = b^y_{m,n-1} - (\sigma_{m,n})^* (\mu_{m,n})^{*-1} f^y_{m,n}$$

$$\mu_{m+1,n} = \mu_{m,n} - (\tau_{m,n}) (\omega_{m,n-1})^{-1} (\sigma_{m,n})$$

$$\omega_{m+1,n} = \omega_{m,n-1} - (\sigma_{m,n}) (\mu_{m,n})^{-1} (\tau_{m,n})$$

$$\sigma_{m,n} = \lambda\sigma_{m,n-1} + (b^y_{m,n-1})^* (f^x_{m,n})' (1-\gamma_{m,n})^{-1}$$

$$\tau_{m,n} = \lambda\tau_{m,n-1} + (f^y_{m,n})^* (b^x_{m,n-1})' (1-\gamma_{m,n})^{-1}$$

$$\gamma_{m+1,n} = \gamma_{m,n} + (b^x_{m,n-1})' (\omega_{m,n-1})^{-1} (b^y_{m,n-1})^*$$

5) Set n=n+1 and go to 3.

For this algorithm $f^x_{p,n}$ is the filter output $g(n)$. For $m<p$, $f^x_{m,n}$ is the optimal output of the prefilter as if an ARMA (m,q) model was used in (13). Therefore, this algorithm automatically provides the optimal filter output for all filters of lower AR order as well. This property can be useful if a good model order is not known apriori.

The second fast recursive algorithm addresses the $t>p$ solution for $A_{N,p}$. There is no currently available fast recursive algorithm to find the exact optimal filter output $g(n)$ using $A_{N,p}$ in (18). However, by generating an approximation of (18) a fast recursive algorithm can be realized. We can express

(18) as

$$A_{N,p} = -\left[Z_{N,p}^{\dagger} X_{N,p}\right]^{-1} Z_{N,p}^{\dagger} \underline{x}_N \tag{19}$$

where $Z_{N,p} = Y_{N,t} Y_{N,t}^{\dagger} X_{N,p}$. If we approximate $Z_{N,p}$ by a matrix $\tilde{Z}_{N,p}$ where

$$\tilde{Z}_{N,p} = \begin{bmatrix} 0 & & \underline{0} \\ \tilde{z}'(q+1) & \ddots & \\ \vdots & \ddots & \tilde{z}'(q+1) \\ \tilde{z}'(N-1) & \cdots & \tilde{z}'(N-p) \end{bmatrix} \tag{20}$$

then the above fast recursive algorithm with $\underline{y}(n) = \underline{\tilde{z}}(n)$ can be used. One effective procedure for obtaining $\underline{\tilde{z}}(n)$ was proposed in [14] for the one-dimensional case. The multidimensional extension is

$$\tilde{z}(n) = [00\ldots01]\, Y_{n,t}\, Y_{n,t}^{\dagger}\, \underline{x}_n \tag{21}$$

The recursive procedure for computing $\underline{\tilde{z}}(n)$ follows by writing

$$\underline{\tilde{z}}'(n) = b_n Y_{n-1,t}^{\dagger} \underline{x}_{n-1} + b_n b_n^{\dagger} \underline{x}'(n) \tag{22}$$

where $b_n = [x'(n-q-1)\ \ x'(n-q-2)\ \ldots\ x'(n-q-t)]$ The update for $\underline{\tilde{z}}(n)$ follows by utilizing the updates

$$\begin{aligned} Y_{k,t}^{\dagger} \underline{x}_k &= Y_{k-1,t}^{\dagger} \underline{x}_{k-1} + b_k^{\dagger} \underline{x}'(k) \\ b_k b_k^{\dagger} &= b_{k-1} b_{k-1}^{\dagger} + x'(k-q-1)x^{*}(k-q-1) - x'(k-q-t)x^{*}(k-q-t) \end{aligned} \tag{23}$$

We have thus presented two block processing and two recursive filtering algorithms. They form a subclass of the 2-D ARMA algorithms presented in the preceeding section. It is important to note that it has never been assumed in these algorithms that the antennas in the array are equally spaced and collinear, so these algorithms can be used <u>with no modification</u> on data obtained from more general array geometries. On the other hand, careful study of equations (17) or (18) reveal that the implicit autocorrelation estimates do not sum lags in the spatial direction, so for collinear, equally spaced arrays the data are not used in the most efficient manner.

VI. RECOMMENDATIONS

We have presented several algorithms for combined time-space prefiltering of antenna array data. Before these algorithms can be incorporated in a communication system, however, many questions remain unanswered. First, the performance of these prefiltering algorithms must be evaluated. Since the purpose of the prefilter is to improve the HF receiver performance, probably the best measure of prefilter performance is the receiver bit error rate. A more autonomous performance measure is the SIR and SNR gains at the prefilter output. Another possible measure is whiteness of the prefilter output array. Comparisons should be made not only among the algorithms presented but also with other prefiltering strategies currently in use. Performance evaluation by computer simulation is currently under investigation.

Selection of the data and prefilter model orders is a second important research topic. The prefilter coefficients can be obtained by assuming either an ARMA data model or an AR data model; it is not known, however, in which model provides the better SIR or SNR improvements. Also, proper AR model order selection is critical. If ρ_1, ρ_2, ρ_3 are too small, ineffective filtering results. If they are too large, the AR coefficient estimation algorithms can become ill-conditioned. Clearly, these questions merit investigation.

Implementation of the algorithms is another area for future research. Effects of finite word length computations on performance is one concern. Also, certain parameters values in these algorithms as presented here increase without bound as the number of data samples (N) increases. The development of normalized

of data samples (N) increases. The development of normalized algorithms is needed to alleviate this effect. In particular, square-root normalized versions of the fast recursive algorithms similar to those in [13] could be derived.

This research effort developed prefiltering methods assuming that the signal direction θ_s was unknown. In some applications, θ_s is either known or may be estimated. Therefore, modifications of these methods that incorporate knowledge of θ_s may provide improved system performance. Also, we have not yet discussed exactly how to choose the weights to combine the M prefilter outputs into one signal which feeds the channel equalizer (see Figure 1). One approach is to feed only one of the M signals to the equalizer. When θ_s is unknown, this is probably the best choice. However, when θ_s is known other choices of these weights should also improve receiver performance. Therefore, investigation of ways to incorporate this knowledge into the weight selection is also recommended.

Finally, recursive algorithms were presented only for a special class of model orders (i.e., AR(M-1) in the spatial direction). Since recursive algorithms are more amenable to implementation, it is important to develop recursive versions of the more general algorithms of Section IV.

REFERENCES

[1] Sylvania Systems Group, "HF Wideband Modem Program", RADC Tech Report CDRL-A008, November 1981.

[2] R. A. Monzingo and T. W. Miller, Introduction to Adaptive Arrays, New York, Wiley-Interscience, 1980.

[3] O. L. Frost, III, "An Algorithm for Linearly Constrained Adaptive Array Processing", Proc. IEEE, Vol. 60, No. 8, August 1972, pp. 926-35.

[4] J. W. Modestino, et. al. "Digital Communications in Spatially Distributed Interference Channels." Rome Air Development Center Tech. Report RADC-TR-82-151 December 1982.

[5] R. A. Wiggins and E. A. Robinson, "Recursive Solution to the Multichannel Filtering Problem," J. Geophysical Research, Vol. 70, No. 8, Apr. 1965, pp. 1885-91.

[6] M. Morf, A. Vieira, D. T. Lee, and T. Kailath, "Recursive Multichannel Maximum Entropy Spectral Estimation," IEEE Trans. Geoscience Electronics., Vol. GE-16, No. 2, Apr. 1978, pp. 85-95.

[7] J. G. Proakis and J. W. Ketchum, "Suppression of Narrowband Interference in Psuedo-Noise Spread Spectrum Systems," Rome Air Development Center Tech. Report RADC-TRL-81-6, Feb. 1981.

[8] Koopmans, The Spectral Analysis of Time Series, New York, Academic Press, 1974.

[9] J. A. Cadzow and R. L. Moses, "Adaptive ARMA Spectral Estimation, Part I" Proc. 1st IEEE Spectral Estimation Workshop, Hamilton, Ont., Aug. 17-18, 1981.

[10] J. A. Cadzow and R. L. Moses, "Adaptive ARMA Spectral Estimation, Part II" Proc. 1st IEEE Spectral Estimation Workshop, Hamilton, Ont., Aug. 17-18, 1981.

[11] J. H. Justice, "A Levinson-Type Algorithm for Two Dimensional Weiner Filtering Using Bivariate Szego Polynomials," Proc. IEEE, Vol. 65, No. 6, June 1977, pp. 882-6.

[12] H. Akaike, "Block Toeplitz Matrix Inversion," SIAM J. Appl. Math, Vol. 24, No. 2, March 1973, pp. 234-41.

[13] B. Friedlander, "Instrumental Variables Methods for ARMA Spectral Estimation," IEEE Trans. Acoustics, Speech, Signal Proc., Vol. ASSP-31, No. 2, April 1983, pp. 404-15.

[14] R. Moses, "Fast Recursive Spectral Estimation From an Overdetermined Set of Extended Yule-Walker Equations," Proc. IEEE Int'l. Conf. Acoustics, Speech, Signal Proc., Boston, MA, April 14-16, 1983.

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

RAMAN SPECTROSCOPY OF INHIBITED AND STIMULATED, NORMAL AND NEOPLASTIC CULTURED HUMAN AND MAMMALIAN CELLS

Prepared by:	Dr. James J. Mrotek
Academic Rank:	Associate Professor
Department, School and University:	Department of Physiology School of Medicine Meharry Medical College
Research Location:	School of Aerospace Medicine Data Sciences Division Biomathematics Modeling Branch Brooks Air Force Base, Texas
USAF Research:	Dr. John Taboada
Date:	October 24, 1983
Contract No:	F49620-82-C-0035

RAMAN SPECTROSCOPY OF INHIBITED AND STIMULATED, NORMAL AND NEOPLASTIC CULTURED HUMAN AND MAMMALIAN CELLS

by

James J. Mrotek

ABSTRACT

Y-1 mouse adrenal tumor cells and two human respiratory tract fibroblast cell lines (HEP-2, transformed, and MRC-5, non-transformed) were used for laser-Raman spectroscopy. One spectral line group observed with non-stimulated Y-1, MRC-5, and HEP-2 cells was similar to bacterial cell lines; the lines may represent laser photon scattering by high energy-containing compounds such as ATP. Comparing spectra from HEP-2, MRC-5 and Y-1 cells, total spectral lines increased in the order: HEP-2, Y-1 and MRC-5. ATP increased adrenal cell steroid production and caused increased numbers of lines in both Stokes and anti-Stokes wavelengths; anti-Stokes activity is seldom observed in Raman spectra. The spectra common to all living cells was masked by the extreme ATP spectral activity. HEP-2 molecules produced significant numbers of tightly grouped, adjacent Raman spectral doublet and triplet patterns not observed with non-transformed fibroblasts. Cultured human and mammalian cell line preliminary results suggested that spectroscopy may produce significant new information about spectral signatures of intracellular molecules and intracellular energy transduction.

Acknowledgement

The scientifically productive and creative summer spent at the School of Aerospace Medicine, Brooks Air Force Base, Texas, would not have been possible without the cooperation of the Air Force Systems Command, the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education. The project described herein was supervised by Dr. Richard Albanese of the Biomathematical Modeling Branch of Data Sciences and conducted in collaboration with a very able and stimulating member of this group, Dr. John Taboada. Our laser-Raman spectroscopy was performed in the Ophthalmology Branch research laboratories. Special mention must be made of the consideration and courtesy of the personnel of the Epidemiology Division, particularly Sgt Dwight Saunier and the members of Dr. Vee Davison's virology laboratory: Jim Dispanet, Cliff Miller, Sgt. Donna Norris and Airman Richard Green. Without their help our work would not have been performed.

I sincerely thank Dr. Bryce Hartman for the well considered matching of my interests and capabilities with those of Dr. Taboada. The interaction which resulted from this matching of our diverse talents offers a better understanding of the often used term, collaboration.

I. INTRODUCTION

Through the use of USAF School of Aerospace Medicine lab director's funds and support, Dr. John Taboada, working in the Clinical Sciences Division of the School of Aerospace Medicine, is developing a research program and experimental facility to investigate the laser-Raman spectroscopy of molecules in living systems. This new and incisive approach to clinical testing is directed toward improving non-invasive, stressed and non-stressed clinical diagnosis of Air Force personnel.

Dr. Taboada initiated the spectroscopy program with a reexamination of the laser-Raman spectroscopy of living bacterial cell systems, recently published by Webb and coworkers (1). His preliminary findings seem to support the existence of useful laser-Raman spectral signatures which are related to life processes.

In the following section the theoretical background for Dr. Tabaoda's research program will be briefly reviewed. This background will discuss the use of laser-Raman spectroscopy to study molecules within living cells and will examine the hypothesis that Raman spectroscopy may be of value in studying energy transduction within cells. The unique features of the Y-1 mouse adrenal tumor cell and other cultured steroidogenic cells will be presented in order to provide a backgroud for the studies conducted during this past summer. This background is necessary for understanding the relation-

ship of intracellular molecules and molecular function in the production of steroids by these cells. In addition, we will investigate the possible ways in which these cells and their molecules could serve as models in the examination of problems of concern to the Air Force. This discussion will be followed by a summary of our preliminary research findings. The influence of these preliminary results contributed significantly to the direction of our future research. We propose continuing the development of methods for spectroscopically examining cultured mammalian cells and initiating studies of model in vitro biochemical species.

THEORETICAL BACKGROUND FOR RAMAN SPECTROSCOPICAL STUDIES:

Because a molecule interacts with the incident photons with which it is irradiated, a small number of photons of shorter (anti-Stokes) or longer (Stokes) wavelengths than those of the original incident light will be scattered by the molecule during the irradiation. Laser-Raman spectroscopy is performed by scanning a range of wavelengths immediately preceeding, and succeeding, the wavelength of the incident photons to detect photons scattered by the irradiated molecules. In theory, a given molecule will produce only a limited number of specific spectra.

Davydov proposed that proteins may transport or store energy within a cell by generating soliton vibrations along the alpha helix of the protein molecule (2). The

soliton is defined as a distinctively non-linear phenomenon in which a conserved quantity propagates without change of temporal shape along a one-dimensional channel as a localized "lump" (3). From the molecular point of view the soliton could be a local, but mobile, region of conformational change; thus, proteins and DNA are natural molecules for soliton formation (4,5). The importance of the Davydov proposal for soliton-generation in biomolecules rests in the fact that solitons represent a mechanism for the transport of energy between its site of generation within a subcellular region and the relatively great cellular distance to the sites where it ultimately is used, or stored (2). Since high molecular weight biomolecules are often located between the site of energy formation and its ultimate utilization, these molecules would be logical candidates in which to generate solitons and transport them. In this regard it is interesting that Raman spectroscopy was used by S. J. Webb and his associates to study molecular vibratory states in living cells (6).

Although laser-Raman spectroscopy was used for a number of years to study inorganic chemical reactions, in 1968 Webb and his associates proposed that this technique could be used to study living cells (7). Using cold shock to synchronize division of the bacterial species, Bacillus megaterium, Webb and coworkers were able to detect characteristic Raman spectra that seemed to be

associated with bacterial cell division (1). In a 1982 study, these same investigators suggested that cells from human breast tumors and the non-tumerous breast of the same individual gave a characteristic laser-Raman line that could not be observed in cells from breasts of non-tumerous patients (8). However, other investigators were unable to confirm these findings.

II. OBJECTIVES

Recently J. Taboada was able to duplicate the observation that cold synchronized Bacillus megaterium did indeed generate characteristic Raman spectra. The objective of the collaboration between Dr. Taboada and me was an attempt to determine whether cultured normal and neoplastic mammalian and human cell types produced Raman spectra and whether these spectra were modified by exogenous stimulants or by transformation of the cells to neoplasias.

Since I was familiar with properties of the Y-1 mouse adrenal tumor cell line and its special culture conditions, we began our studies with these cells. Cultured steroid producing cells can be used for studying the internal and external conditions affecting steroid production, Raman spectra and soliton wave generation because the intracellular molecules and steps involved in steroidogenesis are relatively well defined. Problems slowing our progress developed during adaptation of

the adrenal line to the special conditions required for laser-Raman spectroscopy. We, therefore, expanded our studies to include normal and neoplastic cultured human fibroblast lines in order to utilize my time and expertise most efficiently.

The Adrenal Cell --The background and rationale for the use of the Y-1 adrenal cells in our studies is as follows:

In general for steroid producing cells (9): 1) The stimulating pituitary hormone binds to receptors on the extracellular surface of the cell; 2) Bound hormone activates an enzyme, adenyl cyclase, located on the cytoplasmic side of the surface membrane, causing ATP to be converted into cyclic 3',5' adenosine monophosphate (cAMP); 3) Actin-containing microfilaments are either directly activated by cAMP or by other systems and serve as transducers of information from the cell membrane to the steroid synthesizing organelles, the mitochondria (10-12) [For purposes of discussing soliton wave propagation along alpha helical molecules, it is important to understand that actin microfilaments are composed of single (monomeric) proteins polymerized into two filamentous strands which are alpha-helically wound around each other. The polymers are formed when individual globular actin monomers, themselves containing regions of alpha helix within the monomeric globular protein molecule, are polymerized by ATP (13).]; 4) Activated

actin-containing microfilaments cause cholesterol stored in the cytoplasm of the steroid producing cell to be brought together with mitochondria (11, 12); 5) Within the mitochondria are located the cholesterol side-chain cleavage enzyme complexes (14, 15):

a) The bringing of the cholesterol substrate t. the cleavage enzyme is the rate-limiting step for the synthesis of steroids by tissues such as the placenta, ovary, testis and adrenal cortex; b) This step is ultimately controlled by the pituitary hormone binding to the cell surface; c) The cholesterol side-chain cleavage enzyme complex has a very high molecular weight; d) The active enzyme complex is constructed from three different proteins: non-heme iron, adrenodoxin and cytochrome P-450; e) Each of these molecules is globular in nature and contains sizable alpha helical regions; f) In addition, cytochrome P-450 is a polymeric protein composed of 16 identical monomers; g) The enzyme complex functions when adrenodoxin binds NADPH, cytochrome P-450 binds molecular oxygen and the energy in the NADPH causes the molecular oxygen to be transferred to the side-chain of the cholesterol molecule; h) Because of the large size of this complex it seems possible that solitons may be required to transfer the energy in NADPH the relatively great distance from the adrenodoxin molecule to the region where molecular oxygen is transferred

to cholesterol; i) In order to obtain data examining the hypothesis which predicts that solitons travel along proteins containing alpha helical regions (3), the larger the molecules or molecular complex the easier it will be to investigate these events using Raman spectroscopy. Thus the cholesterol side-chain cleavage enzyme is an ideal model for testing this hypothesis. Since it is projected that future computer systems could use proteins as microswitches, establishing that solitons can be carried along proteins becomes somewhat important;

.. As a result of cholesterol side-chain cleavage the steroid known as pregnenolone is produced (16); 7) This steroid is further modified by enzymes located in the smooth membranes of the cytoplasm and in the mitochondria (9); 8) Once the specific substrates for these later enzymes are present, these enzymes require no further activation to produce and secrete steroid hormones.

III. RESULTS

Adrenal Cells --Using non-stimulated cultures of the cloned Y-1 mouse adrenal tumor cell line, we determined that two populations of spectral lines could be observed. The first population contained spectral lines similar to those appearing in a certain wavelength

region of the bacterial spectra; we speculate that these may represent spectral lines associated with the utilization of high energy compounds such as ATP. Thus, these spectra may represent molecular vibrations generated during activities which are common to living cells such as these mammalian and bacterial cells. The second population of spectral lines observed with the Y-1 adrenal cell seemed to be different from those observed with bacterial cells and may, therefore, represent molecular vibrations that specifically occur in adrenal cell molecules.

Cultured adrenal cells produce small amounts of steroid hormones when they are incubated using standard control conditions (17). If the cultures are exposed to stimulating agents such as the pituitary hormone, adrenocorticotropin (ACTH), the second messenger, cyclic 3',5' adenosine monophosphate (cAMP), ATP, or cholera toxin (17, 18), the cells respond by increasing steroid production at least ten-fold. Because one group of spectral lines observed in non-stimulated adrenal and bacterial cells could result from scattering of laser photons by high energy-containing compounds such as the ATP found in all living cells, we tested whether the addition of excess ATP to increase steroid production by adrenal cells caused alteration of the Raman spectra we observed with the control cells. The number of intense, narrow band peaks recorded for the wavelengths of both the Stokes and anti-Stokes photons increased

dramatically. The effect of ATP on anti-Stokes wavelengths was surprising; cells seldom produce anti-Stokes light scattering events. The total scattering activity was so great that it was difficult to determine the effect of the ATP on those portions of the spectra which seem to be common to all living cells. Using populations of cells obtained from various species, future experiments will be needed to determine the effects on this portion of the spectra by adding increasing concentrations of ATP to these cells.

Fibroblast Cell Lines --In a second group of experiments Dr. Taboada and I studied two different human respiratory tract fibroblast cell lines. Clonal cultures of a transformed human nasal fibroblast carcinoma, HEP-2, were compared to the non-transformed human lung fibroblast cell line, MRC-5. Again, we were able to detect "life-related" spectral bands similar to those observed with the bacterial and the adrenal cells. The spectra scattered by molecules in the HEP-2 cells did not contain as many lines as those produced by molecules contained within irradiated adrenal tumor cells. MRC-5 cell molecules were highly active in producing Raman spectral lines. Preliminary information from the HEP-2 line also suggested that the molecules in these cells produce a significant number of tightly grouped adjacent Raman spectral lines, resulting in recordings of doublet and triplet patterns. Non-transformed fibroblasts did

not emit these multiple signals; instead, numerous, intense narrow band wavelengths were produced. Our observations were complicated by the fact that the fibroblast lines studied were derived from different individuals and from two different regions of the respiratory tract. However, these results are consistent with the results of Webb and coworkers who observed similar spectral differences in Raman scans of normal and cancerous breast tissue (8). Whether our experiments offer a possible method for evaluation of normal and cancerous tissues requires many more observations.

IV. RECOMMENDATIONS FOR FUTURE RESEARCH

In this section procedures are described for solving problems which developed during my ten week summer research fellowship at the School of Aerospace Medicine, Brooks Air Force Base, Texas. Having solved the logistical problems impeding further rapid progress, it is assumed that the techniques will be applied to obtain insights into more fundamental problems of concern to the Air Force. Thus, an additional series of experiments to be conducted over a longer time period are briefly described.

SHORT-TERM EXPERIMENTS DESIGNED TO DEVELOP TECHNIQUES FOR OPTIMIZING RAMAN SPECTROSCOPY OF CULTURED CELLS

In our preliminary work during the summer research fellowship period, several problems were encountered which

are now being solved at Meharry Medical College. These include:

Problems Associated With Using Cell Suspensions Produced By Removing Calcium From The Extracellular Fluid --In order to conduct our experiments it was necessary to adapt the handling of the cells to the existing system for Raman spectroscopy of bacterial cells. Thus, the human and mammalian cells had to be suspended in a non-fluorescent nutrient medium containing the calcium chelator, EDTA, in order to introduce them into a cuvette for scanning. This treatment is traumatic to the physiology of the cell because excess EDTA may cause intracellular calcium to diffuse to the extracellular fluid and alter intracellular activity.

Proposed Approach --An alternative cell handling system for conducting the laser scan modifying techniques I developed to grow cells on negatively charged, carbon-coated microscope slide coverslips for scanning electron microscopy (12). Coverslips will be inverted over a welled microscope slide that contains culture medium in the well (see diagram 1).

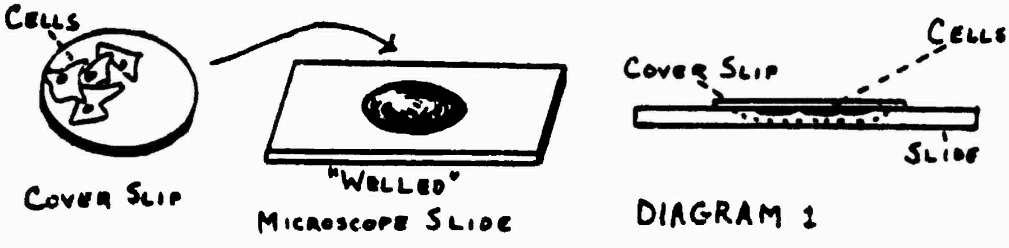


DIAGRAM 1

If suitable modifications were made to the drive system of the Raman spectrometer, the microscope slide containing the inverted, cell-covered coverslip could be

moved through the laser beam to scan the cells.

LONG-TERM INVESTIGATIONS USING HUMAN AND MAMMALIAN CELL LINES FOR RESEARCH IN RAMAN SPECTROSCOPY

Potential Studies Using Steroid Producing Cells --The two steroid producing cell lines which I can culture are mouse adrenal and rat testis tumor cells. The steroid pathways of these cells are well defined and offer an additional advantage that they contain the cholesterol side-chain cleavage enzyme which we wish to eventually study using laser-Raman spectroscopy. Physiological concentrations of various exogenous biochemicals can be used to by-pass any of the above mentioned steps in the steroidogenic pathway and the steroids produced in response to these biochemicals can be conveniently measured using a radioimmunoassay (11, 18-20). As a result, each of these cell lines is useful in defining the step of the pathway affected by environmental toxicants, such as cigarette smoke (19). Because of the success of the cigarette studies, I expect that these cells could also be useful in defining the effects of organophosphates and anaesthetics in these cells. Since others have predicted that certain anaesthetics act by inhibiting soliton generation within membrane proteins, the adrenal may offer a convenient model system for testing this hypothesis. Indeed, studies conducted by Dr. P. F. Hall and myself with the anaesthic Halothane suggest that the ability to stimulate steroidogenesis

is inhibited by this anaesthetic (12). Since the adrenal cell produces the steroid hormones which function to protect an organism against the effects of stress (21), it is important to know how the above mentioned agents affect the production of steroids and the Raman spectra of the cells. Equally important are the insights into effects of toxicants on production of reproductive steroids and the Raman spectra of testicular interstitial cells. Since reproductive steroids are known to affect behavior and efficiency, two activities of concern in Air Force daily activities, cultured testicular interstitial cells will provide convenient models for these studies.

Potential Studies With Fibroblasts --In previously described studies we observed that MRC-5 fibroblasts offer the possibility for evaluating age-related changes occurring within continuously cultured cells. Fibroblasts are the most universal cell found in our body (22); they serve as the framework around which all organs and tissues in our body are constructed. Hayflick, the "father" of cellular gerontology, suggests that understanding the process of senescence in fibroblasts may serve to facilitate understanding of declining function of tissues and organs in the older individual, since continued function of these cells is mandatory for continued function of the organ and tissue (23). The MRC-5 cell line is an extensively characterized, pedigreed human fibroblast cell line used in

studies of aging (24). These cells grow and multiply for 30 generations after which the cells cease dividing. Cells of various ages can be purchased for study of age-related cellular changes. Since the Air Force has as one of its goals development of non-invasive diagnostic techniques to assess age-related reduction in personnel efficiency, MRC-5 cells may serve as a model system to non-invasively determine the characteristic laser-Raman spectral changes that occur as fibroblast cells age. The pattern of these changes may then be sought in tissues and organs of Air Force personnel.

REFERENCES

1. Webb, S.J., and Stoneham, M.E., "Resonances Between Ten To The Eleventh Power And Ten To The Twelfth Power Hz In Active Bacterial Cells As Seen By Laser Raman Spectroscopy," Physics Letters, Vol. 60A, pp. 267-268 (1977).

2. Davydov, A.S., "The Migration Of Energy And Electrons In Biological Systems, "Biology And Quantum Mechanics, (Pergamon Press, Oxford, U.K., 1982), pp. 185-207.

3. Scott, A.C., Chu, E.Y.F., and McLaughlin, D.W., "The Soliton: A New Concept In Applied Science," Proc. IEEE, Vol. 61, pp. 1443-1483 (1973).

4. Bilz, H., Buttner, H., and Froelich, H., "Electret Model For The Collective Behavior Of Biological Systems," Z. Naturforsch., Vol. 36B, pp. 208-212 (1981).

5. Englander, S.W., Kallenbach, N.R., Heeger, A.J., Krumhand, J.A., and Litwin, S., "Nature Of The Open State In Long Polynucleotide Double Helices: Possibility Of Soliton Excitations," Proc. Natl. Acad. Sci. USA, Vol. 77, pp. 7222-7226 (1980).

6. Lomdahl, P.S., MacNeil, I., Scott, A.C., Stoneham, M.E., and Webb, S.J., "An Assignment To Internal Soliton Vibrations Of Laser-Raman Lines From Living Cells," Physics Letters, Vol. 92A, pp. 207-210 (1982).

7. Webb, S.J., and Dodds, D.D., "Inhibition Of Bacterial Cell Growth By 136 gc Microwaves," Nature, Vol. 218, pp. 374-375 (1968).

8. Webb, S.J., Lee, R., and Stoneham, M.E., "Possible Viral Involvement In Human Mammary Carcinoma: A Microwave and Laser-Raman Study," International Journal of Quantum Chemistry, Quantum Biology Symposium 4, (John Wiley and Sons, Inc., New York, 1977), pp. 277-284.

9. Schulster, D., and Schwyzer, R., "ACTH Receptors," Cellular Receptors For Hormones And Neurotransmitters, (John Wiley and Sons, Inc., New York, 1980), pp. 197-217.

10. Mrotek, J.J., and Hall, P.F., "The Influence Of Cytochalasin B On The Response Of Adrenal Tumor Cells To ACTH And Cyclic AMP," Bioch. Biophys. Res Commun., Vol. 64, pp. 891-896 (1975).

Tumor Cells To Adrenocorticotropin: Site Of Inhibition By Cytochalasin B," Biochemistry, Vol. 16, pp. 3177-3181 (1977).

12. Mrotek, J.J., Rainey, W., Lynch, R., Mattson, M., and Lacko, I., "A Scanning And Transmission Electron Microscope Examination Of ACTH-induced 'Rounding Up' In Triton X-100 Cytoskeletal Residues Of Cultured Adrenal Cells," Current Topics In Muscle And Non-muscle Motility (Plenum Press, New York, 1982), Vol. 2, pp. 45-52.

13. Pollard, T., and Weihing, R. R., "Actin And Myosin And Cell Movement," CRC Crit. Rev. Biochem., Vol. 2, pp. 1-80 (1974).

14. Hall, P. F., "Part IV. Mixed-function Oxidases Of The Adrenal Cortex. Properties Of Soluble Cytochromes P-450 From Bovine Adrenocortical Mitochondria," Annals Of The New York Academy Of Sciences, Vol. 212, pp. 195-207 (1973).

15. Shikita, M., and Hall, P. F., "Cytochrome P-450 From Bovine Adrenocortical Mitochondria: An Enzyme For The Side Chain Cleavage Of Cholesterol. II. Subunit Structure," J. Biol. Chem., Vol. 248, pp.5605-5609 (1973).

16. Karaboyas, G. C., and Koritz, S. B., "Identity Of The Site Of Action Of 3', 5'-adenosine monophosphate and adrenocorticotropic hormone in beef adrenal cortex slices," Biochemistry, Vol. 4, pp. 462-468 (1965).

17. Kowal, J., "ACTH And The Metabolism Of Adrenal Cell Cultures," Recent Progress In Hormone Research, Vol. 26, pp. 623-687 (1970).

18. Lee, H.S., "The Effect Of Intermediate Filament Inhibitors On Steroidogenesis And Cytoskeleton In Y-1 Mouse Adrenal Tumor Cells," accepted for publication in Cell Biology International Reports.

19. Morris, P. D., "Cigarette Residues Affect Steroidogenesis In Cultured Y-1 Mouse Adrenal Tumor Cells," M.S. Thesis, Department of Biological Sciences, North Texas State University, Denton, Texas, pp.1-32 (1979).

20. Mattson, M.P., "Effects Of Exogenous Steroid On The Adrenal Plasma Membranes: Alteration Of Steroidogenesis" to be submitted to Steroids.

21. Selye, H., "The Diseases Of Adaptation," Recent Progress In Hormone Research, Vol. 8, pp. 117-142

Progress In Hormone Research, Vol. 8, pp. 117-142 (1953).

22. Ham, A. W., "The Cells Of Loose Connective Tissue And Their Functions," Histology, (J. B. Lippincott Company, Philadelphia, 7th edition, 1974), pp. 220-249.

23. Hayflick, L., "Future Directions In Aging Research," Proceedings Of The Society For Experimental Biology And Medicine, Vol. 165, pp. 175-184 (1980).

24. Hay, R., Macy, M., and Shannon, J., Editors, "ATCC CCL 79, Y-1, Adrenal Tumor, Mouse: Steroid Secreting," American Type Culture Collection Catalog Of Strains II: Cell Lines, Viruses And Antisera, Chlamydiae, and Rickettsiae, (American Type Culture Collection, Rockville, Maryland, Third Edition, 1981), p.76.

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

COMBINED MAGNETIC AND GRAVITY ANOMALIES: A GUIDE TO CRUSTAL TYPE AND TECTONICS IN THE SOUTHEASTERN INDIAN OCEAN AND CARIBBEAN REGIONS

Prepared by:	Dr. Frederick Nagle
Academic Rank:	Professor
Department and University:	Department of Geology University of Miami
Research Location:	Air Force Geophysics Laboratory Terrestrial Sciences Division Geodesy and Gravity Branch
USAF Research:	Dr. A.R. Lazarewicz
Date:	August 15, 1983
Contract No:	F49620-82-C-0035

COMBINED MAGNETIC AND GRAVITY ANOMALIES: A GUIDE TO CRUSTAL TYPE AND TECTONICS IN THE SOUTHEASTERN INDIAN OCEAN AND CARIBBEAN REGIONS

by

Frederick Nagle

ABSTRACT

As a consequence of a literature search to provide a geologic foundation for magnetic and gravity anomalies described by Sailor and Lazarewicz[1,2], it is evident that several large geographic features in the southeastern Indian Ocean may be, at least in part, continental. If any or all of them are continental, then reconstruction models for Antartica- India- Australia will have to be modified.

For one of these areas, the Kerguelen Plateau, there are deep sea core samples in storage, not yet studied in detail, which could help resolve this question. Preparatory steps have been taken for submission of a proposal to the Antartic Branch of the National Science Foundation.

Three model sites are designated for combined analysis of gravity and magnetic data to determine rock types and crustal structure. Ultimately such determinations could be done for other areas from satellite data alone.

Work has begun on a J/p (anomalous magnetization/anomalous density) map for the Caribbean. Since so much detail is known about the rocks and structure of this region, this map could serve as an interpretive guide for anomalies already known in the Indian Ocean or elsewhere.

ACKNOWLEDGMENTS

The author would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for providing the opportunity to spend a rewarding summer with the Terrestriel Sciences Division at the Air Force Geophysics Laboratory, Hanscom AFB, Massachusetts.

Several individuals within the Terrestrial Sciences Division provided counsel, guidence and assistence during the summer; amongst them ere: Dr. Donald Eckhardt, Director; Dr. Thomas Rooney, Chief, Geodesy end Grevity Brench; Jeenne McPhetres, Secretary; and especielly Dr. Andrew Lazarewicz whose work interested me and led to this project. The entire staff at the AFGL Research Library is thanked for courteous help throughout the summer.

Drs. Richard Sailor and Thomas Paar of The Analytic Sciences Corporation (TASC) in Reading, Mass.; Dr. Carl Bowin of Woods Hole Oceanographic Institute, Woods Hole, Mass.; and Dr. Mahlon Ball of the U.S. Geological Survey, Woods Hole all took portions of their own time to discuss verious aspects of this project with the author.

Finally, the author thanks his good friends M. Booth end C. Hamilton who made many personal sacrifices in support of the logistics of this project; and his wife and sons all of whom juggled the duties of their own jobs while helping to keep our home together during my absence.

I. INTRODUCTION:

Accurate magnetic and gravimetric data are essential to the Air Force for many guidance and navigational systems. That same information combined with "ground truth" rock and topographic data from a specific region can help earth scientists gain a three dimensional picture of the structure of that region, as well as to interpret other regions where similar magnetic and gravity signatures occur.

Sailor and Lazarewicz (AFGL)[1] and Lazarewicz and Sailor[2] produced an equivalent source crustal magnetic anomaly map from NASA's Magsat data of a region within the southeastern Indian Ocean (Figures 1, 2, 3, 4). This region contains several major tectonic features, amongst them, the following aseismic ridges and plateaus: Broken Ridge, a portion of the Ninetyeast Ridge, the Naturaliste Plateau and the South East Indian Ridge. All of these features were involved in or caused by the breakup of east Gondwanaland. Figure 3 shows the entire Indian Ocean, earthquake epicenters of the region, and the Deep Sea Drilling sites of the D/V Glomar Challenger.

Sailor and Lazarewicz[1] also produced a gravity anomaly map (Figure 5) and a vertical derivative of the gravity anomaly map (Figure 6). Using Poisson's relation and the maps of Figures 4 and 6, they computed a map of anomalous magnetization/anomalous density (J/p) Figure 7. From their maps one can see a large positive magnetic anomaly over the eastern part of the Broken Ridge and a positive gravity anomaly about 400 km to the west on the same ridge. Neither of these anomalies continues smoothly to the Ninetyeast Ridge. On the J/p map, there is a broad high on the eastern part of the Broken Ridge and a long

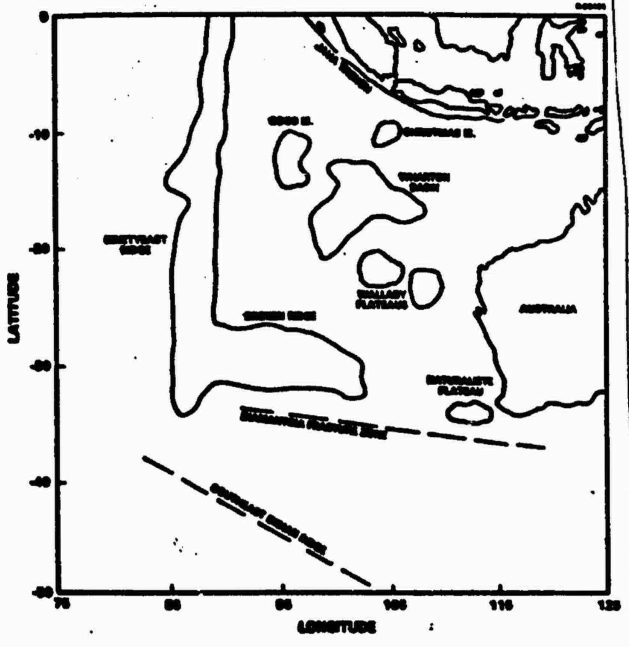


Figure 1. Tectonic features of the southeastern Indian Ocean (from Sailor and Lazarewicz [1]).

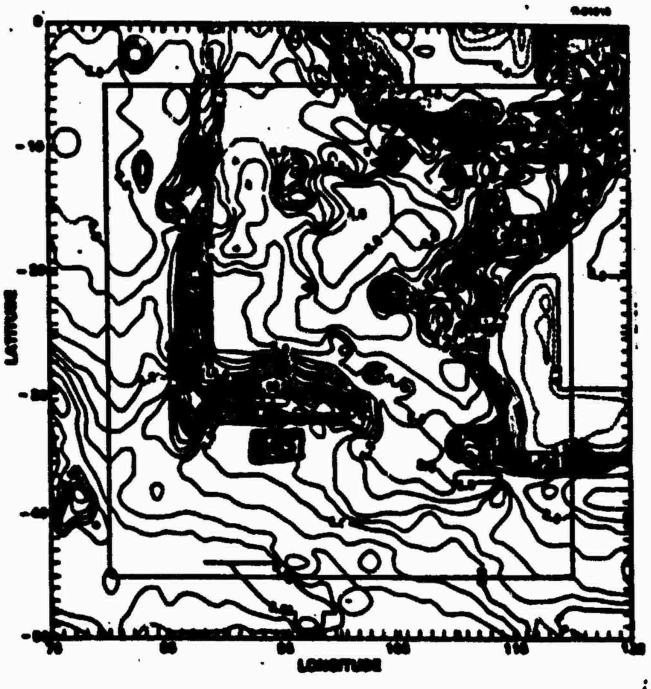


Figure 2. Bathymetry(in km)of southeastern Indian Ocean. Insert box is areacovered by Figure 4 (from Sailor and Lazarewicz [1]).

Figure 3. Principal bathymetric features of the Indian Ocean (Udintsev,et al.,[24]). Triangles mark Deep Sea Drilling Sites and dots the epicenters occurring between 1963 and 1973 from ESSA tape (from Heirtzler,et al.,[16]).

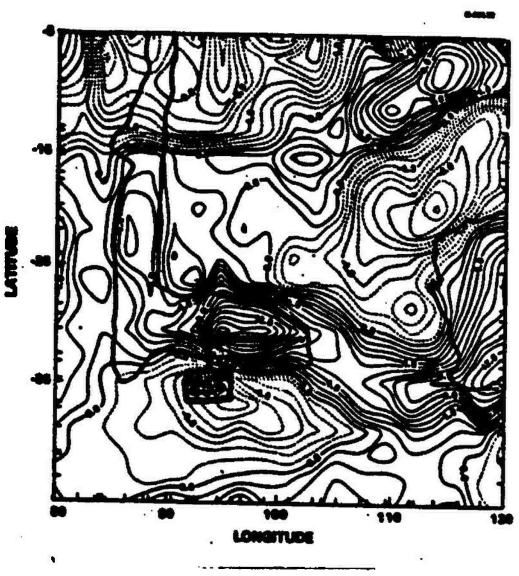


Figure 4. Final equivalent source anomaly map at 350km. Contour interval $\ln T$ (from Sailor and Lazarewicz [1]).

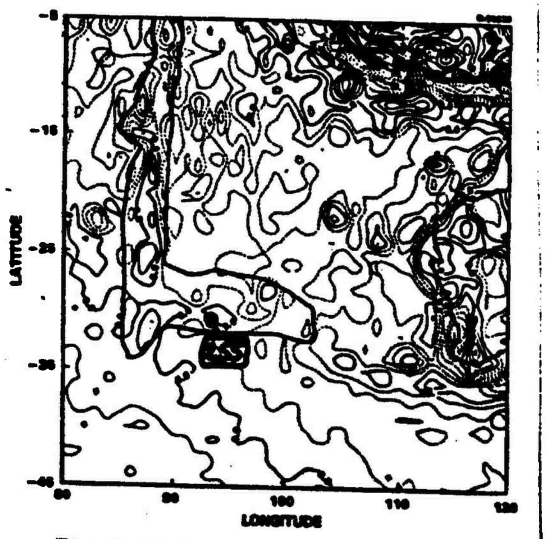


Figure 5. Gravity anomaly map at sea surface. Contour interval 10 mgal. (from Sailor and Lazarewicz[1]).

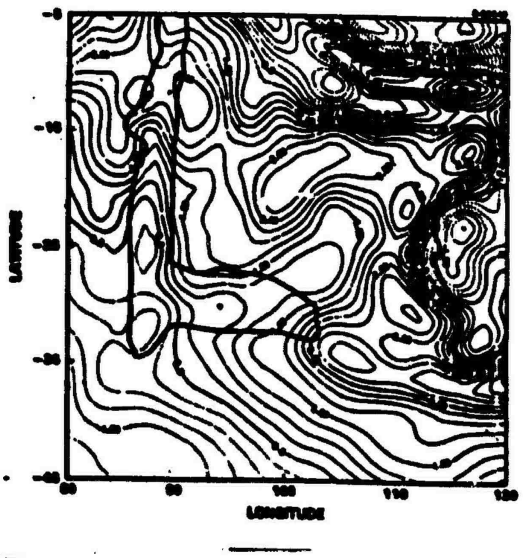


Figure 6. Vertical derivative of gravity anomaly at an altitude of 350 km. (from Sailor and Lazarewicz[1]).

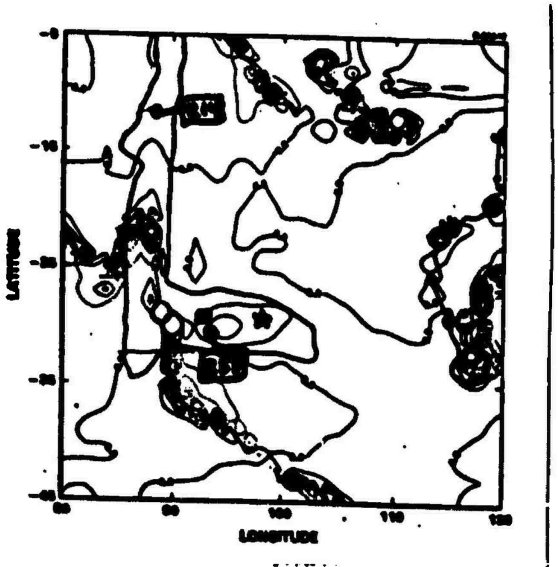


Figure 7. Map of J/p using Poisson's relation and maps of Figures 4 & 6 (from Sailor and Lazarewicz).

linear trend in the southwestern portion of the map which may or may not be associated with a real feature. The authors point out that this linear trend is offset in the area where there are changes in the magnetic and gravity maps (Figures 4 and 5) and that variations in J/p could be due to differences in crustal thickness, rock type, density, susceptibility, or depth to the Curie isotherm.

Combining magnetic and gravity data using Poisson's equation in geophysical studies are not unique to Sailor and Lazarewicz (see references 3, 4, 5, 6). Garland's paper[4], which attempts a determination of the rock type causing the anomalies, is particularly pertinent to this study.

This author knew of considerable geologic information not taken into account in the original Sailor and Lazarewicz papers. Also, he has had both land and sea field experience in the Caribbean region[7,8,9,10]. This region (Figure 8) has several different types of crust within its boundaries: oceanic crust in the Atlantic Ocean, intermediate thickness crust in the Colombian and Venezuelan Basins, island arc crust under the Greater and Lesser Antilles, carbonate platforms under the Bahamas and the Yucatan and continental crust in Venezuela and Colombia. Since many of these areas have been studied in some detail, the Caribbean region is an ideal region to compare to the Indian Ocean area examined by Sailor and Lazarewicz[1,2].

II. OBJECTIVES:

During the course of this study, four objectives are being pursued:

(1) To interpret the geophysical anomalies already discovered in the Indian Ocean by researching the published literature.

(2) To determine the availability of magnetic and gravity data within the Caribbean for the construction of a J/p (anomalous magnetization/

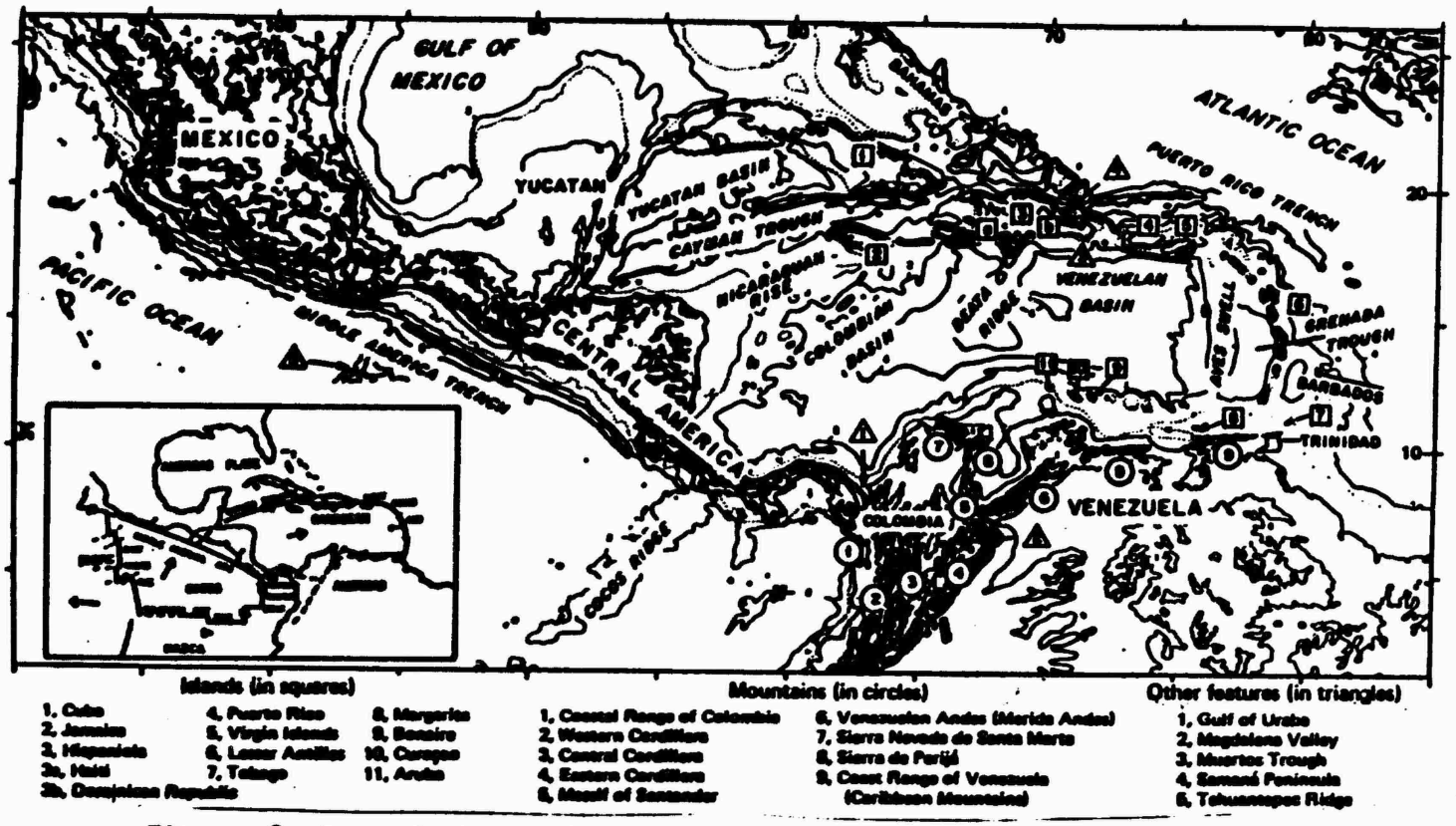


Figure 8. Map of the Caribbean region with principal features. Insert shows plate boundaries as given by Molnar and Sykes, 1969, Fig. 1. (from Bowin [32]).

anomalous density) to compare with the one produced by Sailor and Lazarewicz[1,2] for the southeastern Indian Ocean.

(3) To pick sites in both the Caribbean and southeastern Indian Ocean with well-defined magnetic and gravity anomalies which might be analyzed to define rock type (hence crustal type).

(4) To compare the results of using surface data versus satellite data for objectives (2) and (3). Due to time limitations, this objective was not pursued this summer.

III. GEOLOGY OF THE SOUTHEASTERN INDIAN OCEAN:

Sailor and Lazarewicz[1,2] focused their studies on Broken Ridge, an aseismic ridge which abuts abruptly against the Ninetyeast Ridge. The review below is restricted to features which are joined to this ridge or those which may have been connected to it in the past. Most of the comments are from Schlich's[11] review of the area in the Nairn and Stehli[12] volume. Other publications which give the reader a good geologic review of the region are the Deep Sea Drilling volumes 22, 26 and 28[13,14,15] and the summary volume edited by Heirtzler, et al.[16].

The fact that Broken Ridge is joined at right angles to the Ninetyeast Ridge (Figure 2) is unusual and might make one suspect that the two ridges have had different origins and histories. Broken Ridge is about 1200 km long by 400 km wide at its widest point, and is bounded on the south by a scarp which has a relief of about 4,000 m. Francis and Raitt[17] reported the mantle-crust boundary at 20 km under one portion of the ridge and concluded that the ridge had a "quasi-continental" structure. Drilling at site 255 did not penetrate to basement rocks so the question of whether the ridge has continental or oceanic crust can not be decided, but drilling at that site did reach Cretaceous

limestones overlain with angular unconformity by Eocene sediments (Figure 9). The Cretaceous limestones are considered by Davies, Luyendyk, et al.,[18] to have been deposited in shallow water of a few hundred meters or less. Subsidence began at this site only in the late Eocene. This drilling site, the only Deep Sea Drilling site on Broken Ridge, is located on Figures 2, 3, 4, 5 and 7 for reference. There is one heat flow measurement reported from the ridge (Udinstev[24]) of 2.9 HFU, somewhat higher than the average of 1.6 HFU from aseismic ridges in the Indian Ocean.

As Schlich[11] points out, the Kerguelen-Heard Plateau (Figure 3), separated from the Broken Ridge by the South East Indian Ridge, can be divided at roughly 55° into two distinct domains on the basis of topography. This plateau extends 2,000 km in a northwest-southeast direction and is shallower (mainly above 1,000 m) in the north, deeper (mainly below 1,000 m) with more subdued topography in the south. There are two island groups on the plateau, both with volcanoes which have been active since the early Tertiary. The Kerguelen group on the northern half of the plateau is composed of one large island (160 km across) and several smaller islands totaling approximately 5,000 sq km. Basaltic rocks are dominant and are amongst the oldest at 30my although there are granite differentiates on the island. Heard island, to the south of the Kerguelens, but also on the northern half of the plateau, is about 25 km by 20 km in size and also has mostly basic volcanic rocks as well as some acid differentiates. The oldest rocks here are limestones interbedded with pyroclastics of probable Paleogene age.

None of the rocks exposed on these islands give any indication of continental crust[19,20,21,22], either as xenoliths or as a requirement

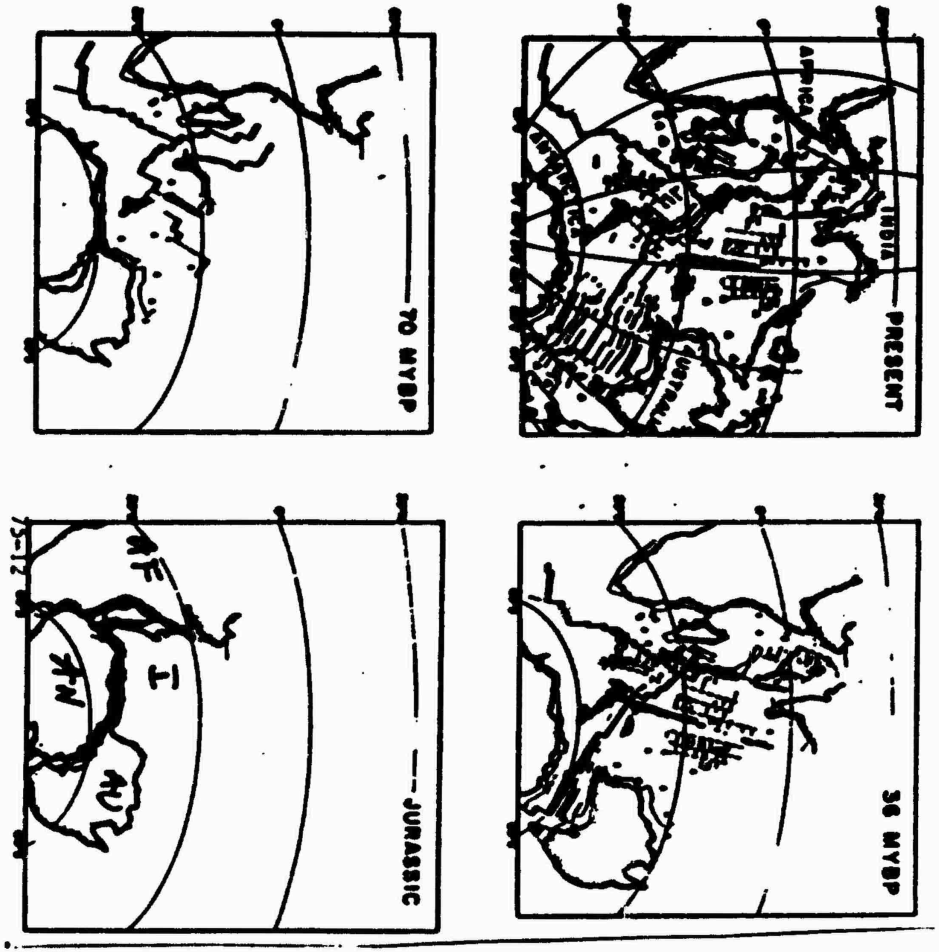


Figure 10. Reconstructions of the Indian Ocean from Jurassic to Present. (Heirtzler, et al.,[16]).

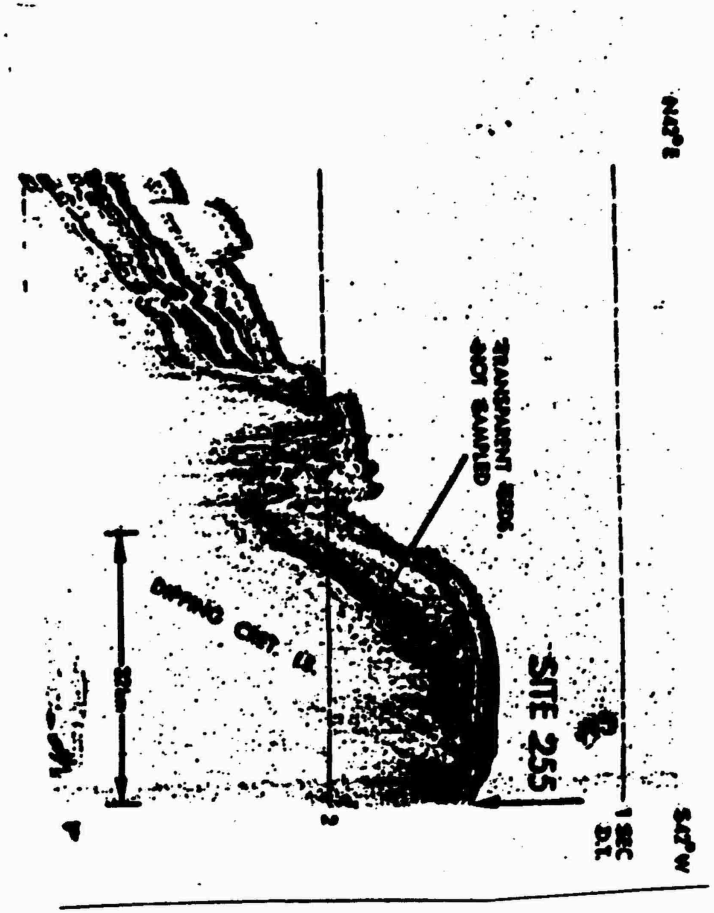


Figure 9. Seismic profile of Broken Ridge (from Naini, et al.,[40]).

for their igneous chemistry, but one should keep in mind that this plateau is immense in comparison to the small islands. The only published information on older rocks south of 55° is by Quilty[23] who reports Cenomanian-Turonian sediments from a single core at about 56°S.

The Naturaliste Plateau is one of the marginal plateaus of western Australia separated from Australia by a deep trough (3,000 m) and extending westward about 300 km (Figure 3). This plateau has morphology very similar to Broken Ridge[11]. Sites 258 and 264 were drilled by D/V Glomar Challenger on this plateau. Neither hole penetrated to basement, but both reached Cretaceous sediments, and the second hole bottomed in volcaniclastic conglomerates containing acid and basic volcanic fragments. Earlier seismic profile and piston core data prompted a number of authors to propose that the plateau has continental crust. Apparently, the same Cretaceous-Tertiary unconformity present on Broken Ridge is also found on the Naturaliste Plateau[11].

Two remarkable features about the Ninetyeast Ridge are certainly its length (about 5,000 km) and its very linear form (Figure 3). Schlich[11] reviews the proposed origins. Following the Deep Sea Drilling Program, it is now generally agreed that the ridge is underlain by basalt and that much of the ridge at one time was at shallow depths as evidenced by the discovery of subaerial conglomerates, glauconite, and lignite on the ridge during the drilling. Also, there is general agreement that the ridge is older toward the north. The basalts on the ridge have chemical affinities with basalts of nearby oceanic islands, in particular, Saint Paul-Amsterdam and Kerguelen. This fact coupled with paleomagnetic data[25], which suggests a northward movement of the ridge some 5,000 km since late Cretaceous, has led several authors

to propose that the ridge is the result of a hot spot from either the Saint Paul-Amsterdam location or one in the Kerguelen area, or both. The latest chemical studies[21,22] favor Kerguelen. The oldest identified magnetic anomaly parallel to the northeastern flank of the Kerguelen Plateau is late Eocene age[26,27]. These same anomalies are the oldest identified south of Broken Ridge. As Schlich[11] points out, it is this symmetry, in addition to the overall geometry, which has led some investigators to propose that the two (and the Naturaliste Plateau) were once joined then separated in the Paleocene or early Eocene when spreading between Australia and Antartica began (Figure 10).

On the basis of new stratigraphic and paleontological data from northern Victoria Land, Laird, et al.,[28] have suggested a modification of the palemagnetic fit between Australia and Antartica. In turn, Norton and Molnar[29] have modified that fit slightly to better fit the magnetic data and fracture zone trends. Their fit shows a large overlap of Kerguelen and Broken Ridges which they believe is due to the fact that the southern part of the Kerguelen Plateau, the Broken Ridge and the Naturaliste Plateau are remnants of an older ridge, since split apart.

In spite of all the work done in the Indian Ocean, the question of whether or not there is continental crust under all or portions of the Kerguelen Plateau, Broken Ridge or the Naturaliste Plateau has not been answered. The odd magnetic anomaly on Broken Ridge (900 km across and offset from a gravity anomaly by several hundreds of kilometers) indicates unusual, possibly continental crust. Drilling on both Naturaliste Plateau and Broken Ridge and a single core from the Kerguelen Plateau south of 55° indicates that there are rocks at least as old as Cretaceous on all of these features. Courtillot and Vink[30] have pointed out that continents

should break up by stretching and fragmenting; not along a sharp line. If their model is correct, there ought to be fragmental material left behind during the breakup of Antartica and Australia. Perhaps Lazarewicz and Sailor[1,2] have found one fragment on the eastern end of Broken Ridge. The other two structures (Kerguelen and Naturaliste) may, in part, be continental. The alternative is that they are all part of an old oceanic ridge formed in the initial stages of breakup of India- Australia- Antartica.

There does not seem to be any geologic feature to explain the linear anomaly seen on the J/p map (Figure 7). Several have noticed that this anomaly follows the zero contour of the vertical derivative of gravity (Figure 6) but what significance this has is unknown. The J/p anomaly on Broken Ridge is probably real. If susceptibility values were available for rocks on the ridge, it might be possible to make a good educated guess at the rock types underlying the ridge from the known magnetic and gravity anomalies using the method of Garland[4].

IV. THE CARIBBEAN:

The geology and geophysics of the Caribbean region have been summarized in several publications[10,31,32]. There is an abundance of seismic, drill hole, dredge and core data on the rock types and structure of the region which make it an ideal modeling area for geophysical studies either from the surface or from satellite measurements.

The gravity of the area from surface measurements has been compiled by Bowin[32,33,34]. These data are available in digital form on a 15' grid. Additional gravity data are available through NASA and DMA.

The magnetic data from surface measurements are compiled in a total intensity map[35], but magnetic anomaly maps have been compiled for only parts of the Caribbean[32,36,37,38]. The surface data have not been

digitized. Currently, we are analyzing Magsat data for the region.

V. SELECTED SITES WITH WELL DEFINED GRAVITY AND MAGNETIC ANOMALIES:

Three sites have been found during this study which might be used to calculate probable rock types causing the anomalies[4]. Broken Ridge is the first site. The eastern end of this feature has a positive 100 mg free air anomaly offset from a 140 gamma magnetic anomaly. The second is Deep Sea Drilling site 214 on the Ninetyeast Ridge where there are basement basaltic rocks of Cretaceous age, a positive 70 mg gravity anomaly and a negative magnetic anomaly of 400 gammas (Figures 11, 12, 13, 14). The third is the Los Roques Trough in the Caribbean at 12.5°N, 67°W (Figures 15, 16) underlain by at least 5 km of sediments on probable basaltic crust. Here, there is a negative 250 mg free air gravity anomaly accompanied by a negative 250 gamma magnetic anomaly.

VI. RECOMMENDATIONS:

This study will result in a proposal for work on deep sea cores from the Kerguelen Plateau, if further discussion turns out to be as encouraging as the preliminary inquiries. The author plans a visit to Florida State University in the fall where the cores are in storage. Catalogue descriptions of these cores, which have not been worked on in detail, are promising in that it is noted that they contain rock fragments and volcanic ash, as well as fossils which could give age, depth, and paleogeographic information. This proposed study, when combined with a magnetic and gravity analysis of the Kerguelen Plateau proposed by Lazarewicz (pers comm), could be definitive works on the Broken Ridge-Kerguelen relationship.

Three sites have been selected for combined analysis of gravity and magnetic anomalies. There is a need to develop the mathematics

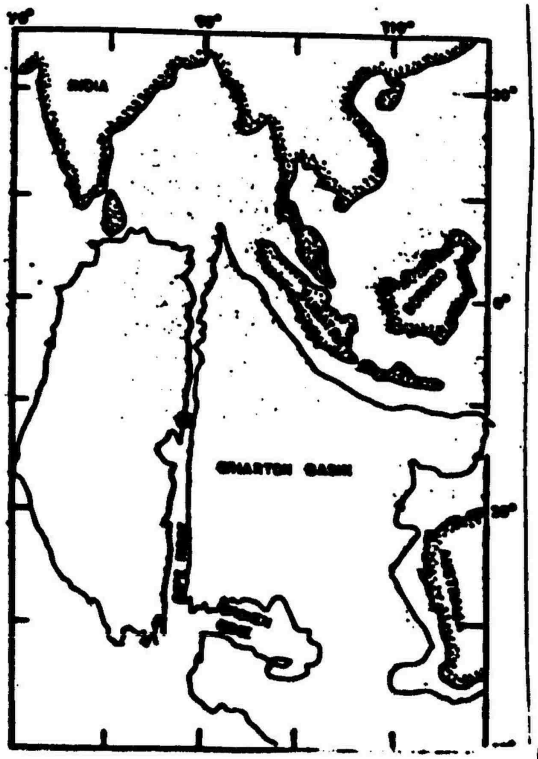


Figure 11. Location of site 214 survey (from Naini, et al., [40]).

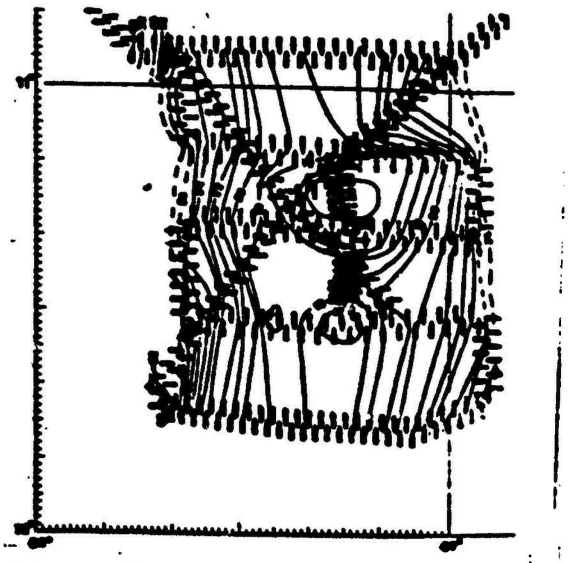


Figure 12. Bathymetry in site 214 area. (from Naini, et al., [40]).

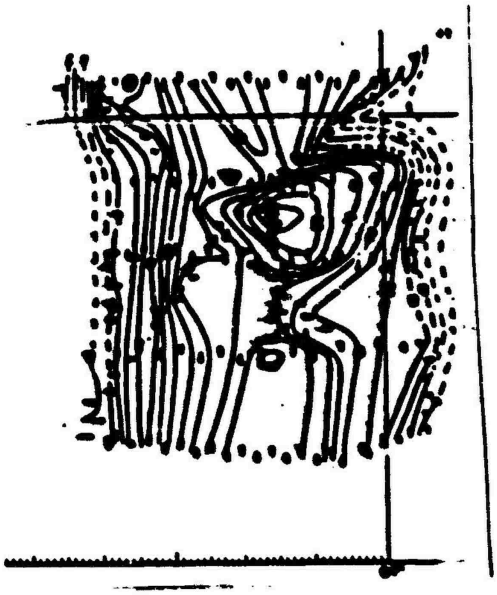


Figure 13. Free-air anomaly in milligals, site 214 area. (from Naini, et al., [40]).

Figure 14. Residual magnetic anomalies at 50 gamma intervals, site 214. (from Naini, et al., [40]).

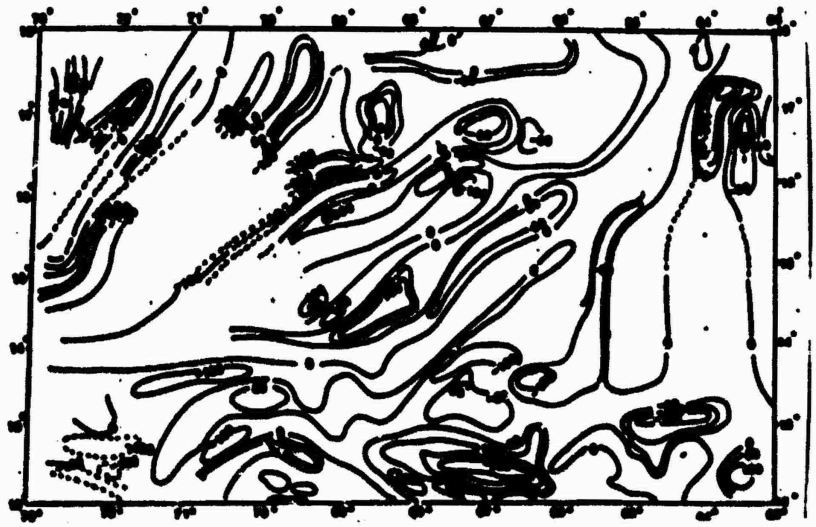


Figure 15. Residual magnetic anomaly map of the Venezuelan Basin, Caribbean. Contour interval 50 nT. Reliable contours in solid lines, less reliable in broken lines (from Hall [61]).

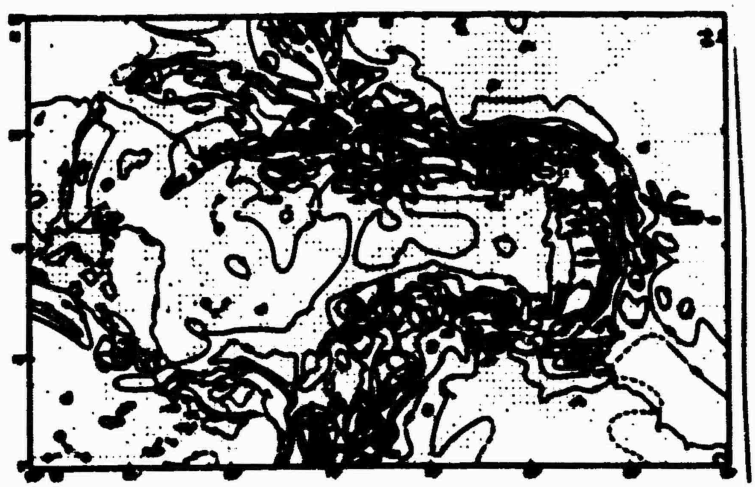


Figure 16. Averaged free air gravity anomaly map (20 min x 20 min) of the Caribbean (from Bowin [32]).

for the procedure suggested by Garland[4] to determine the rock types causing these anomalies. It might be possible to determine at least which rocks are not causing the anomaly under Broken Ridge.

When the difficulties with our processing of Magsat data are unravelled, we will have both magnetic and gravity data in digitized format for the Caribbean. At that time, the original suggestion for the compilation of a J/p map of the Caribbean will be carried out, so that, if there are significant signatures over the known crustal types there, the knowledge of their cause could then be applied to data already collected from the southeastern Indian Ocean and to that which will be analyzed from the Kerguelen Plateau.

Our ability to detect crustal type from analysis of combined satellite gravity and magnetic anomalies, or the ability to predict the anomalies which would be produced from a particular crustal type, could have economic value if, for example, these anomalies are produced from ore deposits or petroleum basins.

REFERENCES

1. Sailor, R.V. and Lazarewicz, A.R., Magsat Investigation of Crustal Magnetic Anomalies in the Eastern Indian Ocean, unpublished report prepared for NASA, Goddard Space Flight Center, Greenbelt, Maryland, 1983.

2. Lazarewicz, A.R. and Sailor, R.V., Two-Dimensional Analysis of Geophysical Data in the Eastern Indian Ocean, in preparation, 1983.

3. Nettleton, L.L. and Elkins, T.A., "Association of Magnetic and Density Contrasts with Igneous Rock Classifications," Geophysics, Vol 9, pp. 60-78, 1944.

4. Garland, G.D., "Combined Analysis of Gravity and Magnetic Anomalies," Geophysics, Vol 16, pp. 51-62, 1951.

5. Kanasewich, E.R. and Agarwal, R.G., "Analysis of Combined Gravity and Magnetic Fields in Wave Number Domain," J. Geophys. Res., Vol 75, pp. 5702-5712, 1970.

6. Cordell, L. and Taylor, P.T., "Investigation of Magnetization and Density of a North Atlantic Seamount Using Poisson's Theorem," Geophysics, Vol 36, pp. 919-937, 1971.

7. Nagle, F., "Blueschist, Eclogite, Paired Metamorphic Belts and the Early Tectonic History of Hispaniola," Geol. Soc. Am. Bull., Vol 85, pp. 1461-1466, 1974.

8. Nagle, F., Stipp, J.J. and Fisher, D.E., "K-A Geochronology of the Limestone Caribbees and Martinique, Lesser Antilles, West Indies," Earth Planetary Science Letters, Vol 29, pp. 401-412, 1976.

9. Nagle, F., "Rocks from Seamounts and Escarpments on the Aves Ridge," Carib. Geol. Conference Trans., No. 6, pp. 409-413, 1972.

10. Nagle, F., "Caribbean Geology, 1970," Bulletin Marine Science, Vol 21, pp. 375-439, 1971.

11. Schlich, R., "The Indian Ocean: Aseismic Ridges, Spreading Centers, and Oceanic Basins," Chapter 2 in Nairn, A.E.M. and Stehli, F.A., The Ocean Basins and Margins, Vol 6, The Indian Ocean, Plenum Press, New York, pp. 51-147, 1982.

12. Nairn, A.E.M. and Stehli, F.G. (eds.), The Ocean Basins and Margins, Vol 6, The Indian Ocean, Plenum Press, New York, p. 776, 1982.

13. Initial Reports of the Deep Sea Drilling Project, Vol 22, Washington, D.C., U.S. Gov't Printing Office, 1974.

14. Initial Reports of the Deep Sea Drilling Project, Vol 26, Washington, D.C., U.S. Gov't Printing Office, 1974.

15. Initial Reports of the Deep Sea Drilling Project, Vol 28, Washington, D.C., U.S. Gov't Printing Office, 1974.

16. Heirtzler, J.R., Bolli, H.M., Davies, T.A., Saunders, J.B. and Sclater, J.G., eds., Indian Ocean Geology and Biostratigraphy: Studies Following Deep Sea Drilling Legs 22-29, Washington, D.C., American Geophysical Union, p. 616, 1977.

17. Francis, T.J.G. and Raitt, R.W., "Seismic Refraction Measurements in the Southern Indian Ocean," J. Geophys. Res., Vol 72, pp. 3015-3042, 1967.

18. Davies, T.A., Luyendyk, B.P., Rololfo, K.S., Kempe, D.R., McKelvey, B.C., Leidy, R.D., Horvath, G.J., Hyndman, R.D., Theirstein, H.R., Herb, R.C., Boltorskoy, E., and Doyle, P., "Site 255" in: Initial Reports of the Deep Sea Drilling Project, Vol 26, pp. 281-294, 1974.

19. Watkins, N.D., Gunn, B.M., Novgier, J. and Baksi, A.K., "Kerguelen: Continental Fragment or Oceanic Island?", Geol. Soc. Am. Bull., Vol 85, pp. 201-212, 1974.

20. Dosso, L., Vidal, P., Cantagrel, J.M., Lamegre, J., Marot, A. and Zimine, S., "Kerguelen: Continental Fragment or Oceanic Island?: Petrology and Isotopic Geochemistry Evidence," Earth Planetary Science Letters, Vol 43, pp. 46-60, 1979.

21. Dosso, L. and Murthy, V.R., "A Nd Isotopic Study of the Kerguelen Islands: Inferences on Enriched Oceanic Mantle," Earth Planetary Science Letters, Vol 48, pp. 268-276, 1980.

22. Mahoney, J.J., Macdougall, J.D., Lugmair, G.W. and Gopalan, K., "Kerguelen Hot Spot Source for Rajmahol Traps and Ninetyeast Ridge?", Nature, Vol 303, pp. 385-389, 1983.

23. Quilty, P.G., "Cenomanian-Turonian and Neogene Sediments from Northeast of Kerguelen Ridge, Indian Ocean," J. Geol. Soc. Australia, Vol 20, pp. 361-371, 1973.

24. Udintsev, A.B., Fisher, R.L., Kanaer, V.F., Laughton, A.S., Simpson, E.S.W. and Zhiv, D.I., Geological-Geophysical Atlas of the Indian Ocean, Academy of Sciences of the USSR, Moscow, p. 151, 1975.

25. Pierce, J.W., "The Northward Motion of India Since the Late Cretaceous," Geophys. J.R. Atr. Soc., Vol 52, pp. 277-311, 1978.

26. LePichon, X. and Heirtzler, J.R., "Magnetic Anomalies in the Indian Ocean and Sea Floor Spreading," J. Geophys. Res., Vol 73, pp. 2101-2117, 1968.

27. Houtz, R.E., Hayes, D.E. and Markl, R.G., "Kerguelen Plateau Bathymetry, Sediment Distribution and Crustal Structure," Marine Geology, Vol 25, pp. 95-130, 1977.

28. Laird, M.G., Cooper, R.A. and Jago, J.B., "New Data on the Lower Palaeozoic Sequence of Northern Victoria Land, Antartica, and its Significance for Australian-Antartic Relations in the Palaeozoic," Nature, Vol 265, pp. 107-110, 1977.

29. Norton, I. and Molnar, P., "Implications of a Revised Fit between Australia and Antartica for the Evolution of the Eastern Indian Ocean," Nature, Vol 267, pp. 338-340, 1977.

30. Courtillot, V. and Vink, G.E., "How Continents Break Up," Scientific American, Vol 249, pp. 43-47, July 1983.

31. Nairn, A.E.M. and Stehli, F.G., The Ocean Basins and Margins, Vol 3, The Gulf of Mexico and the Caribbean, Plenum Press, New York, 1975.

32. Bowin, C., "Caribbean Gravity Field and Plate Tectonics," Geol. Soc. Am. Spec. Paper 169, p. 79, 1976.

33. Bowin, C., Warsi, W. and Milligan, J., Free Air Gravity Map of the World, Geol. Soc. Am. Map and Chart Series, No. MC-45, 1982.

34. Bowin, C., Warsi, W. and Milligan, J., Free Air Gravity Anomaly Atlas of the World, Geol. Soc. Am. Map and Chart Series, No. MC-46, 1983.

35. Matthews, J.E., Map of Total Intensity Geomagnetic Field of the Caribbean Region, U.S. Geol. Survey Map MF-742, 1976.

36. Christofferson, E., Linear Magnetic Anomalies in the Colombian Basin, Control Caribbean Sea, Ph.D. dissert., Kingston Univ., Rhode Island, p. 47, 1973.

37. Hall, S., Residual Anomaly Map of the Venezuelan Basin, unpublished, University of Houston, Texas.

38. Bracey, D.R., "Structural Implications of Magnetic Anomalies North of the Bahama-Antilles Islands," Geophysics, Vol 33, pp. 950-961, 1968.

39. Molnar, P. and Sykes, L.R., "Tectonics of the Caribbean and Middle America Regions from Focal Mechanisms and Seismicity," Bull. Geol. Soc. Am., Vol 50, pp. 1639-1684, 1969.

40. Nairn, B.R. and Eittreim, S.L., "Appendix II Marine Geophysical Survey (Site 214, DSDP) on the Ninetyeast Ridge, Indian Ocean," in Initial Report of the Deep Sea Drilling Project, Vol 22, pp. 843-846, 1974.

41. Heezen, B.C., Lynde, P.L. and Fornari, [illegible] J., Geological Map of the Indian Ocean, Geological World Atlas, UNESCO, place de Fontenoy, 75 Paris-7E, France, 1977.

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

USING ARTIFICIAL INTELLIGENCE IN AVIONIC FAULT ISOLATION

Prepared by:	Dr. Philip D. Olivier
Academic Rank:	Assistant Professor
Department and University:	Department of Electrical and Computer Engineering Louisiana State University
Research Location:	Rome Air Development Center, Reliability and Compatibility Division, Systems Reliability and Engineering Branch, Reliability and Maintainability Engineering Techniques Section
USAF Research	Mr. Dale W. Richards
Date:	August 20, 1983
Contract No.:	F49620-82-C-0035

USING ARTIFICIAL INTELLIGENCE IN AVIONIC FAULT ISOLATION

BY

PHILIP D. OLIVIER

ABSTRACT

The properties of any procedure for isolating faults in avionic equipment are investigated. So that the procedure is suitable for use with artificial intelligence techniques. The application of one such procedure to one device is described. Suggestions for further research are offered.

ACKNOWLEDGEMENT

The author would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research, and the Southeastern Center for Electrical Engineering Education for providing him with the opportunity to study this problem at the Rome Air Development Center, Griffiss AFB, New York. He would like to thank the center, especially the Reliability and Maintainability Engineering Techniques Section for its hospitality.

Finally, he would like to thank Anthony Coppola, Jerome Klion and especially Dale Richards for many interesting and informative discussions.

I. INTRODUCTION

The efficient and accurate diagnosis of faults in avionic equipment is of immense importance to the Air Force. Improvements in this area yield direct and tangible increases in the readiness rate of the fleet. Current techniques are not satisfactory as evidenced by the very large number of unnecessary and/or ineffective maintenance actions that are done.[1] As avionic systems have increased in complexity, these problems have been magnified. Since the electronics industry appears to be introducing more powerful and more sophisticated single chip devices at an accelerating rate the problems will only become bigger. It is necessary to take a critical look at the approach historically taken to diagnose faults in digital avionics and based on this critique describe the attributes of a more appropriate solution that will be applicable to a wide range of devices.

The classical approach is based on circuit tracing techniques whereby an input sequence is engineered so that all faults in a certain subset of "allowed" faults can be distinguished through the use of a prestored fault dictionary. When a Unit Under Test (UUT) is brought into the depot for testing it is exercised by this engineered input, the output is measured and compared with those in the fault dictionary. If the measured output matches one of the fault signatures the job is done. If no match is found the

procedure yields no information about the location of the fault. Further, as systems get larger the number of allowable faults increase, the length of the test input increases, the more time it takes to engineer an appropriate input, the length of the fault signatures that must be stored increase and the time to search this extensive data base as well as the space required to store it become prohibitive. The reason for these problems (other than the size of the circuits) is the basic inflexibility of the test procedure. The entire technique is designed a priori and hence uses the measured test data in a very superficial manner.

The purpose of this research effort is to develop a fault isolation procedure that will be more flexible and thereby, it is hoped, yield some information all of the time. Once flexibility is raised to the level of a design objective it is natural to phrase the problem in terms of the most flexible of man's intellectual constructs, Artificial Intelligence (AI). The discussion in the previous paragraph, however, indicate that a whole new approach to data gathering is necessary before AI techniques can be applied. On the other hand, existing applications of AI to fault isolation (e.g. medical, nuclear reactor systems), rely on a small number (<5,000) of "IF-THEN" rules that are obtained from experienced field engineers. As the first paragraph mentioned, electronics as a whole is progressing far to rapidly to allow ourselves the luxury of developing human experts. We must rely on the mathematical description of our circuits for fault information. Hence, the application of AI to avionic fault isolation is nontrivial and cannot be achieved by simply borrowing the techniques of other fields.

Because of the extreme complexity of todays digital circuitry the fault isolation problem is taking on many of the attributes of the analog

fault isolation problems, i.e., fault dictionaries are practically, if not truly, infinite. Also as the circuit size increases the "single fault" assumption becomes more and more academic. Once multiple faults and even intermittent faults are targets for isolation the size of the fault dictionary becomes truly infinite.

Techniques that can be applied to both analog and digital circuits should be highly prized because most circuits have some sort of power supply that acts as the interface between the two worlds. Also the truly hybrid circuits currently in use are on the most expensive printed circuit boards ($5,000 each).

It is the author's opinion that AI techniques can only be applied to avionic fault isolation within the context of a given fault isolation technique. Given the inflexibility of the current technique the first order of business is to develop a new way to isolate faults in digital circuits. Section three discusses one such new procedure that: a) uses a very compact fault dictionary (compact in the sense of the number of entries as well as the size of each entry), b) can yield fault isolation information even when a match is not found, c) can use inputs other than the one that produced the fault dictionary to aid in th: fault isolation process, d) can be used within an "expert" system, and, e) is related to an analog fault isolation procedure which should allow for the treatment of hybrid boards.

II. OBJECTIVES:

The objectives of this project are twofold. First, a fault diagnosis procedure that is inherently flexible was sought. Second, ways to integrate Artificial Intelligence into this fault diagnosis procedure were to be investigated. Only the first has been accomplished. Comments regarding the second can be found in the section on recommendations.

To obtain a flexible procedure, the a priori simulation-before-test approach was abandoned. A simulation-after-test procedure that was developed for analog circuits[2] was chosen as a likely starting point. The first goal was to develop a computer program that would 1) simulate a digital circuit, 2) allow for the insertion of faults, and 3) compare the expected output (from the unfaulted circuit) to the test output (from the faulted circuit).

In order to guarantee maximum flexibility many inputs must be allowed. This led to choosing a technique for generating these inputs based on maximal sequences from linear shift registers[3, 4] (these are described later).

Such a program was developed for two sample circuits. The procedure's ability to discriminate between faults was investigated. Based on this experience it was decided that a small fault dictionary could be very useful in the early steps. So the next goal was to decide on a fault dictionary that would have minimal storage requirements and (instead of being exhaustive) that would direct the program toward the fault. Techniques for building such a dictionary need to be investigated.

Now that a fault isolation procedure has been "tailored" to AI it is necessary to tailor AI to this procedure. This can be done in several ways. First, if the fault is not in the dictionary then making inferences based on a comparison of the measured fault syndrome to the fault signatures should narrow the ambiguity set. Once the ambiguity set is narrowed use of the history of the device along with reliability information about the still suspect components can be used to prioritize a search. AI techniques can then be used to "engineer" an input that will test a certain suspect component.

III. A FAULT DETECTION PROCEDURE

The fault detection procedure used is a variation of one due to Wu et al (2). It has much more fault isolation capability than some other techniques because it uses the test data in a more sophisticated manner. It is based on a description of the UUT in terms of the component connection model.[5] The component connection model for a UUT consists of two parts.

The first part consists of the component equations

$$b_i = f_i(a_i) \qquad i = 1, \ldots n \tag{1a}$$

or

$$b = f(a) \tag{1b}$$

where a_i, b_i and f_i are the input vector sequence, output vector sequence and functional description of machine i and $b = \text{col}\,(b_i)$, $a = \text{col}\,(a_i)$ and f are defined by

$$b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \qquad a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}, \qquad f(a) = \begin{bmatrix} f_1(a_1) \\ f_2(a_2) \\ \vdots \\ f_n(a_n) \end{bmatrix} \tag{1c}$$

The second part consists of the connection equations which are

$$a = L_{11}b + L_{12}u \tag{2a}$$

$$y = L_{21}b + L_{22}u \tag{2b}$$

where u and y are the system input and output vector sequences and the L_{ij} are matrices of zeros and ones that describe the interconnection structure of the system.

A test is defined to be the application of an input u and the measurement of the resulting output y. Now, if u and y are both known (i.e. measured) they can be used as inputs to the fault isolation procedure, and it stands to reason that they should be processed in some way that is determined by the UUT to produce information concerning the location of the faults.

To process the test results the components are partitioned into two groups (renumbering the components where necessary) so that the CCM can be rewritten in expanded from as

$$\underline{b}^1 = f^1(\underline{a}^1),\ \underline{b}^2 = f^2(\underline{a}^2) \qquad (3a)$$

$$\underline{a}^1 = L_{11}^{12}\underline{b}^1 + L_{11}^{12}\,\underline{b}^2 + L_{12}^{1}\,\underline{u} \qquad (3b)$$

$$\underline{a}^1 = L_{11}^{21}\,\underline{b}^1 + L_{11}^{12}\,\underline{b}^2 + L_{12}^{2}\,\underline{u} \qquad (3c)$$

where the _ emphasizes that these variables are generated during a test. Equation (3b) can now be solved for b^2 (provided L_{21}^2 has a left inverse (and this is a precondition on the choice of the partition) and this b^2 substituted in (3b) and (3c) to yield. (See figure 1)

$$\underline{a}^1 = K_{11}\underline{b}^1 + K_{12}\begin{bmatrix}\underline{u}\\ \underline{y}\end{bmatrix}$$

$$\begin{bmatrix}\underline{a}^2\\ \underline{b}^2\end{bmatrix} = K_{21}\underline{b}^1 + K_{22}\begin{bmatrix}\underline{u}\\ \underline{y}\end{bmatrix}$$

which together with $\underline{b}^1 = f^1(\underline{a}^1)$ forms the CCM fpr a "pseudo circuit" with input col($\underline{u}$, $\underline{y}$) and output col($\underline{a}^2$, $\underline{b}^2$). This measured $\underline{b}^2$ can now be compared with the expected outputs obtained from $\underline{b}^2 = f^2(\underline{a}^2)$.

Advantages of this procedure include:

(1) The ability to handle several input

(2) The ability to tailor an input to a suspected fault;

(3) The ability to further verify your conclusion as to which component is faulty;

(4) A relatively small data base

(5) The ability to find fualts that do not have prestored signatures;

(6) There is no need to pre-engineer the input.

Before continuing it is appropriate to make a comment about the differences between this procedure applied to digital circuits and it applied to analog circuits. For analog circuits there are usually many partitions that yield loft invertible L_{21}^{2} 's. This is because the outputs of several devices are added together before they are fed into another device. Such linear "fan in" situations are not as common in digital circuits.

IV. Generating Inputs

The procedure described above is capable of using arbitrary inputs. In fact the procedure is capable of using field inputs therefore we felt it unwise to "design" an input for the test. To obtain an input that is independent of the fault we were searching for, we used a sequence that was generated by a linear sequential shift register with feedback. These sequential machines are shown in Figure 2.

Each of the n registers initially contains a 1. At each clock change the contents of register $i \geq 2$ is replaced by the contents of register i-1. The contents of register 1 is replaced by the binary sum of registers n and j. By judicious choice of j and n you can obtain a sequence of length 2^n-1 that contains no repeating subsequences. Such sequences are called maximal. The generation of maximal sequences for a given n sometimes requires feedback from multiple registers.

It is reasonable to separate the inputs into two types-data and instruction. The instruction inputs should mimic, at least statistically, the actual operating conditions. To be resonable we cycled through the different instructions, switching instructions every ten clock cycles. We also ran the program when the instructions were selected "randomly" by the maximal shift register sequence.

It is important to note that if the sequence $b^2-\underline{b}^2 = d = 0$ then no fault has been detected. If the input can be "engineered" then it is possible to choose an input that will detect any fault (except in pathological cases).

V. Fault Signatures

A significant problem is finding short signatures that can give direction in further fault isolation. The goal is not to uniquely isolate the fault based on its signature, but rather, to limit the number of faults that must be further investigated.

This fault isolation procedure can be implemented in two extremely different ways or in any compromise between the extremes. It can be used to construct a very efficient fault dictionary and thereby yield a knowledgeable, but dumb, procedure. On the other hand it can be implemented in such a way as to use no fault dictionary and thereby yield an ignorant, but intelligent, procedure. The second procedure is the one sought but the first is more completely developed.

For the sample circuit that is a modified Am 2914 chip it was possible to construct a fault dictionary that consisted of 85 vectors each with eight components that isolated the fault to ambiguity sets of at most 3 machines. This was with an input that cycled through the instruction set (each instruction "on" for 10 clock cycles), but otherwise left the input unengineered. The ambiguities that remained could be removed by a few further simulations with different inputs (at most 11 additional simulations).

The intelligent procedure would keep track of the discrepancies for a certain number of clock cycles after the first discrepancy is noted. The fault syndrome would look like (again for the modified Am 2914)

$$(C_{11}, C_{12}, C_{21}, C_{22}, C_{31}, C_{32}, C_{41}, C_{42})$$

where C_{11} notes the clock cycle of the first discrepancy for machine i and C_{12} notes the clock cycle of the second discrepancy for machine i. To determine where a fault occurred it is necessary to note that a discrepancy

in machine i could occur because machine i contains the fault or because a machine "upstream" of machine i contains the fault. This observation along with the fact that the procedure uses measured outputs as inputs leads to two rules for the construction of an ambiguity set.

Rule 1: All machines "upstream" from any machine that shows a discrepancy is in the ambiguity set. This "upstream chain" is broken at any measured output.

Rule 2: All machines that show no discrepancies are, in fact, good.

To priortize further simulations those machines that are upstream from 2 or more output machines that have shown discrepancies are more suspect than others. Reliability information and maintenance history can be used to further prioritize the order of further simulations. These further simulations can be carried out under any number of single or multiple fault assumptions.

Since this intelligent, but ignorant, procedure uses no fault dictionary it can be implemented in Built-in-Test Equipment (BITE) to work on in-flight data. The processing may take longer than is customary, but it does not require taking the UUT "off line" for periodic checks.

For depot implementaiton a judicious mixture or compromise between the two approaches is obviously the preferred approach.

As an example of this procedure without any fault dictionary, consider the fault syndrome (0, 0, 30, 70, 70, 270, 0, 0). This indicates that the outputs from machines 10 and 9 were always the expected outputs, whereas the outputs from Machines 8 and 6 differed from expectations (between the 31st and 30th and again between 61st and 70th clock cycle for machine 8 and between the 61st and 70th and again between the 261st and 270th for machine 6). If we make the working assumption that only one machine is faulted,

then the fult must lie upstream from both machine 6 and 8. Further, under the single faulted machine hypothesis, machines 6, 8, 9 and 10 are assured to be good. Finally, since the first discrepancy occurred relatively early (between clock cycle 21 and 30) it should be physically close to machine 8. Without additional information this syndrome indicates that simulations should be done in the following order - first machine 3 then either 4 or 5 followed by 7 then 11 or 2 and finally 1. This syndrome was produced by the first input of machine 3 stuck at 0.

Problems occur with this procedure when a gate is stuck at the value that it almost always assumes. In these situations the discrepancies can show up in very late clock cycles but be caused by a physically close fault. In a depot situation these anamolies could be taken care of by a very small dictionary in which signatures of only these exceptions are stored.

VI. Conclusions.

A fault isolation procedure for digital circuit has been developed that is not totally dependent on a fault dictionary. The procedure can be integrated within an "expert", "knowledge based" or "intelligent" fault diagnosis procedure because of its inherent flexibility. It can also be incorporated into a built-in-test environment because it can analyze in-flight data as easily as depot generated data. Because of the flexibility of the procedure it can tailored to different capacity computers, hence the same basic procedure can be used by many installations resulting in overhead economies.

Another important fact is that once a component is declared faculty it is possible, because of the built in capacity to simulate the UUT, to verify that it was indeed the culprit.

Finally, any truly useful diagnostic tool should be easily transported not only from base to base but from plane to plane on the flight line. This requires a relatively small and powerful computer. When comparing different tools preference should be given to the one which is physically smaller.

VII. Recommendations.

Several areas still need to investigated. These include:

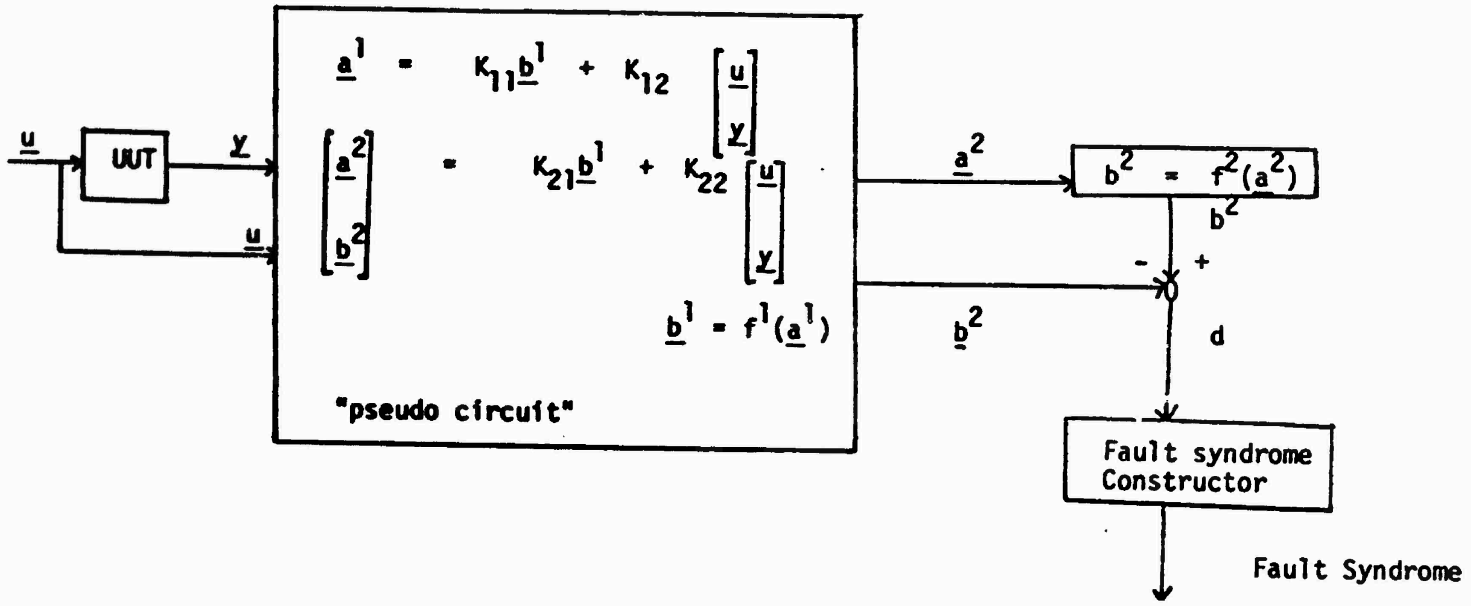
1. Precisely how this procedure can be integrated with artificial intelligence techniques;

2. What are the best fault signatures to minimize storage requirements and maximize information?

3. Can "statistical" signatures be developed for use on in-flight data? By this I mean, is it possible to construct a signature based on a Monte-Carlo simulation of outputs due to different inputs for a given fault that will help in analyzing data obtained from BITE?

4. Can this procedure be implemented on a microprocessor for realistic systems?

FIGURE 1:


u
UUT
y
u
$\underline{a}^1 = K_{11}\underline{b}^1 + K_{12}\begin{bmatrix}\underline{u}\\ \underline{y}\end{bmatrix}$
$\begin{bmatrix}\underline{a}^2\\ \underline{b}^2\end{bmatrix} = K_{21}\underline{b}^1 + K_{22}\begin{bmatrix}\underline{u}\\ \underline{y}\end{bmatrix}$
$\underline{b}^1 = f^1(\underline{a}^1)$
"pseudo circuit"
$\underline{a}^2$
$\underline{b}^2$
$b^2 = f^2(a^2)$
b^2
-
+
d
Fault syndrome Constructor
Fault Syndrome

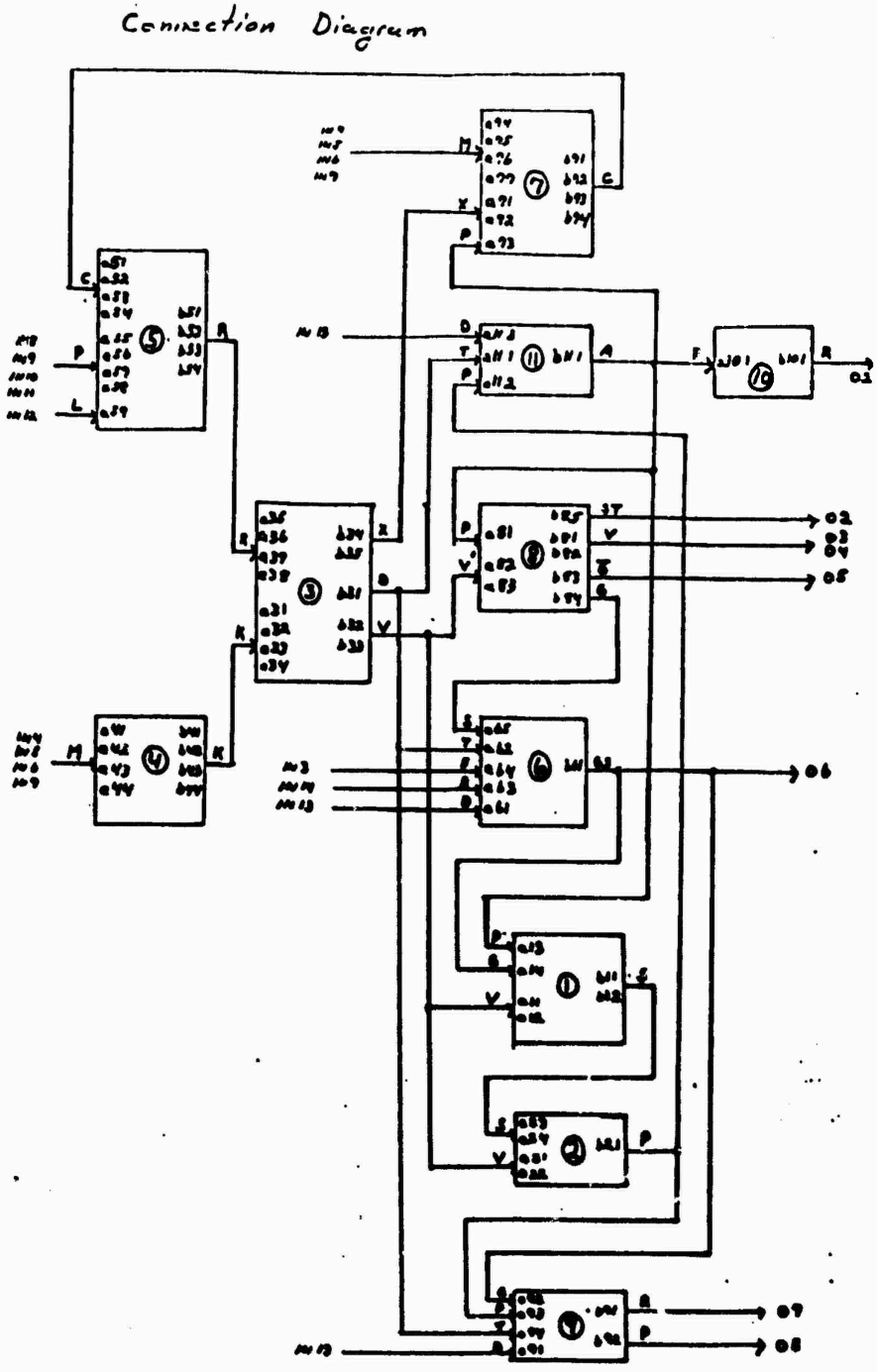


Figure 3a: Modified AM2914

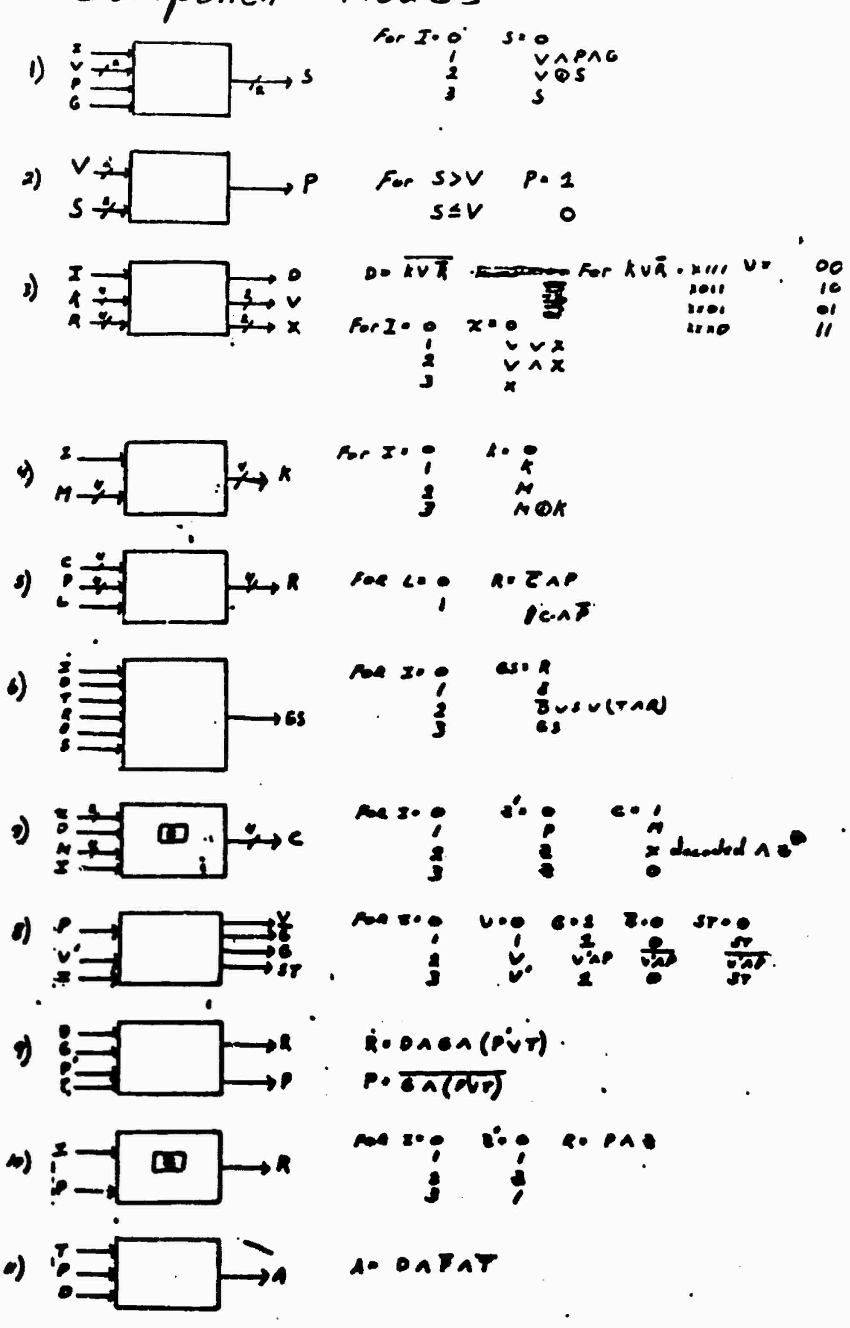
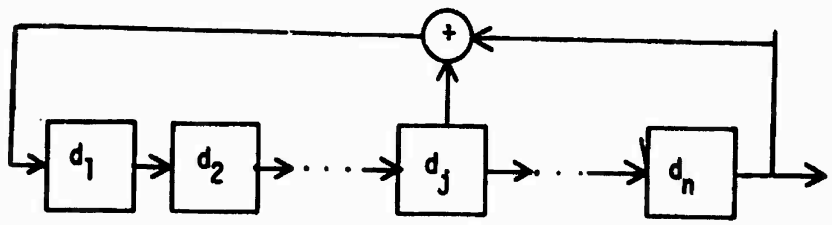


Figure 3b: Modified AM2914



$$d_1' = d_j \oplus d_n$$

$$d_k' = d_{k-1} \qquad k = 2,\ldots,n$$

Initially $d_i = 1 \qquad i = 1,\ldots,n$

Figure 2: Linear Shift Register

REFERENCES

1. Anthony Coppola, AI Application to Maintenance - A report for the OSD/IDA R&M Study, April 15, 1983

2. C-C Wa, K. Nakajima, C-L Wey, and R. Salks, "Analog Fault Diagnosis with Failure Boards", IEEE CAS-29, No. 5, May 1982, pp 277-284.

3. R. G. Brown, Introduction to Random Signal Anlysis and Kalman Filtering, John Wiley & Sons, New York 1983, pp 101-102.

4. R. C. Dixon, Spread Spectrum Systems, New York, Wiley 1976, Table 3.6.

5. R. A. DeCarlo & R. Saeks, Interconnected Dynamical Systems, Mew York: Marcel Dekker, 1981.

ABSTRACT

ORGANIC REACTIONS IN ROOM TEMPERATURE CHLOROALUMINATE MOLTEN SALTS

Albert L. Payton
Hampton Institute

Mixtures of some organic chloride salts and aluminum chloride form new salts that are liquid below room temperature. The mixture of 1-methyl-3-ethylimidazolium chloride (MeEtImCl) and aluminum chloride ($AlCl_3$) is one such ionic liquid. It is liquid substantially below room temperature, has good thermal stability at high temperature, has a low vapor pressure, and dissolves both organic and inorganic compounds. Earlier work at the Frank J. Seiler Research Laboratory showed that some of the MeEtIm chloroaluminate melts promoted Friedel-Crafts reactions. The Lewis acidity of the melts presumably catalyzed the formation of reactive alkylating and acylating intermediates, which resulted in high yields of alkyland acyl-substituted aromatic compounds. In this project the scope and optimimum conditions for Friedel-Crafts reactions in the MeEtIm chloroaluminate melts were examined. Also, the utility of the reactions was tested by preparing polyalkylated aromatic compounds with large alkyl substituents, which may have some use as synthetic lubricants or hydraulic fluids.

The optimization of reaction conditions involved a study of the effect of temperature, reactant composition, reaction time and melt compostion on the rate of reaction, distribution of products and yields of desired products. Briefly, the best results were obtained by running the reactions at low temperature for longer times to reduce side products. The rate of reaction could also be controlled by adjusting the proportion of $AlCl_3$ in the melt, as long as the $AlCl_3$ was in excess over the MeEtIm. An example of a large alkyl polysubstuted benzene was synthesized by acylating benzene with octanoyl chloride followed by Wolf-Kishner reduction of the carbonyl. This monosubstituted product was used as the starting material for two subsequent acylations and reductions, resulting in a 1,2,4-substituted product. The physical properties of the product were not examined due to time limitations.

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

A METHOD OF SENSING SMALL CHANGES IN THE ANGULAR SEPARATION OF CROSSED LASER AND NEUTRAL PARTICLE BEAMS

Prepared by:	Dr. David J. Pegg
Academic Rank:	Professor
Department and University:	Department of Physics University of Tennessee, Knoxville
Research Location:	Air Force Weapons Laboratory, NTYP Branch
USAF Research Contact:	Lt. Col. Raymond Zazworsky
Date:	September 2, 1983
Contract No:	F49620-83-C-0035

A METHOD OF SENSING SMALL CHANGES IN THE ANGULAR SEPARATION OF CROSSED LASER AND NEUTRAL PARTICLE BEAMS

by

David J. Pegg

ABSTRACT

Small changes in the angular separation of crossed laser and particle beams can be sensed by monitoring the strength of the interaction between the two beams. This monitor signal will be proportional to the population of the upper level of the absorbing transition. The sensitivity of the method depends on the signal-to-noise ratio and width of the monitor signal. Optical and non-optical monitor signals have been considered.

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

A METHOD OF SENSING SMALL CHANGES IN THE ANGULAR SEPARATION OF CROSSED LASER AND NEUTRAL PARTICLE BEAMS

Prepared by:	Dr. David J. Pegg
Academic Rank:	Professor
Department and University:	Department of Physics University of Tennessee, Knoxville
Research Location:	Air Force Weapons Laboratory, NTYP Branch
USAF Research Contact:	Lt. Col. Raymond Zazworsky
Date:	September 2, 1983
Contract No:	F49620-83-C-0035

A METHOD OF SENSING SMALL CHANGES IN THE ANGULAR SEPARATION OF CROSSED LASER AND NEUTRAL PARTICLE BEAMS

by

David J. Pegg

ABSTRACT

Small changes in the angular separation of crossed laser and particle beams can be sensed by monitoring the strength of the interaction between the two beams. This monitor signal will be proportional to the population of the upper level of the absorbing transition. The sensitivity of the method depends on the signal-to-noise ratio and width of the monitor signal. Optical and non-optical monitor signals have been considered.

ACKNOWLEDGEMENT

The author would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for providing him with the opportunity to spend a very worthwhile and interesting summer at the Air Force Weapons Laboratory, Kirtland AFB, N.M. He would like to acknowledge the laboratory, in particular the Neutral Particle Beams branch, for its hospitality and excellent working conditions.

Finally, he would like to thank Lt. Col. Raymond Zazworsky for suggesting this area of research and for his collaboration and guidance, and he would like to acknowledge many helpful disscussions with Capt. Michael Norris.

I. INTRODUCTION:

Intersecting laser and particle beams will resonantly interact when the radiation frequency, as "seen" by the moving particles, matches a transition frequency, ν_0, characteristic of the particle. When this resonant condition occurs, the particles will absorb radiation from the laser field. The Doppler equation relates the frequency of the laser radiation, as seen in the particle frame, ν', to the actual frequency of the laser radiation, as seen in the laboratory frame, ν ;

$$\nu' = \nu \gamma (1 - \beta \cos\theta) \qquad (1)$$

where θ is the angle between the intersecting laser and particle beams, as measured in the laboratory frame, $\beta (= v/c)$ is related to the speed, v, of the beam particles and $\gamma = (1 - \beta^2)^{-1/2}$ is the usual relativistic factor. In the laboratory frame, the resonance condition can therefore be stated as:

$$\nu = \nu_0 / \gamma (1 - \beta \cos\theta) \qquad (2)$$

If the particle beam was monoenergetic and the particle and laser beams were perfectly unidirectional the resonance would occur at a single laser frequency since there is only one velocity class of particles. In reality, of course, the particle beam suffers some energy dispersion and both beams diverge. This implies that there are spreads in the quantities β and θ and so a resonance can be tuned over a finite range of frequencies since the laser can resonate with different particles at different frequencies. Doppler tuning usually refers to changing the frequency of the laser, as seen by the particles, by varying either the particle beam speed, β , or the angular seperation of the two beams, θ . In this report we are interested in the change in the strength of a signal used to monitor the amount of absorption as a function of angle-detuning of a resonance from its

maximum.

Rohringer[1] pointed out that a crossed laser-particle beam arrangement such as that just described could be used to detect small changes in the angular separation of the two beams or, alternatively, to sense small changes in the direction of a particle beam relative to a reference direction established by a laser beam axis. When the angular separation of the two beams is changed slightly from the angle which satisfies the maximum resonance condition for a given beam energy and fixed laboratory frame laser frequency, the corresponding detuning of the absorption resonance in the particle frame is signatured by a decrease in a laboratory frame signal used to monitor the amount of radiation absorbed by the particles. This change in monitor signal could, of course, be used as an error signal in a feedback loop designed to correct the small angular excursion. The sensitivity of this particle beam direction-sensing method depends rather critically on the width of the resonance signal as well as the signal-to-noise ratio of the monitor signal. Rohringer showed that the resonance could be narrowed considerably if the angle between the two intersecting beams at resonance could be maintained at the laboratory frame angle, $\theta = \cos^{-1}\beta$. This "magic" angle corresponds, in the particle frame, to the perpendicular irradiation of the particles leading to the elimination of a significant part of the first-order Doppler broadening contribution to the resonance width. Under this condition the resonance linewidth is determined primarily by the combined divergences of both the particle and laser beams.

The particle beam direction-sensing technique proposed by Rohringer has been further developed by TASC[2]. This group has proposed to use a neutral particle beam consisting of metastable hydrogen atoms in the 2S state which will resonantly absorb photons on the 2S-3P transition when they

pass through a laser field tuned to the appropriate frequency. The resonant absorption process is to be monitored in the laboratory frame by studying the strength of the Ly_β fluorescence signal i.e. radiation spontaneously emitted on the 3P-1S transition. Since the laser radiation is first absorbed and later remitted in this technique it is often referred to as the resonant scattering of laser radiation by metastable H-atoms. Figure 1 is a particle energy level diagram for the H-atom showing the relevant absorption and emission lines.

II. OBJECTIVES

The project involved a theoretical study of a method used to sense small changes in the direction of a beam of fast-moving particles. The method, which has been outlined in Section I, is based on monitoring, in the laboratory frame, a resonant absorption interaction between intersecting laser and particle beams. Changes in the monitor signal reflect changes in the angular separation of the two beams if the particle velocity and the laser frequency in the laboratory frame remain constant.

The principal objectives of this project were:

(1) To analyze the possible homogeneous and inhomogeneous broadening mechanisms that could contribute to the linewidth of the signal used to monitor the absorption resonance between the intersecting laser and particle beams.

(2) To make a critical analysis of the version of the particle beam direction-sensing method proposed by TASC. The proposal involves the use of a hydrogen atom beam prepared in the metastable 2S state. Resonant absorption occurs on the 2S-3P transition and fluorescence monitoring of this resonance is done on the 3P-1S transition. This analysis involved pointing out potential problems, making suggestions

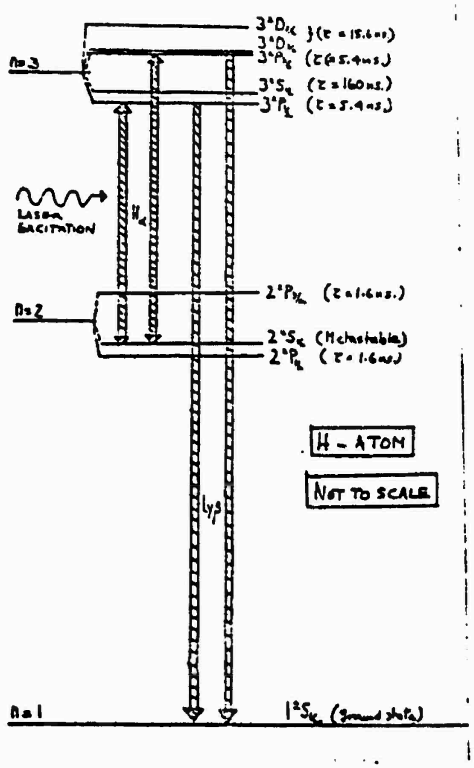


Fig. 1

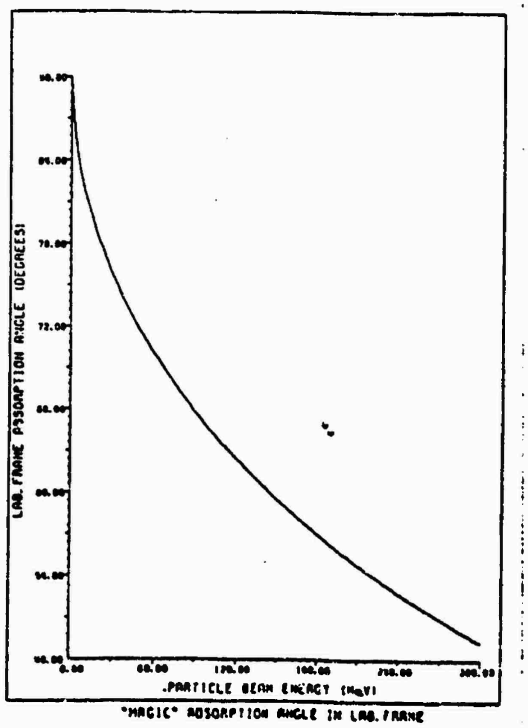


Fig. 2

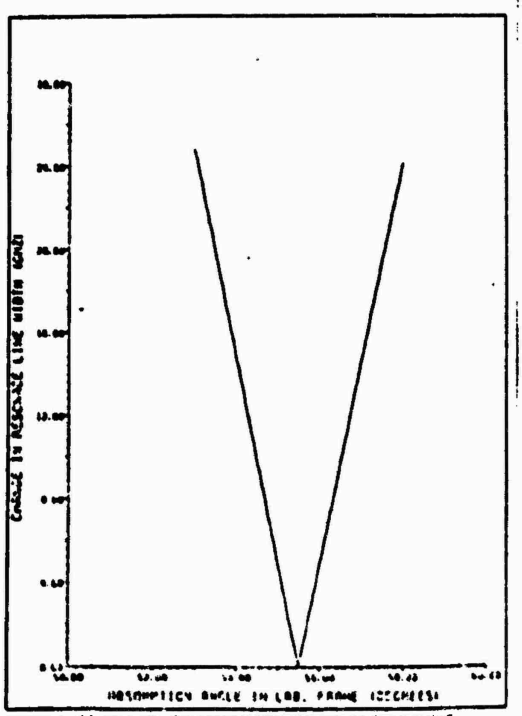


Fig. 3

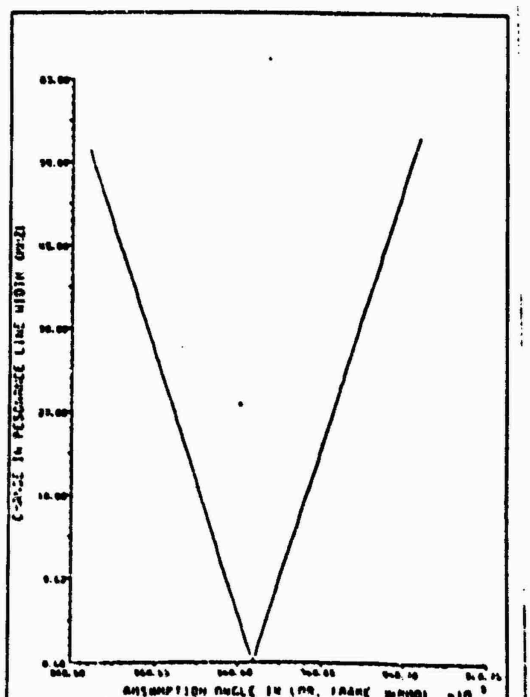


Fig. 4

for possible improvements where appropriate, and further developing some areas of the proposal.

(3) To propose alternative schemes for monitoring the resonant absorption interaction between intersecting laser and particle beams.

All of these objectives are consistent with the overall objective of designing a future experiment to test the principles behind the particle beam direction-sensing method and to experimentally investigate the parameters which most critically determine the sensitivity of the method. Throughout the report mention is made of design parameters for this experiment.

III. CONTRIBUTIONS TO THE RESONANCE WIDTH

In this section a study is made of the mechanisms that contribute to the width of the resonance signal used to monitor the amount of absorption of laser radiation by the moving beam particles. This quantity is of considerable interest since the sensitivity of the method of sensing such changes in the direction of a particle beam with respect to an intersecting laser beam depends rather critically on how wide the resonance is. Clearly, the narrower the resonance, the larger the change in the detected signal corresponding to a small change in the angular separation of the beams.

Let us begin by considering how a single beam particle interacts with laser radiation. Resonant absorption will occur when the frequency of the laser radiation, as seen in the particle frame, matches some transition frequency, ν_0 , characteristic of the particle. As a result of the finite lifetimes of the upper and lower levels of the transition there will be exist of small spread, $\delta\nu_0$, about ν_0 over which the particle can absorb laser radiation. This particle transition frequency linewidth is the result of the fact that the particle only interacts with the laser field

for a finite amount of time leading to a uncertainty in the energies of the levels of the transition. In addition, other processes which limit the interaction time between the particle and the laser radiation will contribute to this type of line broadening which is called homogeneous line broadening since it is essentially the same for all particles independent of their velocity class. Homogeneously-broadened line profiles are Lorentzian in shape. From the particle frame it would appear that the laser output is spread over a range of frequencies determined by the duration of the interaction. Such broadening could, for example, arise when one uses a pulsed laser instead of a CW laser. It is more likely, however, that in the crossed-beam arrangement used in the beam direction-sensing method the dominant homogeneous broadening mechanism will be due to the transit time of the fast-moving beam particles through the finite length laser field. This process is called transit-time broadening. In the proposed experiment care should be taken that the transit time is large enough so as not to cause the homogeneous broadening to dominate the overall linewidth of the monitor signal. It is assumed that power broadening of the resonance due to strong laser fields can be rendered neglible in the proposed experiment.

Up until now we have considered a single particle interacting with a radiation field. This lead to several possible homogeneous broadening mechanisms. We now wish to extend the analysis to include the whole ensemble of particles making up the particle beam. In the ideal case of the particle beam being monoenergetic and unidirectional and the laser beam also being unidirectional, all the particles would resonate at the same laser frequency in the laboratory frame. This is because both the longitudinal component of the particle velocity, β , and the angle between the direction of motion of a particle and the propagation direction of

the photons of the laser field, Θ, would be the same for all particles i.e. only one class of velocities exist. The width of the resonance associated with the whole ensemble of particles would then be the same as for a single particle i.e., the resonance will be homogeneously broadened. Unfortunately, in a real situation, the beam particles will exhibit some energy dispersion which leads to a Gaussian-distributed spread $\delta\beta$ in β about some mean value $\overline{\beta}$ determined by the nominal beam energy. In addition, both the laser and particle beams will have some finite divergence leading to a Gaussian-distributed spread, $\delta\Theta$, in the angle between interacting particles and photons about a mean value, $\overline{\Theta}$, determined by the angle between the particle and laser beam axes. The angular spread, $\delta\Theta$, is the quadrature sum of individual dispersions in both the particle and laser beam directions. We can think of the whole ensemble of beam particles as being divided into a large number of different velocity subgroups each of which is homogeneously broadened. When the laser is tuned it will resonate with the particles in the different velocity subgroups at different resonance frequencies. The width of the resonance signal is now determined by the range of these differenct resonance frequencies or equivalently the spreads in β and Θ. This type of line broadening is called Doppler broadening and it is an example of the more general broadening termed inhomogeneous broadening. If the homogeneous widths are kept small, the resonance signal linewidth will essentially be determined by the Doppler broadening. Inhomogeneously broadened line profiles are Gaussian in shape. In the proposed experiment care must be taken to minimize beam divergences and the particle beam energy dispersion in order to reduce Doppler broadening of the resonance signal.

We will now investigate how the spreads $\delta\beta$ and $\delta\theta$ effect the width of the resonance signal used to monitor the absorption process in the laboratory frame. To understand better the relative magnitudes of the first-order contributions to Doppler broadening of the monitor signal brought about by spreads in β and θ, we expand, in a Taylor series, the expression for the laser frequency in the laboratory frame, $\nu(\beta,\theta)$ about the mean value $\nu(\bar{\beta},\bar{\theta})$. Thus to first-order we have:

$$\nu(\beta,\theta) = \nu(\bar{\beta},\bar{\theta}) + \left.\frac{\partial\nu}{\partial\beta}\right|_{\bar{\beta}}(\beta-\bar{\beta}) + \left.\frac{\partial\nu}{\partial\theta}\right|_{\bar{\theta}}(\theta-\bar{\theta}) + \cdots \qquad (3)$$

The first-order dispersion coefficients can be expressed in terms of the resonant transition frequency, ν_o, as:

$$\left.\frac{\partial\nu}{\partial\beta}\right|_{\bar{\beta}} = \frac{\nu_o\bar{\gamma}\bar{\beta}\left(1-\frac{1}{\beta}\cos\bar{\theta}\right)}{(1-\bar{\beta}\cos\bar{\theta})^2} \qquad (4)$$

$$\left.\frac{\partial\nu}{\partial\theta}\right|_{\bar{\theta}} = \frac{\nu_o\bar{\beta}\sin\bar{\theta}}{\bar{\gamma}(1-\bar{\beta}\cos\bar{\theta})^2} \qquad (5)$$

Equation (4) shows that when $\cos\bar{\theta} = \bar{\beta}$ the first-order contribution in $\beta-\bar{\beta} \approx \delta\beta$ vanishes, i.e. at this laboratory-frame angle, which we shall call the "magic" angle, we become immune to the energy dispersion of the beam particles to first-order and the Doppler broadening of the resonance signal is significantly reduced. Physically, this laboratory frame angle corresponds to the perpendicular irradiation of particles in the most populated $\bar{\beta},\bar{\theta}$ velocity subgroup in the particle frame. Thus, in the particle frame, a considerable amount of the first-order Doppler broadening is eliminated. Figure 2 illustrates how the "magic" angle varies with particle beam energy. The question arises as to what happens to the resonance width if we are detuned slightly from the "magic" angle. Figure 3 shows that the first-order contribution in $\delta\beta$ to the Doppler broadening of the resonance increases rather dramatically if one is detuned

by only about 0.1° from the "magic" angle. This plot is for a particle beam energy of 200 MeV for which the "magic" angle is 55.53°. Thus if the mean angular separation of the laser and particle beam axes is not maintained to within a few hundred microradians of the "magic" angle, the resonance signal will broaden rapidly and soon disappear under the noise level (assuming the area under the resonance peak stays constant). Figure 4 shows that for very small angular excursions from the "magic" angle, the broadening due to the energy dispersion in the particle beam is comparable to or less than other sources of broadening such as the first-order contribution in $\delta\theta$, the beam divergences. This situation can be tolerated. In the proposed experiment we must therefore be able to set to high precision, and maintain, the "magic" angle for the particular particle beam energy to within about 0.01°. At the "magic" angle first-order contributions in $\delta\beta$ disappear and second-order contributions in $\delta\beta$ can be shown to be negligibly small compared with other sources of broadening. It remains to consider the magnitude of the first-order term in $\delta\theta$ in the Taylor expansion i.e.:

$$\left[\frac{\gamma_0 \bar{\beta} \sin\bar{\theta}}{\bar{\gamma}(1-\bar{\beta}\cos\bar{\theta})^2}\right]\delta\theta \qquad (6)$$

At the "magic" angle $\bar{\theta} = \cos^{-1}\bar{\beta}$ this reduces to;

$$\left[\frac{\gamma_0 \bar{\beta}}{1-\bar{\beta}^2}\right]\delta\theta \qquad (7)$$

where $\delta\theta = (\delta\theta_p^2 + \delta\theta_L^2)^{1/2}$ is the quadrature sum of the spread in directions of particles in the particle beam and photons in the laser beam. At the "magic" angle, for which the first-order term in $\delta\beta$ is eliminated, the contribution due to the beam divergences (Equation 7) is the major cause of Doppler broadening of the resonance signal. Equation 7 shows that typical particle beam divergencies of about a milliradian would produce

a resonance width of about 400 GHZ at a beam energy of 200 MeV. A resonance this wide would render the method of sensing small changes in the angular separation of the particle and laser beams very insensitive. Care must be taken in the proposed experiment to reduce both the particle beam and laser beam divergencies to well below the milliradian range. This, of course, can be achieved by expanding the beams or, at the expense of intensity, collimating them.

IV. MONITORING THE ABSORPTION RESONANCE WITH Ly_β FLUORESCENCE

Figure 5 is a schematic of the essential features of the experimental arrangement proposed by the TASC group to sense changes in the direction of a H-atom beam. The apparatus can be broken down into three spatially separated zones; the production zone, the excitation zone and the deexcitation zone. The fluorescence signal per laser pulse, S, can be written as;

$$S = N(2S) \times P(2S\text{-}3P) \times P(3P\text{-}1S) \times P_{DET} \qquad (8)$$

where N(2S) is the number of H-atoms in the metastable 2S state leaving the charge-exchange cell, P(2S-3P) is the probability an atom in the 2S state will be excited to the 3P state in the laser field. P(3P-1S) is the probability that the excited 3P state will deexcite via Ly_β fluorescence to the 1S ground state in deexcitation zone and P_{DET} is the probability that a Doppler-shifted Ly_β photon, emitted in the deexcitation zone will be collected and detected. This factor involves the probability of emission at a particular angle in the laboratory frame, the solid angle of the collection optics, the reflectivity of the collection optics and the efficiency of the detector.

In the production zone the H-atom beam is formed by charge-exchange and prepared in the metastable 2S state. This will probably be done via the

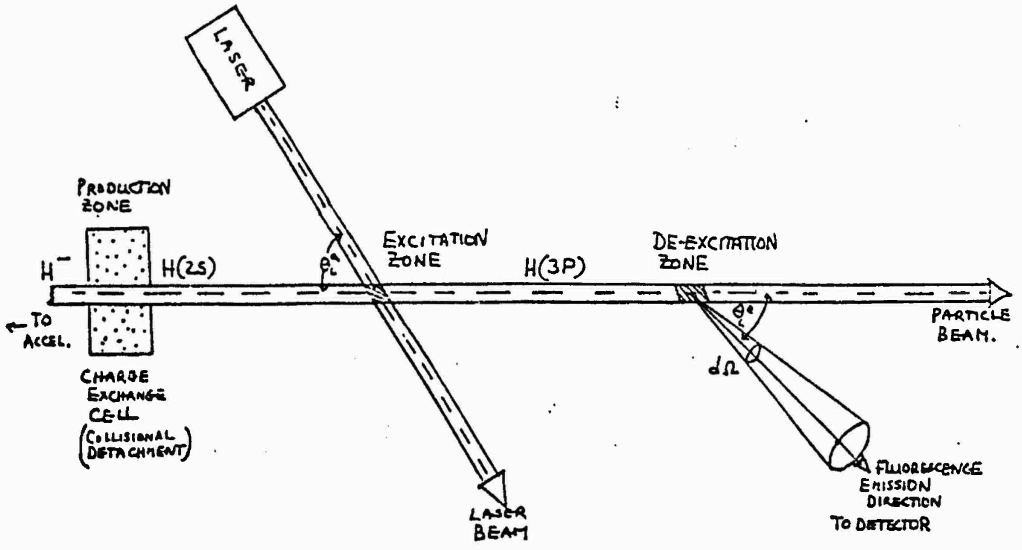

LASER
PRODUCTION ZONE
H⁻
H(2S)
EXCITATION ZONE
H(3P)
DE-EXCITATION ZONE
PARTICLE BEAM.
← TO ACCEL.
CHARGE EXCHANGE CELL
(COLLISIONAL DETACHMENT)
LASER BEAM
FLUORESCENCE EMISSION DIRECTION
TO DETECTOR

Fig. 5

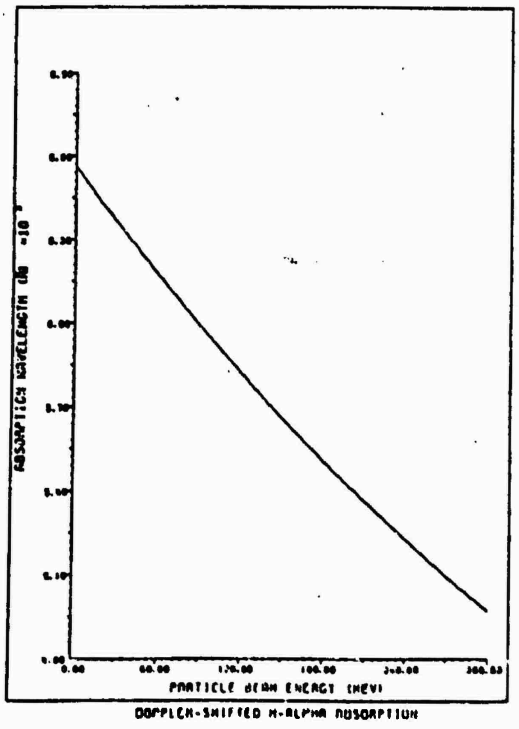

PARTICLE BEAM ENERGY (MEV)
DOPPLER-SHIFTED H-ALPHA ABSORPTION

Fig. 6

collisional detachment process;

$$H^- + X \longrightarrow H(2S) + X + e^- \qquad (9)$$

where X signifies a gaseous specie. At particle beam energies in the hundreds of MeV one cannot expect much more than 50% of the H^- beam tc be converted to H-atoms. Most of these atoms will be in the ground state. We can expect somewhere close to 10% of them to be in the 2S metastable state. Once produced we must be careful that stray electric or magnetic fields are not present. Such fields could quench the 2S state by mixing it with the 2P states. This same process could however be used intentionally further downstream in order to monitor the N (2S) as a function of time thus allowing us to normalize to any fluctuations in the particle beam intensity. Since the production process is energy unselective we can also expect the higher-lying states of the H-atom to be populated also. The probability of production of these states by collisional detachment should follow a n^{-3} scaling rule[3] leading to a $\sim$ 5% contribution. No measurements seem to have been made at high beam energies. Some of the higher-lying levels are long-lived and if they are populated reasonably strongly they could cascade into the 3P state and produce a noise contribution to the fluorescence monitor signal which is uncorrelated with the laser absorption. The nS and nD states with $n \geq 4$ will cascade directly into the 3P state and the nS states, in particular, are long-lived. For example, the 6S state has a lifetime of 570 ns., which is more than a hundred times longer than the 3P state lifetime. This state would be able to repopulate the 3P state by cascades for quite some time after its production in the charge-exchange cell. Making the separation of the production and excitation zones large would help to some extent. Indirect cascading from even longer-lived Rydberg states is also possible. Fortunately, cascading into the 3P state from the

highly excited and exceptionally long-live "yrast" states with $l = n - 1$ is not possible. The role of cascading in populating the 3P state in the beam is an area that could be further investigated. It would be useful to have data on the probability of the production of high-lying Rydberg states in the H-atom following collisional detachment of high energy H beams. Branching ratios for cascading into the 3P state are readily available in the case of the H-atom.

Selective population of the 2S state by photodetachment would clearly be preferable to collisional detachment and would eliminate entirely the possibility of cascading. The cross section for photodetachment of H^- producing H in all states reaches a maximum value of approximately 4×10^{-17} cm^2 at a wavelength of 8300Å in the particle frame. In the laboratory frame this wavelength could be Doppler shifted using angle-tuning to a more convenient laser wavelength, for example, within the bandwidth of a powerful laser such as a Nd^+: YAG laser. No measurements have been made to determine the cross section for production of the H-atom in the 2S state but it is suspected that this cross section is so small that when combined with the available fluxes of the particle and laser beams it would fail to produce a significant number of atoms in the 2S state. The cross section for photodetachment of H^- via intermediate resonances is somewhat higher[4] and it is conceivable that some mechanism which preferentially populates the 2S state of H could be found. Both the collisional and photodetachment process with beams are highly anisotropic production mechanisms and so, in general, one could expect the 2S state to be oriented or aligned. Alignment is not possible, however, for a state with $J = \frac{1}{2}$ and the state cannot be oriented in the collisional detachment process due to the cylindrical symmetry of the particle beam.

In the excitation region the TASC group propose to photoexcite atoms in the beam from the 2S state to the 3P state using a pulsed tunable laser. The mean angle between the laser and particle beam directions must be maintained at the "magic" angle for the particular beam energy chosen. Figure 6 shows the laboratory frame laser wavelength that must be used so that it is Doppler-shifted into resonance with the 2S-3P (H_α) transition in the particle frame. For example, for a beam energy of 200 MeV the output wavelength of the laser should be 5411Å which is just on the edge of the useful range of Rhodamine dyes. One possible pump laser for this dye might be a copper vapor laser which has two discrete output lines at 5100Å and 5780Å. Plasma Kinetics make a model which has an average power of 50W at 5100Å with a repetition rate of about 5KHZ. The duty cycle of 2×10^{-4} is poor but typical of pulsed lasers. Unless the number of atoms in the 2S state is kept high the poor duty cycle of a pulsed laser source will mean long intergration times. It would be preferable to use a CW laser for excitation of the beam particle but it would be very difficult to match the peak powers available with pulsed sources. The time structure of a pulsed laser may, however, be used to advantage in discriminating against noise which is uncorrelated with the laser. It will probably be useful to expand the cylindrical-shaped output beam of the laser to reduce divergence and then use a sheet-shaped portion (with a rectangular cross-section) so that all the beam particles traverse the same length of laser field. This will produce uniform transit-time broadening. It would also be preferable to use the single TEM_{oo} mode so that wavefronts are uniform.

Consideration should be given to how the shape of the resonance is effected by the unresolved hyperfine structure associated with each fine-structure component of the H_α line. The two fine-structure components, which are

3.25GHZ apart, are the result of the $2^2S_{1/2}$ - $2^2P_{1/2}$ and the $2^2S_{1/2}$ - $2^2P_{3/2}$ transitions (see Figure 1). The operation of the pulsed dye laser in a single mode configuration should ensure complete resolution of the two fine-structure components. Exact theoretical lineshape studies should be made on each component using the expected relative spacings and intensities of the hyperfine multiplets to see if any serious lineshape asymetries exist.

The output of lasers are usually linearly polarized but this can be changed to circular polarization if desired. Linearly polarized radiation will, in general, align the excited state while circularly polarized radiation with orient it. Normal isotropic excitation would populate all the magnetic substates of the excited state equally. Alignment means that magnetic substates of opposite sign are equally populated but that the population is a non-uniform function of $|M_J|$. The fluorescence from an aligned state will be linearly polarized and have a anisotropic radiation pattern even in the particle frame. Similarily, orientation implies that a single magnetic substate is overpopulated with respected to the rest. Fluorescence from an oriented state would be circularly polarized and have an anisotropic radiation pattern in the particle frame. The fact that the 3P state is anisotropically prepared by the laser could be used to our advantage. For example, if the plane of polarization of a linearly polarized laser was adjusted in such a way in the laboratory frame that, in the particle frame, it produced a dipole radiation pattern with the lobes directed toward the detectors, one could improve on the amount of fluorescence detected at a set emission angle even in the laboratory frame where kinematic effects reshape the particle frame radiation pattern so that it peaks in the forward direction. Polarization-sensitive detection of the polarized fluorescence could also be employed in order to discriminate against unpolarized noise contributions. If should be remembered that

$J = \frac{1}{2}$ levels cannot be aligned but can be oriented. This interesting area, not considered by TASC, seems to be worthy of further development. TASC has made a detailed study, however, of the photoexcitation probability, P(2S-3P), without consideration of the polarization of the laser. They conclude that it is a complicated function of the frequency and power of the laser as well as the transit-time of particles through the laser field. Care must be taken in choosing a particular combination of these variables if one wants efficient production of the 3P state.

After the excited atoms in the 3P state leave the excitation zone they begin to decay in flight in an exponential manner as the result of the spontaneous emission of radiation. The natural lifetime of the 3P state is 5.4 ns. There exists two different decay channels out of the 3P state: the 3P-1S and the 3P-2S transitions. These correspond to the emission of Ly_β and H_α fluorescence respectively. Since the branching ratio for the former route is 0.89, the Ly_β emission line is some 7.5 times stronger than the H_α emission line. For this reason TASC chose the Ly_β fluorescence to monitor the absorption resonance. In this case, since the Ly_β collection optics and detector is insensitive to the laser radiation directly there seems to be no reason (other than possible cascading effects) not to keep the separation of the excitation and deexcitation small. This avoids losing a significant fraction of the 3P state population by exponential depletion. It is also very important that the 3P state be allowed to spontaneously radiate free of external perturbations such as stray electric or magnetic fields (a fast-moving particle "sees", in its own frame, a motional electric field as well as a magnetic field when it passes through a laboratory frame magnetic field). The earth's magnetic field is an example of a laboratory frame field which is always present. A stray electric field, in the particle

frame, could Stark-mix the 3P states with other opposite parity states of the n = 3 manifold. If, for example, an appreciable amount of mixing occured between the 3P and 3S states, the Ly_β emission line should be reduced in strength since once in the 3S level the atoms can only decay to the 2P levels emitting H_α radiation (later, in a discussion of alternative schemes for monitoring the 3P state population, mention is made of enchancing the H_α line strength by purposefully mixing the 3P and 3S states).

The Ly_β fluorescence, which TASC propose to use to monitor the resonance absorption process, will be, in the laboratory frame, both Doppler - shifted and anisotropically peaked in the forward direction as a result of being emitted by fast-moving atoms. Figure 7 illustrates how the particle frame wavelength of the Ly_β fluorescence, 1026Å, is Doppler-shifted in the laboratory frame for various beam energies and emission angles. Figure 8 is an expanded version of Figure 7 for the acute emission angles only. Figure 7 shows that for acute emission angles the Ly_β fluorescence is blue-shifted (shorter wavelengths) and for obtuse emission angles it is red-shifted (longer wavelengths). All curves represent the sum of first-and second-order Doppler shifts except for the case of 90° emission in the laboratory frame where the first-order Doppler shift disappears and only the second-order shift remains which is always a red-shift. The second-order shift, often called the relativistic shift, is the result of time dilation effects. Figure 9 shows the effect of kinematics on the angular distribution of the emitted Ly_β fluorescence. This plot assumes that the radiation is emitted isotropically in the particle frame and that azimuthal symmetry exists. Clearly the higher the beam energy, the more the emitted radiation is peaked in the forward direction. For example, for a particle beam energy of 200 MeV, the

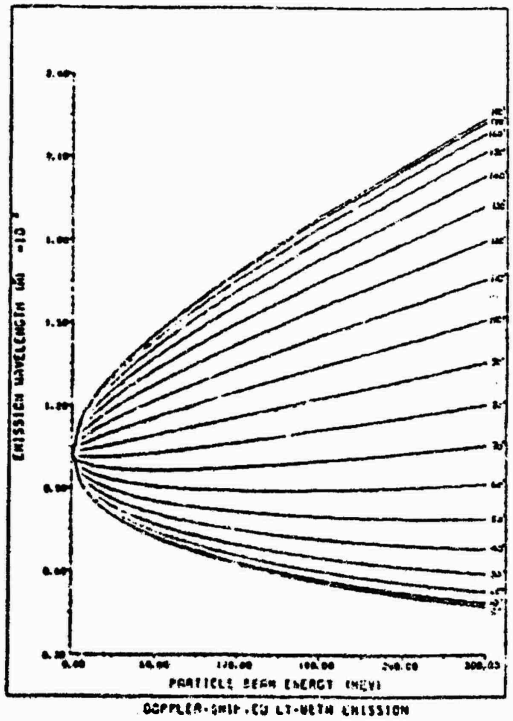


Fig. 7

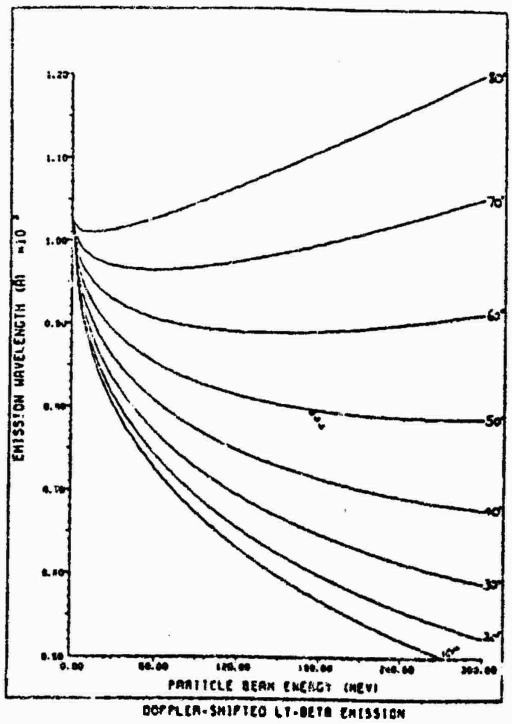


Fig. 8

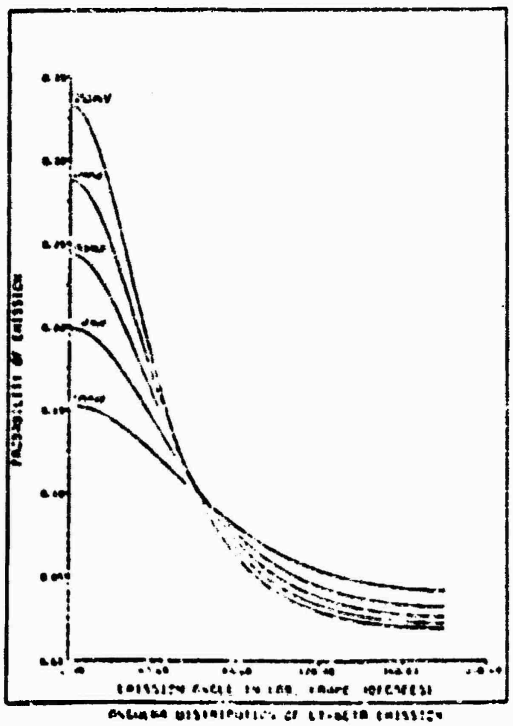


Fig. 9

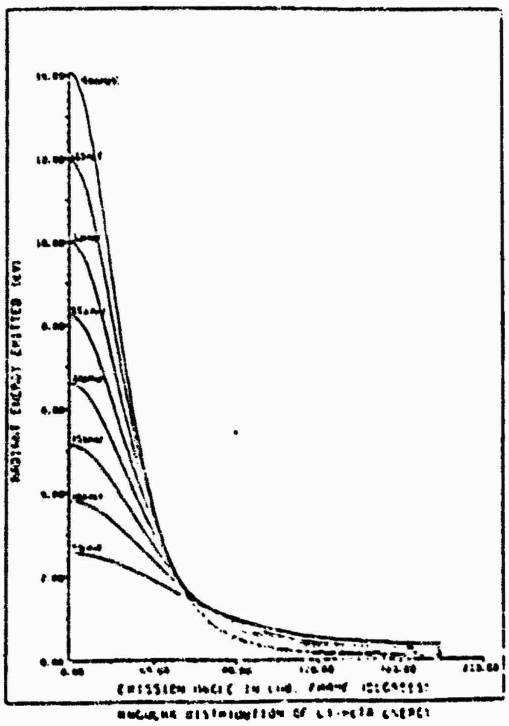


Fig. 10

probability of emission drops to half of what it is in the straight-on ($\theta = 0°$) direction for an emission angle of about 40°. Figure 10 shows how the radiant energy is distributed with angle of emission. This quantity is the product of the energy per photon which increases with beam energy for a given emission angle and the probability of emission at that angle which decreases as the beam energy increases. Figure 9 indicates that , at 200 MeV, we must use an acute emission angles equal to or less than about 40° in order to enchance the probability of emission whereas Figure 7 shows that at these rather small acute emission angles the fluorescence is blue-shifted into the wavelength range of approximately 550 to 700 Å. Unfortunately, this is a difficult wavelength region to work in experimentally since lenses can no longer be used to collect the emitted photons, mirrors coatings have rather poor reflectivities at near-normal incidence and detector efficiencies are rather low. The details of a collection optics and detection schemes has not, as yet, been proposed by TASC.

A number of second and third series transition metals have been found to be useful as reflecting coatings for mirrors below 1000Å. They are Pt, Ir, Os, Re, W, Rh and Au. Hass and Hunter[5] show reflectivity curves for most of these coatings. There appears to be a dependence of reflectivity on coating thickness for most coatings. The reflectance of all the coatings generally decreases toward shorter wavelengths but most of them show rather sharp increases over a small range deep in the UV. It would be to our advantage to arrange for the Doppler-shifted Ly_β fluorescence to fall at the wavelength of one of these sharp increases. In the wavelength range $\lesssim$ 500-800Å the reflectivities are, in general, no better than about 0.20 to 0.30. Os, which has a sharp maximum of 0.35 at 600Å, and Re, which has a similar maximum of 0.36 at 630Å, appear to be good

candidates for coatings on collection optics mirrors. Since the reflectivities are generally poor in this wavelength region we should try to design the collection optics to have a single reflection only. For example, one could use an off-axis paraboloidal reflector to focus the Ly_β fluorescence photons emitted from the beam. Collection schemes based on two reflectors such as a Cassegrain system with a primary ellipsoid and a secondary hyperboloid mirror would be less efficient in terms of overall reflectivity. Grazing incidence optics would improve the reflectivity somewhat but would probably restrict the solid angle of collection to a small value. The solid angle of collection should be kept as high as possible to help off-set the losses due to poor reflectivities.

The best available detectors in the optical range $\lesssim$500-800Å appear to be channel electon multipliers (channeltrons). The detection efficiency in this range, however, is rather poor, typically 0.10-0.18. Arrays of roughly 10^4-10^7 miniature electron multipliers called microchannel plates are also commercially available[6] and it might be possible to "wrap" many such plates in a cylindrical form around the beam and forego the use of mirrors altogether. This would also improve the solid angle of collection but the collected photons would have a range of wavelengths since there would be little restriction on emission angle. Only those photons whose wavelength fell within the spectral response of the dectectors would be detected.

From this discussion it should be apparent that a considerable amount of work remains to be done on the design of an efficient collection and detection scheme for the Doppler-shifted Ly_β fluorescence chosen by TASC to monitor the absorption resonance. In light of the obvious difficulties in working in the $\lesssim$500-800Å spectral region I have spent some of my time trying to devise alternative schemes for monitoring the population of the 3P state and, through it, the resonant absorption process.

V. ALTERNATIVE SCHEMES FOR MONITORING THE RESONANCE

In the previous section an analysis was made of the proposed TASC method of detecting small changes in the direction of a H-atom beam. It involved the use of the intrinsically strong Ly_β fluorescence to monitor the resonant absorption of radiation from a laser. It was shown that when high energy particle beams are used, the fluorescence to be detected is preferentially emitted in the forward directions and suffers considerable Doppler blue shifting at the small acute emission angles that we are forced to use. The particle frame wavelength of 1026Å was shifted into the range of approximately 500-800Å depending on beam energy and emission angle. Photons in this wavelength range are difficult to collect and detect efficiently. In this section we discuss alternative ways of monitoring the 3P state population after the absorption process.

One obvious method would be to monitor the H_α fluorescence instead of the Ly_β fluorescence even though it is intrinsically about an order of magnitude weaker than the Ly_β fluorescence. Advantages may be gained by the fact that, even after Doppler-shifting, the H_α photons emitted in the laboratory frame fall in the more accessible visible region of the electromagnetic spectrum where collection and detection efficiencies can be made quite high. Figure 11 illustrates how the Doppler-shifted photons change in wavelength with both particle beam energy and emission angle. Of course, the wavelength will still be preferentially directed into a small range of angles around the direction of motion of the beam. Let us, for example, choose a particle beam energy of 200 MeV and an emission angle of 30°. Figure 11 shows that the H_α fluorescence, whose particle frame wavelength is 6563Å, will be Doppler blue-shifted down to 4052Å in the laboratory frame. For this wavelength, reflectivities can be around 0.90 or more and, in addition, lenses can be employed to

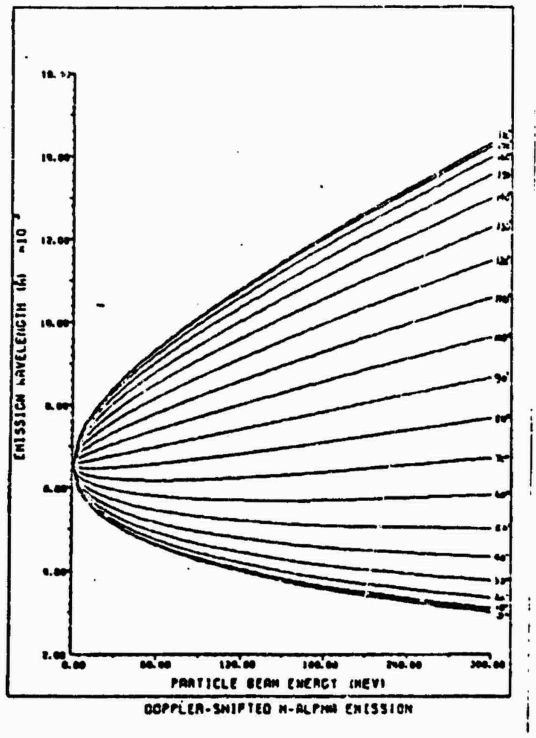

EMISSION WAVELENGTH (Å) ×10
PARTICLE BEAM ENERGY (KEV)
DOPPLER-SHIFTED H-ALPHA EMISSION

Fig. 11

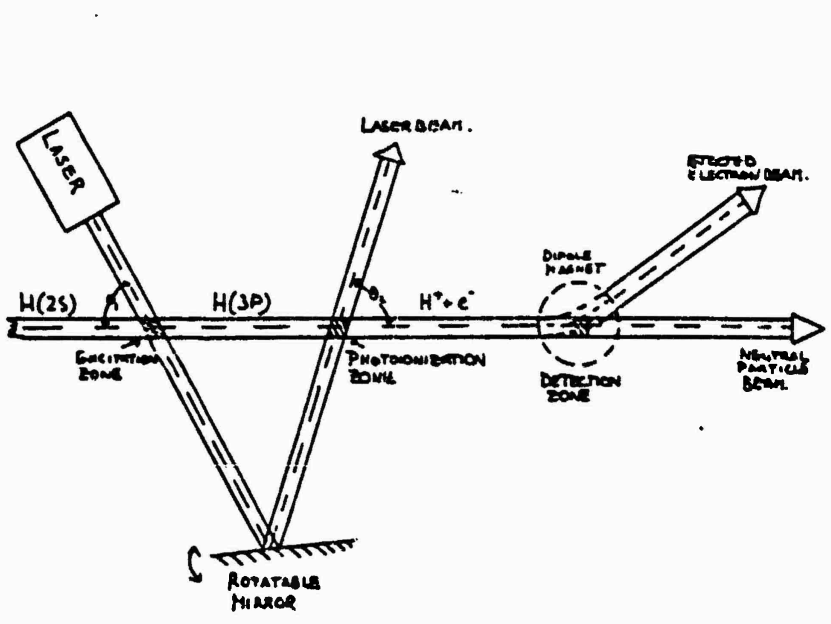

LASER
LASER BEAM
H(2S)
H(3P)
H⁺+e⁻
EXCITATION ZONE
PHOTOIONIZATION ZONE
DIPOLE MAGNET
DETECTION ZONE
NEUTRAL PARTICLE BEAM
ROTATABLE MIRROR

Fig. 12

focus the emitted radiation. The quantum efficiency of commercial photomultiplier detectors can be as high as about 0.30 at this wavelength although the intrinsic detector noise will be somewhat higher than for the channel electron multipliers which we proposed to use to detect the Ly_β fluorescence. It could be that the overall gain in collection and detection efficiencies using H_α rather than Ly_β fluorescence would more than offset the signal loss due to sampling a much smaller fraction of the 3P state population when monitoring with H_α .

I have also thought about ways of enhancing the strength of the fluorescence at the expense of the Ly_β fluorescence. It seems that this could be realized, for example, by intentionally Stark-mixing the 3P and 3S states in an applied electric field. The 3S state can only decay to the 2P levels emitting H_α radiation and so the enhancement of the (3P-2S) H_α signal by the (3S-2P) decay would be at the expense of the (3P-1S) Ly_β fluorescence. If this enhancement could be made appreciable then the monitoring of the H_α fluorescence would certainly become a viable alternative to monitoring the Ly_β fluorescence. Again, as in the case, it might be possible to improve the H_α signal further by adjusting the angular distribution pattern in the particle frame using a linearly or circularly polarized laser for absorption.

Thus far, only optical schemes for monitoring the 3P state population have been considered. Non-optical schemes could perform the same job and perhaps more efficiently. For example, inspection of the H-atom energy level diagram (Figure 1) shows that the 3P state lies only 1.51 eV below the ionization limit, i.e., the excited electron is only bound by 1.51 eV. Thus the threshold wavelength for photoionization, in the particle frame, is 8211Å. For a given beam energy, the angle between

the particle beam and the laser beam used for photoionization could be chosen so as to have the laser wavelength, in the laboratory frame, fall within the operating range of a powerful laser. Alternatively, the laser beam used to photoexcite the 2S-3P transition could be directed back onto the particle beam in order to photoionize atoms in the 3P state. A possible experimental arrangement is shown schematically in Figure 12. We could, for example, choose a particle beam energy of 200 MeV which corresponds to the "magic" angle of 55.53° and a laboratory frame wavelength of 5411Å for the laser performing the photoexcitation in the upstream position. If we then redirect this same laser beam (used a rotatable mirror) back onto the particle beam at an angle of 36.21° the particles will "see" a wavelength of 8211Å, the threshold wavelength for direct photoionization (into the continuum). The cross-section for this process at threshold has been calculated by Volkov[7] to be 2.6×10^{-17} cm^2. Since cross-sections for bound-bound transitions are typically a few orders of magnitude higher than for bound-free transitions, it would probably be more efficient to indirectly photoionize the atoms in the 3P state by first photoexciting them to a high-lying, but bound, Rydberg state. The field ionization of such states with $n \geq 20$ can be made about 100% efficient. The products of photoionization of the H-atom in the 3P state will be a proton and an electron which will have a small kinetic energy in the particle frame equal to the difference in the energy of the photon used for photoionization and the binding energy of the electron which, in this case, is 1.51 eV. This energy will always be very small compared to the translational energy of the particle frame itself so that the ejected electrons will travel with essentially the same speed and direction as the remaining H-atoms and the liberated protons. Their kinetic energy and momentum, however, will be

considerably smaller than those of the heavier particles. An electron which has, for example, been ejected from a 200 MeV particle beam by the process of photoionization will have an energy of about 109 KeV in the laboratory frame. Calculations also show that all the ejected electrons will be emitted in the forward direction into an angle of just a few milliradians around the particle beam axis. Figure 12 shows a dipole magnetic field being used to sweep the electrons, which have relatively small momenta, out of the main particle beam. This secondary beam can then be focused onto a detector such as an open-ended photomultiplier tube. The noise associated with non-optical detection schemes in generally much lower than for optical detection. The detected electron signal will be proportional to the population of the 3P state at any time and so can be used as a monitor of the absorption process. Further calculations on expected signal rates based on photoionization probabilities and collection and detection efficiencies should be made to see if such a scheme could become an alternative to the optical detection schemes mentioned earlier.

Another non-optical scheme that has possibilities is one based on detection by magnetic deflection. The famous experiments of Rabi showed that any change in the magnetic substate of an atom could be detected by a change in deflection in an inhomogeneous magnetic field. The amount of deflection is proportional to the transverse field gradient, the length of interaction, the magnetic moment and inversely proportional to the kinetic energy of the particle. This type of detection has been used for a long time in magnetic resonance experiments and more recently in high resolution spectroscopy experiments on atomic beams[8]. The inhomogeneous field can be provided by a six-pole magnet. In this configuration the field is zero on the axis and increases with distance from the axis. Thus the magnet will have focusing properties

for atoms with positive M_J values and defocusing properties for the other sublevels. If an upstream polarized laser is used to optically pump a transition and create an unequal population among the magnetic substates of an excited state or ground state this change could be detected by passing the atoms through a six-pole magnet. Calculations on the amount a hydrogen atom in the 3P state would be deflected by a given field need to be made to see if such a technique could be used to monitor the laser absorption process.

VI. SUMMARY

An analysis has been made of a method that could be used to sense small changes in the direction of a particle beam. It is based on studying the response of a signal used to monitor the strength of an interaction between intersecting laser and particle beams to a detuning from the resonance maximum brought about by a small change in the angular separation of the beams. The monitor signal can be any quantity proportional to the population of the upper level of the absorbing transition. Both optical and non-optical methods of monitoring this population have been proposed. The ability of the method to sense very small changes in the particle beam direction depends both on the signal-to-noise ratio and the linewidth of the resonance monitor signal. Narrow signals can be obtained by maintaining the mean angular separation of the two beams at the so-called "magic" angle given by $\bar{\theta} = \omega_0^{-1}\bar{\beta}$. The resonance width is then immune to Doppler broadening arising from first-order terms in $\delta\beta$, the spread in the speeds of the beam particles. It was found that the resonance broadens dramatically as one detunes appreciably from the "magic" angle. At the "magic" angle the resonance width is determined primarily by the combined divergences of the laser and particles beams assuming the homogeneous width due to transit-time broadening and power broadening is

kept small. The signal-to-noise ratio of the monitor signal determines the integration time needed to accumulated a certain amount of counting statistics. The strength of the signal used to monitor the population of the upper level of the absorpting transition is proportional to the product of the following quantities which must stay constant during the measurements: The number of atoms prepared in the lower level of the absorbing transition; the probability that the absorption takes place in the laser field; the probability of detecting the change in population of the upper level of the absorbing transition somewhere downstream from the resonant interaction region. If fluorescence is used to monitor the upper level population the latter factor depends on the angular distribution of photon emission in the laboratory frame, the angle of emission, the solid angle of collection of the emitted photons, and the efficiencies of the collection optics and detectors. Clearly all these factors must be optimized to produce a workable signal. Noise will be generated by any process which can populate the upper level without being correlated with the laser absorption process. Another noise source and possibly the most serious in a real experiment, could be due to random fluctuations in the number of beam particles and the direction of the particle beam axis. It would be advisable to continually monitor the population of the lower level of the transition in this case. The noise in the signal and the signal strength itself will depend on the method chosen to monitor the resonance signal.

VII. RECOMMENDATIONS

Based on my analysis of the method of sensing small changes in the direction of a particle beam, I make the following recommendations for the proposed experiment to be designed to investigate the feasability of the method and detect potential problems. The measurements will probably be made on an unexpanded fast-moving beam of H-atoms prepared in the 2S state and photoexcited to the 3P state. The monitor signal, at least initially, will be based on

fluorescence. The success of the experiment will depend to a great extent on the quality and stability of the particle beam chosen for the investigation.

It is clear that considerable attention must be devoted to maintaining the mean angle between the crossed beams at the "magic" angle for a given beam energy in order to preserve the sensitivity of the method. It is likely that a control loop will have to be used to achieve this. Additional problems could arise if the mean energy of the particle beam fluctuates during the course of the experiment. If the angular separation of the beams is maintained at the "magic" angle the sensitivity of the method will then depend on the ability to reduce the divergencies of both the particle and laser beams. The laser beam should be expanded and then collimated to the desired shape. The particle beam will probably remain unexpanded in the initial experiment but will be collimated by an amount depending to a great degree on the intensity of the particle beam that is available at the accelerator facility.

I recommend that the spacing between the excitation and deexcitation or sensing zone be held quite small in order to avoid exponential depletion of the population of the 3P state. At the high beam energy that will be employed, the Ly_β fluorescence will be Doppler blue-shifted by a considerable amount and will be emitted preferentially in the forward direction. At 200 MeV, I would recommend an angle of emission of about 20° which corresponds to a Doppler-shifted Ly_β fluorescence wavelength of about 580Å. At this wavelength several reflecting coatings for mirrors have reflectivities of 0.30. Tungsten or osmium, for example, might be good candidates for mirror surfaces. Since the reflectivities in this region of the spectrum are generally low I would recommend an optical collection system which involved a single reflection if possible. An off-axis paraboloidal reflector seems to be a good choice. The mirror diameter should be quite large so as to enhance the solid angle of collection. A large solid angle will encompass a range of emission angle

and photons with Doppler-shifted wavelengths. As long as these photons fall within the spectral response curve of the detector they will be useful. Channel electron multipliers will have to be used for detection even though their efficiencies are only about 0.15.

In addition to the proposed experiment I would recommend that a theoretical investigation be made of the following areas:

(1) The role of cascading in populating the 3P state in the H-atom following collisional detachment of H^-.

2) Analysis of the effect of the hyperfine structure on the shape of the 2S-3P fine structure line used in absorption.

3) The effect of polarized laser radiation on the angular distribution and polarization, in the particle frame, of the fluorescence emitted in the decay of the 3P state.

4) The effect of Stark-mixing the n = 3 states of the H-atom on the relative strengths of the H_α and Ly_β fluorescence lines.

5) An investigation of the possibility of photoionizing the H-atom in the 3P state and using the ejected electrons as a monitor signal for the 2S-3P absorption process.

6) A study of the possibility of using magnetic deflection of optically-pumped and oriented H-atoms in the 3P states by a six-pole magnet.

Any of the above processes may lead to an increase in the signal-to-noise of the signal used to monitor the strength of the interaction between crossed laser and particle beams and hence improve the sensitivity of the particle beam direction-sensing method.

REFERENCES

1. G. Rohringer, "Particle Beam Diagnostics by Resonant Scattering," General Research Corp. Report No: CR-1-773 (1977).

2. TASC (The Analytic Science Corp.), "Neutral Particle Beam Sensing and Control Concepts Study," Report No: TR-3286-1 (1981).

3. J. Risley et. al, J. Phys. B11, 1783 (1978); D.R. Bates et al, Proc. Phys. Soc. A70, 306 (1951)

4. H.C. Bryant et. al, Phys. Rev. A27, 2889 (1983).

5. G. Hass and W.R. Hunter, "New Developments in VUV Reflecting Coatings for Space Astronomy," in Space Optics, published by the National Academy of Science, Washington D.C., pp. 525-553 (1974).

6. E.A. Kurz, American Laboratory, March 1979; J.L. Wiza, Nucl. Instr. and Methods, 162, 587 (1979).

7. N. Volkov, International Atomic Energy Commission, Report No: IAE-873 (1965).

8. H.T. Duong et. al., Opt. Comm. 7, 371 (1973).

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

EVALUATION OF NAPHTHALENE TOXICITY USING SHORT TERM EXPOSURE AND THE AMPHIPOD, GAMMARUS TIGRINUS

PREPARED BY:	Dr. Gerald V. Poje
Academic Rank:	Assisitant Professor
Department and University:	Department of Zoology Miami University
Research Location:	Aerospace Medical Research Laboratory, Toxic Hazards Division, Toxicology Branch
USAF Research	Mr. Jeffrey Fisher
Date:	September 12, 1983
Contract No:	F49620-82-C-0035

Report Approval: Jeffrey Fisher

Mr. Jeffrey Fisher
Effort Focal Point
Toxicology Branch
Toxic Hazards Division
Wright-Patterson AFB
Ohio 45433

EVALUATION OF NAPHTHALENE TOXICITY USING SHORT TERM EXPOSURE AND THE AMPHIPOD, GAMMARUS TIGRINUS

by
Gerald V. Poje

ABSTRACT

Pulsed releases of toxic chemicals, such as occur with accidental spills, are poorly modelled by traditional bioassay techniques. The toxicity of naphthalene, an important constituent of jet fuels, is examined using short term exposures typical of a spill event. Median lethal concentration over the range of short term exposure durations are described for the crustacean, Gammarus tigrinus. Sublethal impact upon growth and reproductive parameters are quantified. Short term exposure to naphthalene resulted in altered feeding and mating behaviors. Morphological alteration of gill tissue using scanning electron microscopy is discussed. Sugestions for further research in this area is offered.

ACKNOWLEDGEMENT

The author would like to thank the Air Force Systems Commmand, the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for providing him with the opportunity to participate in toxicological research at the Toxic Hazards Division of the Aerospace Medical Research Laboratory at Wright-Patterson AFB. In particular I would like to acknowledge the Toxicology Branch for extending every effort to insure that this project could be completed during the allotted time.

The encouragement of Col. Roger Imman, division chief, and Col. Carl Olson, branch chief, were greatly appreciated, as were the technical assistance of Mr. Don Baird, Ms. Patty Hochwalt, and AFC Dave Miller. Finally I would like to thank Mr. Jeffrey Fisher for his collaboration, counsel and discussion of the research.

I. Introduction.

Presently, toxicological testing for aquatic hazard evaluation relies primarily upon two bioassay techniques to assess median lethality. In theses studies, organisms are exposed to various concentrations of suspect chemicals either for short periods of time (acute bioassay: 24, 48, 96 hours duration) or for a period of time typical of the species' life span (chronic bioassay). The advantage of the former is the economy of time and money for testing. The advantage of the latter is that exposure to the most sensitive life history stages can be assured, thereby providing greater understanding of a species tolerance for the suspect chemical. However, certain pollution events such as chemical spills or pulsed-releases occur with relatively high frequency. These are poorly modelled by these standard exposure and observation methods.

Important effects upon individuals and populations can occur after exposure to sublethal concentrations (1,2). As a result, many studies have been conducted using low levels of persistant chemicals such as PCBs and DDT. However, widespread anthropogenic processes such as power plant operations, scheduled chemical plant releases, and accidental spills result in frequent challenges to the aquatic ecosystems with pulses of toxic chemicals that rapidly degrade and detoxify. With chlorination events, the traditional bioassay techniques would indicate non-hazard for organisms so broefly exposed (3): yet when observation

continues far beyond the immediate exposure period, adverse impacts upon growth and reproduction have been discovered (4).

The lower molecular weight aromatic hydrocarbons are important components of many fuels. They have sufficiently high aqueous solubility and toxicity that they must be considered primary toxic agents when oil spills occur (5). Wells (6) in a literature review cited evidence that naphthalene is a potent feeding rate inhibitor of crustaceans during the time that it is rapidly accumulated. The amphipod, Gammarus oceanicus, showed increased toxicity with light crude oils as compared to heavy crude; this effect was more pronounced in the youngest animals where reduced growth rates were observed during sublethal exposures (7). Anderson and colleagues (8) characterized these light crudes as rich in low molecular weight aromatics such as naphthalene. Similar light crudes can alter precopulatory behavior in amphipods and decrease their fecundity (9); additional evidence cites decreased nesting behavior and altered lipid metabolism (10).

The genus Gammarus is recognized as a very sensitive species for toxicity testing (11). Most recently, field studies have shown this genus to be a most sensitive marker among other macroinvertebrates for projecting adverse levels of industrial activity (12). Furthermore, toxicity testing of persistant effects of short term exposure (1 hour) with agricultural chemicals have been reported (13).

II. Research Objectives.

The general objective of this study was to evaluate the aquatic toxicity of naphthalene to a sensitive aquatic organism when the exposure situation was analagous to a spill of fuel oils. Both lethal and sublethal impacts of naphthalene were investigated to determine rapid, yet ecologically-relavent assays. Gammarus tigrinus proved ideal for this study because: a) it grows rapidly in culture and has a high fecundity (14); b) methods for quantifying reproductive capacity and growth rates have been described (4); and c) the species dominates the epibenthic macrozooplankton fauna of several American (15) and European (16) aquatic ecosystems which indicates its importance as a test organism.

The specific objectives of the study were five:

a. The median lethal responses to naphthalene were determined. Replicate groups of newly-released young amphipods were exposed to one of an increasing series of naphthalene concentrations so that the responses 48 hours after exposure ranged from total survival to total mortality.

b. Subpopulations of young amphipods were exposed to sublethal concentrations of naphthalene and observed at weekly intervals until reproductive maturity was achieved. Sexual composition of these animals and several morphometric parameters were recorded.

c. Mated pairs of adult amphipods were exposed to sublethal

concentrations of naphthalene. Reproductive competence was assessed over several mating cycles.

d. An essential reproductive behavior for this species is the clasping of the female by the male for several days prior to copulation (precopulatory amplexus). Individual pairs in amplexus were exposed to various concentrations of naphthalene. Persistance of this behavior during exposure and success in reestablishing this behavior after exposure were assessed.

e. During the initial portion of the study observations of exposed amphipods disclosed a high frequency of morphological alteration of gill tissue. A separate exposure series was generated to supply animals for scanning electron microscopy. The techniques for animal preparation and examination of gills and excretory ducts were completed.

III. Median Lethal Concentrations.

Lethal responses of amphipods to naphthalene exposure followed the expected trend of increased sensitivity with increasing exposure duration (Table 1). In general, as the exposure duration increased by a factor of three, the median lethal concentration decreased by a factor of two. Confidence intervals surrounding these LC50s were greatest for the longest exposure durations. At very high concentrations of naphthalene (>5ppm) all animals demonstrated initial hyperactivity for approximately 3-5 minutes followed by complete cessation of swimming movements. At 2-5 ppm animals remained hyperactive

throughout exposure until either recovery or death. Below approximately 2 ppm, amphipods showed no altered activity in comparison to controls.

Table 1. Median lethal concentrations (ppm) and confidence intervals for various exposures to naphthalene.

Exposure Duration	LC50	95% Confidence Interval
30 minutes	*	
54.7 minutes	40.05	19.64 and infinity
100 minutes	13.26	10.89 and 17.73
173.4 minutes	6.54	2.67 and 15.82
300 minutes	5.69	5.27 and 6.12
1000 minutes	2.11	1.83 and 2.42
3000 minutes	1.65	1.26 and 2.06

* less than 5% mortality at the limits of naphthalene solubility.

IV. Growth Rates and Reproductive Capacity

Three groups of newly-released young amphipods were exposed to either: ambient water for 300 minutes; 1.0 ppm naphthalene for 100 minutes; or 0.5 ppm naphthalene for 300 minutes. Both naphthalene exposures represented approximately 10% of the LC50 value for those exposure durations. Survival was noted during weekly observations. At biweekly intervals until reproductive maturity was achieved the following parameters were noted on each of ten individuals: head length; head-peraeon length; length of peduncular segments on the first and second antennae; number of flagellar segments on each antennae. No significant differences in any of these morphological characteristics were noted among the three groups at any observation period.

In general, head length and number of flagellar segments provided the most reliable indicators of the population growth. The sexual ratio (female/male) of the control group at maturity was 1.17; for the 0.5 ppm group it was 0.65; and for the 1.0 group it was 0.58. This would suggest that naphthalene had exerted a impact on either female survival or on the development of female sexual characteristics.

For studying the impact of naphthalene exposure on adult amphipod reproduction, groups of 10-15 individual pairs of animals were exposed for a single 100 minutes duration to either: ambient water; 1.2 ppm; 2.5 ppm; or 5.0 ppm naphthalene. After exposure pairs were observed daily to identify mating behavior, status of the females' brood pouch, and release of young animals. There were no significant differences in the durations of precopulatory amplexus (1-5 days), in the duration of the intermolt period (6-9 days), or in the frequency of molting and ovigery. However, only in the control group was there a high frequency of immediate post exposure release of young (7/10) and subsequent release of young (8/10). None of the naphthalene exposed groups demonstrated an early release of young; subsequent releases were 3/15, 2/12, and 1/10 for the 5, 2.5 and 1.0 exposure groups respectively.

V. Amplexus Behavior

Individual adult amphipods are dispersed widely in the aquatic environment. However, females are available for copulation during only a brief period (hours) after molting.

To insure successful reproduction males will clasp females (amplex) for several days prior to her molting. Preliminary experiments indicated the efficacy of quantifying the temporal aspects of precopulatory amplexus during naphthalene exposure and the recovery of this behavior after cessation of chemical stress. Individual pairs of amplexed amphipods were exposed for 100 minutes duration to one of the following: ambient water; 1.3 ppm; 3.0 ppm; 6.5 ppm; or 14.1 ppm naphthalene. Animals were observed during the exposure period (100 minutes), and during the 180 minute recovery period. All pairs were separated physically prior to the recovery period observations. The results are summarized in Table 2.

widespread (but non-persistant) chemicals. One of particular concern to the Air Force would be the degreasing agent, trichloroethylene.

Pulsed exposure to toxicants may have a generic impact upon the respiratory competence of amphipods. Certainly, a developmental profile of the gills of these crustaceans and the influence of the molting cycle on surface features would be essential for interpertation of scanning electron micrographs. In addition, I would suggest that this information be coupled with a physiological study of the respiratory rates of these animals. Gammarus is well suited to studies using small chamber, Warburg-like respirometers.

Table 2.

A. Percent of exposure time (100 minute maximum) that amphipods remained amplexed.

Concentration	n (*)+	Mean + S.D.	Range	
			minimum	maximum
14.1	10 (10)	0.19 + 0.1	0.03	0.4
6.5	14 (14)	1.82 + 2.7	0.09	8.4
3.0	10 (10)	30.2 + 29.0	1.1	84.0
1.3	10 (3)	85.1 + 25.7	27.3	100.0
0.0	20 (0)	100.0 + 0	100.0	

+ n (*) = number of exposed animals (number which broke amplexus)

B. Percent of recovery period (180 minute maximum) that amphipods remained amplexed.

Concentration	n (*)++	Mean + S.D.	Range	
			minimum	maximum
14.1	10 (9)	0.78 + 2.5	0	7.78
6.5	9 (4)	36.5 + 36.6	0	89.3
3.0	10 (1)	70.1 + 33.2	0	96.4
1.3	10 (0)	87.9 + 28.2	7.7	99.5
0.0	20 (0)	97.3 + 1.8	92.8	99.4

++ n (*) = number of exposed animals (number not recovered)

Gammarus pairs showed sensitivity to naphthalene with

this reproductive behavior even at concentrations as low as 1.3 ppm: 30% broke amplexus, while all pairs recovered after cessation of hydrocarbon stress. At the higher naphthalene concentrations all pairs abandoned amplexus, and higher percentages failed to recover the precopulatory stance.

VI. Scanning Electron Microscopy

Naphthalene has been demonstrated as a respiratory toxicant for some crustaceans (17, 18). During preliminary experiments, examination of adult amphipods showed that a large percentage of exposed animals posessed deformed, swollen gills. Optimum preparation of amphipods was accomplished with a 24 hour exposure to one-half strength Karnovsky's Fixative and subsequent maintanence in pH 7.4 cacodylate buffer. All animals were dehydrated in graded solutions of ethanol and amylacetate; critical point dried with liquid CO2; sputter-coated with gold; and observed on an ETEC scanning electron microscope. Gills were examined at 70X; 500X; and 2000X. The gland cone, an antennal excretory organ, were examined at 500X and 2000X.

Animals exposed to elevated concentrations of naphthalene showed erosion of the membrane-like sheath surrounding the surface of the gills. High power resolution disclosed extensive elimination of surface ruffling on the gills. This loss of intimate contact with the subsurface cell layer could increase the diffusion path for oxygen across the gills, and thereby explain the decrease in respiratory rates. At the highest naphthalene concentrations

(ones which proved ultimately lethal to exposed animals) several animals demonstrated a pronounced occlusion of the excretory pore for the gland cone; this was not consistant for all exposed animals.

VII. Recommendations.

The information derived from this study can provide a valuable framework for the interpertation of chemical spill impact on aquatic ecosystems. The lethal responses of a sensitive macroinvertebrate to the important fuel oil constituent, naphthalene, have been mapped for several short term exposures. Sublethal exposures as low as 1.3 ppm for 100 minutes have been demonstrated to adversely affect the reproductive capacity and reproductive behaviors of these amphipods. Although growth rates were not altered by exposure to 1.2 and 0.5 ppm naphthalene for 100 and 300 minutes, sexual compositions of the exposed groups were biased towards higher proportions of males and therefore lower capacity for population increases.

This study enhances the validity for examination of short term exposure with long term observations. In addition, it presents some methods (amplexus behavior, growth rates, scanning electron microscopy) for relatively rapid and sensitive evaluation of the aquatic toxicity of many short-lived contaminants. To date such investigations have focused upon only three stressors: chlorine biocides; thermal elevations; and naphthalene. I would recommend that this line of research be extended into studies with other

REFERENCES

1. Sprague, J.B. 1971. Measurement of pollutant toxicity to fish III. Water Res. 5:245-266.

2. Perkins, E.J. 1972. Some methods of assessment of toxic effects upon marine invertebrates. Proc. Soc. Analyt. Chem. 9(5):105-114.

3. Ginn, T.C. and J.M.O'Connor. 1978. Response of the estuarine amphipod Gammarus daiberi to chlorinated power plant effluent. Estuarine Coastal Mar. Sci. 6:459-469.

4. Poje, G.V., S.R.Riordan, and J.M.O'Connor. 1982. Power Plant Chlorination: immediate and persistant effects of sublethal concentrations on an estuarine crustacean. IN: Jolley, R.J. (ed.) Water Chlorination: Environmental Impact and Health Effects, Vol.4, Book 1, Chemistry and Water Treatment. Ann Arbor Science, Ann Arbor, Michigan.

5. Connell, D.W. and G.V.Miller. 1980. Petroleum hydrocarbons in aquatic ecosystems- behavior and effects of sublethal concentrations: Part 1. CRC Critical Rev. Environm. Control. 11(1):37-104.

6. Wells, P.G. 1982 Zooplankton. IN: J.B.Sprague, J.H. Vandermuelen and P.G.Wells (eds.). Oil and Dispersants in Canadian Seas- Research Appraisal and Reccommendations. Environment Canada En 46-3/82-2E.

7. Linden, O. 1976a. Effect of oil on the amphipod Gammarus oceanicus. Env. Pollut. 10:239-250.

8. Anderson, J.W., J.M. Neff, B.A. Cox, H.E. Tatem, and G.M. Hightower. 1974. Characteristics of dispersion and water-soluble extracts of crude and refined oils and their toxicity to estuarine crustaceans and fish. Mar. Biol. 27:75-88.

9. Linden, O. 1976b. Effects of oil on the reproduction of the amphipod, Gammarus oceanicus. Ambio 5:36-37.

10. Lee, W.Y., S.A.Macko and J.A.C.Nicol. 1981. Changes in the nesting behavior and lipid content of a marine amphipod (Amphithoe valida) to the toxicity of No.2 fuel oil. Water, Air, and Soil Pollut. 15:185-195.

11. Arthur, J.W. 1980. Review of freshwater bioassay procedure for selected amphipods. IN: A.L.Buikema, Jr. and J.Cairns (ed.) Aquatic Invertebrate Bioassays ASTM STP 115. American Society for Testing and Materials. Philadelphia, PA.

12. Pieper, H. and M.P.D.Meijering. 1981. Zur situation der gattung Gammarus in abflussgeblet der oberen fulda. Beitr. Naturkde. Osthessen. 17:61-69.

13. Bluzat, R., O.Vonot and J.Seuge. 1982. Acute toxicity of thiram in Gammarus pulex: effect of one-hour contamination and degradation of an aqueous suspension. Bull. Environm. Contam. Toxicol. 29:248-252.

14. Poje, G.V., S.R.Riordan and J.M.O'Connor. In Press. Food habits of the estuarine amphipod, Gammarus tigrinus and the effect of diet upon its growth and reproduction. IN: Smith, C.Lavett (ed.) Hudson River Fisheries Conference. Hudson River Environmental Society. New York, NY.

15. Wilhm,J., H.J.Harmon and J.Y.Crider. 1982. Effects of naphthalene on the haemoglobin concentration and oxygen uptake of Daphnia magna. Bull. Environm. Contam. Toxicol. 28:52-57.

16. Holland, D.G. 1976. The inland distribution of brackish water Gammarus species in the area of the Mersey and Weaver River Authority. Freshwater Biol. 6:277-285.

17. Buikema, A.L. and J.G.Geiger. 1981. Oxygen consumption and filtering rate of Daphnia pulex after exposure to water soluble fractions of naphthalene, phenanthrene, No.2 fuel oil, and coal tar creosote. Bull. Environm. Contam. Toxicol. 27:783-789.

18. Wilhm, J., H.J. Harmon and J.Y. Crider. 1982. Effects of naphthalene on the haemoglobin concentration and oxygen uptake of Daphnia magna. Bull. Environm. Contam. Toxicol. 28:52-57.

ABSTRACT

In this report, a technique for the simplification of those nonlinear systems whose nonlinearities can be represented by continuous functions is presented. The method is based upon constructing an equivalent nonlinear system in which the linear and nonlinear signals appear as orthonormal signals. The simpler model is selected from this equivalent system. The technique is illustrated by a numerical example.

ACKNOWLEDGMENTS

I would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research and Southern Center for Electrical Engineering Education for providing me with this opportunity of creatively interacting with the Air Force Flight Dynamics Laboratory, Wright Patterson Air Force Base, Ohio under the Summer Faculty Research Program. I would particularly like to thank Dr. Robert Schwanz, Group Leader, Flying Qualities Group, for having provided me with excellent working conditions. I also wish to thank Dr. Bob Schwanz and Mr. Frank George for several stimulating and useful conversations.

TABLE OF CONTENTS

SECTION		PAGE
I	Introduction	1
II	Formulation of the Problem	2
III	Definitions and Notation	3
IV	Derivation of the Algorithm	4
V	An Illustrative Example	8
VI	DESROCHERS and AL-JAAR Algorithm	10
VII	Summary	13
VIII	Recommendations	14
IX	References	15

I. INTRODUCTION

With the emergence of large scale systems, model-reduction techniques have assumed tremendous importance for the purposes of design and analysis. The essence of a model-simplification technique is to replace a given large order system by a simpler system in such a way that the simpler model retains the significant properties of the given system. In the case of linear time invariant systems, the model-reduction problem has been extensively researched [1, 2, 4, 5].

In the case of nonlinear systems, very often, the system is first linearized about the operating point and then a simpler model is obtained by using a model simplification technique for linear systems. This situation is not quite satisfactory for the following reasons:

1. It involves two steps of approximation. First, a nonlinear system is approximated by a linear system and then the linear system is approximated by a simpler model.

2. Linearization of the nonlinear system is often not very accurate when one moves away from the operating point.

Thus, there is an urgent need for developing model-simplification techniques in the case of nonlinear systems. This problem is complicated in case of those systems by virtue of the nature of the systems.

Recently, Desrochers and Al-Jaar [3] developed an algorithm for the simplification of those non-linear systems in which the nonlinearities are continuous functions of the states and the inputs. Their technique is based on selecting a subset of the

nonlinearities through the minization of a suitable defined error between the given and the simpler systems.

In this report, the given non-linear system is first replaced by an equivalent system in which the signals (both linear and non-linear) are replaced by an orthonormal system. Then a subset of this orthonormal set is selected which will minimize a suitably defined error between the given and the simplified systems. A simple numerical example is presented to illustrate the technique.

II. FORMULATION OF THE PROBLEM

Let

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \cdot \\ \cdot \\ \dot{x}_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \cdots a_{1\ell} \\ a_{21} & a_{22} \cdots a_{2\ell} \\ - & - \; - \; - \\ - & - \; - \; - \\ a_{n1} & a_{n2} \cdots a_{n\ell} \end{bmatrix} \cdot \begin{bmatrix} f_1(\underline{x},u) \\ f_2(\underline{x},u) \\ \cdot \\ \cdot \\ f_\ell(\underline{x},u) \end{bmatrix} \qquad (1)$$

be a given system where a_{ij}'s are real constants, $f_i(\underline{x},u)$ are continuous functions of $\underline{x}$ and u, $\underline{x} = (x_1(t), x_2(t), \ldots x_n(t)$ are the states,, u(t) is the input. Equation (1) can be written as:

$$\dot{\underline{x}} = A.\underline{F} \qquad (2)$$

where

$$A = \begin{bmatrix} a_{11} & a_{12} \cdots & a_{1\ell} \\ a_{21} & a_{22} \cdots & a_{2\ell} \\ - - - - - - & & \\ - - - - - - & & \\ a_{n1} & a_{n2} \cdots & a_{n\ell} \end{bmatrix}$$

$$\underline{F} = (f_1(\underline{x},u), f_2(\underline{x},u), \ldots f_\ell(\underline{x},u))^T \qquad (3)$$

The system (2) is to be replaced by a simpler system given by

$$\dot{\underline{x}} = B.F_{\underline{r}} \tag{4}$$

where

$$B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ - & - & - & \\ - & - & - & \\ b_{n1} & b_{n2} & \cdots & b_{nm} \end{bmatrix} \tag{5}$$

$$\underline{F_r} = (f_{r_1}(\underline{x},u),\ f_{r_2}(\underline{x},u),\ \ldots\ ,\ f_{r_m}(\underline{x},u))^T \tag{6}$$

and $\{r_1, r_2, \ldots, r_m\}$ is a subset of $\{1,2,3,\ldots,\ell\}$. In (4), b_{ij}'s are unknown real constants to be determined in such a way that a suitable error between (2) and (4) is minimized. We assume that the relative dominance of the signals f_i's in (2) is known from physical considerations and they are arranged in the diminishing order of dominance in (2).

III. Definitions and Notation

Def: For signals $f_i(\underline{x},u)$, $f_j(\underline{x},u)$,

$$\text{define } \langle f_i, f_j \rangle = \int_{t_1}^{t_2} f_i(\underline{x},u)\, f_j(\underline{x},u)\, dt \tag{7}$$

where $[t_1,t_2]$ is the time-interval of interest over which the approximation is desired.

As usual,

$$\|f_i\|^2 = \langle f_i, f_j \rangle \tag{8}$$

Note that if the signals f_i, f_j are discrete, then $\langle f_i, f_j \rangle$ will denote the ordinary dot product of discrete vectors.

IV. DERIVATION OF THE ALGORITHM

Define the error E between the given system (2) and the simplified system (4) as follows:

$$E = \sum_{k=1}^{n} \left\| \sum_{j=1}^{\ell} a_{kj} f_j - \sum_{j=1}^{m} b_{kj} f_{rj} \right\|^2 \qquad (9)$$

Let $\Delta' = \{r_1, r_2, \ldots r_m\}$

and $\Delta = \{1,2,\ldots \ell\} - \Delta'$ (10)

Note that $\Delta \cap \Delta' = \emptyset$, empty set.

To determine the optimal parameters by minimizing E, we first replace f_i's by an orthonormal set to obtain an equivalent system. For the moment, we assume that f_i's are orthonormal and prove the following proposition to determine b_{ij}'s.

Proposition

If the signals $f_1, f_2, \ldots, f_\ell$ are orthonormal, then the error E is minimum if

$$b_{ij} = a_{ir_j}, \quad 1 \le i \le n, \quad 1 \le j \le m$$

and the minimum value of E is given by $\sum_{k=1}^{\ell} \sum_{j \in \Delta} a_{kj}^2$

Proof:

From (9), we have

$$E = \sum_{k=1}^{n} \Big\langle \sum_{j\in\Delta'} (a_{kj} - b_{kj}) f_{rj} + \sum_{j\in\Delta} a_{kj} f_j, \sum_{j\in\Delta'} (a_{kj} - b_{kj}) f_{rj} + \sum_{j\in\Delta} a_{kj} f_j \Big\rangle$$

$$= \sum_{k=1}^{n} \Big\langle \sum_{j\in\Delta'} (a_{kj} - b_{kj}) f_{rj}, \sum_{j\in\Delta'} (a_{kj} - b_{kj}) f_{rj} \Big\rangle$$

$$+ 2 \Big\langle \sum_{j\in\Delta'} (a_{kj} - b_{kj}) f_{rj}, \sum_{j\in\Delta} a_{kj} f_j \Big\rangle$$

$$+ \Big\langle \sum_{j\in\Delta} a_{kj} f_j, \sum_{j\in\Delta} a_{kj} f_j \Big\rangle$$

$$= \sum_{k=1}^{n} \sum_{j\in\Delta'} (a_{kj} - b_{kj})^2 + 2.0 + \sum_{j\in\Delta} a_{kj}^2 \qquad (11)$$

because f_i's are orthonormal and $\Delta \cap \Delta' = \phi$. From the form of E given by (11), it is clear that E is minimum if

$$a_{kj} = b_{kj}, \quad j \in \Delta', \quad 1 \leq k \leq n$$

Now if $a_{kj} = b_{kj}$, $j \in \Delta'$, $1 \leq k \leq n$,

then E is given by

$$E = \sum_{k=1}^{n} \sum_{j\in\Delta} a_{kj}^2 \qquad (12)$$

Q.E.D.

The author is not sure if the proposition is already known explicitly.

NOTE

The above simple proposition is extremely useful as, in case f_i's are orthonormal, the error E has a simple and useful form. A careful examination of E given by (12) shows that for the purposes of approximation, the ith column of the matrix A in (2) can be thought of a 'weight' of f_i.

We formalize it below.

DEFINITION

If f_i's in (2) are orthonormal, define

$$\text{weight } (f_j) = \sum_{i=1}^{n} a^2_{ij},\ 1 \leq j \leq \ell \tag{13}$$

Continuing with algorithm: if the signals in (2) are orthonormal, then using the above proposition, we have the 'weight' of each of the signals represented by the columns of A. We can pick the signals whose weights are maximum: i.e. we delete those which give us the error within certain reasonable limits.

In general, for a given system, the signals $f_1, \ldots f_\ell$ are not orthonormal. In that case, we replace f_i's by Ψ_i's by the Gram-Schmidt's orthogonalization process such that Ψ_i's are orthonormal.

By the standard Gram-Schmidt process, there exists a lower triangular $\ell x \ell$ matrix C such that

$$\begin{bmatrix} \Psi_1 \\ \Psi_2 \\ \vdots \\ \Psi_\ell \end{bmatrix} = C \cdot \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_\ell \end{bmatrix} \tag{14}$$

i.e.
$$\begin{bmatrix} f_1 \\ f_2 \\ \cdot \\ \cdot \\ f_\ell \end{bmatrix} = C^{-1} \cdot \begin{bmatrix} \Psi_1 \\ \Psi_2 \\ \cdot \\ \cdot \\ \Psi_\ell \end{bmatrix} \qquad (15)$$

Using (15), (2) can be written as

$$\dot{\underline{x}} = A.C^{-1} \cdot \begin{bmatrix} \Psi_1 \\ \Psi_2 \\ \cdot \\ \cdot \\ \Psi_\ell \end{bmatrix} \qquad (16)$$

Now in (15), Ψ_i's are orthonormal and, therefore, the columns of $A\ C^{-1}$ represent the 'weights' of Ψ_i's. Using the proposition, we can select a subset of Ψ_i's which will satisfy the error specifications.

Since each Ψ_i is a linear combination of f_j, $1 \leq j \leq i$, we have a subset of f_i's which will satisfy the error specification.

Summarizing, we have the following steps in the algorithm:

1. Replace the given set of signals $f_1, f_2, \ldots f_\ell$ in (2) by a corresponding set $\Psi_1, \Psi_2, \ldots, \Psi_\ell$ of orthonormal signals obtained by the Gram-Schmidt process. This will give us a new system equivalent to the given one.

2. Using the proposition, select a subset $(\Psi_{r_1}, \Psi_{r_2}, \ldots \Psi_{r_n})$ which will satisfy the error constraints.

3. Replace the selected subset of Ψ_i's in terms of f_i's to get

an approximating system in terms of f_i's.

V. An Illustrative Example

The above technique is illustrated by a discrete numerical example.

Consider the discrete system:

$$\begin{bmatrix} 1 & 2 & -1 \\ 2 & 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix} \tag{17}$$

where $f_1 = (1, 1, 0)$, f_2 $(1, 0, 2)$, $f_3 = (1, 1, -1)$.

Let Ψ_1, Ψ_2, Ψ_3 be an orthonormal set obtained from f_1, f_2, f_3 by using the Gram-Schmidt orthogonalization procedure.

We have:

$$\Psi_1 = \frac{f_1}{\sqrt{2}}$$

$$\Psi_2 = \frac{-\sqrt{2}}{6} f_1 + \frac{\sqrt{2}}{3} f_2 \tag{18}$$

$$\Psi_3 = \frac{-11}{3} f_1 + \frac{4}{3} f_2 + 3f_3$$

Equations (18) can be written as

$$\begin{bmatrix} \Psi_1 \\ \Psi_2 \\ \Psi_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & 0 \\ \frac{-\sqrt{2}}{6} & \frac{\sqrt{2}}{3} & 0 \\ \frac{-11}{3} & \frac{4}{3} & 3 \end{bmatrix} \cdot \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix} \tag{19}$$

or

$$\begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix} = \begin{bmatrix} \sqrt{2} & 0 & 0 \\ \frac{1}{\sqrt{2}} & \frac{3}{\sqrt{2}} & 0 \\ \sqrt{2} & \frac{-2\sqrt{2}}{3} & \frac{1}{3} \end{bmatrix} \cdot \begin{bmatrix} \Psi_1 \\ \Psi_2 \\ \Psi_3 \end{bmatrix} \qquad (20)$$

Using (20), (17) can be written as

$$\begin{bmatrix} 1 & 2 & -1 \\ 2 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} \sqrt{2} & 0 & 0 \\ \frac{1}{\sqrt{2}} & \frac{3}{\sqrt{2}} & 0 \\ \sqrt{2} & \frac{-2\sqrt{2}}{3} & \frac{1}{3} \end{bmatrix} \cdot \begin{bmatrix} \Psi_1 \\ \Psi_2 \\ \Psi_3 \end{bmatrix}$$

$$= \begin{bmatrix} \sqrt{2} & \frac{11.\sqrt{2}}{3} & \frac{-1}{3} \\ \frac{7.\sqrt{2}}{2} & \frac{5.\sqrt{2}}{6} & \frac{1}{3} \end{bmatrix} \cdot \begin{bmatrix} \Psi_1 \\ \Psi_2 \\ \Psi_3 \end{bmatrix} \qquad (21)$$

From (21), we have

$$\text{weight}\left(\Psi_1\right) = (\sqrt{2})^2 + \left(\frac{7.\sqrt{2}}{2}\right)^2 = 51$$

$$\text{weight}\left(\Psi_2\right) = \left(\frac{11.\sqrt{2}}{3}\right)^2 + \left(\frac{5.\sqrt{2}}{6}\right)^2 = \frac{509}{18}$$

weight $\left(\Psi_3\right) = \left(\frac{-1}{3}\right)^2 + \left(\frac{1}{3}\right)^2 = \frac{2}{9}$.

Thus the error will be $\frac{2}{9}$ if Ψ_1, Ψ_2 are selected to approximate (17).

Thus the approximating model is

$$\begin{bmatrix} \sqrt{2} & \frac{11.\sqrt{2}}{3} \\ \frac{7.\sqrt{2}}{2} & \frac{5.\sqrt{2}}{6} \end{bmatrix} \cdot \begin{bmatrix} \Psi_1 \\ \Psi_2 \end{bmatrix}$$

$$= \begin{bmatrix} \sqrt{2} & \frac{11.\sqrt{2}}{3} \\ \frac{7\sqrt{2}}{2} & \frac{5.\sqrt{2}}{6} \end{bmatrix} \cdot \begin{bmatrix} \frac{f_1}{\sqrt{2}} \\ -\frac{\sqrt{2}}{6} f_1 + \frac{\sqrt{2}}{3} f_2 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{-2}{9} & \frac{22}{9} \\ \frac{29}{9} & \frac{5}{9} \end{bmatrix} \cdot \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} \qquad (22)$$

VI. DESROCHERS AND AL-JAAR ALGORITHM

In this section, we explain the recently published algorithm developed by Desrochers and Al-Jaar. [3].

Let the given system be given by (2). The goal is to select a subset $\{f_{r_1}, f_{r_2}, \ldots f_{r_m}\}$ of $\{f_1, f_2, \ldots, f_\ell\}$ so that the error as defined by (9) between the given and simplified systems is minimum. Since an m-element subset of $\{f_1, f_2, \ldots, f_\ell\}$ can be chosen in

$\frac{m!}{\ell!\,(\ell - m)!}$ ways, it is very cumbersome to select an optimal subset consisting of m elements. To overcome this difficulty, Desrochers and Al-Jaar developed an algorithm of selecting the subset iteratively.

Let

$$\begin{bmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{n1} \end{bmatrix} \cdot \begin{bmatrix} f_i \end{bmatrix} \tag{23}$$

be a one-term approximant of (2) where $b_{11}, b_{21}, \ldots b_{n1}$ are unknown real parameters to be optimally determined. The error between (2) and (23) can be written as

$$E_i = \sum_{k=1}^{n} \langle (a_{k1} - b_{k1}) f_i, (a_{k1} - b_{k1}) f_i \rangle + 2 \langle (a_{k1} - b_{k1}) f_i, \sum_{j \in \Delta_i} a_{kj} f_j \rangle + \langle \sum_{j \in \Delta_i} a_{kj} f_j, \sum_{j \in \Delta_i} a_{kj} f_j \rangle \tag{24}$$

where $\Delta_i = \{1, 2, \ldots, \ell\} - \{i\}$ (25)

Equating to zero the partial derivative of E_i w.r.t. to b_{r1} and after algebraic manipulation, we get

$$b_{r1} = \frac{\sum_{j=1}^{\ell} a_{rj} \langle f_j, f_j \rangle}{\langle f_i, f_i \rangle}, \quad 1 \leq r \leq n \tag{26}$$

Substituting (26) into (24), we get the minimum value of E_i denoted by E_i^* and given by

$$E_i^* = \sum_{k=1}^{n} \left\| \sum_{j=1}^{\ell} a_{kj} f_j - \frac{\sum_{j=1}^{\ell} a_{kj} < f_i, f_j >}{< f_i, f_i >} f_i \right\|^2 \tag{27}$$

Similarly if (2) is approximated by one term model f_r, the corresponding optimal error E_r^* will be given by

$$E_r^* = \sum_{k=1}^{n} \left\| \sum_{j=1}^{\ell} a_{kj} f_j - \frac{\sum_{j=1}^{\ell} a_{kj} <f_r, f_j>}{<f_r, f_r>} f_r \right\|^2 \tag{28}$$

If f_i is the one term optimal approximant, we have

$$E_i^* \leq E_j^* \tag{29}$$

After algebraic manipulation, (29) can be written as

$$\sum_{k=1}^{n} \frac{\left(\sum_{j=1}^{\ell} a_{kj} < f_i, f_j > \right)^2}{\|f_i\|^2} \geq \sum_{k=1}^{n} \frac{\left(\sum_{j=1}^{\ell} a_{kj} <f_r, f_j> \right)^2}{\|f_r\|^2} \tag{30}$$

Equation (30) is extremely important in the algorithm of Desrochers and Al-Jaar in as much as it decides the optimal one term approximant

Define $$\beta_i = \sum_{k=1}^{n} \frac{\left(\sum_{j=1}^{\ell} a_{kj} \langle f_i, f_j \rangle \right)^2}{\|f_i\|^2} \qquad (31)$$

If $b_{11}, b_{21}, \ldots b_{n1}$ are the optimal parameters when f_j is the one term optimal approximant, to choose the second approximating signal, they first consider the system:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1\ell} \\ a_{21} & a_{22} & \cdots & a_{2\ell} \\ - & - & & - \\ - & - & & - \\ a_{n1} & a_{n2} & \cdots & a_{n\ell} \end{bmatrix} \cdot \begin{bmatrix} f_1 \\ f_2 \\ \cdot \\ \cdot \\ f_\ell \end{bmatrix} - \begin{bmatrix} b_{11} \\ b_{21} \\ \cdot \\ \cdot \\ b_{n1} \end{bmatrix} \cdot \begin{bmatrix} f_i \end{bmatrix} \qquad (32)$$

To choose the second approximating signal f_r, they work with (32). First they replace $f_1, f_2, \ldots f_{i-1}, f_{i+1}, \ldots f_\ell$ by signals $g_1, g_2, \ldots, g_{i-1}, g_{i+1}, \ldots g_\ell$ such that each g is orthogonal to f_i. Then they choose the second approximating signal using the criterion (30). This iteration goes on till the error constraint is satisfied.

VII. SUMMARY

In this report, a technique of simplifying a class of nonlinear systems in which the signals are continuous functions is presented. The technique is based on obtaining a new system equivalent to the given system by replacing the given signals by an orthonormal set of signals. The columns of the associated matrix in the new

equivalent system, acting as weights of the orthonormal signals, help decide the selection of a subset of the orthonormal set that will approximate the given system. The technique is illustrated by a numerical example.

The related simplification technique of Desrochers and Al-Jaar is also explained.

VIII. RECOMMENDATIONS

Since nonlinear systems are of great use to the U.S.A.F., it is important that efficient and accurate techniques of obtaining simpler nonlinear models of nonlinear systems be available. Not much work has been done in this direction. This summer, I have made a good start in developing a simplification technique for nonlinear systems. But due to the short period of time, it has not been possible to apply the technique to harder examples of interest. We, therefore, make the following recommendations:

1. Suitable software be developed so that the technique developed in this report be applied to the simplification of examples of interest.

2. The technique as presented in this report assumes that the relative dominance of the nonlinearities is known from physical considerations. Thus, besides the simplification technique, there is a need for more work to be done to analytically determine the relative dominance of the nonlinearities.

3. There is a need to apply the technique to design compensators for nonlinear practical systems of interest.

REFERENCES

1. C.F. Chen and L.S. Shieh, "A Novel Approach to Linear Model Simplifications", Int. J. Contr., vol. 8, pp 561-570, 1968.

2. E.J. Davison, "A Method for Simplifying Linear Dynamic Systems", IEEE Trans. Automatic Control, vol. AC-11, pp 93-101, 1966.

3. A.A. Desrochers and R.Y. Al-Jaar, "A Case For Nonlinear Model Simplification in the Design of Flight Control Systems," A.C.C. Conference, 1983, pp 788-793.

4. R. Genesio and M. Milanese, "A Note on the Derivation and Use of Reduced Model", IEEE Trans. Automatic Control, vol. AC-21, pp 118-122, 1976.

5. S.V. Rao and S.S. Lamba, "A New Frequency-Domain Technique for the Simplification of Linear Dynamic Systems", Int. J. Contr., vol 20, pp 71-79, 1974

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

The Effect of Large Deformation on the Fracture Mechanics of Solid Propellants

Prepared by:	David L. Questad
Academic Rank:	Assistant Professor
Department and University:	Department of Engineering Science and Mechanics The Pennsylvania State University
Research Location:	Edwards Air Force Base, Rocket Propulsion Laboratory, Solid Propellant Division, MKPB.
USAF Research:	Dr. C. T. Liu
Date:	8/22/83
Contract No.:	F49620-82-C-0035

The Effect of Large Deformation on the Fracture Mechanics of Solid Propellants

by

David L. Questad

ABSTRACT

Most solid propellants consist of an elastomeric binder containing a high volume fraction of filler particles. The propellants, therefore, are composite materials which, under deformation, are susciptible to microstructural damage, crack growth and fracture. The damage occurs mostly due to debonding at the filler/binder interface, and fracture may result from a single large flaw propagating through the damaged propellant or from the coalescence of many smaller cracks which have been induced by debonding. The deformation of a solid propellant involves nonlinear viscoelastic effects as well as irreversible deformation. A general theory of deformation, failure and fracture is impracticle, so a number of assumptions and simplifications are made in order to develop workable theories. This paper reviews some of the current thinking on failure and fracture in solid propellants.

ACKNOWLEDGMENTS

I would like to acknowledge and thank the Air Force Systems Command, Air Force Office of Scientific Research for their sponsorship of this program. I would also like to thank the people of the Solid Propellant Division of the Rocket Propulsion Lab and especially Dr. C. T. Liu, Mr. D. I. Thrasher and Mr. R. Biggers for their help prior to and during my summer research period.

Introduction

Many solid propellants are composite materials consisting of relatively rigid filler particles embedded in an elastomeric binder. These propellants are highly filled, containing perhaps 60 to 90% filler, and as a result the filler/matrix bond is the primary factor governing mechanical behavior and failure properties [1]. Newer, high elongation propellants without filler particles are capable of strains up to several hundred percent.

Experiments show that solid propellants are nonlinear, viscoelastic materials exhibiting permanent memory effects [1-3]. Two types of approaches to the characterization of nonlinear materials have been developed: 1) a rational approach in which a mathematical theory is developed, and from the theory appropriate experiments are suggested in order to completely characterize the material; and 2) an empirical approach in which experimental evidence suggests the form of the constitutive equations [4]. The first approach makes use of multiple integral equations of the form suggested by Green and Rivlin [5], but the number of experiments to fully characterize such materials is usually prohibitive [6]. For this reason, and because of the complexity of solid propellants, the latter approach is normally taken.

In service, solid propellants experience a wide variety of stress-strain and temperature histories. These histories may be constant over time or variable at different frequencies. The complexity of the theory needed to describe the propellant will depend on the complexity of the stress and strain states to which the propellant is exposed in service.

Most analytical studies of solid propellants use linear elastic constitutive equations in finite element programs such as TEXGAP [7]. Linear viscoelastic analysis has also been used [8,9], but the limitations on linear viscoelastic analysis as applied to propellant grains are not well defined. Laboratory tests show that moderate to large deformations produce a nonlinear response in solid propellants, but the extent to which these circumstances exist in propellant grains in not completely clear [10].

The major cause of nonlinear behavior in solid propellants is associated with mechanical damage [1,3,11,12]. This damage is usually

attributed predominantly to filler-binder dewetting (Mullin's Effect [13]) [1,3,15], although molecular bond scission may be important as well [15]. Accumulated damage ultimately leads to propellant failure.

Proper characterization of the damage has been recognized as an important step in characterizing nonlinear behavior and permanent memory effects [1,3]. This, in turn, will affect the failure properties as interpreted by fracture mechanics theories.

Objectives

The objective of this article is to examine current research on the effects of large deformation and nonlinear viscoelastic behavior on failure and fracture of solid propellant materials. As already mentioned, a general theory for nonlinear viscoelastic materials is too difficult to be practical, so the existing nonlinear theories have grown out of previous linear theories. This means that the nonlinear fracture theories are subject to certain limitations (e.g. strains which are not too large away from the crack tip, thermorheologically simple viscoelastic behavior, etc.). These assumptions are often necessary to make the theory tractable, but the assumptions must be weighed when applying the theories to real materials. More will be said about this in the Suggestions section.

Failure and Fracture

If a material contains no macroscopic flaws or cracks, then one normally speaks of a failure criterion for which some quantity reaches a critical value. This quantity may be a function of stress, strain, temperature, aging, etc. If a single, macroscopic flaw ultimately propagates and causes the structure to fail one speaks of fracture, and fracture mechanics is used to analyze this situation.

Smith [16] described tensile stress-strain failure properties of elastomers using a failure envelope which was independent of time and temperature. The Smith failure envelope has been shown to work well in describing strain rate and temperature dependence of failure in solid propellants [17]. The Smith failure envelope assumes that damage accumulates as a function of both stress and strain and a critical

amount of damage is reached at the boundary of the envelope. Strain-cumulative damage criteria have also been used to predict tensile failure in propellants [17,18]. In these criteria damage is assumed to increase linearly with applied strain. A stress-cumulative damage criterion, however, did not properly predict tensile failure in solid propellants [17]. Damage energy has also been used to predict uniaxial failure of solid propellants [15].

In order to encompass general states of stress in a failure criterion the concept of damage must be generalized. Schaeffer [19] proposed that failure would occur when the strain energy density, W, reached a critical value. This leads to an expression:

$$\frac{\tau_{OCT}}{\tau^{T}_{OCT}} = \left[\frac{1 + \frac{1}{3\beta(\alpha-1)^2}}{1 + \frac{\sigma_{OCT}/\sigma^{T}_{OCT}}{3\beta(\alpha - \sigma_{OCT}/\sigma^{T}_{OCT})^2}} \right]^{1/2} \qquad (1)$$

where α and β are constants, σ_{OCT} is the octahedral, normal stress:

$$\sigma_{OCT} = \frac{1}{3}(\sigma_1 + \sigma_2 + \sigma_3)$$

τ_{OCT} is the octahedral shear stress:

$$\tau_{OCT} = \frac{1}{3}\sqrt{(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2}$$

σ_1, σ_2 and σ_3 are principle stresses. σ^{T}_{OCT} is the octahedral normal stress in uniaxial tension (i.e. $\sigma^{T}_{OCT} = \sigma_t/3$), and τ^{T}_{OCT} is the octahedral shear stress in uniaxial tension (i.e. $\sqrt{2}\,\sigma_t/3$). Time and temperature effects can be taken into account by evaluating through the Smith failure envelope [19].

An expression for the Smith failure envelope to be used with Eq. (1) was developed from the experimentally observed relationship between crack growth rate, $\dot{a}$, and the mode one stress intensity factor, K_I [19]:

$$\dot{a} = A K_I^{B} \qquad (2)$$

where A and B are constants and A is rate and temperature dependent. Schapery has derived an expression of the form of Eq. (2) for linearly viscoelastic materials (see Eq. (95) ref. [20]) in which B is a material constant. The constancy of B has been demonstrated experimentally [21]. A similar expression has also been derived in terms of the J integral [22,23] (see Eq. (30):

$$\dot{a} = f_1 J^k \tag{3}$$

where f_1 and k are similar to A and B in Eq. (2). Following earlier work by Farris and Fitzgerald [24], Schapery used Eq. (3) to obtain another damage function as follows.

For proportional loading of nonlinear, power law materials J is proportional to $|\sigma|^{M+1}$, where σ is the time-varying magnitude of the stress tensor (i.e. $\sigma_{ij} = \sigma \hat{\sigma}_{ij}$ where $\hat{\sigma}_{ij}$ is time invariant). For isolated flaws, J is also proportional to the crack length, a. With these assumptions $J = f_3\, a\, |\sigma|^{M+1}$ [12], and integration of Eq. (3) yields:

$$\left[a_o^{K-1} (K-1)\, f_3^K \right]^{-1} = \int_0^{t_f} f_1 \, |\sigma|^{(M+1)K} dt \tag{4}$$

where a_o is the initial flaw size of a local flaw and t_f is the local failure time of that flaw. The right hand side of Eq. (4) is a measure of damage which must reach a certain value for a flaw to propagate. For a solid propellant with a distribution of particle (i.e. flaw) sizes Eq. (4) can be generalized. For the ith flaw in the material the damage parameter, S_i, is:

$$S_i = \int_0^{t_i} f_1 \, |\sigma|^{(M+1)K} dt \tag{5}$$

or

$$S_i = \int_0^{t_i} f_1 \, |\sigma|^{q} dt \tag{5a}$$

with $q = (M+1)k$.

For a wide distribution of flaws the S_i becomes essentially continuous and a continuous function, S, can replace the S_i in Eqs. (5) and (5a). It is more convenient to define a new damage parameter as the qth root of S or:

$$S_q = \left(\int_0^t f_1 |\sigma|^q \, d\tau \right)^{\frac{1}{q}} \qquad (6)$$

This is the weighted Lebesque norm of $|\sigma|$. Expressions for damage parameters similar to Eq. (6) can be obtained in terms of the strains as well as the stress. This was originally done by Farris and Fitzgerald [24].

In order to use this concept of a damage parameter to predict failure, it must be incorporated into a constitutive relation. This has been done to obtain nonlinear constitutive equations for solid propellants [1,25-27].

When accounting for damage-induced failure a limited number of flaws will ultimately grow to produce final fracture. By following the growth of these cracks, the concepts of fracture mechanics can be used to predict failure times [26].

Schapery's Theory

Schapery [26] uses the damage function in Eq. (6) to derive an expression for the probability of failure of a nonlinear viscoelastic material. The result is:

$$P(\Psi_f) = P_0 \left\{ \int_0^{\Psi_f} M^{\frac{1}{n}} \sigma^q \, d\Psi \right\}^{-\frac{1}{q}} \frac{M^{\frac{1}{n}} \sigma^q}{q \left\{ \int_0^{\Psi_f} M^{\frac{1}{n}} \sigma^q \, d\Psi \right\}^{\left(\frac{1+q}{q}\right)}} \qquad (7)$$

where ξ_f is the reduced failure time (i.e. $\xi_f = \int_0^{t_f} (1/a_T)\,dt$), n and q are constants ($q = 2(1 + 1/n)$). M is a function of aging (M=1 for no aging) and p_g represents a distribution function for the "worst" flaw which exists in each specimen of a large number of specimens. By worst flaw is meant that flaw which ultimately leads to failure. In order to calculate failure times one must evaluate n, p_g, M and a_T from experimental data (e.g. uniaxial creep tests) and then apply Eq. (7) for a particular probability, $p(\xi_f)$.

Theory of Bills

Bills and coworkers proposed a semiempirical theory for fracture of nonlinearly viscoelastic propellants [14,28]. This theory is based on the experimentally observed relationship between relaxation modulus, E(t) and fracture stress, σ_f. It produces an equation of the following form:

$$\frac{\sigma(\xi_f) - \sigma_\infty}{\sigma_g} = \left[\frac{E(\xi_f/q) - E_e}{E_g}\right]^{1/2} \qquad (8)$$

where ξ_f is, again, the reduced failure time (t_f/a_T), $\sigma(\xi_f)$ is the uniaxial failure stress, σ_∞ is the stress below which failure does not occur and σ_g is the glassy (i.e. low temperature) failure stress which is essentially constant. $E(\xi_f/q)$ is the relaxation modulus at time ξ_f/q ; q is an empirical shift factor; E_e is the equilibrium modulus and E_g the glassy modulus.

Eq. (8) can be expressed in terms of fracture mechanics parameters by expressing σ_g in terms of the Griffith expression as:

$$\sigma_g = \frac{K\,\Gamma^{1/2}}{c^{1/2}}\,E_g^{1/2} \qquad (8a)$$

where K is a constant, Γ the fracture energy per unit area of new crack surface and C is the crack length. Then Eq. (8) becomes

$$\sigma(\xi_f) - \sigma_\infty = \frac{K\,\Gamma^{1/2}}{c^{1/2}}\left[E(\xi_f/q) - E_e\right]^{1/2} \qquad (8b)$$

Eq. (8b) is a viscoelastic fracture criterion which contains the Griffith criterion as a special case.

Fracture Mechanics

When a material contains a macroscopic flaw and the applied stresses and strains cause this flaw to propagate to fracture, then the principles of fracture mechanics are more applicable to the failure process. The fracture mechanics approach attempts to characterize the mechanical state of the material at the crack tip and relate this to the propagation of the crack.

One can characterize this region in terms of the stress magnitude as measured by the stress intensity factors, K_I, K_{II} and K_{III} [29] or in terms of the energy needed to propagate the crack. The latter method takes a number of different forms such as energy release rates [30], a thermodynamic power balance [31] and the J integral [32,33].

Linear elastic analyses of solid propellants usually make use of stress intensity factors [21,34]. Linear viscoelastic analyses have used the thermodynamic power balance [31] and stress intensity factors [20]. For nonlinear analyses of propellants subjected to large deformations the thermodynamic power balance approach has been used [35] and, more recently the J integral method [22,23].

Hufferd's Theory

Hufferd and coworkers [35] proposed a theory for crack growth in nonlinear viscoelastic materials obtained from the thermodynamic power balance approach. This theory is based on a nonlinear, permanent memory constitutive equation [36] and results in an integro-differential equation relating crack growth rate to the state of stress and strain in the vicinity of the crack and to the specific fracture energy. For general stress states and crack orientations the governing equation is difficult to analyze. However, by making certain simplifications to the theory a simple crack growth formula results:

$$\lambda(t) = \frac{a(t)}{a(0)} = \cosh(Pt) \qquad (9)$$

where a(t) and a(o) are the crack lengths at times t and zero, respectively and P is the fracture dissipation power. In principle, P can be evaluated from the theory, but in practice it is an empirical quantity.

The theory was generalized to account for more complicated states of stress by introducing a damage parameter, d, incorporating the loading history, s:

$$d(t) = \max_{0 \le s \le \infty} \left[5 I_1^2(t-s) + 6 I_2(t-s) \right] \qquad (10)$$

where I_1 and I_2 are first and second strain invariants of the infinitessimal strain tensor.

In terms of d, Eq. (9) can be written:

$$\lambda = \cosh \left[C \left(\sqrt{d} - \sqrt{d_c} \right) \right] \qquad (11)$$

where d_c is a critical value of the damage parameter and C is an empirical quantity depending on the relaxation modulus and the specific fracture energy. The theory has been quite successful in predicting constant strain rate crack growth in uniaxial and biaxial tests on solid propellants [35].

Schapery's Theory

R. A. Schapery has published perhaps the most extensively developed theories of fracture in nonlinear viscoelastic materials. For this reason his theory will be outlined in some detail.

Schapery applied the J integral method to crack growth in nonlinear viscoelastic materials subjected to potentially large deformations [22,23]. The nonlinear theory involves the use of pseudo strains in order to solve the viscoelastic problem in terms of a reference elastic problem. The pseudo strain, ε_{ij}^e, is defined in terms of the actual strain, ε_{ij}, through a hereditary integral as follows:

$$\varepsilon_{ij}^e = \frac{1}{E_R} \int_0^t E(t-\tau, t) \frac{\partial \varepsilon_{ij}}{\partial \tau} \, d\tau \qquad (13)$$

where the terms in Eq. (17) refer to Fig. 1. C_1 is a curve connecting the bottom crack surface (pt.1) counterclockwise to the top crack surface (pt.2); ds is an increment of length along C_1; Φ is the pseudo strain energy density; T_i are the surface tractions applied across C_i and U_i^R are the pseudo displacements:

$$U_i^R = \frac{1}{E_R} \int_0^t E(t-\tau, t) \frac{\partial u_i}{\partial \tau} d\tau \qquad (18)$$

J_v is path independent [32] so it must be equal to another integral, J_f defined as:

$$J_f = \int_{C_2} \left(\Phi \, dx_2 - T_i \frac{\partial u_i^R}{\partial x_1} ds \right) \qquad (19)$$

where C_2 is a curve from pt. 1 to pt. 2 immediately adjacent to the crack. For a relatively long, slender crack the first term in Eq. (19) will be negligible compared to the second, so that for an opening mode stress state

$$J_f = -\int_{C_2} \sigma_f \frac{\partial \Delta u_2^R}{\partial x_1} ds = \int_0^{\alpha} \sigma_f \frac{\partial \Delta u_2^R}{\partial \xi} d\xi \qquad (20)$$

where ξ defines a new set of axes relative to the crack tip, and α is the length of the crack (i.e. the damage zone) (see Fig. 1). In Eq. (20) σ_f is the displacement which initially adjacent material points undergo as the crack advances.

In order to determine the time to initiate a crack, t_i, Schapery assumes that the failure stress distribution, σ_f, is constant along the crack and can be denoted by σ_m. Then Eq. (20) becomes:

$$J_f = \sigma_m \Delta u_{2\alpha}^R \qquad (21)$$

where $\Delta u_{2\alpha}^R$ is the opening displacement of the reference elastic problem at $\xi = \alpha$. Since $J_f = J_v$ then $\Delta u_{2\alpha}^R = J_v / \sigma_m$ and the time dependent crack opening displacement is given by the inverse of Eq. (18):

$$\Delta u_{2\alpha} = E_R \int_0^t D(t-\tau, t) \frac{\partial \Delta u_{2\alpha}^R}{\partial \tau} d\tau \qquad (18a)$$

where $E(t-\tau,t)$ is the relaxation modulus which may depend on aging as indicated by the second t in parentheses and E_R is a constant termed the reference modulus. The inverse of Eq. (13) gives the actual strain in terms of the pseudo strain:

$$\varepsilon_{ij} = E_R \int_0^t D(t-\tau,t) \frac{\partial \varepsilon^e_{ij}}{\partial \tau} d\tau \qquad (13)$$

where $D(t-\tau,t)$ is the creep compliance.

A consideration of a stress relaxation experiment reveals the usefulness of the pseudo strain concept. For stress relaxation (i.e. ε_{ij} constant after t = 0) Eq. (13) becomes:

$$\varepsilon^e_{ij} = \frac{E(t)}{E_R} \varepsilon_{ij} \qquad (14)$$

but

$$\sigma_{ij} = E(t)\,\varepsilon_{ij} = \varepsilon^e_{ij}(t)\,E_R \qquad (15)$$

Eq. (15) shows that the pseudo strain has the same time dependence as the stress. Eq. (13), therefore, has the effect of separating the viscoelastic effects from the constitutive equation. At any given time the stress, σ_{ij}, is related to the pseudo strain, ε^e_{ij}, by the same nonlinear constitutive equation. One can then solve a nonlinear viscoelastic problem by first solving the reference, nonlinear elastic problem in terms of the pseudo strains. The nonlinear, elastic problem can be solved by introducing the pseudo strain energy density, Φ analagous to the strain energy density in a purely elastic solution:

$$\sigma_{ij} = \frac{\partial \Phi}{\partial \varepsilon^e_{ij}} \qquad (16)$$

A theory for the fracture of nonlinear viscoelastic material can then be developed by incorporating the reference elastic solution into the J integral.

The J integral as originally defined by Rice [32,33] applied to materials which were linearly elastic away from the vicinity of the crack tip. For a viscoelastic material Schapery uses the J integral evaluated for the reference elastic solution. The result is:

$$J_v = \int_{C_1} \left(\Phi\, dx_2 - T_i \frac{\partial u^e_i}{\partial x_1} ds \right) \qquad (17)$$

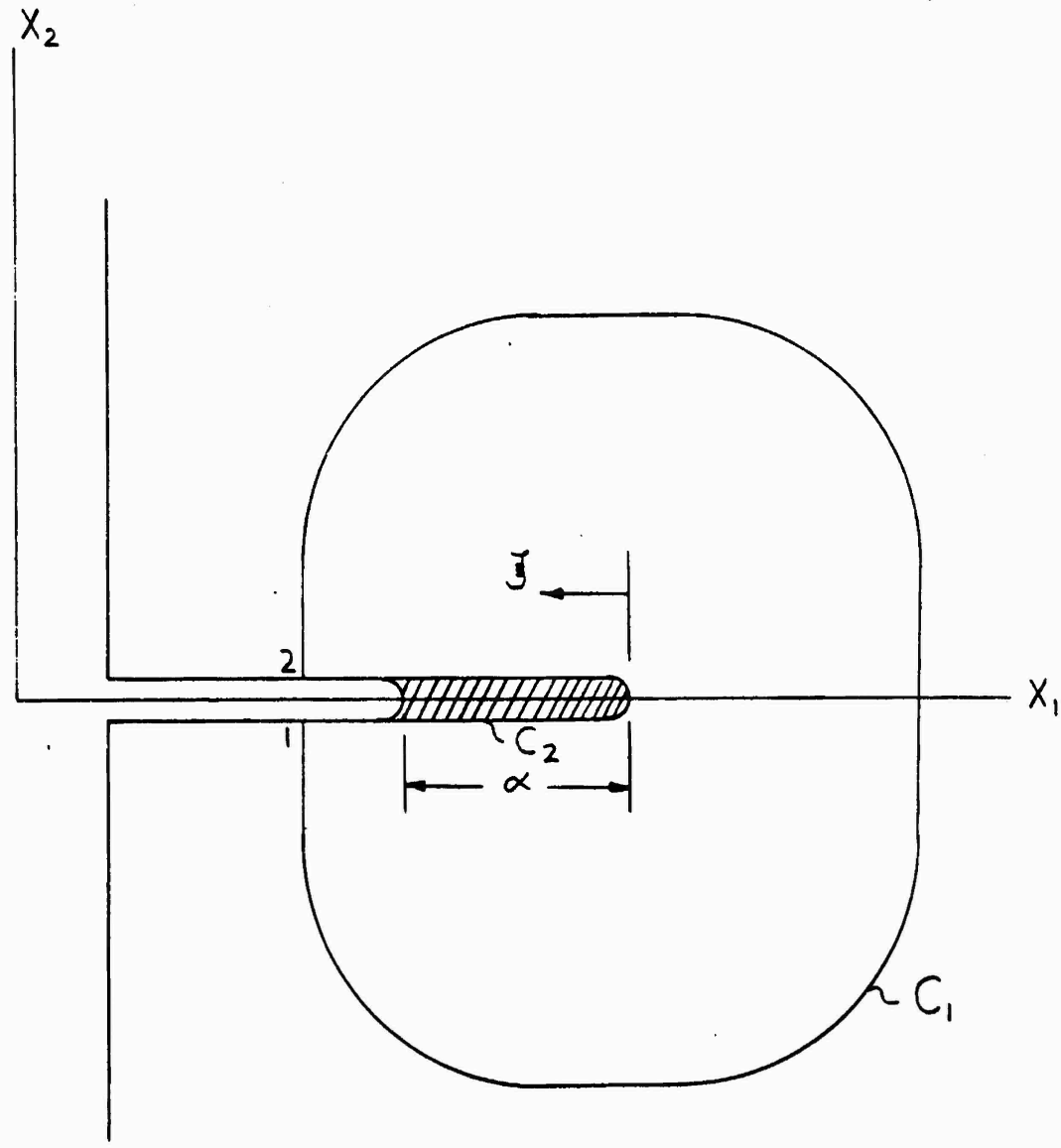


Figure 1 - Idealized crack tip for J integral calculations.

or $$\Delta U_{2\alpha} = E_R \int_0^t D(t-\tau, t) \frac{\partial (J_v/\sigma_m)}{\partial \tau} d\tau \qquad (18b)$$

If a crack opening criterion governs crack initiation then the time to initiate failure, t_i, can be calculated by setting $\Delta U_{2\alpha}$ to a critical crack opening deformation, ΔU_c:

$$\Delta U_c = E_R \int_0^{t_i} D(t-\tau, t) \frac{\partial (J_v/\sigma_m)}{\partial \tau} d\tau \qquad (22)$$

If, however, crack initiation is governed by an energy criterion then one has:

$$\int_0^{t_i} \sigma_{f\alpha} \frac{\partial \Delta U_{2\alpha}}{\partial \tau} d\tau = 2\Gamma_i \qquad (23)$$

where $\sigma_{f\alpha}$ and $\Delta U_{2\alpha}$ are the stress and deformation of a column of material at $\xi = \alpha$. When the work done per unit area on this column equals a critical value, $2\Gamma_i$, the crack will propagate. For $\sigma_{f\alpha} = \sigma_m$ and using Eq. (18b), Eq. (23) becomes:

$$\int_0^{t_i} \sigma_m \frac{\partial}{\partial t} \left[E_R \int_0^t D(t-\tau, t) \frac{\partial (J_v/\sigma_m)}{\partial \tau} d\tau \right] dt = 2\Gamma_i \qquad (24)$$

Crack growth velocity can also be determined using an approach similar to crack initiation. The crack will propagate when the work done on a column of material ahead of the crack tip reaches a critical value. This results in an expression similar to Eq. (23):

$$\int_0^{\alpha} \sigma_f \frac{\partial \Delta U_2}{\partial \xi} = 2\Gamma \qquad (25)$$

where the integration is now over the failure zone, α, and the fracture energy for propagation is 2Γ. Once again using Eq. (18) one finds the time-dependent crack opening displacement, ΔU_2, substitutes this into Eq. (25) to predict the time to propagate the crack a distance, α.

Schapery simplifies this approach considerably by assuming that σ_f and $\dot{a}$, the crack growth velocity, are constant during the time interval $\alpha/\dot{a}$; and that ΔU_2 in Eq. (18) simplifies to:

$$\Delta U_2 = E_R D(k\alpha/\dot{a}, t)\, \Delta U_2^R \qquad (26)$$

where $k = 1/3$

with these assumptions Eq. (25) becomes:

$$E_R D(k\alpha/\dot{a}, t) \int_0^{\alpha} \sigma_{\xi} \frac{\partial \Delta U_2^R}{\partial \xi}\, d\xi = 2\Gamma \qquad (27)$$

However, from Eq. (20) this becomes:

$$E_R D(k\alpha/\dot{a}, t)\, J_v = 2\Gamma \qquad (28)$$

For a power law creep compliance (i.e. $D = D_0 t^n$) and no aging Eq. (28) results in an expression for $\dot{a}$ as follows:

$$\dot{a} = \left(\frac{E_0 D_0}{2\Gamma}\right)^{\frac{1}{n}} K\alpha\, J_v^{\frac{1}{n}} \qquad (29)$$

or

$$\dot{a} = A\, J_v^{q}$$

The exponent, q, depends on the relation between J_v and α the constancy of A depends on the constancy of the fracture energy, Γ. Eq. (30) is the same as Eq. (3) which forms the basis of the nonlinear damage function.

Although the general theory is fairly well developed, its applicability depends on a sizeable amount of experimental characterization. More will be said about this in the final section.

Conclusions and Suggestions

As a result of studying the effects of large deformation on fracture in solid propellants a number of conclusions and suggestions seem appropriate. These are summarized as follows.

1. Nonlinear, large deformation fracture mechanics theories should continue to be developed and applied. This is especially true considering the development of high elongation propellants.
2. Experimental studies should be continued i) to provide necessary data for nonlinear, viscoelastic analysis, ii) to evaluate the correctness of existing nonlinear theories and iii) to dilineate the boundary between the simpler linear theories and the more complex

nonlinear ones. In many situations linear viscoelastic and even linear elastic analysis may be sufficient. Experimental evidence from multiaxial testing will indicate what type of theory will be required in individual cases. In this context linear fracture theories should also continue to be developed and applied.

3. As part of the experimental examinations some analysis should be made to determine to what extent solid propellents actually experience nonlinear, large, viscoelastic deformations in rocket motor configurations. Some analog failure studies have begun to make these analyses [37-39].

4. In depth investigations should be made to characterize the failure zone in front of the crack tip. Since this zone is directly responsible for crack growth, it is important to understand its behavior and morphology. Such studies may involve microscopy, ultrasonics, x-ray analysis, etc. Attempts to characterize the failure zone using ultrasonics have already begun [40].

5. Nonlinear, large deformation finite element programs (e.g. TEXLESP) should be extended to analyze fracture in such materials.

6. Aging effects should be incorporated into the fracture mechanics approach. This is especially true when large deformations are involved. The sensitivity of elastomers to environmental aging is greatly increased by strain, and this will be an important factor affecting crack growth.

Nonlinear viscoelastic theories for fracture analysis in materials subjected to large deformations are quite well developed. The theory of Schapery is probably the most highly developed. It appears, however, that the biggest obstacle limiting the application of such theories is the lack of experimental characterization of the materials.

Characterization of crack growth in nonlinear, viscoelastic materials over large strain ranges involves a considerable experimental program, but until such characterization is done, existing theories cannot be applied or refined. In light of the trends toward high elongation propellants such studies are clearly important.

REFERENCES

1. Farris, R. J. and Schapery, R. KA., "Development of a Solid Rocket Propellant Nonlinear Viscoelastic Constitutive Theory", Vol. 1 - Final Report, AFRPL-TR-73-50 (1973).

2. Carlton, C. H. and Francis, E. C., "Propellant Evaluation for Nonlinear Constitutive Theory Development", JANNAF Structures and Mechanical Behavior Subcommittee Meeting, (PIA Pub. No. 368 (1982) p. 143.

3. Gutierrez-Lemini, D. and Francis, E. kC., "Material Characterization and Stress Response of Nonlinear Solid Propellants", JANNAF Structures and Mechanical Behavior Subcommittee Meeting, CPIA Pub. No. 368 (1982) p. 371.

4. Lockett, F. J., "Experimental Characterization of Nonlinear Viscoelastic Materials", in Mechanics and Chemistry of Solid Propellants, ed. by A. C. Eringen, H. Liebowitz, S. L. Koh, J. M. Crowley, Pergamon Press, Oxford, New York, (1967), p. 333.

5. Green, A. KE. and Rivlin, R. S., Arch. Ration. Mech. Anal., 1 (1957), p. 1.

6. Arridge, R. G. C., Mechanics of Polymers, Clarendon Press, Oxford (1975), p. 77.

7. Francis, E. C., Carlton, C. H. and Thompson, R. E., "Predictive Techniques for Failure Mechanisms in Solid Rocket Motors", AFRPL-TR-79-87 Vol. 4, (1980).

8. Knuk, M. P. and Knauss, W. G., "Delayed Fracture in Viscoelastic-Plastic Solids", Int. J. Solids Structures, Vol. 6, (1970), p. 995.

9. Cheung, M. C. and Knauss, W. G., "Crack Propagation in Filled Elastomers", AFRPL-TR-70-143, CHECIT PL-70-3, part II, (1970).

10. Hufferd, W. L. and Bennett, S. J., "On the Development of an Aging Constitutive Theory", JANNAF OS and SMBWG, CPIA Pub. No. 264 (1975) p. 63.

11. Schwarzl, F. R., "On Mechanical Properties of Unfilled and Filled Elastomers", in Mechanics and Chemistry of Solid Propellants, ed. by A. C. Eringen, H. Liebowitz, S. L. Koh, J. M. Crowley, Pergamon Press, Oxford, New York (1967), p. 503.

12. Schapery, R. A., "On Viscoelastic Deformation and Failure Behavior of Composite Materials with Distributed Flaws", in Advances in Aerospace Structures and Materials, S. S. Wang, W. J. Renton, eds., ASME (1981) p. 5.

13. Mullins, L. J., "Effect of Stretching on the Properties of Rubbers", J. Rubber Res., 16, (1947), p. 275.

14. Bills, K. W. and Blatz, P. J., "Non-Linear Fracture Mechanics", Naval Weapons Center Report No. TP5684, (1975).

15. Lepie, A. H. and Adicoff, A., "Energy Balances and Uniaxial Failure of Propellants", JANNAF OS and SMBWG, CPIA Pub. No. 264 (1975) p. 207.

16. Smith, T. L., "Ultimate Tensile Properties of Elastomers. I. Characterization by a Time and Temperature Independent Failure Envelope", J. Polym. Sci., Part A, 1 (1963) p. 3597.

17. Nelson, J. M. and Kruse, R. B., "Evaluation of Failure Criteria for HTPB Propellants", JANNAF OS and SMBWG, CPIA Pub.No. 264 (1975) p. 233.

18. Anderson, G. P., Laheru, K. L., Christiansen, A. G., "Analog/Failure Integration Program Status, JANNAF Structures and Mechanical Behavior Subcommittee Meeting, CPIA Pub. No. 368 (1982) p. 195.

19. Schaeffer, B., "Fracture Criterion for Solid Propellants", in Advances in Research on the Strength and Fracture of Materials, Vol. 3B, ed. by D. M. R. Taplin, Pergamon Press, New York (1977), p. 1145.

20. Schapery, R. A., "A Theory of Crack Growth in Viscoelastic Media", Texas A & M University, Report No. MM2764-73-1 (1973).

21. Liu, C. T., "Crack Propagation in an HTPB Solid Propellant", JANNAF Structures and Mechanical Behavior Subcommittee 16th Mtg., Vol. 1, CPIA Pub. No. 311, (1979) p. 193.

22. Schapery, R. A., "Continuum Aspects of Crack Growth in Time Dependent Materials", Texas A & M Report No. MM4665-83-2 (1983).

23. Schapery, R. A., "Correspondence Principles and a Generalized J Integral for Large Deformation and Fracture Analysis of Viscoelastic Media", Texas A & M Report No. MM4665-83-7 (1983).

24. Farris, R. J. and Fitzgerald, J. E., "Deficiencies of Viscoelastic Theories as Applied to Solid Propellants", JANNAF, Mech. Behavior Working Group, 8th Mtg., CPIA Pub. 193 (1970).

25. Hufferd, W. L., Lindsey, G., Parmerter, R., Blatz, P., Knauss, W., Murch, S. A., Briggs, W. E., Schapery, R. A., Gurtin, M., Fitzgerald, E., "Predictive Techniques for Failure Mechanisms in Solid Rocket Motors", Vol. 5, AFRPL-TR-79-87 (1980) p. 955.

26. Schapery, R. A., "A Nonlinear Constitutive Theory for Particulate Composites Based on Viscoelastic Fracture Mechanics", CPIA Pub. No. 253 (1974) p. 313.

27. Schapery, R. A., "Models for Damage Growth and Fracture in Nonlinear Viscoelastic Particulate Composites", Proc. Ninth U.S. National Congress of Appl. Mech., ASME (1982).

28. Bills, K. W., "Non-linear Fracture Mechanics", JANNAF OS and SMBWG, CPIA Pub. No. 264, (1975) p. 219.

29. Sih, G. C. and Liebowitz, H., "Mathematical Theories of Brittle Fracture", in Fracture Vol. II, ed. by H. Liebowitz, Academic Press, New York (1968) p. 67.

30. Williams, J. G., Stress Analysis of Polymers, 2nd ed., Halstead Press, New York (1980) p. 291.

31. Williams, M. L., "Initiation and Growth of Viscoelastic Fracture", Int. J. Frac. Mech., 1(4) (1965) p. 292.

32. Rice, J. R., "A Path Independent Integral and the Approximate Analysis of Strain Concentration by Notches and Cracks", J. Appl. Mech. (1968) p. 379.

33. Rice, J.. R., "Mathematical Analysis in the Mechanics of Fracture", in Fracture Vol. II, ed. by H. Liebowitz, Academic Press, New York (1968) p. 191.

34. Liu, C. T., "Investigating the Statistical Distribution Function of Crack Growth Data", JANNAF Structures and Mechanical Behavior Subcommittee Meeting, CPIA Pub. No. 368, (1982) p. 153.

35. Hufferd, W. L., Jacobs, H. R. and Laheru, K. L., "On the Fracture of Viscoelastic Materials", JANNAF OS and SMBWG, CPIA Pub. No. 264 (1975) p. 339.

36. Liem, K., Hufferd, W. L. and Fitzgerald, J. E., "Memory Function Norms for Solid Propellants: Further Studies on the Character of the Damage Function", UTEC 74-082, University of Utah (1974).

37. Christiansen, A. G., Anderson, G. P., Laheru, K. L., "Propellant Wedge and Conical Bond Specimen for Failure Analogs", JANNAF SMBSM, CPIA Pub. No. 368 (1982) p. 39.

38. Hufferd, W. L. and Francis, E. C., "Failure Theory Evaluation with Analog Devices", JANNAF SMBSM, CPIA Pub. No. 368, (1982) p.195.

39. Anderson, G. P., Laheru, K. L. and Christiansen, A. G., "Analog/Failure Integration Program Status", JANNAF SMBSM, CPIA Pub. No. 368 (1982) p. 195.

40. Martinson, R. H. and Hartog, J. J., "Acoustic Imaging Applied to Crack Propagation Studies", JANNAF SMBSM, CPIA Pub. No. 311 (1980) p. 321.

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

THE APPLICATION OF AN EXTENDED KALMAN FILTER TO THE DESIGN OF A BANK-TO-TURN MISSILE AUTOPILOT

Prepared by:	Dr. Dallas W. Russell
Academic Rank:	Professor
Department or University:	Department of Electrical Engineering Auburn University
Research Location:	System Analysis and Simulation Branch Guided Weapons Division Armament Laboratory Eglin Air Force Base, Florida
USAF Research:	Capt. Randy Liefer
Date:	August 26, 1983
Contract No.:	F49620-82-C-0035

THE APPLICATION OF AN EXTENDED KALMAN FILTER TO THE DESIGN OF A BANK-TO-TURN MISSILE AUTOPILOT

by

Dallas W. Russell

ABSTRACT

In the past, missile autopilots have been designed using classical control techniques. The bank-to-turn missiles will require improved performance. Modern control and estimation theory are promising approaches to achieving improved performance. An Extended Kalman Filter that estimates system states and time-varying parameters has been designed for a low-order model with good results. It is recommended that this effort be continued beginning with simpler mathematical models of a generic missile; a set of equations, with numerical values for all the parameters is included.

Acknowledgement

This has been a very enjoyable and rewarding work experience. I appreciate the opportunity to have been a part of the USAF-SCEEE Summer Faculty Research Program.

The group I worked with this summer, the Research and Technology Section of the System Analysis and Simulation Branch, Armament Laboratory, Eglin Air Force Base, Florida, were very hospitable. In particular I wish to thank the following members of the section for their helpful suggestions and assistance: Capt. Randy Liefer, the Effort Focal Point, Mr.. Norman Speakman, Lt. L.C. Shelton, Mr. Jim Reed.

Mr. Norman Tew, the Auburn University graduate student who accompanied me, contributed much to the studies we undertook this summer.

I. INTRODUCTION

Current trends in air-to-air warfare indicate a need for highly maneuverable bank-to-turn (BTT) missiles [1]. The autopilot for such missiles must generate coordinated roll and acceleration responses while subjected to widely varying engagement conditions. To the present time there has not been a concentrated effort to apply modern control theory to the design of the autopilot. Modern control theory requires the accurate measurement or estimation of the system states. The Extended Kalman Filter (EKF) and similar filters have proven to be quite useful in state and parameter estimation [2], [3].

The design of most missile autopilots has been based on small perturbation theory and a gain scheduling technique to insure stability under the rapidly changing flight conditions. An overview of one such adaptive scheme is given in Appendix A. However some researchers [4], [5] have begun to combine the optimal and classical approaches. Their experience has been that modern control theory works well for coupled systems but high frequency effects can cause problems. An earlier study [6] of modern control techniques for the design of missile autopilots recommended the EKF for parameter estimation in conjunction with adaptive pole assignment as a viable approach.

II. OBJECTIVES

The overall objective was to study the design of autopilots for BTT missiles, in particular the application of modern control and estimation theory. The approach taken to achieve this objective was:

- Make a literature research on current practices and design.

- Evaluate a state-of-the-art autopilot to determine problem areas.
- Study the mathematical model of a typical missile.
- Apply optimal control and estimation theory to improve the design.

A key to design using modern control theory is an accurate estimate of the system states. The Kalman Filter is the most used technique to obtain accurate estimates of a system's states. Also, in missile autopilots, the fin effectiveness gains, which change with flight conditions, must be known rather accurately to ensure stability under all flight conditions. This is a parameter estimation problem for which the Kalman filter is well-adapted. Thus the research effort was concentrated on how to best apply a Kalman Filter to estimate the missile states and certain critical parameters.

If successful the Kalman Filter could replace the adaptive autopilot network described in Appendix A. Further, the states would be available for the feedback required in applying modern control theory. However this is not a simple task since the missile equations of motion are nonlinear and the parameters change as the missile maneuvers.

III. THE EXTENDED KALMAN FILTER (EKF)

There are a number of textbooks describing the EKF; for example, references [7]-[11]. Reference [3] is a special issue of the IEEE Transaction on Automatic Control that is dedicated to the applications of Kalman filters. The notation used herein is taken from the text by Gelb (Reference 7).

The system is modeled as follows:

$$\underline{x}(t) = \underline{f}(\underline{x}(t), t, \underline{u}(t)) + \underline{w}(t)$$

where $\underline{w}(t) \sim N(\underline{0}, Q)$

$\underline{x}(0) \sim N(\underline{x}(0), P_0)$

The measurement at discrete times is:

$$\underline{z}_k = \underline{h}(\underline{x}(t_k), t_k) + \underline{v}_k$$

where $\underline{v}_k \sim N(\underline{0}, R_k)$

To apply the filter it is necessary to model the system as a linear, discrete system. First, to linearize the equations:

$$\dot{\underline{x}}(t) = F(t)\underline{x}(t) + \underline{w}(t)$$

$$\underline{z}_k = H_k \underline{x}_k + \underline{v}_k$$

where $$F(t_k) = \left.\frac{\partial \underline{f}}{\partial \underline{x}}\right|_{\substack{\underline{x} = \underline{x}(t_k) \\ t = t_k}}$$

$$H_k = \left.\frac{\partial \underline{h}}{\partial \underline{x}}\right|_{\substack{\underline{x} = \underline{x}(t_k) \\ t = t_k}}$$

Next the continuous equation is converted to a difference equation,

$$\underline{x}_{k+1} = \Phi_k \underline{x}_k + \underline{w}_{k+1}$$

The Φ_k matrix as well as the process noise covariance matrix,

$$Q_k = E[\underline{w}_k \underline{w}_k^T]$$

are needed in the filter equations. A number of techniques to compute Φ_k and Q_k were evaluated. The simplest is the one step Euler Method.

$$\Phi_k = I + \Delta t\, F(t_k)$$

$$Q_k = \Delta t\, Q(t_k)$$

Several problems arise in applying the filter. All mathematical models are in error since some effects must be neglected so the system model can be reduced to a manageable set of equations. Some parameters may be known only approximately. Statistical information on the noise present, particularly the process noise, is usually not available. The presence of these inaccuracies cause degradation in the performance of the filter, possibly to the extent that the errors increase without bound.

The references list some corrective actions that can be taken to improve performance but emphasize that what works in one case may not apply to another system; some tuning or "fiddling" will be required. Also engineering judgement based on familiarity with the system will be required if pitfalls are to be avoided.

Some of the corrective actions to avoid poor performance by the filter are listed.

- change the process noise statistics
- adaptive adjustment of noise
- finite memory - over weight the most recent data
- exponential weighting of the data
- parameter estimation (and thus state estimation) is improved by appending the parameter as a state
- change measurement noise statistics

In particular Gelb [7] suggests that the variance associated with a parameter be set at

$$\frac{[\text{change expected in parameter in } \Delta t]^2}{\Delta t}$$

There are also variations of the filter equations such as square root filtering and more complex models.

IV. APPLICATION OF THE FILTER TO MISSILES

Initially the Kalman filter equations will be applied to a simplified set of missile equations. Computer time for a run will be short and the results can be compared to a "truth" model. The simplified models will be used to determine an appropriate sampling time and to tune the filter; that is adjust the initial state estimates and the process noise and measurement noise statistics for best performance. As satisfactory performance is achieved more complex system models will be used.

Step 1. A simplified set of longitudinal missile equations of motion will be programmed. To avoid table look up the stability derivatives will be approximated as linear functions of the missile velocity. The equations are given in Appendix B.

Step 2. Same as Step 1 except that the lateral equations of motion will be programmed. These are also given in Appendix B.

Step 3. The six-degree-of-freedom (6 DOF) missile equations of motion will be used. The stability derivatives will be approximated as before and angular rates will be kept small so body transformations to inertial coordinates are not required.

Step 4. The Armament Division in-house BTT missile simulation, flying a typical trajectory, will be the model for application of the Kalman Filter.

V. SIMULATIONS

An EKF was designed for a second-order linear system with one time varying coefficient. This coefficient was appended to the states resulting in a third-order, time varying, nonlinear system. The accompanying report [12] by Mr. Norman Tew, "Parameter Estimation via Kalman Filtering for Use in Bank-to-Turn Missile Autopilot Design", gives the results of these simulations.

VI. RECOMMENDATIONS

The simulations made using a low-order time varying system indicate that the addition of the time varying parameter as a state improves the performance of the Extended Kalman Filter. Follow-on research recommended is to proceed according to the steps outlined in Section IV of this report. A complete set of equations with numerical values is given in Appendix B. The procedure is to program the equations and make simulations to tune the filter. The effort should begin with the "truth" system model and the model used in the filter being in exact agreement. Then mismatches can be introduced and the resulting degradation in performance and possible corrective action can be studied. As successful designs are achieved the "truth" model can be made more complex.

REFERENCES

1. Modern Control Theory for Autopilot, R&D Statement of Work, DLMA83-43, Guided Weapons Division, AF Armament Lab, Eglin AFB, FL, 15 Nov. 1982.

2. Singer, A.E., Stefani, R.T., "Discrete Kalman Filtering Applied to Aerospace Vehicles with Widely Varying Parameters", Douglas Paper No. 4121, Control Theory Symposium, George C. Marshall Space Flight Center, Huntsville, Alabama, September, 1966.

3. Sorenson, H.W. (ed.), "Special Issue on Applications of Kalman Filtering", IEEE Trans. Automatic Controls, Vol. AC-28, No. 3, March, 1983.

4. Nesline, F.W., Zarchan, P., "A Classical Look at Modern Control for Missile Autopilot Design", AIAA Guidance and Control Conference, San Diego, California, pp. 90-104, August, 1982.

5. Nesline, F.W., Wells, B.H., Zarchan, P., "Combined Optimal/Classical Approach to Robust Missile Autopilot Design", J. Guidance and Control, Vol. 4, No. 3, pp. 316-322, May-June, 1981.

6. Price, C.F., "Advanced Concepts for Guidance and Control of Tactical Missiles", AD762593 NTIS, U.S. Dept. of Commerce, June, 1973.

7. Gelb, Arthur, Applied Optimal Estimations, Cambridge, MASS., M.I.T. Press, 1974.

8. Jazwenski, A.H., Stochastic Processes and Filtering Theory, New York City, Academic Press, 1970.

9. Brown, R.G., Introduction to Random Signal Analysis and Kalman Filtering, New York, John Wiley and Sons, 1983.

10. Maybeck, P.S., Stochastic Models, Estimation, and Control, Volume 2, New York, Academic Press, 1982.

11. Bierman, G.J., Factorization Methods for Discrete Sequential Estimation, New York, Academic Press, 1977.

12. Tew, N., "Parameter Estimation via Kalman Filtering for Use in Bank-to-Turn Missile Autopilot Design", SCEEE Graduate Student Summer Support Program, Eglin AF Base, 1983.

13. Ehrich, R.D. et al, "Detailed Stability and Control Investigations of a Bank-to-Turn (BTT) Configuration", AFATL-TR-78-10 (unclassified), AF Armament Laboratory, Eglin AF Base, Florida, January, 1978.

APPENDIX A

Overview of Rockwell Autopilot

Adaptive Network

The bank-to-turn missile autopilot described in [13] uses an adaptive network to maintain the gains at the crossover frequency of each control loop at a constant or almost constant value, thus achieving stability at all flight conditions. This is accomplished by adding a sinusoidal input (dither signal) to the yaw rate loop at the crossover frequency. The open-loop gain is measured at the crossover frequency and this is used to generate a signal labeled K_r. Linear functions of Kr are used to schedule all the loop gains in the autopilot.

The inputs to the adaptive network are the yaw rate loop error signal and the yaw rate gyro output. These two signals are fed to identical band-pass filters to recover the sinusoidal component. Each is then fed to a full-wave rectifier and the difference through a noise filter to a lag network. Let this signal be r_e-r_g where

r_e = constant x yaw rate error signal

r_g = constant x yaw rate gyro output

K_r is generated with the lag network shown in Figure A-1.

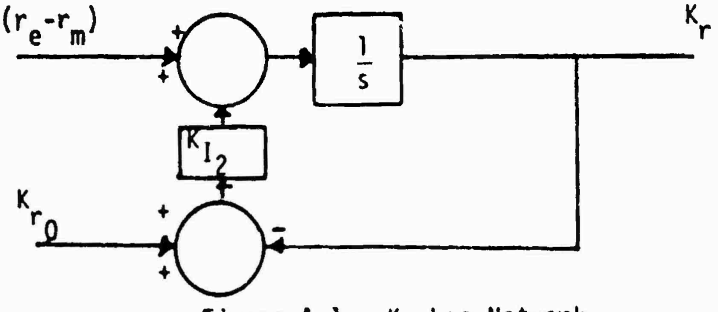


Figure A-1. K_r Lag Network

If the dynamics are neglected then the inputs to the integrator must sum to zero. This yields

$$K_r = \frac{r_e}{K_{I_2}} \left(1 - \frac{r_g}{r_e}\right) + K_{r_0} \qquad \text{(A-1)}$$

where r_g/r_e is recognized as the open-loop gain at the crossover frequency. Let $\frac{r_g}{r_e} = K_r K_c K_m$ where

K_c = the fixed control gains at the crossover frequency

K_m = the missile gain at the crossover frequency

The purpose of the adaptive scheme is to generate K_r so $K_r K_m$ is a constant. Solving (A-1) for K_r,

$$K_r = \frac{K_{r_0} + r_e/K_{I_2}}{1 + \frac{K_c r_e}{K_{I_2}} K_m} \qquad \text{(A-2)}$$

If the second denominator term is much larger than unity, then K_r is proportional to $1/K_c K_m$, the desired objective.

Based on small perturbation theory the control effectiveness gains at the crossover frequencies are

$$K_m(\text{yaw}) = \frac{r}{\delta r} = \frac{C_{n_{\delta q}} q_\infty S\ell}{\omega_c I_y}$$

$$K_m(\text{pitch}) = \frac{q}{\delta q} = \frac{C_{m_{\delta q}} q_\infty S\ell}{\omega_c I_y}$$

$$K_m(\text{roll}) = \frac{p}{\delta p} = \frac{C_{\ell_{\delta p}} \; q_\infty \; S\ell}{\omega_c \; I_x}$$

and

$$K_r = \frac{1}{K_m(\text{yaw})} = \frac{\omega_c \; I_y}{C_{n_{\delta q}} \; q_\infty \; S\ell}$$

As can be seen in Figure A-2, $|C_{n_{\delta r}}|$ and $|C_{m_{\delta q}}|$ are quite similar in form with $|C_{\ell_{\delta p}}|$ somewhat different. As a consequence the multiplier in the roll loop is a more involved function of K_r than that for the pitch loop.

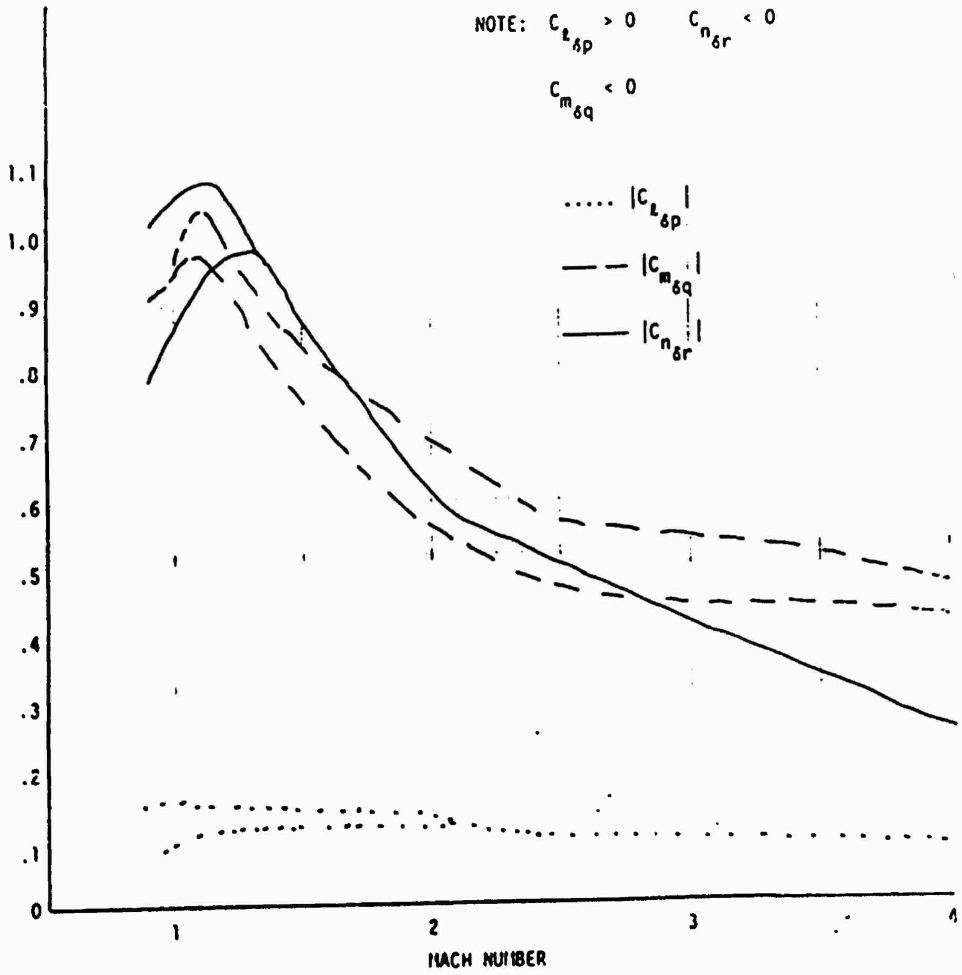


Figure A-2. Control Effectiveness Gains

APPENDIX B

Simplified Equations of Motion for the

Bank-to-Turn Missile

The simplified longitudinal equations of motion are based on the three equations,

$$m[\dot{U} + wq] = ma_x = (C_X + C_{XT})q_\infty S + \text{Thrust}$$

$$m[\dot{w} + Uq] = ma_x = (C_{z_\alpha}\ \alpha + C_{z_{\delta q}}\ \delta q)\ q_\infty S + mg$$

$$I_y \dot{q} = (C_m + C_{m_{\delta q}} + \frac{C_m \ell}{2V_m} q)\ q_\infty S\ell$$

which are taken from [1]. The notation is that in general use with the standard set of units.

Initial conditions:

$U(0) = 900$; $h(0) = 10{,}000$; $\rho_0 = .1756 \times 10^{-2}$

$\alpha(0) = 5°$; $w(0) = -40$; $q(0) = 0$;

$\delta q(0) = -1°$

Parameters: $m = m(0) = 5.75$

$I_y = I_y(0) = 34.1$

The C's are approximated as linear functions of V_m using data from the tables in [1]. Thrust = 9250

$0 \leq t \leq .6$ sec

The "truth" model for the longitudinal dynamics

$$\dot{U} = -wq + [-1800 + .35\ V_m](.04)(10^{-6})V_m^2 + 1610$$

$$\dot{w} = Uq + [(-44000+3\ V_m)\frac{w}{U} + (-200+.04\ V_m)\delta_q](.04)10^{-6}V_m^2+32.17$$

$$\dot{q} = [-1300+.35\ V_m+(-1200+.2\ V_m)\delta_q+\frac{(-4400+.8V_m)}{V_m}q](.0045)(10^{-6})V_m^2$$

where $V_m^2 = U^2 + w^2$; $\alpha = \frac{w}{U}$

and $\delta_q(t)$ is an input to be specified.

Measurements, the longitudinal dynamics

$$\underline{z} = \begin{bmatrix} a_x \\ a_z \\ q \end{bmatrix} = \begin{bmatrix} (-1800+.35\ V_m)(.04)(10^{-6})V_m^2+1610 \\ [(-44000+3\ V_m)\frac{w}{U}+(-220+.04\ V_m)\delta q](.04)(10^{-6})V_m^2+32.17 \\ q \end{bmatrix}$$

The simplified lateral equations of motion are also taken from [1].

$$m[\dot{U} - Vr] = ma_x = (C_x+C_{xT})q_\infty S + \text{Thrust}$$

$$m[\dot{V} + Ur] = ma_x = (C_{y_B} B + C_{y_{\delta r}} \delta_r)q_\infty S$$

$$I_x\dot{p} = [C_{\ell_B} B + C_{\ell_{\delta p}} \delta_p + \frac{C_{\ell_p}\ell}{2V_m} p]q_\infty S\ell$$

$$I_z\dot{r} = [C_{n_B} B + C_{n_{\delta r}} \delta_r + \frac{C_{n_r}\ell}{2V_m} r]q_\infty S\ell$$

Initial conditions:

$U(0) = 900$; $h(0) = 10{,}000$; $\rho_0 = .1756\times10^{-2}$

$\alpha(0) = 0$; $\delta p(0) = 0$; $\delta r(0) = 0$; $B(0) = 0$

$V(0) = 0$; $p(0) = 0$

Parameters:

$$m = m(0) = 5.75$$

$$I_z = I_z(0) = 34.1$$

$$I_x = I_x(0) = .34$$

As for the longitudinal set of equations, the C's are approximated as linear functions of V_m using data from the tables in [1]. Thrust = 9250

$$0 \leq t \leq .6 \text{ sec}$$

The "truth" model for the lateral dynamics

$$\dot{U} = +Vr + [-1800+.35\ V_m](.04)10^{-6}V_m^2+1610$$

$$\dot{V} = -Ur + [215-.04\ V_m](.04)(10^{-6})V_m\ \delta r$$

$$\dot{p} = [(188-.028\ V_m)\delta p + \frac{(-1715-.12\ V_m)\ p}{V_m}](.45)(10^{-6})V_m^2$$

$$\dot{r} = [(82100-21.2\ V_m]\ \frac{v}{U} + (-1188+.22\ V_m)\ \delta r$$

$$+(-4620+09\ V_m)\ \frac{r}{V_m}](.0045)(10^{-6})V_m^2$$

Measurements, the lateral dynamics

$$\underline{z} = \begin{bmatrix} a_x \\ a_y \\ p \\ r \end{bmatrix} = \begin{bmatrix} (-1800+.35\ V_m)(.04)(10^{-6})V_m^2+1610 \\ (215-.04\ V_m)(.04)(10^{-6})V_m\ \delta_r \\ p \\ r \end{bmatrix}$$

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

ANALYSIS OF PILOT SELECTION DATA

Prepared by:	Herman F. Senter
Academic Rank:	Associate Professor
Department and University:	Department of Mathematical Sciences Clemson University Clemson, South Carolina
Research Location:	Manpower and Personnel Division Human Resources Laboratory Brooks Air Force Base, Texas
USAF Research Contact:	Dr. Jeffrey E. Kantor
Date:	July 22, 1983
Contract No:	F49620-82-C-0035

ANALYSIS OF PILOT SELECTION DATA

by

Herman F. Senter

ABSTRACT

Historical data on test performance scores and the individual characteristics of student pilots were analyzed to identify measures most highly associated with success in pilot training school. Groups of scores from psychomotor tests, from flight training lessons and from the Air Force qualifying examinations were studied separately as well as jointly for their relationship to training school attrition. A regression equation to predict the likelihood of success in training based on selected grades and characteristics of individuals was developed.

ACKNOWLEDGMENTS

Throughout the project the investigator worked jointly with graduate student Steve Ramsier from the Department of Mathematical Sciences at Clemson University. Steve made significant contributions to all aspects of the project; especially noteworthy was his rapid mastery of the computer system at the Air Force Human Resources Laboratory (AFHRL) that expedited use of those facilities.

Dr. Jeffrey Kantor and Major Paul Bordelon of AFHRL provided valuable guidance and assistance for the project. Major Bordelon was actively involved in specific data analyses, offering practical insights and explanations based on his knowledge of Air Force procedures.

Lt. Richard Brown of the Manpower and Personnel Division at AFHRL was most generous with advice and assistance on use of the AFHRL computing facilities.

The investigator deeply appreciates the contributions of each of the above individuals to this project. In addition he wishes to thank the personnel of AFHRL and Brooks AFB for supporting the Summer Faculty Research Program, and for an enjoyable and productive ten-week period. He also acknowledges with gratitude the sponsorship of the Air Force Systems Command and the Air Force Office of Scientific Research.

I. INTRODUCTION:

This report summarizes the findings of an exploratory analysis of student pilot selection data conducted for the Aircrew Selection and Classification Function of the Air Force Human Resources Laboratory (AFHRL) at Brooks AFB, Texas. A detailed discussion of the procedures and findings of the entire investigation would be quite lengthy and is not given; rather an overview of the study methods is provided, supported by specific results which typify the procedures used and results obtained. The considerable volume of numerical results generated by the exploratory approach cannot be displayed in the report. Instead, computations are contained in several unattached appendices consisting of computer printouts. Contents of these indexed appendices, listed at the end of this report, constitute a "trail" of the overall study by which the steps of the analysis may be followed. Throughout the report references to relevant items in the appendices will be made parenthetically as an appendix letter designation followed by an item number; e.g., the notation "(C-7)" directs the reader to Appendix C, item 7.

Data used in this study were compiled and edited at AFHRL as part of Phase I of the Pilot Selection and Classification project. Ample documentation of variable codes, names and labels is on-hand at AFHRL and will not be reproduced in this report. The variable labels used here are identical to those assigned at AFHRL. Any departures from established names, codes or labels are noted; new variables are defined as they first appear in the report.

The items used in this study were taken from two subfiles -- the CPF and PPF files -- of the presently incomplete Phase I file, and consisted of 6,647 individual cases with fifty-two variables per case (exclusive of missing values). The statistical analysis was performed using SPSS, Version 9, on the UNIVAC system at AFHRL.

II. OBJECTIVES OF THE RESEARCH EFFORT:

This study was undertaken as part of AFHRL's ongoing Pilot Selection and Classification Project to develop and evaluate criteria for choosing candidate pilots and assigning them to specific training programs. Under Phase I of the project, historical data on more than 6,600 student pilots have been gathered, with certain items edited and recorded in two subfiles, designated as the CPF and PPF files. The goal of this investigation was to develop a multiple linear regression model for predicting student pilot success (or attrition) in Undergraduate Pilot Training (UPT) based on test performance scores and individual characteristics available from the CPF and PPF files.

Objectives of the study required several subanalyses of groups of test scores, of individual variables, and of the relationships of these variables with one another and with UPT attrition. Specific objectives were as follows:

(1) an analysis of the interrelationships of five psychomotor scores and their predictive validity relative to UPT attrition;

(2) an investigation of FSP data to develop a single measure or index of FSP performance based on class ranks and to identify those lesson grades most indicative of potential for UPT success;

(3) an examination of the relationship to UPT attrition of the AFOQT scores and other predictor variables; and,

(4) specification and computation of a regression equation for estimating the likelihood of UPT success from selected predictor variables.

The regression model derived is called the "interim model" for two reasons: it will be updated and improved as additional Phase I data are available, and the criterion variable -- UPT attrition -- will be replaced by a measure of actual performance as a operational pilot. Analysis of psychomotor test scores was of particular interest since a new series of tests (the BAT tests), to be administered by electronic digital devices, are being formulated and evaluated.

III. ANALYSIS OF PSYCHOMOTOR SCORES:

Psychomotor test scores consist of five ratio-scale values, coded as variables V15 to V19 in the Composite Predictor File (CPF). Two of these -- V15 and V16 -- are measures of horizontal and vertical tracking deviations recorded from the first of two tests, and are also denoted 1X and 1Y. The remaining three scores from the second test -- V17 to V19 -- consist of horizontal and vertical error sums associated with a joystick control and with rudder (foot) pedals, denoted 2X, 2Y, and 2Z. All scores represent accumulated root mean-square error from perfect tracking measured during the tests; thus, larger scores correspond to poorer performance. All scores are positive, ranging from the low-hundreds to the 70-thousands. All five scores were available for 2,527 individuals, most of whom went on to pilot training after testing.

III.I Univariate Distributions of Psychomotor Scores

Summary statistics for the five scores are given in Table 1 (A-1). Scores 1X and 1Y from the first test are similar in scale and distribution; mean scores from the second test are consistently smaller than those of the first test but are comparable among themselves. All five distributions exhibit strong positive skewness; the extremely long right tails are evident in the range of values above the mean.

TABLE I SUMMARY STATISTICS FOR PSYCHOMOTOR SCORES

(N = 2527)

VARIABLES	TEST SCORE	MEAN	STD. DEV.	MIN.	MAX.
V15	1X	14719	5407	428	65268
V16	1Y	16749	5546	5687	57292
V17	2X	4513	6520	326	71040
V18	2Y	3858	6445	486	76320
V19	2Z	5576	6015	167	71040

Each of the five test measures produced some extremely large as well as a few exceptionally small scores; however, the very large values were much more dispersed as indicated by the long right tails of the distributions (A-3, A-4).

To indicate the frequencies of very large scores, the upper 1%, 5% and 10% values (percentiles) are given in Table II. In some instances individuals with one large score, say 1X, would have a corresponding extremely small 1Y score (see scatterplots, A-5). Distributions of V17 and V18 show "spikes" -- a cluster of a dozen scores -- on their extremes right tails.

Table II. DISTRIBUTION OF LARGE SCORES

(N = 2527)

VARIABLE	TEST SCORE	$\bar{x} + 4s$	1%	5%	10%
V15	1X	36347	31000	25000	22000
V16	1Y	38933	34000	27000	24000
V17	2X	30593	29000	14000	9500
V18	2Y	29638	24000	11000	7000
V19	2Z	29636	30000	18000	13000

An analysis of extreme scores was made to seek an explanation for these exceptional values. There was no evidence of failure of the testing devices or of data recording or editing mistakes. Some

plausible explanations for the deviate scores are that individuals may have adopted a strategy of focusing attention on one direction of control while ignoring the other; that an individual might lack eye-motor coordination of a partricular sort; that an individual may not have made a sincere effort to perform well.

III.2 Bivariate Distributions

The relationships among test scores were examined by computing pairwise correlations (Pearson's) and constructing scatterplots (A-5). Scores within each of the two tests were much more highly correlated than were scores between the tests as shown in Table III.

TABLE III. PEARSON CORRELATION COEFFICIENTS FOR PSYCHOMOTOR SCORES

VARIABLE	V15	V16	V17	V18	V19
V15	1.00	.84	.17	.15	.21
V16	.84	1.00	.21	.16	.25
V17	.17	.21	1.00	.39	.55
V18	.15	.16	.39	1.00	.46
V19	.21	.25	.55	.46	1.00

The scatterplot of V15 versus V16 exhibited the strongest linear relationship. Plots of V15 and of V16 against V17, V18 and V19 were all quite similar showing considerable dispersion and little evidence of a relationship. Plots of V17, V18, V19 against one another were similar, giving evidence of a general positive relationship with substantial scatter. None of the scatterplots indicated non-linearity in the pairwise relationships of scores. However the presence of small clusters of points with one extremely large score suggested that the slope of a fitted regression line could be significantly affected by these few observations and that omission of exceptionally large scores should be considered.

III.3 UPT Criterion

A main objective of the psychomotor test analysis was to determine how well the scores predicted student pilot performance measured in terms of completing Undergraduate Pilot Training (UPT). Using training outcome as a criterion, performance was broken down into three categories: passed UPT (V06 =1), flying deficiency failure (V06 = 0) and "other elimination" from UPT (V06 = 9). Means for each of the five scores show significant differences among those passing and failing UPT (A-6). In every test the means of those passing UPT were substantially less than for those that failed; and, except in one test (V17) the mean scores of the flying deficiency failures (V06 = 0) were greater than the corresponding means for other eliminations (V06 = 9). Such large differences in mean scores of passing and failing students suggests that psychomotor test results are indicative of success potential.

Further examinations of psychomotor scores against the UPT criterion V06 were made using crosstabulations of scores by training outcome (A-6 to A-11). Particular attention was given to cases with exceptionally large scores on one or more tests. Elimination rates for individuals with very large scores on at least one test were substantially above those with no extreme scores. Figures in Table IV indicate the differences in passing rates.

TABLE IV. UPT SUCCESS RATES BY TEST SCORE SIZE

VARIABLE	TEST SCORE	PASSED UPT (%)	No. OF CASES
V15	under 25000	78.6	1677
	over 25000	62.2	74
V16	under 25000	78.7	1618
	over 25000	68.4	133
V17	under 10000	79.3	1619
	over 10000	60.0	130
V18	under 10000	79.0	1676
	over 10000	52.1	73
V19	under 10000	79.5	1539
	over 10000	66.2	210

An investigation of the relatively few number of cases with one or more very small test scores showed no apparent departure in mean UPT passage rate for these students from the norm.

III.4 Extreme Values

A thorough study of individual cases with unusually large or small scores was conducted. The variation and the inconsistency in UPT success associated with very large scores led to the conclusion that such values were not reliable indicators of success potential, and that in future data gathering, a policy of automatically retesting students with very large scores be considered. Values which lie approximately four standard deviations above the mean test scores are recommended as cutoff points for defining outliers or establishing retesting criteria. Thus, cutoff scores of 40,000 for V15 and V16 and of 30,000 for V17 to V19 are suggested. These values are used in some of the regression analyses to improve the quality of predictive relationships (B-10, C-3). No basis for analogous cutoffs on the low end of the scale was evident, and therefore, no lower end cutoffs are recommended.

III.5 Regression of UPT on Psychomotor Scores

Stepwise (forward) regressions were run using both UPT pass/elim (all eliminations) and UPT pass/fail (flying deficiency failures) as the dependent variable against the five psychomotor scores (V15 to V19). The simple additive models are referred to as "main effects" models (Appendix B). Since extremely large scores were suspected of having unwarranted influence on the estimated coefficients, runs were made first on all cases (B-1) then deleting cases with scores in the upper 1%, 5%, and 10% of the distributions (B-3 to B-5). The SPSS stepwise procedure automatically terminated when entry of any additional independent variables would produce insignificant t-values at the .05 level. Cases with a missing score were deleted.

Prior considerations suggested potential interaction among test scores for the five tests. Thus stepwise regressions with all simple interactions were run against the two dependent criteria (B-2, B-6 to B-8). Finally regressions were performed omitting all cases with at least one test score more than 4 standard deviations above the mean (B-10).

Among the regressions run, some variation in the variables entering and in the order of entry was observed. In general, however, most of the variance was explained by the first two entering variables, consisting of one score from each of the two types of tests or an interaction of two such scores. Variable V15 was the best predictor from the first test and, with outliers excluded, V18 was best among the second test scores.

Multiple-R values on the order of .20 were the general rule with higher values for predicting UPT pass/fail than UPT pass/elim. This suggests psychomotor tests may be somewhat better predictors of flying deficiency failure than of all types of eliminations.

With outliers excluded the first variable entering the main effects model was V18 followed by V15. The interaction terms V1518 contributed little additional information; thus the simple main effects model with independent variables V15 and V18 was preferred, with estimated coefficients computed excluding outliers (i.e., V15, V16 less than 40,000 and V17 to V19 less than 30,000).

III.6 Transformed Psychomotor Scores

The extremely long right tails of the distributions of psychomotor scores suggested that logarithmic transformations might produce more nearly normal distributions. Indeed this was the case as indicated by the results in Appendix C-1 for the natural logarithms of the scores (denoted LnV--). However, pairwise correlations between transformed scores and UPT pass/elim were no better than for raw scores, and regressions using the transformed variables (C-2, C-3) produced results quite close in terms of multiple-R and standard error of residuals to those for the raw scores. Thus, untransformed psychomotor scores were used in subsequent analyses.

III.7 Factor Analysis of Psychomotor Scores

An examiniation of the five test scores in the absence of a criterion variable was done using principal factoring with iteration (SPSS procedure PA2) followed by three types of rotations. The analyses (D-1 to D-3) indicate two significant factors, one weighted heavily on the scores of the first type of test -- V15 and V16 -- and the other weighted mainly on scores from the second test, V17, V18 and V19.

III.8 Sources of Variation in Psychomotor Scores

Means of psychomotor scores at each level of several categorical variables were computed (using SPSS procedure BREAKDOWN) to determine variables and interactions that should be included in an initial model specification for predicting UPT. A number of interesting questions were prompted by these analyses which could not be addressed in the limited time frame of this study; more extensive analyses are recommended for follow-on research.

Table V shows for each of the five scores whether significant differences in mean scores at the various levels of the variables were significant (+) or not (o) with $\alpha \leq .05$. The number of active cases for each variable is indicated (N). Some of the main results and differences observed are briefly outlined below:

FSP Status, V05. For the first test scores (V15, V16) FSP failures, V05 = 0, were lowest in mean score while V05 = 5 had the highest means. On the second test, V05 = 3 scored highest on V17, V18 and V19 with V05 = 1 lowest on V18 and V19.

FSP Experimental Group, V30. Despite a relatively small number of cases per group, differences in means among the groups were significant at the .05 level on all five scores in the directions indicated below:

Group 1 - highest or second highest scores
Group 2 - lowest on four of five scores (except V15)
Group 3 - middle range on all
Group 4 - performed well on V15, V16, poorly on V17-19
Group 5 - highest on V15, V16; moderate on V17-19.

Test Date, V14. Covering years 1979 to 1981, no significant time trend in average scores was observed. This suggests that calibration problems resulting from equipment age were not a factor in test scores.

Test Devices, V20 Device #2 produced significantly larger scores both in mean and variance for all five tests. This could possibly explain the differences in mean scores by source of commission, and also suggests that psychomotor tests may be more effective predictors of student pilot performance if between-machine variation can be eliminated.

Prior Flying, V13. In general the average scores decreased with additional flying experience to about 25 hours and then increased. Those with 0 prior flying hours had the highest scores while those with 16-25 hours scored lowest. Means for the 362 cases with prior flying were below the general population means on V15-17 and above on V18 and V19. Additional analysis is needed to better understand the relationships.

Age, V14. No significant difference among mean test scores was observed for the various age groups. (The five oldest students, ages 28-33, scored consistently low on all tests.)

Sex, V40. Scores were available for very few women (56) relative to the number of men(2395). Womens' scores averaged more than those

of men on all five tests with substantial differences (p = .001) on V15-17.

Race, V41. Very few non-Caucasians -- 45 Blacks and 34 Others -- were in the study. While non-Caucasians' mean scores were higher, only for V16 were differences significant. There is not substantial evidence of a race effect based on this limited sample.

Degree Level, V42. There were relatively few individuals with other than a baccalaureate degree. In general precollege graduates showed the highest scores followed by those with a baccalaureate degree. Individuals with post-graduate work or degrees scored lower.

Degree Type, V43. On each test, holders of non-tech degrees had higher average scores. Differences were not significant on V15 and V16 but were highly significant on V17-19.

Married Status, V45. On each test, the scores of married students averaged less but differences were not substantial.

Source of Commission, V46. For the first test (V15, V16) differences in scores were highly significant with AFA highest and ROTC lowest. On the second test, ROTC scores were significantly lower with AFA and OTS higher and approximately the same.

SUMMARY. Psychomotor scores show strong differences by FSP Status, FSP Experimental Group, Test Device Prior-Flying Hours, and Source of Commission; less substantial differences occur by Degree Type, Degree Level and Sex.

TABLE V. SUMMARY OF COMPARISON OF PSYCHOMOTOR SCORE MEANS AT DIFFERENT LEVELS OF SPECIFIED VARIABLES (SIGNIFICANT DIFFERENCES (+) WITH ($\alpha \leq$ .05)

SPECIFIED VARIABLE	PSYCHOMOTOR SCORE V15	V16	V17	V18	V19	N
V05, FSP STATUS	+	+	+	+	+	697
V06, UPT STATUS	+	+	+	+	+	1748
V20, TEST DEVICE	+	+	+	+	+	217
V14, TEST DATE	0	0	0	0	0	1010
AGE	0	0	0	0	0	1748
V13, PRIOR FLYING	0	0	0	0	+	362
V30, FSP EXP. GP.	+	+	+	+	+	393
V40, SEX	+	+	+	0	0	2528
V41, RACE	0	+	0	0	0	2528
V42, DEGREE LEVEL	0	+	0	0	0	2234
V43A, DEGREE TYPE	0	0	+	+	+	2208
V45, MARRIED STATUS	0	0	0	0	0	2163
V46, COMMISSION	+	+	+	0	+	2234

IV. ANALYSIS OF FSP DATA:

The FSP data consists of twenty-one measures of student achievement during FSP (primarily for OTS candidates, with twelve exceptions) including lesson grades, retest grades, class rank and an FSP completion code. The main purpose of this analysis was data reduction, with two objectives: (1) to develop an FSP performance index based on class rank, and (2) to determine those lesson grades which, individually or in combination, are most indicative of UPT success.

IV.1 Univariate Analysis of Lesson Grades

Grades for the twelve lessons (integer values 0 to 5) averaged between 2.5 and 3.0 with a general trend of higher grades on the early lessons. Within-lesson variance is smallest for Lesson 1 grades and tends to increase thereafter, with Lesson 8, 9.1 and 9.2 showing the greatest variability (G-1.)

IV.2 Bivariate Analysis of Lesson Grades

Lesson 9.1, 9.2 and 10 were the most highly correlated, each pairwise coefficient exceeding .70 (G-2). FSP class size was not highly associated

with any of the lesson scores. There was a noticeable pattern of steadily increasing correlations between grades for Lessons 1 to 10 with Lesson 11 (final check-ride).

Pairwise correlation (G-2) between lesson grades and FSP pass/fail was highest for Lesson 9.1 (r = .74). Generally, UPT pass/fail was more highly correlated with later lesson grades and was highest at Lesson 11 (r= .31).

IV.3 UPT Criteria

Grades on early tests, especially Lessons 1 to 3, are quite uniform and averages differ very little for the groups that passed and failed UPT (G-3). However, as the lesson number increases, the mean grades of those passing UPT are significantly higher than those of the failing group. Differences are greatest at Lesson 8, 9.1 and 11.

Crosstabs of lesson grades by UPT status show that the groups with better grades have a correspondingly higher UPT passing rate, most strongly evident in Lesson 11. Lesson 8 showed an interesting and uncharacteristic drop in passing percentage for those receiving proficiency advancement, suggesting further investigation may be useful. FSP completion code appears to be moderately predictive of UPT success, but FSP class size shows at best a weak relationship with UPT status.

Percentile class rank (rank adjusted by class size), available for only three FSP classes, was used as the criterion for developing an FSP performance index based on lesson grades, class size and number of sorties. Using stepwise regression (G-5) Lesson 11 was the most significant predictor of rank, followed by Lesson 9.1. The two-variable model gives a multiple-R^2 = .78; a four-variable model, adding Lesson 5 and class size, increases R^2 to .84. When the predicted percentile rank -- the FSP performance index -- for each of the two models was regressed on FSP pass/fail, the two variable model with Lessons 9.1 and 11 was the better predictor of success (G-6). The regression equation for the two-variable performance index, denoted PRED, is

PRED = 1.1720 - .1625 * V3112 - .0852 * V3109

where V3109 and V3112 are grades for Lessons 9.1 and 11, respectively.

The FSP performance index PRED (from the two-variable model) is a better single predictor of UPT success than is FSP pass/fail (G-7). However, for the sole purpose of predicting UPT pass/fail, a model with Lessons 8 and 11 does somewhat better than the FSP performance index alone. The FSP performance index does show that class rank is a good measure of potential UPT success and, with adequate data, could be an excellent predictive variable.

IV.4 UPT Status Versus FSP Measures and Psychomotor Scores

Prediction of UPT success was examined using stepwise regression, first against the twelve FSP lesson grades and the five psychomotor scores, then against the FSP performance index and psychomotor scores. Multiple-R values on the order of .40 were obtained. The model based on lesson grades contained, in order of entry, Lesson 11, psychomotor score IX (V15) and Lesson 9.2 as independent variables.

A two independent variable model with FSP performance index and psychomotor score IX performed almost as well for predicting UPT pass/flying deficiency. For the UPT pass/all eliminations case, the inclusion of a third variable -- Lesson 9.2 -- was necessary to achieve the same degree of predictability.

IV.5 Summary of FSP Analysis

FSP lesson grades together with psychomotor scores provide good predictions of UPT success. Among the FSP measures, the later grades are better predictors of success, notably Lessons 11, 9.1 and 9.2. An FSP performance index, determined from class ranks, also correlates well with UPT status. Class ranks appear to be good indicators of success potential, and should be recorded as measures of FSP achievement. There is no obvious class size effect in FSP. Students receiving a proficiency advance on Lesson 8 showed an unexpectedly large failure rate in UPT, suggesting that an examination of this phenomenon may be useful in reducing UPT attrition. The available FSP data are almost exclusively from OTS classes; similar information from the AFA and ROTC programs would likely prove useful predicting UPT success for those students.

V. ANALYSIS OF AFOQT SCORES:

A complete analysis of the AFOQT test data was not possible in the time period of this study, nor was it mandated. An investigation of the five AFOQT test scores as indicators of student pilot performance and as they relate to other predictor variables is proposed as a topic for follow-on research.

Summary statistics (F-1) show that the mean test scores vary substantially, as much as ten points between tests, but that within-test dispersion of scores is about the same for all five tests. Pairwise correlation coefficients range from just under .20 to nearly .80. In terms of association the tests appear to divide into two groups: Pilot-Navigator and Academic-Verbal-Quantitative.

Against UPT pass/fail and UPT pass/elim criteria, the Pilot test was most highly corelated (r = .16) followed in decreasing size of "r" by Navigator, Quantitative, Academic and Verbal tests. The Verbal test showed very little association with UPT success (r = - .015 and .015).

Except on the Verbal test, the mean scores of those passing UPT were significantly higher than those of the UPT failure groups, with differences up to ten points (F-1). The "flying deficiency" failure group had the lowest mean scores on Pilot and Navigator tests, while on the remaining three tests the "other elimination" group averaged lowest.

In a stepwise regression of UPT pass/elim on AFOQT scores, the Pilot test variable entered first followed by Quantitative, Verbal and Navigator tests. Against UPT pass/fail, the order of entry was Pilot, Navigator, Verbal then Quantitative. Multiple-R values of approximately .20 were achieved in each case (F-2).

VI. ASSOCIATION OF OTHER VARIABLES WITH UPT STATUS:

An examination of simple pairwise associations of potential predictor variables with UPT status (V06) was conducted primarily using SPSS procedures BREAKDOWN and CONDESCRIPTIVE. Numerical results, given in Appendix F, are summarized below.

VI.1 Continuous Variables Age, Prior Flying Hours, and UPT Class Size

The mean age of those passing UPT was 23.8 years in contrast to a mean of 24.4 years for individuals not passing (F-3). Using seven age groups,

the UPT passing percentage was 82.2 for the youngest students, 20 to 22 years; 85.0 for those between 22 and 23; and then declined steadily to 66.1 for the oldest students, 28 to 33. Thus, there is a clear association between age and UPT success.

Prior flying hours were reported for 351 individuals with known UPT status, and included 99 cases listing zero hours (F-4). Divided into seven groups, those with zero prior flying hours had a relatively low UPT passing rate, 59.6%. Groups of individuals with some flying time had passing rates of 71.4% or higher, showing a weak trend toward higher rates with greater experience. Additional data on prior flying hours are needed to establish a clear relationship between amount of experience and likelihood of passing UPT.

The mean class size of those passing UPT was 268 compared to a mean of 285 for individuals who did not pass (F-5), indicating some association between UPT class size and UPT success. However there is also a strong association between class size and source of commission, because of a common time relationship. Most of the AFA data were collected during the first years of the study when UPT classes were relatively small; these students have a significantly higher passing rate. Thus the apparent association between class size and UPT success may be a spurious one.

VI.2 Categorical Variables Sex, Race, Degree Level, Type Degree, Married Status and Source of Commission

Crosstabulations of categorical variables by UPT status (V06) are contained in Appendix F-6; the results are condensed in Table VI below. With the exception of Married Status, V45, differences between categories were significant (Chi-square test) with $\alpha \leq .05$ for each variable. For variable V42, Degree Level, the category "Graduate work" in Table VI is a grouping of cases with Post Graduate Studies, Masters Degree or Professional Degree (V42 = G, M or P). With regard to Race, V41, when "Black" and "Other" categories are combined (N = 137) the UPT passing rate for this non-Caucasian group is 54.7%.

The crosstabulation of FSP Experimental Group (V30) by UPT status shows significant, and expected, differences in UPT success rates (F-6). An interesting and possibly unanticipated finding is the exceptionally

high passing rate of the Group 2 students, those who received six additional flight hours. This suggests that a few extra hours of flight training may significantly reduce UPT attrition.

TABLE VI SUMMARY: CROSSTABULATIONS OF CATEGORICAL VARIABLES BY UPT STATUS

		UPT Status: percentage of cases			
Variable	Category	Passed (V06=)	Flying Elim. (V06=0)	Other Elim. V06=9)	Cases N
Sex (V40)	Female	71.6	15.3	13.1	176
	Male	80.0	11.0	9.0	4483
Race (V41)	Caucasian	80.4	10.6	9.0	4521
	Black	44.2	34.5	16.3	86
	Other	72.5	21.6	5.9	51
Degree Level (V42)	Precollege grad	78.9	10.9	9.1	114
	Bac. degree	80.0	11.4	9.6	4365
	Graduate work	71.4	18.3	10.3	175
Type Degree (V43A)	Tech.	85.0	8.6	6.5	1868
	Non-tech.	76.1	13.0	10.9	2674
Married Status (V45)	Married	79.1	10.6	10.3	1426
	Single.	80.3	11.3	8.4	3121
Source of Commission (V46)	AFA	89.0	6.2	4.8	812
	ROTC	78.2	12.4	9.4	2493
	OTS	75.6	12.3	12.1	1116

VII. THE INTERIM PREDICTION MODEL:

Computer printouts of regression analysis runs leading to a linear regression model for predicting the likelihood of an individual successfully completing UPT are contained in Appendix H. After examining the many variables of the CPF and PPF files and the suspected interactions among these for their predictive ability, a tentative interim model was specified containing the potentially significant predictor variables. The aim of the subsequent regression analyses was to obtain a reliable predictive model without overfitting the sample data. Thus a relative parsimonous model was sought to avoid including spurious effects which would diminish the stability of the model. Based on the earlier study of the variables and the series of regression runs a final equation was obtained (Appendix H).

A thorough validation of the derived model, a study of its stability, and an investigation of it predictive behavior were not completed prior to this writing because of time constraints. A continuation of this work to update and improve the interim model as additional data from Phase I are available is recommended as a project for follow-on research.

VIII. RECOMMENDATIONS:

Specific recommendations and suggestions made in earlier sections of this report are briefly summarized:

(1) Psychomotor testing should be continued since scores are useful in screening candidate pilots and in predicting UPT performance;

(2) Class ranks of FSP students are good measures of FSP performance and should be recorded;

(3) The use of "proficiency advancements" on Lesson 8 in FSP should be examined;

(4) FSP-type information should be gathered for AFA and ROTC students;

(5) Additional "prior flying hours" data are needed to determine the incremental effects of increased hours; and,

(6) The six additional hours of flight training given FSP Group 2 students appears to be associated with lower UPT attrition; this relationship should be investigated.

The work begun during this brief research appointment period has potential for large savings in pilot training costs through reduced attrition at various stages of training. Several general recommendations for continuation of this work which are made below have also been discussed with personnel at AFHRL.

A number of test performance variables for which data are available, such as ZITA and APAMS scores, should be analyzed for their relationship to UPT attrition. Information from such analyses should be utilized in formulating and evaluating the new BAT tests.

As additional data from Phase I become available on file, the analysis of the CPF and PPF variables should be expanded to include this information.

The binary model can serve as a basis for establishing cutoff scores on screening tests for admission to pilot training as well as for determining satisfactory progress during UPT. Training and attrition costs together with predicted attrition can be used to develop cutoff criteria.

The bineary criterion variable in this research study --UPT pass/fail -- is a gross measure of student performance. A continuous-scale performance index or grade could assist in predicting student UPT success and should be developed from the Phase I data.

A longer term goal is to examine the relationship of selection variables to operational pilot performance. As additional information becomes available under Phase II this work can be undertaken as an extension of the present efforts.

CONTENTS

APPENDIX A

PSYCHOMOTOR SCORES, V15 to V19

A-1. Summary Statistics
A-2. Distributions and Histograms
A-3. Distributions and Histograms: Central Portions
A-4. Distributions and Histograms: Right Tails
A-5. Pairwise Correlations and Scattergrams
A-6. Breakdown of Scores by V06
A-7. Crosstabs of Scores by V06
A-8. Crosstabls of Right Tails by V06
A-9. Crosstabs of Large V15, V16 Scores (over 25000) by V06
A-10. Crosstabs of Large V17, V18, V19 Scores (over 10000) by V06
A-11. Crosstabs of V15 by V06

APPENDIX B

REGRESSIONS V06 ON PSYCHOMOTOR SCORES

B-1. Main Effects: All Cases
B-2. Interactions: All Cases
B-3. Main Effects: Deleting Upper 1% Scores
B-4. Main Effects: Deleting Upper 5% Scores
B-5. Main Effects: Deleting Upper 10% Scores
B-6. Interactions: Deleting Upper 1% Scores
B-7. Interactions: Deleting Upper 5% Scores
B-8. Interactions: Deleting Upper 10% Scores
B-9. Main Effects: Deleting Upper 1% V18 Scores Only
B-10. Main Effects, Interactions: Outliers Deleted

APPENDIX C

NATURAL LOG TRANSFORMS OF PSYCHOMOTOR SCORES,

LNV15 TO LNV19

C-1. Summary Statistics

C-2. Regression UPT on LnV15 to LnV19: Main Effects, All Cases

C-3. Regression UPT on LnV15 to LnV19: Main Effects, Interactions, Outliers Deleted.

APPENDIX D

FACTOR ANALYSES OF PSYCHOMOTOR SCORES

D-1. Principal Components Analysis, Varimax Rotation

D-2. Principal Components Analysis, Quartimax Rotation

D-3. Principal Components Analysis, Equimax Rotation

APPENDIX E

BREAKDOWN OF PSYCHOMOTOR SCORES

E-1. By Test Device (V20), Test Date (V14).

E-2. By FSP Status (V05), By FSP Status (V05) and UPT Status (V06)

E-3. By Age

E-4. By prior Flying (V13), FSP Exp. Group (V30), Sex (V40), Race (V41), Degree Level (V42), Degree Type (V43A), Married Status (V45), and Source of Commission (46).

APPENDIX F

RELATIONSHIP OF OTHER VARIABLES TO UPT STATUS

F-1. AFOQT Scores: Summary Statistics, Correlations, Breakdowns by UPT Status, and by Commission, Factor Analysis

F-2. Regressions of UPT pass/fail on AFOQT Scores

F-3. Age: Summary Statistics, Breakdown by UPT Status, Crosstabs by UPT Status

F-4. Prior Flying (V13): Distribution, Crosstabs by UPT Status

F-5. UPT Class Size: Breakdown by UPT status, by V46, Crosstabs by UPT Status

F-6. Crosstabs of Categorical Variables Sex (V40), Race (V41), Degree Level (V42), Degree Type (V43A), Married Status (45) and Source of Commission by UPT Status.

F-7. Crosstabs: V06 by V30, V40 by V46, V41 by V46, V06 by V40 by V46, V06 by V1 by V46.

APPENDIX G
ANALYSIS OF FSP PERFORMANCE MEASURES

G-1. Summary Statistics

G-2. Pairwise Correlations including FSP Class Size and UPT Status (V06)

G-3. Breakdown of Lesson Grades by UPT Status (V06)

G-4. Crosstabs of Lesson Grades, FSP Class Size by UPT Status

G-5. Regressions of Rank on Grades, Class Size and Sorties; Factor Analysis of Grades.

G-6. Regressions of FSP Pass/Fail, UPT Pass/Fail on Performance Index.

G-7. Regressions of UPT Status on Performance Index, FSP Completion Code.

G-8. Regressions of UPT Status on Grades and Class Size.

G-9. Regressions of UPT Status on Grades, Class Size and Psychomotor Scores.

G-10. Regressions of UPT Status on Performance Index and Psychomotor Scores.

APPENDIX H
INTERIM MODEL REGRESSION RUNS

H-1. Benchmark Model: AFOQT and FSP pass/fail.

H-2. Interim Model Runs.

1983 USAF SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

EFFECT OF JET FUEL JP-4 FRACTIONS ON FISCHER 344 MALE RATS

Prepared by:	Dr. M. Paul Serve'
Academic Rank:	Professor
Department and University:	Department of Chemistry, Wright State University
Research Location:	Air Force Aerospace Medical Research Laboratory, Toxic Hazards Division, Toxicology Branch, Wright-Patterson AFB, OH 45433
USAF Research Colleague:	Dr. Carl T. Olson and Dr. Roger C. Inman
Date:	March 20, 1983
Contract No:	F49620-82-C-0035

EFFECT OF JET FUEL JP-4 FRACTIONS ON FISCHER 344 MALE RATS

by

Dr. M. Paul Serve'

ABSTRACT

The effects of jet fuel JP-4 on the kidneys and livers of male Fischer 344 rats were studied in order to ascertain whether all hydrocarbons produce similar effects. JP-4, after separation into boiling range fractions, was administered intragastrically to the rats. Weight loss and urine production were studied for 48 hours. Gross pathological as well as histopathological examination of the kidneys and livers of the rats 2 days and 10 days post-exposure to the JP-4 indicated that kidney and liver damage was maximized early. The rats appeared to be able to repair any damage by day 10. The higher boiling fractions of JP-4 appeared to be the most toxic.

ACKNOWLEDGEMENT

The author would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for providing him with he opportunity to spend a worthwhile and interesting summer at the Air Force Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio. He would like to acknowledge the laboratory, in particular the Toxicology Branch for its hospitality and excellent working conditions.

Finally he would like to thank Dr. Carl T. Olson for suggesting this area of research and his assistance in pursuing this project. He would also like to thank Dr. Roger C. Inman and Dr. Richard H. Bruner for their encouragement and cooperation in the development of this project.

I. INTRODUCTION

The study of the toxic properties of jet fuels is one of the main thrusts of the Aerospace Medical Research Laboratory, Toxic Hazards Division. The Air Force jet fuel identified as JP-4 has been shown by gas chromatography to be composed of over 200 hydrocarbons of various aliphatic and aromatic structures[1]. It has been reported that when aromatic hydrocarbons are inhaled at high concentrations, humans suffer disturbed vision, dizziness, tremors, cardiac stress, central nervous system (CNS) depression, confusion, and coma[2]. Likewise, the liquid aliphatic hydrocarbons have been shown to promote CNS depressant and neurotoxic effects when inhaled at high concentrations[3]. Recently, it has been reported that subchronic vapor inhalation studies with a number of jet fuel mixtures have shown a pattern of kidney damage in male rats[4]. Phillips has shown that repeated exposure for 8 weeks to Stoddard Solvent (composed of many high boiling hydrocarbons) vapors caused renal cell exfoliation in Fischer 344 male rats[5]. A recent study by the Toxic Hazards Research Unit of the Air Force Aerospace Medical Laboratory identified a few of the urine metabolites of rats who had been exposed to inhalation experiments using jet fuel JP-4[6]. The metabolites were glucuronic acid conjugates of alcohols and ketones containing no more than 8 carbons. This implied that either hydrocarbons of boiling point less than 150° C were not a major portion of the inhaled vapor or that the analytical technique used did not pick up any of the higher molecular weight hydrocarbon metabolites.

Until now, there has been no attempt made to see whether all hydrocarbons cause this kidney effect or whether only hydrocarbons of a sufficiently high molecular weight and boiling point can bring about the kidney changes. It should be noted that rats should have greater difficulty in outgassing higher boiling hydrocarbons than lower boiling ones.

II. OBJECTIVES

The overall objective of this study was to examine the effects of various fractions of the jet fuel JP-4 on Fischer 344 rats rats. The effects to be looked at include:

a) The weight loss by the animals.
b) The volume and coloration of the urine produced by the rats.
c) The changes in the rat liver and kidney 2 and 10 days after administration of the JP-4 fraction.

III. The Separation of JP-4 into Boiling Point Fractions

The jet fuel JP-4 was divided into the following fractions using the technique of fractional distillation using a glass helicied packed distillation column.

Fraction	Boiling Point Range (°C)
A	70-100
B	101-110
C	111-130
D	131-145
E	146-160
F	161-175
G	176-190
H	191-220

The number of degrees in the boiling point range was varied for the different fractions of JP-4 so that a sufficient fraction volume could be obtained for the dosing experiment. A minimal volume of 20 ml for each fraction was deemed necessary.

IV. Dosing of Male Fischer 344 Rats Followed by Weight Loss and Urine Output Studies

After weighing, groups of 6 male Fischer 344 rats were given 2 ml of the desired JP-4 fraction by intragastric administration. The rats were individually kept in Fischer metabolism cages for 48 hours. During that period of time, the rats were given rat chow and water ad libium. At the end of the 48 hour period, the rats were reweighed and the urine volume was measured. The results of the two day metabolic JP-4 studies are listed in Table I.

TABLE I

Effect of JP-4 Fractions on the Weight Loss and Urine Output for Fischer 344 Male Rats

Fraction of JP-4	Weight Before Dosing (gm)	Weight Loss (gm)	Urine Output (ml)
H_2O	304	0 ± 4.9	16.5 ± 3.7
A	257	16.5 ± 4.4	14.0 ± 4.2
B	324	14.3 ± 3.2	15.7 ± 2.9
C	312	14.0 ± 4.2	16.0 ± 3.2
D	300	13.0 ± 6.8	23.4 ± 9.0
E	331	23.6 ± 3.1	38.8 ± 12.2
F	348	24.1 ± 3.2	25.4 ± 3.8
G	305	24.3 ± 7.8	31.5 ± 11.0
H	289	24.6 ± 6.8	25.7 ± 7.0

A group of 6 rats dosed with 2 ml of water was used as the control group. With the exception of the fraction A and H studies, all the rats utilized weighed at least 300 gm. There was a significant weight loss in all the rats studied relative to those rats administered water. The weight loss was more pronounced for those groups of rats given JP-4 fractions boiling above 146° C. Urine output volume by the rats given JP-4 fraction A, B, and C were noted to be approximately the same as the urine output for rats administered water. Fractions of JP-4 with boiling points above 130° C showed a significant increase in urine volume output.

An explanation for the increased urine production could be the reduced ability for the rats to exhale the higher boiling hydrocarbons leaving urine excretion as the only method by which the rats could rid themselves of the unwanted hydrocarbons.

It was also noted that the urine from the rats administered the higher boiling fractions of JP-4 had a deeper, darker coloration and a more pungent odor.

V. Attempted Recovery and Identification of JP-4 Rat Urine Metabolites

The rat urine from the two day metabolic studies for each JP-4 fraction was pooled. After adjusting its pH to 6.8 the urine was treated with 0.5 ml of β -Glucuronidase-aryl sulfatase (Calbiochem number 34742) in order to effect the hydrolysis of any glucuronide or sulfate derivative of the hydroxylated metabolites of the JP-4 fraction. After incubation at 37° C for 24 hours, the rat urine was filtered through a Clin Elut column with methylene chloride as the eluting solvent. After evaporation of the excess methylene chloride, the samples were examined by gas chromatography-mass spectral (GC-MS) analysis using a Hewlett-Packard 5985 GC-MS system.

Unfortunately, because the GC-MS system was not equipped with a capillary column mo ification, the gas chromatograph could not effect an efficient separation of all the peaks so that information on the metabolites could be obtained with reasonable certainty.

VI. Gross Pathological Examination of the Kidneys and Livers Two Days and Ten Days Post-Exposure to JP-4 Boiling Fractions

The gross pathological examination reports on the kidneys and livers of the male Fischer 344 rats two days and ten days after exposure to the various JP-4 fractions indicated that the gross effects were seen two days after exposure. Reticulated kidneys were noted for JP-4 boiling point fractions 111-130° C, 131-145° C and 176-190° C. These effects were all resolved within ten days. This indicated that the rat had the ability to overcome the kidney effect of a massive dose of hydrocarbons. Of note, however, is that the fraction 176-190° C was the fraction which gave the most highly discolored urine.

VII. Histopathological Examination of Kidneys and Livers Two Days and Ten Days Post Exposure to JP-4 Boiling Fractions

The histopathological examination of the kidneys and livers two days and ten days after exposure to the JP-4 fractions indicated that for most rats two days post-exposure there was a moderate accumulation of hyaline droplets in the tubular cells of the kidneys. This was most noticeable in the boiling ranges 100-110° C, 146-160° C,

. Ten days after exposure the hyaline droplet production had dropped to a minimal level, once again indicating kidney repair. Mild hepatocellular fatty changes were noted in the livers of rats two days after exposure for the boiling fractions 100-110° C, 146-160° C, 161-175° C, 176-190° C, and 191-220° C. These changes disappeared ten days after exposure.

VIII. Multiple Intragastric Administrations of JP-4 Boiling Point Fraction 175-190° C

Three male 344 Fischer rats were intragastrically administered 2 ml of JP-4 boiling point fraction 175-190° C on three separate occasions each 9 days apart. The results, listed in Table II, are those based on a 48 hour study.

TABLE II

Repetitive Intragastric Administration of JP-4 Fraction (Boiling Point 175-190° C).

Day of Administration of JP-4 Fraction	Loss of Weight (gm)	Urine Volume (ml)
1	20.1 ± 2.1	27.0 ± 3.3
10	13.7 ± 3.8	48.3 ± 8.6
19	10.8 ± 3.1	45.5 ± 9.7

During the experiment, it was noted that the rats regained back a portion of the weight that they had lost in the first 48 hours after administration of the JP-4 fraction. However, even before the administration of the second and third doses of the JP-4 fraction, the rats still had not reached their initial weight on day 1. The weight loss by the rats appeared to be at a maximum during the first administration of the JP-4 fraction. This, along with the increased

urine production during the second and third administrations of JP-4 fraction, would indicate that the rats metabolic system was gearing up to metabolize the JP-4 and get rid of it rapidly in the urine. The reduced weight loss durng the second and third administrations of the JP-4 fraction could be explained by the increased amount of water ingested by the rats during the metabolic study.

IX. RECOMMENDATIONS

The study of the effects of hydrocarbons on the kidneys showed that urine production was increased with an accompanied loss of weight by the male rats. In addition gross pathological examination showed changes in the kidneys, while histological studies demonstrated the production of unusual hyaline fluid droplets in the kidneys. The effects are more pronounced with the higher boiling fractions of hydrocarbons.

It is recommended that future studies start to hone in on compounds in these higher boiling ranges and attempt to collect them according to structural similarities. Individual hydrocarbons can then be looked at to see if the kidney effects are due to a certain hydrocarbon structure.

An examination of the type of metabolites formed in JP-4 fractions should also be undertaken. This can readily be accomplished via an adaptation of the GC-MS so that capillary columns can be used. The effects on the kidneys are most likely due to metabolites, since

hydrocarbons are water insoluable and would not be expected to reach the kidneys in the free or unbound to albumin state. Thus, a particular type of metabolites structure which can yield very reactive carbocations could be the culprit in the kidney changes.

REFERENCES

1. Hayes, R. C. and Pitzer, E. W., "Kovats Indices as a Tool in Characterizing Hydrocarbon Fuels In Temperature Programmed Glass Capillary Gas Chromatography", Air Force Wright Aeronautical Laboratories, Wright-Patterson AFB, OH 45433, AFWAL-TR-81-2102.

2. Gitelson, S., Aladjemoff, L., Ben-Hador, S., and Katznelson, R., "Poisoning by a Malathion-Xylene Mixture", J. Am. Med. Assoc. 197, 819-821, 1966.

3. Sandmeyer, E. E., in Pattys Industrial Hygiene and Toxicology, Vol 2B (G. D. Clayton and F. E. Clayton, Eds. John Wiley and Sons, 1981), pp 3176-3177.

4. MacEwen, J. D. and Vernot, E. H., "Toxic Hazards Research Unit Annual Technial Report: 1981", Air Force Aerospace Medical Research Laboratory, Wright-Patterson AFB, OH 45433, AFAMRL-TR-81-126.

5. Phillips, R. D., "Effects of Stoddard Solvent on Kidney Fruncton and Structure of Fiscner 344 and Sprague-Dawley Rats", presented at the 13th Annual Conference on Environmental Toxicology, Dayton, OH, 16-18 November 1982.

6. MacEwen, J. D. and Vernot, E. H., "Toxic Hazards Research Unit Annual Technical Report: 1982", Air Force Aerospace Medical Research Laboratory, Wright-Patterson AFB, OH 45433, AFAMRL-TR-82-62.

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

SECONDARY MUZZLE FLASH IN RAPID FIRE CANNONS

Prepared by:	Dr. Robert K. Sigman
Academic Rank:	Senior Research Engineer
Department and University:	School of Aerospace Engineering Georgia Institute of Technology
Research Location:	Air Force Armament Laboratory, Direct Fire Weapons Division, Ballistics Branch
USAF Research:	Mr. Otto K. Heiney
Date:	September 8, 1983
Contract No.:	F49620-82-C-0035

SECONDARY MUZZLE FLASH IN RAPID FIRE CANNONS

by

Robert K. Sigman

ABSTRACT

The problem of secondary muzzle flash in rapid fire aircraft cannon has been investigated. An additional mechanism for initiation of secondary flash has been discovered, which is not present in single fire cannons or in single shot firings, which are generally used to test for secondary flash. This initiation mechanism is associated with blow-by or leakage of hot propellant gas past the rotating band. Computer programs have been prepared for prediction of the flow field within the barrel ahead of the projectile and for the interior ballistics of a leaking gun. Recommendations are offered for further research in this area.

Acknowledgement

The author would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for giving him the opportunity to spend an interesting summer at the Air Force Armament Laboratory, Eglin AFB, Florida. He would like to thank the Armament Laboratory, and particularly the Ballistics Branch, for providing excellent working conditions and genuine hospitality.

Finally, he would like to thank Mr. Otto K. Heiney for his helpful suggestions and guidance. The helpful conversations with Mr. Bert K. Moy are also appreciated. The technical advice and debates on the merits of the Braves vs. the Dodgers with Lts. Terry W. Shiver and James S. Schoeneman will long be remembered.

I. INTRODUCTION

Secondary muzzle flash occurs when the fuel rich residual propellant gas from the discharge of a firearm is expelled from the gun barrel, mixes with air, and ignites. This flash is highly objectionable, since it can reveal the location of static gun positions, cause loss of night vision of gun crews, and obstruct the vision of aircraft pilots. Clearly, a stoichiometric propellant would eliminate such a flash, but the overriding demand for acceptable physical properties requires that solid propellants for rockets and guns be fuel rich.

The alternative is to try to prevent ignition of the fuel-air mixture by the addition of chemical suppressants. Addition of such suppressants results in a decrease of gun efficiency, and the residue of these suppressants coats the windshield of aircraft obstructing the pilot's vision. Thus, it is desirable to minimize the amount of added suppressant, and this requires an understanding of the mechanisms which cause secondary flash.

Studies of firings of single shot rifles and cannons reveal the principal source of ignition. After shot ejection, the fuel rich propellant gas in the barrel is at a pressure on the order of 3000 atmospheres while the velocity is sonic at the muzzle. The geometry of the flow field of the expanding gas is shown in Figure 1. An explanation of the physical reasons for such a flow field can be found in Reference 1. It is sufficient to note that the fuel rich propellant gas mixes with air and then passes through a normal shock wave or "Mach disk." This normal shock provides rapid heating of the combustible gas mixture and frequently produces a small, brief flash known as primary flash even in the presence of suppressants. Although this primary flash is acceptable, the shock heating and primary flash often lead to secondary flash. Thus, analyses of secondary flash usually concentrate on the gasdynamic aspects of the flow field after shot ejection with varying degrees of emphasis on the mixing of the propellant gas with air, shock heating of the fuel-air mixture, and ignition of the hot fuel-air mixture. Of particular interest is the analysis of Heiney and West,[2] which uses an interior ballistic code to calculate the temperature and velocity of the propellant gas leaving the muzzle and a simple one-dimensional mixing model. All shock heating and ignition criteria are extracted from the experimental results of Cartagno.[3] This analysis and other more complex analyses[4,5] provide reasonable estimates for the amount of suppressant which must be added to a propellant. Single shot firings

of the propellant-suppressant mixture are then used to verify these predictions and to determine the optimum amount of suppressant.

Such techniques were used to reduce the amount of suppressant in the propellant for the GAU-8, 30 mm rapid fire aircraft cannon. Although no secondary flash was observed in single shot firings, flash was evident in rapid fire tests. Because of the previous success of such analyses and experiments, it was believed that an alternative ignition source was present.

II. OBJECTIVES

The objective of the research during the summer period was to identify the mechanism for secondary flash in rapid fire aircraft cannons. As previously noted, the mechanism for ignition of secondary flash in single fire cannons and small arms as described in the Introduction was not believed to be responsible, since single shot firings did not produce flash. Once the mechanism was identified, attempts would be made to remove the ignition source or alter the analytical model to predict the required amount of suppressant. Specifically, the objectives were to:

(1) Identify the mechanism for ignition of secondary flash in rapid fire cannons.

(2) Eliminate the source of ignition if possible.

(3) Investigate analytical models to predict the differences between ignition mechanisms for single shot firings and rapid firings.

(4) Prepare computer codes for (3), compatible with the interior ballistics codes and secondary flash codes in use at Eglin AFB.[2]

Identification of the ignition mechanism would be accomplished by evaluation of data from firings consisting primarily of high speed motion pictures.

III. IDENTIFICATION OF THE IGNITION SOURCE

The search for the ignition source generally proceeded along two paths. Data from the initial rapid-firings was sketchy and it was hypothesized that ignition of a fuel-air mixture might be caused by a subsequent firing. High speed motion pictures of later firings revealed the presence of blow-by as a possible ignition source or, at least, a concurrent event. Each of the paths will be discussed with primary emphasis on the blow-by aspect.

A. Ignition by Subsequent Firings.

No motion pictures were made of the initial firings. The only data was that:

1) secondary flash was present, and 2) burning slivers of propellant were being ejected from the barrel.

Secondary flash in single shot cannons must be self-igniting, i.e., the mechanism for ignition must be within the dynamics of a single shot. However, in rapid fire cannons, it is conceivable that firing of one barrel could ignite the propellant gas-air mixture from the previous barrel. In this case, the burning particles could act as ignition sources for the gas cloud.

For static firings of rapid fire cannons, the propellant gas accumulates in front of the cannon. On the other hand, if the cannon is mounted on an aircraft in flight, the gas cloud will be swept back over the aircraft. Thus, the initial analysis was to prepare a simple flow field model to predict the location of the propellant gas cloud with time.

It is known that the gas cloud from a static single shot cannon behaves like a moving fluid dynamic source.[6] The source term produces a spherical expansion of the gas cloud. The movement of the center of the source means that the spherical cloud will be convected forward in the case of static firings, or backward in the case of in-flight firings. Rather than develop a complete model, the moving source model was used with two free constants: the source strength and the velocity of the source center. Realistic values of these two constants were obtained from photographs of static firings. The velocity of the source center for firings from an aircraft in flight was assumed to be the difference between the aircraft flight velocity and the static source center velocity. The static and flight models are shown in Figure 2.

The results were not surprising and could have been extrapolated from the experimental data of Reference 7. The initial expansion (source term) from a 1 inch diameter cloud to a 3 foot diameter cloud occurs on the order of 10^{-6} sec. During the period $10^{-6} < t < 5 \times 10^{-3}$ sec this cloud of nearly constant diameter convects forward or backward very slowly. For subsonic aircraft using high rate of fire cannons, the gas cloud from one barrel will remain in the path of subsequent firings. Thus, if ignition is caused by burning slivers, the aircraft motion will not alleviate the problem.

B. Ignition by "Blow-by."

A second series of rapid firings, consisting of two 24 shot bursts, one 59 and one 94 shot burst, was conducted on June 24, 1983. Two high speed motion picture

cameras were used to record the firings.

Each of the 201 shots was analyzed for primary flash, secondary flash, ejecta (burning slivers), and ignition of previous gas cloud. Both films were used to cross check the data analysis.

Two conclusions were evident from the data:

(1) Some barrels flashed frequently, while other seldom flashed, and

(2) Secondary flash was always preceded by a large "pre-flash."

Close examination using frame-by-frame analysis revealed that the "pre-flash" was not a primary flash which occurs after the projectile is ejected, but occurs before the projectile is ejected. This normally indicates that hot propellant gas is leaking past the rotating bands, and precedes the projectile out of the barrel. Such leakage is generally associated with worn barrels, failed rotating bands, or hot expanded barrels. In this case, it was determined that the barrels were quite worn and a final series of firings was scheduled using new barrels.

A final series of firings was conducted on July 12, 1983. For comparison, one 10 shot and one 25 shot burst of "high" suppressant (standard issue) ammunition were fired. One 24 shot, one 59 shot, one 94 shot, and one 275 shot burst of "low" suppressant (nitramine) ammunition also were fired. Two high speed motion picture cameras were again used with one camera set up close to the muzzle.

In short bursts (less than 52 shots), no secondary flash was observed. For longer bursts, blow-by was observed after about 44 shots, and secondary flash was observed after 52 shots. It was concluded that as the barrels heat up, they expand and gas leaks past the projectile so that a "flash" of hot gas precedes the projectile out of the barrel. This flash of hot, radiating gas is quite evident in the close-up movies; and it appears that after shot ejection, the propellant gases expand through this hot gas.

Thus, it was decided that an analysis should be prepared for modeling the interior ballistics of leaking guns with consideration given to the gasdynamics of the flow field ahead of the projectile.

IV. ANALYSIS OF LEAKING GUNS

Interior ballistics is defined as the branch of applied mechanics which deals with the motion and behavior characteristics of projectiles while under the influence of gases produced by the propellant. Because the pressure of the gas behind the projectile dominates the acceleration of the projectile in the barrel, the

gas ahead of the projectile (precursor gas) is generally ignored.

Although the state of the precursor gas is unimportant insofar as the motion of the projectile is concerned, it is necessary to develop an analysis of this region since it is believed that this region could be responsible for igniting secondary flash in rapid fire cannons. Further, it is necessary to re-evaluate the interior ballistics analysis because the leakage reduces the effective pressure in the chamber.

Corner presents an analysis for the interior ballistics of leaky guns, but it is not presented in the formalism of current analyses. Thus, a new analysis for the interior ballistics of a leaky gun is developed using the nomenclature and approximations of Reference 2. Presumably, the results of this analysis should be equivalent to Corner's results.[8]

A. Precursor Analysis

As a non-leaking projectile accelerates down the barrel, compression waves are generated which coalesce to form a shock wave. This precursor region is shown in Figure 3, for a typical gun condition. A description for this flow field is given by the classical piston problem in Liepmann and Roshko.[9]

The question arises as to how this region changes when high pressure, hot propellant gases leak past the projectile. It is clear that the velocity at the "leak" is sonic and the flow into the region ahead of the projectile is supersonic. At this point numerous isentropic models such as expansion to the full barrel area, and expansion to atmospheric or "shocked" pressure, were considered but rejected due to discrepancies. For example, isentropic expansion to the full barrel area would give a gas with a high velocity but sub-atmospheric pressure.

A realistic (and consistent) model is shown in Figure 4. In addition to the usual shock generated by the motion of the piston, a second shock is formed at some point on the projectile.

Referring to Figure 4, the high pressure, high temperature gas from Region 1 leaks past the rotating band. The flow in Region 2 is expanding supersonically with decreasing temperature and pressure. The flow then shocks to Region 3, expands subsonically with decreasing velocity and increasing pressure to a final region, 4, at the full barrel area. The final velocity (Region 4) is slightly higher than the projectile velocity and the temperature is nearly the stagnation temperature of Region 1. Regions 4 and 5 are separated by a contact surface, so that the pressure and velocity in Regions 4 and 5 must be identical. Note that the temperatures,

densities and gases are different. The final moving shock separating Regions 5 and 6 lowers the pressure to atmospheric pressure and drops the velocity to zero.

For actual calculations, the location of the projectile shock must be determined by trial and error. For given stagnation conditions, leakage area and projectile velocity; calculations proceed from Region 1 to Region 6 using one-dimensional compressible flow relations[10] with an assumed projectile shock location. If the gas in Region 6 is computed to have atmospheric pressure, the correct projectile shock location has been determined.

Two computed solutions are shown in Figure 4 for large, and small leakage. It is satisfying to note that as the leakage decreases, the flow field approaches the no leak flow field shown in Figure 3. Further, at the contact surface between the hot propellant gases (Region 4) and the warm air (Region 5), combustion (which is visible in motion pictures) is possible.

B. Interior Ballistics Analysis

The interior ballistics codes used by the Air Force have been developed for non-leaking projectiles and must be modified to account for leakage. The changes involve a restatement of the energy equation and the burning rate equation.

Let the leakage flow rate be denoted by W. Then it can be shown that[8,10]

$$W = \rho_t S_t V_t = \psi p_r S_t (RT_r)^{-\frac{1}{2}} \tag{1}$$

where S_t is the leakage area, ρ_t and V_t are the throat density and sonic velocity, p_r and T_r are the stagnation pressure and temperature in Region 1 and ψ is a function of γ, the specific heat ratio. Thus, for an assumed leakage area, the flow rate is dependent only on the current chamber conditions.

Let ϕ be the mass of propellant gas generated by combustion and N be the mass of propellant gas remaining in the chamber. Thus,

$$\frac{d\phi}{dt} = \frac{dN}{dt} + W \tag{2}$$

and as usual

$$\frac{d\phi}{dt} = \rho_p S_p r \tag{3}$$

where ρ_p is the propellant density, S_p the exposed surface area, and r the burning rate.

Consider the control volume containing the combustion chamber as shown in Figure 5. For this volume, the energy equation is:[11]

$$-Q + \int_{\mathcal{V}} \frac{\partial}{\partial t}\left\{ \rho e + \frac{1}{2}\rho V^2 \right\} d\mathcal{V} + \int_{S} \left\{ h + \frac{1}{2}V^2 \right\} \rho V \cdot dS = 0 \qquad (4)$$

Since the volume is independent of time, the time derivative can be taken outside of the volume integral. Within Region 1, the volume can be considered to consist of three subvolumes of "gas." Subvolume a consists of a true gas of mass N, volume $\mathcal{V}_a$, and temperature T. Subvolume b consists of a dense pseudo gas of mass $\left(\frac{C_w}{g} - \phi\right)$, temperature T_o and volume $\mathcal{V}_b$. Volume b thus is a gaseous representation of the burning propellant. Volume c, representing the projectile, is a very dense gas of mass M_p, volume $\mathcal{V}_c$, negligible temperature and uniform velocity V_p. The volume integral can thus be written as

$$\frac{\partial}{\partial t}\left\{ Nc_vT + \left(\frac{C_w}{g} - \phi\right) c_vT_o + \frac{W_p}{g} \cancel{c_vT_p}^{\,0} + \int_{\mathcal{V}_a+\mathcal{V}_b} \frac{1}{2}\rho V^2 A dx + \frac{1}{2} M_p V_p^2 \right\} \qquad (5)$$

and the volume integral over a portion of Region 4 is also neglected. Note that the kinetic energy of the gas (the integral over volumes a and b) consists of mass $\frac{C_w}{g} + N - \phi$. If we neglect the loss in mass, the kinetic energy expression from Reference 2 can be used. Thus Equation (5) is written as

$$\frac{\partial}{\partial t}\int_{\mathcal{V}} \left[\rho e + \frac{\rho V^2}{2} \right] d\mathcal{V} =$$

$$c_v \frac{dNT}{dt} - c_vT_o \frac{d\phi}{dt} + M_A V_p \frac{dV_p}{dt} \qquad (6)$$

and M_A is the "apparent mass" of the projectile. Noting that the temperature T is considered to be the "stagnation" temperature, and the mass flux through the control volume is W, the surface integral in Eqation (4) is

$$\int_S \left[h + \frac{1}{2} V^2 \right] \rho V \cdot dS = c_p TW = \frac{\gamma RTW}{(\gamma-1)M_W} \tag{7}$$

where M_W is the molecular weight and R is the universal gas constant.

From Reference 2, the heat loss is

$$-Q = \beta M_A V_p \frac{dV_p}{dt} \tag{8}$$

Thus, Equations (4), (6), (7), and (8) can be combined to give the energy equation for a leaking gun:

$$\frac{R}{M_W} \frac{dNT}{dt} = \frac{RT_o}{M_W} \frac{d\phi}{dt} - (1+\beta)(\gamma-1) M_A V_p \frac{dV_p}{dt}$$

$$- \gamma \psi p_r S_t (RT)^{1/2} \tag{9}$$

which agrees with Corner's[8] expression.

Note also that if $S_t = 0$, Equations (2) and (3) reduce to

$$\frac{dN}{dt} = \rho_p S_p r \tag{10}$$

and Equation (9) reduces to

$$\frac{R}{M_W} \frac{dNT}{dt} = \frac{RT_o}{M_W} \frac{dN}{dt} - (1+\beta)(\gamma-1) M_A V_p \frac{dV_p}{dt} \tag{11}$$

Equations (10) and (11) are used in Reference 2. Thus Equation (9) is consistent with References (2) and (8).

V. IGNITION MODEL

At this point, computer codes are available for computing the flow field ahead of a leaking projectile and for computing the interior ballistics of a leaking gun. The interior ballistics program[2] uses a time-dependent procedure which tracks the build-up of pressure in the chamber and motion of the projectile with time. Assuming equilibrium at each time step, the flow field ahead of the

projectile could also be developed with time.

However, as a first attempt at obtaining an ignition criteria, a simple flow model was employed. It was assumed that the leak is very small, and that the hot gas bleeds into the region ahead of the projectile at the instantaneous stagnation temperature and at the instantaneous projectile velocity. Thus, at any time, the gas ahead of the projectile has an average temperature given by:

$$T = \frac{\int_0^t TW\,dt}{\int_0^t W\,dt} \tag{12}$$

where, W, defined in Equation (1), is computed at each time step.

After the projectile leaves the barrel, the precursor gas temperature and the propellant gas temperature are computed. For non-leaking guns, the propellant gas is assumed to mix with cool air to form a fuel-air mixture with mixture ratio, r, and temperature, T. This mixture is compared with experimental ignition temperatures, T_{ign}, at various mixture ratios to determine whether the mixture will, may, or will not flash.

For leaking guns, the simple flow model compares both the precursor temperature and the propellant temperature with experimental ignition temperatures. As an indication of the approximate leakage area, S_t, it is known that leakage decreases the muzzle velocity by 200 ft/sec. Calculations with the leaking interior ballistics codes indicate that the corresponding leakage area is 0.0005 in^2. This is reasonable since, if the rotating bands were not present, the leakage area would be about 0.005 in^2.

As would be expected, the leaking gas decreases the temperature of the propellant gas and slightly reduces the tendency of the propellant gas to flash (See Fig. 6). However, the precursor gas temperature predicted by Equation (12) is quite high (~ 1850°K), so that the precursor will always flash. In fact, motion pictures show it to be ignited within the barrel. Thus the simple precursor ignition model does not seem to indicate a simple criterion for ignition by the precursor.

VI. RECOMMENDATIONS

The problem of secondary flash in rapid fire aircraft cannons has been investigated. Motion pictures revealed that concurrent with the occurrence of

secondary flash is the appearance of "blow-by" or hot propellant gas leaking past the projectile. This hot gas appears to be the ignition source for the remainder of the propellant gas which is expelled after shot ejection. This event is not generally encountered in single fire cannon tests, which are used to test for secondary flash.

Analyses of the gasdynamics of the regions ahead of and behind the projectile for a leaking gun have been developed. A limiting case for the region ahead of the projectile has been combined with the"leaking" interior ballistics codes to predict ignition of the precursor gas. This model did not produce a reliable criteria for predicting secondary flash, due to the simplicity of the precursor model.

It is recommended that the complete precursor gas model, including expansion of the precursor region, be interfaced with the interior ballistics codes to produce more comprehensive calculations. Further, the shock heating model used to predict heating of the fuel-air mixture should be adjusted to account for the lower pressure (weaker shock) of the precursor gas.

In single shot testing, "good" barrels are normally used to prevent facility damage by erratic projectile trajectories. If measures are employed to prevent such damage, worn barrels could be used to allow "blow-by" on single shot firings and to test for "optimum" suppressant amounts.

REFERENCES

1. Abbett, M., "Mach Disk in Underexpanded Exhaust Plumes," AIAA Journal, Vol.9, No. 3, March 1971, pp. 512-514.

2. Heiney, O. K. and J. R. West, "Interior Ballistics, Muzzle Flash and Gas Gradients of Aircraft Cannon," Air Force Armament Laboratory, Rept. No. AFATL-TR-76-34, March 1976.

3. Carfagno, S. P., Handbook of Gun Flash, The Franklin Institute, November 1961.

4. May, I. W. and S. I. Einstein, "Prediction of Gun Muzzle Flash," CPIA Publication No. 292, Vol. II, 14th JANNAF Combustion Meeting, August 1977, pp. 229-239.

5. Yousefian, V., May, I. W. and J. M. Heimerl, "Modeling the Occurrence of Muzzle Flash in Guns," CPIA Publication No. 329, Vol. II, 17th JANNAF Combustion Meeting, November 1980, pp. 125-140.

6. Erdos, J. I. and P. D. Del Guidace, "Calculation of Muzzle Blast Flowfields," AIAA Journal, Vol. 13, No. 8, August 1975, pp. 1048-1055.

7. Schmidt, E. M. and D. D. Shear, "Optical Measurement of Muzzle Blast," AIAA Journal, Vol. 13, No. 8, August 1975, pp. 1086-1091.

8. Corner, J., Theory of the Interior Ballistics of Guns, John Wiley & Sons, Inc., New York, 1950.

9. Liepmann, H. W. and A. Roshko, Elements of Gasdynamics, John Wiley & Sons, Inc., New York, 1957

10. Shapiro, A. H., The Dynamics and Thermodynamics of Compressible Fluid Flow, Vol. I, The Ronald Press Company, New York, 1953-1954.

11. Zucrow, M. J. and J. D. Hoffman, Gas Dynamics - Volume I, John Wiley & Sons, Inc., New York, 1976-1977.

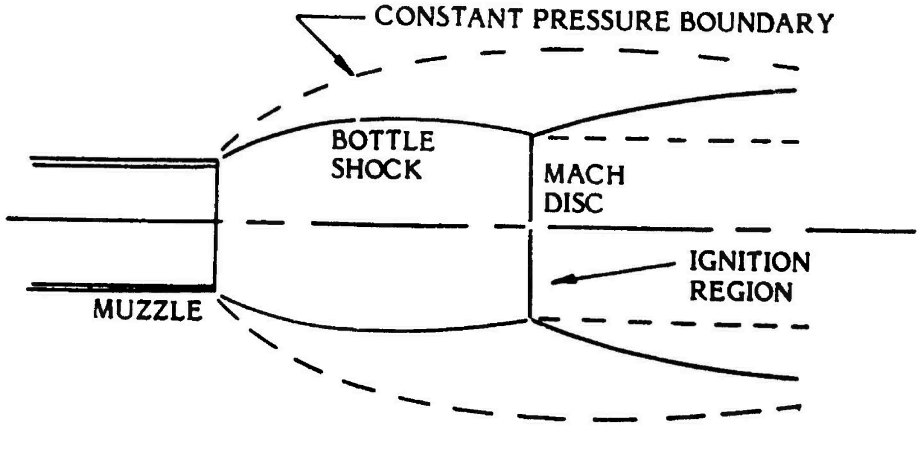


FIGURE 1 - GEOMETRY OF CANNON FLOW FIELD

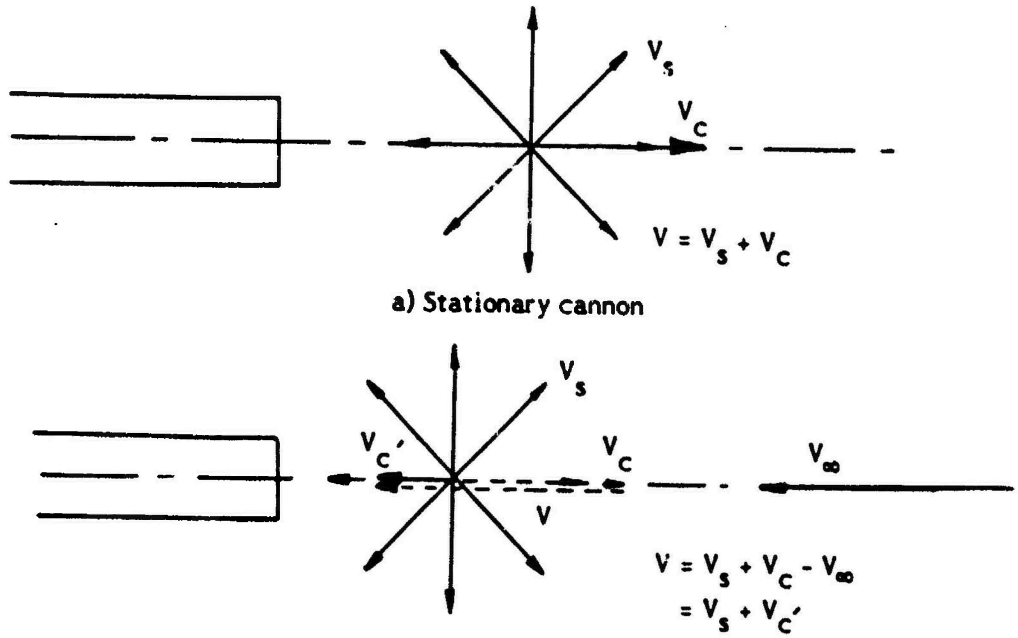


FIGURE 2 - FLOW FIELD MODELS (moving source)

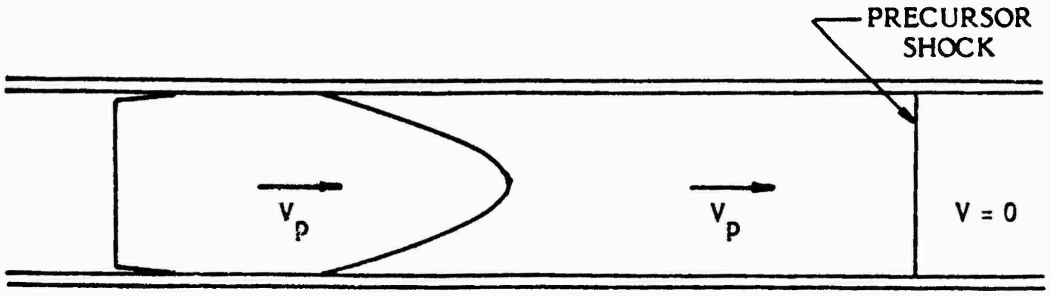
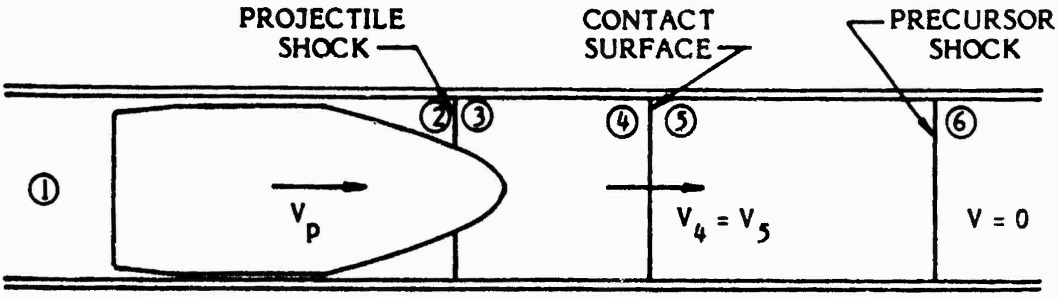


FIGURE 3 - PRECURSOR FLOW FIELD OF NONLEAKING PROJECTILE



BORE AREA = 1 in^2

P_1 = 477677.7 psi

T_1 = 2363°K

V_p = 775 ft/sec

S_t	=	0.001 in^2	S_t	=	0.00001 in^2
P_4	=	94.4 psi	P_4	=	36.9 psi
T_4	=	2337.4°K	T_4	=	2363°K
V_4	=	1864.5 ft/sec	V_4	=	803.2 ft/sec
	=	V_p + 1089 ft/sec		=	V_p + 28.2 ft/sec
T_5	=	607°K	T_5	=	398°K
P_6	=	14.6 psi	P_6	=	14.8 psi

FIGURE 4 - PRECURSOR FLOW FIELD OF LEAKING PROJECTILE

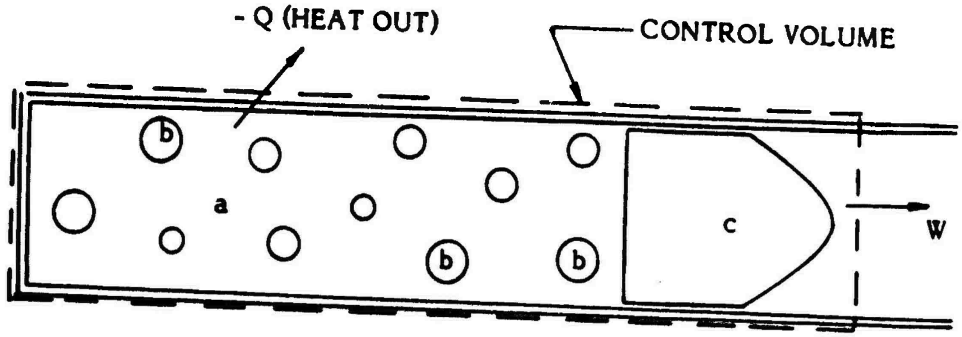


FIGURE 5 - CONTROL VOLUME FOR ANALYSIS

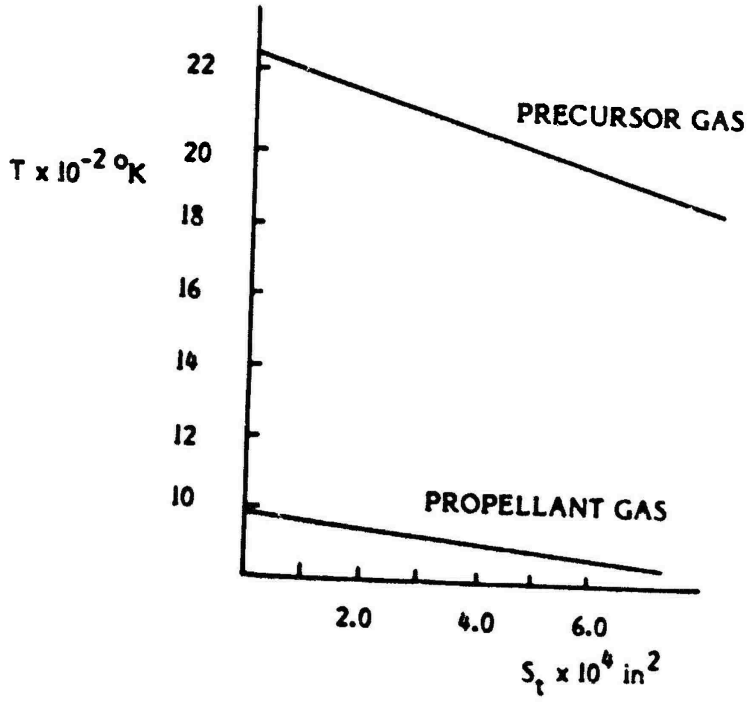


FIGURE 6 - TEMPERATURE AT MUZZLE

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

POROELASTIC MODELS OF THE INTERVERTEBRAL DISK

Prepared by:	Dr. Bruce R. Simon
Academic Rank:	Associate Professor
Department and University:	Aerospace and Mechanical Engineering The University of Arizona
Research Location:	Air Force Aerospace Medical Research Laboratory, Biodynamics and Bioengineering Division; WPAFB, OH.
USAF Research:	Dr. Leon E. Kazarian
Date:	August 24, 1983
Contract:	F49620-82-C-0035

POROELASTIC MODELS OF THE INTERVERTEBRAL DISK

by

Bruce R. Simon

ABSTRACT

Analytical, experimental, and finite element models were developed for spinal motion segments including the vertebra and the intervertebral disk (IVD). The cortical and cancellous bone were assumed to be elastic, whereas the disk tissue (nucleus and annulus) were assumed to behave "poroelastically" as deformable porous solids through which a fluid can flow. Simple one-dimensional poroelastic models were used to estimate mechanical parameters for creep response data for rhesus IVD's (L_2 - L_3 level). A prototype axisymmetric finite element model of the rhesus IVD (L_2 - L_3 level) was developed. Deformation, internal stress, pore fluid pressures, and pore fluid motion results were obtained for spinal motion segments in axial compression subjected to static, step (creep), and steady-state (sinusoidal) loading. The results of this project demonstrate that meaningful poroelastic models can be developed which will allow a detailed consideration of the mechanics of spinal motion segments. Models of this type will be applicable to the study of aircraft ejection phenomena; crash injury; disk degeneration, enucleation, and herniation processes; as well as provide a fundamental understanding of the relation of possible fluid motion in the disk to internal stresses, pressures, and pathways for transport of materials to and from the avascular regions of the IVD.

ACKNOWLEDGEMENT

I would like to thank the Air Force Systems Command, Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for providing the opportunity to carry out this research at the Air Force Aerospace Medical Research Laboratory/ Biodynamics and Bioengineering Division at Wright Patterson Air Force Base, Dayton, Ohio. In particular, I wish to acknowledge Dr. Leon Kazarian for his scientific inputs to this effort. Capt. Paul France and Clarence Oloff deserve special thanks for their invaluable assistance during the summer. Other individuals who were especially helpful in providing technical support were SSgt. Mike Jonopulos and K. C. Smith.

I. INTRODUCTION

The spine is a complicated, dynamically loaded system composed of a regular pattern of spinal motion segments (SMS). Each SMS has an intervertebral disk (IVD) acting as a relatively soft tissue structure between the more rigid boney vertabral bodies. Structural models for the SMS are of practical importance in the following studies: aircraft ejection events and ejection seat design, long-term structural fatigue in the spine, spinal crash injury analysis, lower back pain syndromes, and (IVD) herniation and enucleation. Such models will also allow the study of the fundamental mechanics and behavior of "normal" and "degenerated" disks. Finite element models developed in conjunction with experimental models can be effective in predicting deformations and stresses in the spine. Various analytical, experimental, and finite element models have been proposed for the SMS that were based on elastic and viscoelastic material response models. Such models are somewhat phenomenological and do not give a unified, physically based description of observed disk characterics such as incompressible nuclear material behavior, fluid motion in the disk, and long-term creep. In this report the IVD was viewed as a poroelastic, two-phase material composed of a deformable porous solid phase through which a fluid phase is allowed to flow. This "biphasic" model gives a physical basis for the observed history-dependent SMS response by associating relative fluid flow losses with hydraulic permeability in the IVD tissues. Finite element models based on this poroelastic allow detailed representation of materials, geometry, and boundary conditions and can predict deformation, stresses, and in particular pore fluid motion in the IVD. What is needed now are (1) a quantitative description of the

materials in the SMS (including a poroelastic material model for the IVD) and (2) a concurrent development of the corresponding methods for finite element analysis of the SMS.

II. OBJECTIVES

The AFAMRL has been active in the development of experimental, analytical, and finite element modeling of the spine and SMS's. The objective of the research described here was to expand these efforts in the following areas:

A. Mechanical parameter identification and quantification. Available experimental creep data was considered using a poroelastic model developed to describe the soft tissues of the nucleus and annulus. Creep response data for rhesus monkey SMS's (L_2 - L_3 level) was evaluated using several types of analytical poroelastic models and mechanical properties were identified and quantified.

B. Finite element analysis. A prototype axisymmetric finite element model (fem) was generated for the L_2 - L_3 level rhesus SMS. The fem represented the cortical and cancellous bone of the vertebral bodies as elastic materials, whereas a poroelastic material model (based on the results of A.) was used to represent the nucleus and annulus of the IVD. This fem was used to predict the static, transient creep, and steady-state structural reponse of the SMS.

III. SUMMARY OF THE THEORETICAL BASIS

The poroelastic models described here are based on the theory of Biot[1] which can be expressed in a matrix notation in terms of the displacement of the solid phase $\underset{\sim}{u}$, and the relative fluid displacement vector $\underset{\sim}{w}$. The coupled, dynamic equilibrium equations for the two phases are

(no body forces)

$$\underset{\sim}{L}^T\underset{\sim}{\sigma} = \rho\ddot{\underset{\sim}{u}} + \rho_f\ddot{\underset{\sim}{w}} \text{ and } \underset{\sim}{\nabla}\pi_f = \rho_f\ddot{\underset{\sim}{u}} + \frac{\rho_f}{n}\ddot{\underset{\sim}{w}} + k^{-1}\dot{\underset{\sim}{w}} \tag{1}$$

where the solid strain $\underset{\sim}{e} = \underset{\sim}{L}\underset{\sim}{u}$ and fluid strain $\zeta = \underset{\sim}{\nabla}^T\underset{\sim}{w}$ are related to the total stress $\underset{\sim}{\sigma}$ and pore fluid pressure π_f via a linear, isotropic constitutive law of the form

$$\underset{\sim}{\sigma} = \hat{\underset{\sim}{D}}\underset{\sim}{e} + \alpha Q\underset{\sim}{m}\zeta \quad \text{and} \quad \pi_f = \alpha Q\underset{\sim}{m}^T\underset{\sim}{e} + Q\zeta \tag{2}$$

In the above $\underset{\sim}{L}$, $\underset{\sim}{\nabla}$, and $\underset{\sim}{m}$ are matrix forms for operators of spatial partial derivatives, and the del operator, and the Kroneker delta, respectively, and superposed dot implies a time partial derivative. Equation (2) can be written as

$$\underset{\sim}{\sigma} = \underset{\sim}{D}\underset{\sim}{e} + \alpha\underset{\sim}{m}\,\pi_f \quad \text{where} \quad \underset{\sim}{D} = \hat{\underset{\sim}{D}} - \alpha^2 Q\underset{\sim}{m}\underset{\sim}{m}^T \tag{3}$$

Here $\underset{\sim}{D}$ is the "drained" material stiffness matrix containing the drained elastic modulus and Poisson ration, E and ν. Parameters α and Q are additional material constraints (described by Biot[1,2] and Simon, et al.[3,4]). The remaining material properties are the permeability k, the porosity n, and the overall density $\rho = (1-n)\rho_s + n\rho_f$ where ρ_s and ρ_f are the densities of the solid and fluid phases. Boundary and initial conditions complete the description of the boundary value problems.

A fem can be utilized to discretize the above equations. Spatial interpolations using shape functions $\underset{\sim}{N}_u$ and $\underset{\sim}{N}_w$ for displacements are $\underset{\sim}{u} = \underset{\sim}{N}_u\bar{\underset{\sim}{u}}$ and $\underset{\sim}{w} = \underset{\sim}{N}_w\bar{\underset{\sim}{w}}$(in terms of nodal displacements $\bar{\underset{\sim}{u}}$ and $\bar{\underset{\sim}{w}}$) yielding discrete dynamic equilibrium equations in each finite element of the form

$$\underset{\sim}{M}\ddot{\underset{\sim}{r}} + \underset{\sim}{C}\dot{\underset{\sim}{r}} + \underset{\sim}{K}\underset{\sim}{r} = \underset{\sim}{R} \tag{4}$$

where $\underset{\sim}{r}^T = [\underset{\sim}{\bar{u}}^T | \underset{\sim}{\bar{w}}^T]$ and $\underset{\sim}{M}$, $\underset{\sim}{C}$, $\underset{\sim}{K}$, and $\underset{\sim}{R}$ are consistent mass, damping, stiffness, and load matrices (see Ghaboussi and Wilson[5]). The assembled form of equation (4) is a very large set of coupled, ordinary differential equations that can be integrated in time given $\underset{\sim}{R}(t)$ and appropriate boundary and initial conditions. Solutions are obtained for (a) <u>static analysis</u> with $\underset{\sim}{R} = \underset{\sim}{R}_o$ by solving $\underset{\sim}{K}\underset{\sim}{r} = \underset{\sim}{R}_o$; (b) for quasi-static transient <u>creep analysis</u> with $\underset{\sim}{R} = \underset{\sim}{R}_o 1(t)$, step loading, by solving $\underset{\sim}{C}\dot{\underset{\sim}{r}} + \underset{\sim}{K}\underset{\sim}{r} = \underset{\sim}{R}_o 1(t)$; and (c) <u>steady-state analysis</u> with $\underset{\sim}{R} = \underset{\sim}{R}_o e^{j\omega t}$, oscillatory loading and displacements $\underset{\sim}{r} = \underset{\sim}{r}_o e^{j\omega t}$, by solving $[-\omega^2\underset{\sim}{M} + j\omega\underset{\sim}{C} + \underset{\sim}{K}]\underset{\sim}{r}_o = \underset{\sim}{R}_o$ with $j \equiv \sqrt{-1}$ and ω = frequency. Full dynamic solutions for transient analysis are obtained by applying appropriate time integrators (explicit, implicit, staggered, etc.) to equation (4) for given $\underset{\sim}{R}(t)$.

IV. REPRESENTATIVE RESULTS

A. Creep models. Equations (1) - (3) can be written in the form of a single diffusion equation for a one-dimensional model of a creep test on the SMS. A series solution following Biot[2] was obtained and used as an analytical model for the observed creep in three rhesus SMS's (L_2 - L_3 level). The experimental data (provided by E.P. France[6]) was analyzed using this model and a minimization search program (based on the method of Powell[7]) determined optimal values (in the least squares sense) for material parameters E, k, and the fluid bulk modulus K_f for given values α, ν, and n. Here note that $Q = [n/K_f + 3(1-2\nu)(\alpha-n)(1-\alpha)/E]^{-1}$. Figure 1-A shows the analytical fit of experimental data for one rhesus SMS (#0AIAA). Figure 1-B depicts the creep response (model and/or experimental) for all three SMS's considered here. Analytical and experimental response curves were virtually identical after search. The average values for material properties were $\overline{E} = 0.106 \times 10^7$ N/m^2, $\overline{k} = 0.275 \times 10^{-15}$ m^4/Ns, and $\overline{K}_f = 0.140 \times 10^8$ N/m^2 (with $\alpha = 0.9$, $\nu = 0.45$, n = 0.7 and step loading of 66.72 N = 15 lb for all disks). These values are similar to values reported in the literature by Urban and Maroudas[8] and by Mow, et al.[9] permeability and elastic modulus of soft tissues. However, the value for $\overline{K}_f$ is well below that of water implying that the pore fluid in these disks may be compressible. The values of these parameters are useful in interpreting one-dimensional test results, but must be reconsidered (see below) before being utilized in two-(or three) dimensional structural models of a SMS.

B. Finite Element Model. Figure 2 shows a prototype axisymmetric fem of a representative rhesus SMS at the L_2 - L_3 level. The fem is com-

posed of 121 nodes and uses 99 four-noded quadrilateral elements plus one triangular element. At the outset, some of the results reported here will be improved by using more refined grids in future fem's. The vertebrae is composed of an elastic cancellous core surrounded by a thin elastic cortical shell. The nucleus and annulus were represented as poroelastic materials. The material properties of A. were used as first estimates for the response of the IVD fem. Values for these one-dimensional fit parameters were then adjusted in the fem until reasonable agreement was obtained between fem and experimental creep response for the three rhesus SMS (see Figure 1-B). The final adjusted poroelastic properties of the annulus were $E_A = 0.201 \times 10^7$ N/m^2, $k_A = 0.692 \times 10^{-15}$ m^4/Ns, $K_{f_A} = 2.21 \times 10^9$ N/m^2, $\nu_A = 0.45$, $\alpha_A = 0.9$, $n_A = 0.7$. Here K_f is elevated significantly. Further research should be initiated to determine values for K_f in models of this kind. The nucleus was assumed to have identical properties except the elastic modulus was reduced to $E_N = 0.201 \times 10^5$ N/m^2 to simulate a more "fluid-like" nuclear behavior. Elastic properties of the vertebra were $E_{can} = 1.75 \times 10^7$ N/m^2 and $E_{cort} = 3.3 \times 10^8$N/m^2 for the cancellous and cortical bone, respectively and $\nu_{can} = \nu_{cort} = 0.25$. Symmetrical displacement boundary conditions were assumed for the solid and fluid phases in the fem(see Figure 2). The pore fluid was allowed to flow freely between the nucleus and annulus and the disk and vertebrae. No flow was allowed at the "outer" boundary of the disk. Static, step (creep), and oscillatory uniform traction was applied at the "top" (mid-centrum plane) of the fem and deformation, stress, and pore fluid motion were calculated in each case using the solution schemes discussed above. Only a representative sample of the results will be described here with

more details to appear in subsequent publications.

The creep analysis using the fem revealed an interesting transient history of deformation, stress, and fluid motion associated with the long-term creep of the SMS. Figure 1-B shows the fem creep displacement vs. time results. Figure 3 shows a typical deformed shape of the rhesus SMS and the corresponding relative flow field in the IVD after 84 minutes of creep have transpired. Bending in the cortical shell and deformation of the cancellous core in the vertebra, bulging of the disk and deformation of the nuclear-annular border associated with creep are apparent. The fluid motion (relative to the deforming solid) in the disk indicates that the creep response is associated with flow from the disk into the vertebral body as well as with the flow radially outward from nucleus to annulus. The model indicates that after 84 minutes, minimal flow has developed in the central interior region of the nucleus due to the very low nuclear permeability. Greater flow is present in the annulus at this time (even though annular permeability is also very low) due to deformation of the tissue and its interaction with the flow in the pores.

Figure 4 is representative of the large amount of stress analysis data obtained. Figure 4-A shows σ_{zz} contours associated with the static analysis and Figure 4-B shows the "effective" von Mises stress contours (after 84 minutes of creep) as an indicator of spatial stress level variations in the SMS. Here "effective" stress is taken in the soil mechanics usage (see Biot[1]) for the poroelastic disk, i.e. effective stress $\underset{\sim}{\sigma}'$ is the portion of the total stress $\underset{\sim}{\sigma}$ in excess of the local fluid pressure (here $\underset{\sim}{\sigma}' = \underset{\sim}{\sigma} - \underset{\sim}{m}\pi_f$) and is associated with failure analysis in porous media. The von Mises formula is then applied to the components

of σ' to give a single number to represent effective stress level at a point in the fem. Inspection of Figure 4 reveals regions where stress levels are elevated, e.g. at t = 84 minutes a number of areas in the vertebral cortical shell are highly stressed due to bending. A region of elevated stress occurs at the outer mid-line of the annulus. Comparison of effective von Mises stresses with material strength may give an indication of possible areas where failure is likely in the SMS, e.g. the location of stress concentration at the side of the fem corresponds to observed failure sites in the SMS.

It is important to note that transient analyses of impact load conditions (similar to the creep study above) can be carried out using this type of fem and time integration of equation (4). The computer program used here has been applied successfully to such problems in soil mechanics where dynamic effects and stress waves are of importance. Short term transient applied stress histories (associated with ejection, crash, etc.) could now be applied to the present fem and the corresponding deformation and stress time sequences could be calculated.

Other results were also produced using this fem. The static analysis revealed displacements and stresses similar to the creep study above with the exception of the fluid flow field. The relatively low values for permeability yield an "undrained" steady-state pore fluid flow field where $\underset{\sim}{w} = \underset{\sim}{0}$ in the "normal" disk over a wide range of frequency. This condition corresponds to the observed "incompressible" response of the nucleus. Steady-state fem results exhibited impedance and phase angle versus frequency trends that were very similar to experimental data reported by Kazarian[10] for SMS's - again supporting the validity of a

poroelastic soft tissue model. Pore fluid pressure contour plots suggest that positive values for π_f (i.e. tension) may develop in the IVD. The possible significance of this is currently being studied.

V. RECOMMENDATIONS

The sample results presented here demonstrate the potential and the feasibility of carrying out structural analyses of SMS's using analytical, experimental, and finite element models based on a poroelastic material model for the IVD. Recommendations for continuation of this project are in the following specific avenues of research:

1. Experimental and the associated data analysis modeling should be expanded. The initial linear views of the observed material response in a SMS can now be broadened to include directional effects (anisotropy), nonlinearity (both geometric and material), as well as inclusion of fluid viscosity and linear and nonlinear viscoelasticity in the solid phase in addition to the incorporation of nonlinear flow losses (k not constant) in the porous media material model. Further development of creep tests, volumetric compliance tests, jacketed and unjacketed material tests, (see Biot[1,2]), etc. should be undertaken. Analytical and experimental programs should be developed in parallel to provide more realistic and meaningful structural models. Further, attention should be devoted to producing more accurate two-(or perhaps three-) dimensional data analysis schemes. The application of the finite element model in an "inverse sense" to determine material properties (see Iding, et al.[11]) is a promising approach to this aspect of analytical-experimental modeling efforts.

2. Further work should be initiated to extend the prototype fem

described here when the more detailed and accurate material descriptions of 1. above become available, anisotropy, nonlinearity, viscoelasticity, fluid viscosity, etc. can be introduced into an improved fem. Current linear models can readily be extended to study shorter duration, transient events (related to ejection, crash, parachute opening, etc.) Further development of steady-state and transient creep fem's should provide information regarding possible structural fatigue in the spine. Possible mechanisms for IVD degeneration might be simulated and studied by varying mechanical parameters in the fem of a SMS. Kazarian's data[12] suggests a relation between degeneration and elevated disk permeability k. Values for k could be increased in the fem and the resulting changes in structural response (deformation, stress, pore fluid flow) could be predicted and assessed. Similar fem's with "normal" mechanical properties will provide insight regarding the "normal" mechanics of a SMS. The prediction of flow fields in a normal IVD will yield an insight into possible paths for transport of materials to cells resident in the avascular interior disk regions. Simulation of possible mechanical changes in these paths in a fem would be useful in the study of pathological conditions in a SMS. Future fem's could be developed to study the disk herniation process or the prediction of the structural effects of disk enucleation (either by surgical or chemical means). Eventually, three-dimensional fem's could be developed for consideration of problems where a very detailed prediction of the structural response of the spine (or its individual SMS's) is required. Such three-dimensional fem's can only be generated by building upon the experience gained from two-dimensional fem's and the associated data analysis and experimental modeling efforts.

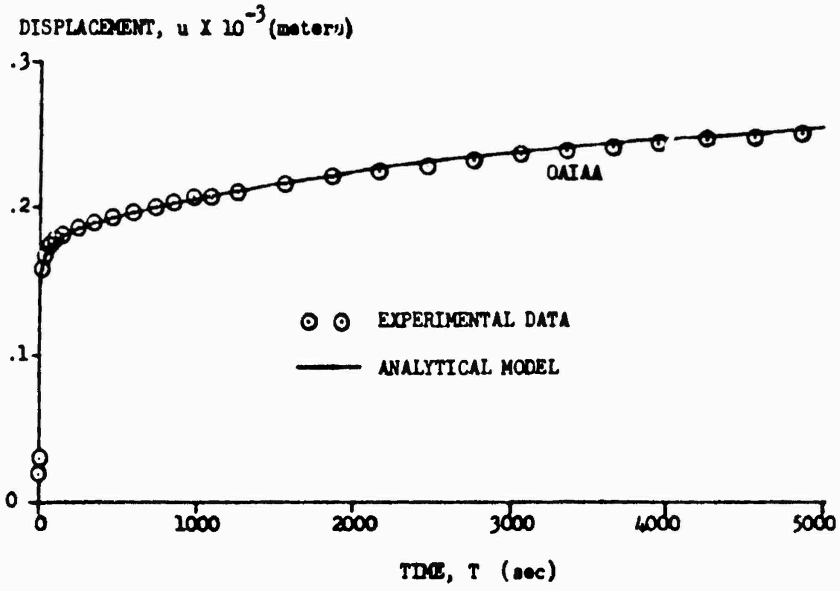


FIGURE 1 - A. CREEP DISPLACEMENT VS. TIME. ANALYTICAL (1-D,BIOT) FIT OF EXPERIMENTAL DATA (RHESUS IVD, L_2-L_3, #0AIAA)

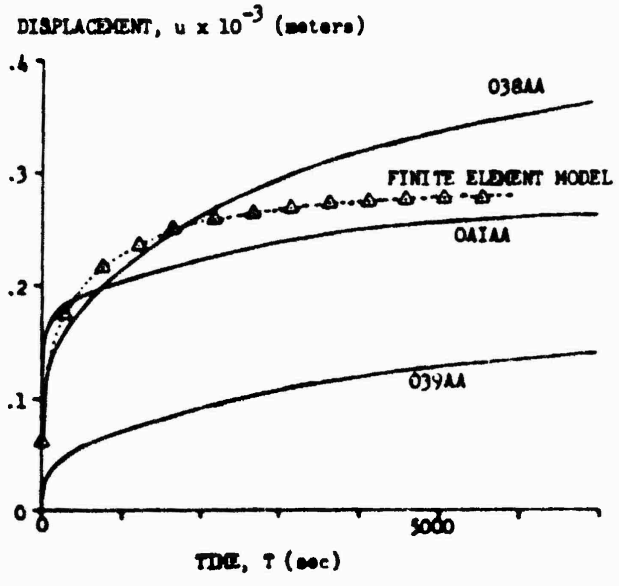


FIGURE 1 - B. CREEP DISPLACEMENT VS. TIME. FINITE ELEMENT MODEL (AXISYMMETRIC) AND EXPERIMENTAL DATA (THREE RHESUS IVD'S, L_2-L_3,#038AA,#0AIAA,#039AA)

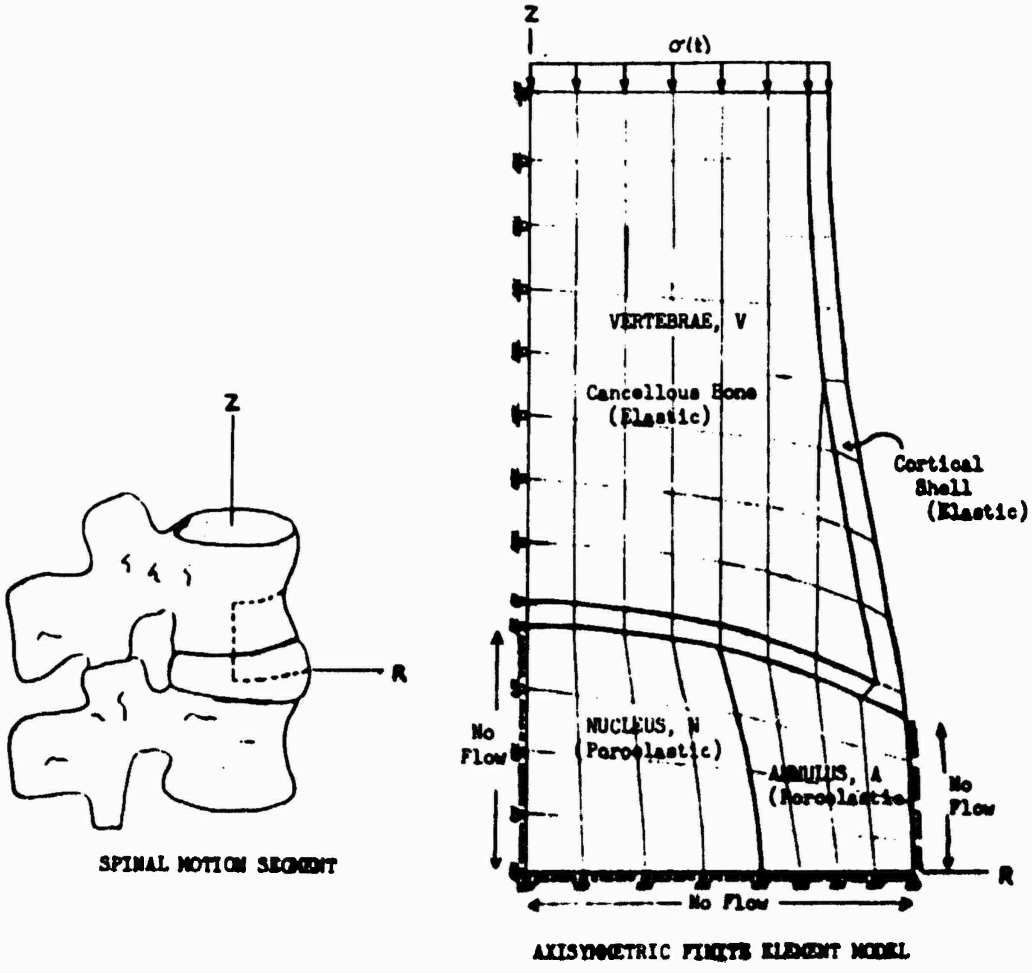


FIGURE 2. AXISYMMETRIC FINITE ELEMENT MODEL OF A RHESUS SPINAL MOTION SEGMENT (L_2-L_3)

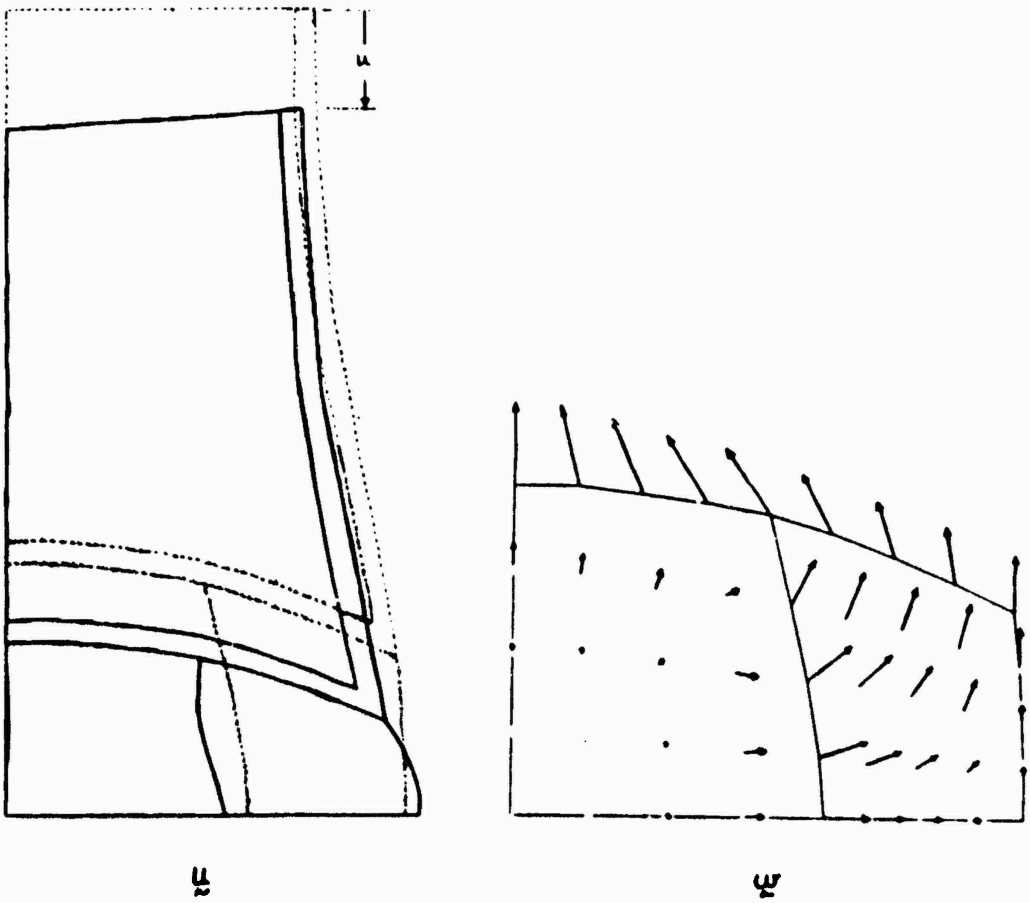


FIGURE 3. AXISYMMETRIC FINITE ELEMENT RESULTS FOR $\underset{\sim}{u}$ = SOLID DISPLACEMENTS AND $\underset{\sim}{w}$ = RELATIVE FLUID DISPLACEMENTS AFTER TRANSIENT CREEP UNTIL TIME = 84 MINUTES.

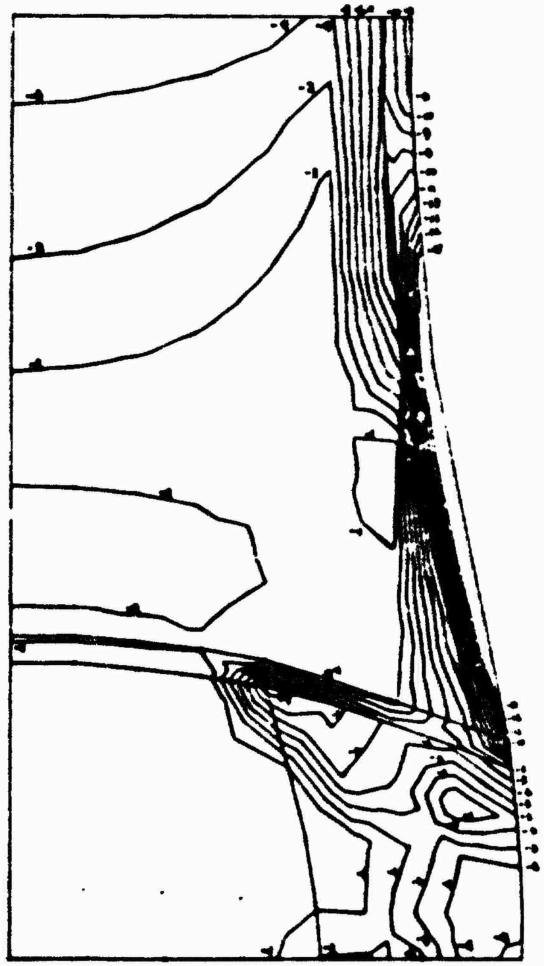


FIGURE 4 - A. AXISYMMETRIC FEM RESULTS FOR σ_z = AXIAL STRESS FOR "STATIC" ANALYSIS

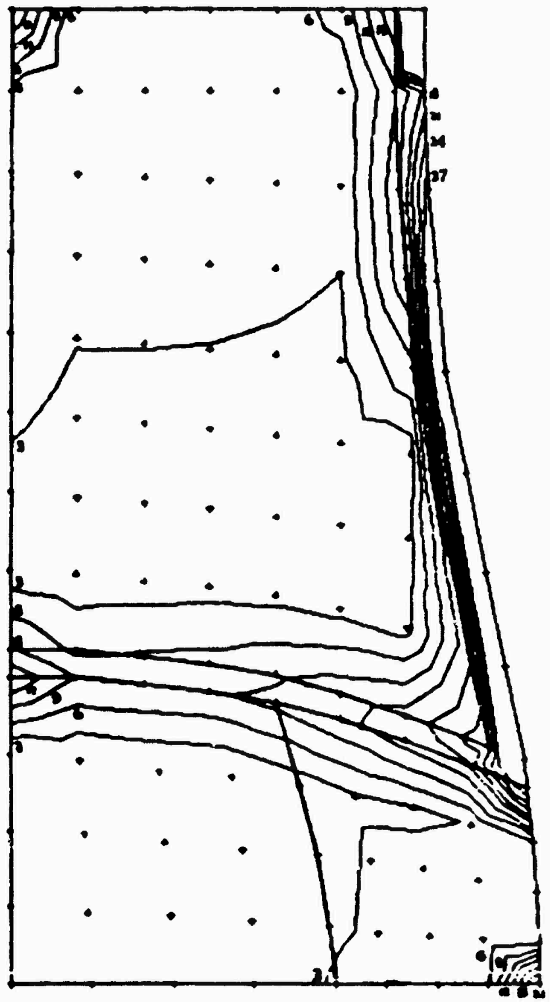


FIGURE 4 - B. AXISYMMETRIC FEM RESULTS FOR σ'_{VM} = "EFFECTIVE" VON MISES STRESS AFTER TRANSIENT CREEP UNTIL TIME = 84 MINUTES.

REFERENCES

1. Biot, M.A., "Mechanics of Deformation and Acoustic Propagation in Porous Media," J. Applied Physics, 33, 4, 1962, pp. 1482-1498.

2. Biot, M.A., "General Theory of Three-Dimensional Consolidation," J. Applied Physics, 12, 1941, pp.155-164.

3. Simon, B.R., O.C. Zienkiewicz, and D. K. Paul, "An Analytical Solution for the Transient Response of Saturated Porous Elastic Solids" to appear in Intl. J. Num. & Anal. Meth. Geomechanics.

4. Simon, B.R., J.H. Evans, and J.S.S. Wu, "Poroelastic Models for the Intervertebral Disk," abstract to be presented at ASME/WAM, Boston, November 1983.

5. Ghaboussi, J. and E. L. Wilson, "Variational Formulation of Dynamics of Fluid-Saturated Porous Elastic Solids," J. Engr. Mech. Div.,ASCE, EM4, 1972, pp. 947-963.

6. E.P. France, personal communication.

7. Powell, M.J.D., "A Method for Minimizing a Sum of Squares of Nonlinear Functions Without Calculating Derivatives," Computer J., 7, 1965, pp. 303-307.

8. Urban, J. and A. Maroudas, "Measurement of Swelling Pressure and Fluid Flow in the Intervertebral Disc with Reference to Creep," I. Mech. Engr. Publ. 1980-2, 1980, pp. 63-69.

9. Mow, V.C., S.C. Kuei, N.M. Lai, and C.G. Armstrong, "Biphasic Creep and Stress Relaxation of Articular Cartilage in Compression: Theory and Experiments," J. of Biomech. Engr., 102, 1973, pp. 73-84.

10. Kazarian, L.E., "Dynamic Response Characteristics of the Human Vertebral Column," Acta Ortho. Scand. Suppl. No. 146, 1972.

11. Iding, R.H., K.S. Pister, and R.L. Taylor, "Identification of Nonlinear Elastic Solids by a Finite Element Method," Comp. Meth. Appl. Mechs Engr., 4, 1974, pp. 121-142.

12. Kazarian, L.E., "Creep Characteristics of the Human Spinal Column," Ortho. Cl. North Am., 6, 1975, pp. 3-18.

1983 USAF-SCEEE GRADUATE STUDENT SUMMER S[illegible] PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

EFFECTS OF FLUID SHIFTS AND HYPOVOLEMIA IN INDIVIDUALS WITH DIFFERENT WORKING CAPACITIES WHILE RESTING AT A FIVE DEGREE DECLINATION

Prepared by:	Dr. William G. Squires
	Debra K. Rotto
	Diane M. Rotto
Academic Department:	Physiology
University:	Texas Lutheran College and Graduate School of Biomedical Sciences, University of Texas Health Science Center at San Antonio
Research Location:	Aerospace Physiology Laboratory, Biomedical Division, Crew Technology Branch
USAF Research Contact:	Dr. Sarah Nunneley
SFRP Supervising Faculty Member:	Dr. William G. Squires
Date:	August 1, 1983
Contract No.:	7930-14-04

EFFECTS OF FLUID SHIFTS AND HYPOVOLEMIA IN INDIVIDUALS WITH DIFFERENT WORKING CAPACITIES WHILE RESTING AT A FIVE DEGREE DECLINATION

by

Debra K. Rotto
Diane M. Rotto
William G. Squires

ABSTRACT

Data from actual space flight studies have demonstrated functional abnormalities and changes of the cardiovascular system immediately post-flight. Interest lies in what occurs to the body during exposure to zero gravity that could account for these post-flight alterations in cardiovascular functions. During zero gravity or head-down rest, the hydrostatic intra- and extravascular pressure gradients that are normally present in the upright position are abolished or minimized. This causes a headward shift of body fluids from the lower portions of the body cavity. This massive fluid shift induces adaptive changes in other body systems, such that reexposure to normal gravitational forces produces signs of of orthostatic intolerance. These adaptive changes include the following: inhibition of the renin-angiotensin and ADH systems, diuresis, decreased blood volume, weight loss associated with diuresis and decreased blood volume, inhibition of sympathetic activity due to decreased levels of circulating catecholamines and increased activity of the carotid sinus nerve, decreased stroke volume, no change (or slight increase) in arterial pressure, cardiac output, or the contractile state of the heart, compensatory bradycardia, increased right and left atrial filling pressures and central venous pressure, increased left ventricular end-diastolic volume and left ventricular ejection fraction, decreased leg volumes and increased forearm volumes. Preliminary studies were done on a dog and the expected changes in the cardiovascular parameters were found. Contractility of the heart increased, left ventricular pressure increased, left ventricular end-diastolic pressure increased, heart rate decreased, renal blood flow and carotid blood flow increased, and systemic arterial pressure increased after the dog was tilted five degrees head down for one hour.

Acknowledgement

The authors would like to thank the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for the opportunity to work on an exciting and mind-stimulating project for the summer. We are grateful for the many things we learned and the interesting experiences we had at Brooks AFB/SAM while honing our research skills. In particular, we would like to acknowledge the Aerospace Physiology Division of the Crew Technology Branch for its hospitality, guidance, patience and excellent working conditions.

Finally, we would like to thank Dr. Carter Alexander for suggesting this area of research and we would like to acknowledge the collaboration and excellent guidance of Dr. Russell Burton and Dr. Sarah Nunneley.

I. INTRODUCTION:

The major focus of the experiment will be to compare responses to zero gravity and orthostatic tolerance post-tilt between individuals with different fitness levels. Understanding the mechanisms involved requires elucidation of changes in plasma renin-angiotensin levels, plasma catecholamine levels, body fluid shifts, and segmental volume changes. As a result, further significance could be drawn from this study regarding the use of the head-down tilt method as a means of pre-flight preparation for the pilot.

With the advent of manned space flight new questions have arisen as to the effects of zero gravity on human cardiovascular function. The adaptive responses of man to this new type of environment involve many interrelated and complex organ systems. Data from actual space flight studies have shown no signs of severe cardiovascular impairment but functional abnormalities and changes of the cardiovascular system have been demonstrated immediately post-flight.[6,9,24,30,31,32] Interest lies in what occurs to the body during exposure to zero gravity that could account for these post-flight alterations in cardiovascular function.

Because of time, method, and other limitations during actual space flight studies, it is necessary to develop ground-based simulations of weightlessness. Bed rest and water immersion both produce the same human responses as those seen with prolonged exposure to zero gravity. However, recent Russian and American studies have demonstrated that a slightly head-down position (from -4° to -6°) produces effects qualitatively similar to those of bed rest except that the adaptation is accelerated. During short-term experiments head-down tilt produces a greater degree of cardiovascular deconditioning than horizontal bed rest of equal duration. Water immersion has also been used to simulate zero gravity, though there are problems inherent in this technique. Prolonged immersion is logistically difficult because of skin lesions due to maceration and presence of hydro static pressures alters respiratory mechanisms. The temperature of the bath is also critical. Thermoregulatory mechanisms disturb the hydrostatic effects if the bath is not thermoneutral, i.e., kept at 33-35° C. In general, the comparisons of the effects of space

flight and head-down tilt support the use of tilt as a simulation method for cardiovascular studies since it, rather than bed rest and water immersion, most closely resembles conditions during weightlessness.

During normal erect posture most blood volume is usually maintained below the level of the heart. Under zero gravity or head-down rest there is a headward shift of body fluids form the lower portions of the body. (5,6,10,18,40,44) Therefore, blood tends to pool in the thoracic cavity. This massive fluid shift induces adaptive changes in other body systems. These changes are referred to in Figure 1 below.

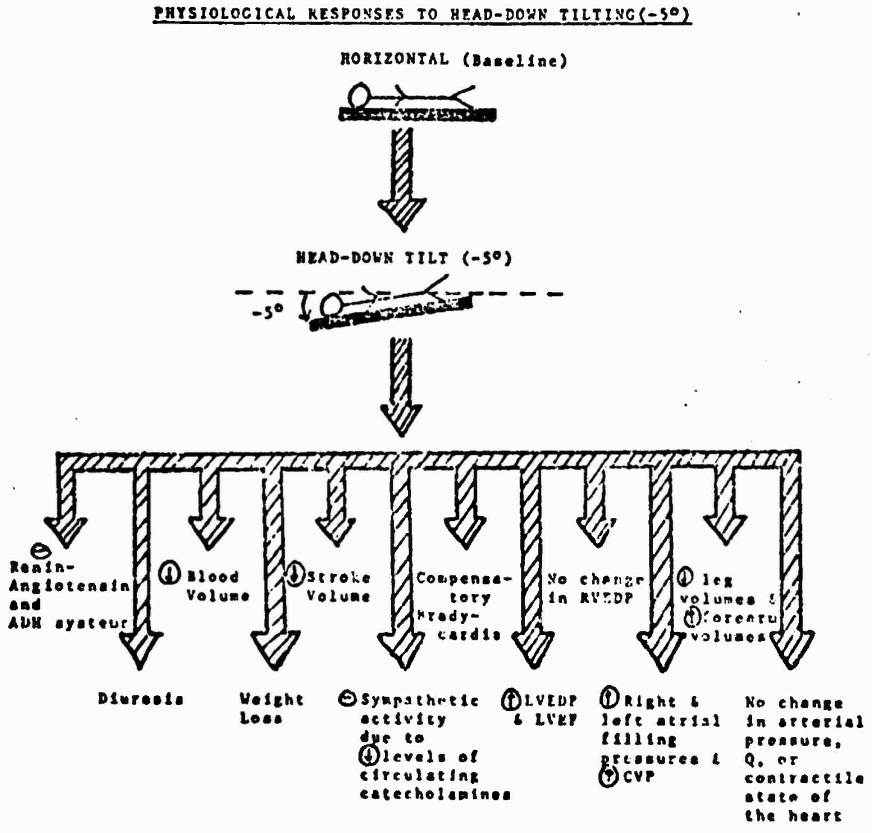


FIGURE 1

Due to the ambitious nature of this project, time limitations and necessary approval of this protocol from the Human Use Committee, this study was not carried out this summer, but will be done next summer under

the same USAF-SCEEE Summer Support Program sponsored by the Air Force Office of Scientific Research. What follows is the proposed protocol for this experiment.

However, a preliminary study involving dog models was undertaken this summer and results showed the expected relative changes in the cardiovascular system of the dog when it was tilted 5° head-down.

A study similar to the one outlined here using trained and untrained dogs instead of human subjects as the experimental models is currently being written for an AFOSR mini-grant and if approved it will be conducted in the fall and winter of 1983.

II. OBJECTIVES:

The objectives of this experiment are to determine if individuals of different working capacities respond differently to head-down tilting and to elucidate the time course of volume changes in the right and left heart, so that a correlation between these changes with other biochemical and physiological changes may be made. In this way a clearer understanding of the mechanisms involved with volume regulation might be brought forth.

III. PROCEDURES:

A.) Human Study: The experiment will use a radionuclide cardiovascular imaging technique that will measure blood volume changes in both the right and left heart. Individuals undertaking this protocol will undergo and baseline stress thallium scan and baseline multiple gated radionuclide angiogram (MUGA). MUGA's performed during the experiment will require injections of technetium (Tc^{99m}) which will be administered periodically throughout the testing period if the nuclear tag elutes rapidly from the individual's red blood cells. The total expected dose of radiation will be from 1.845 RADS to 2.295 RADS. One week prior to the actual declination study the subject will perform a maximal exercise stress test on the bicycle ergometer so that aerobic capacity can be determined and a baseline stress thallium scan and MUGA can be performed in order to study right and left heart functioning. The attenuation factor will also be determined.

The subject will be asked to report to the bedrest facility at 0800 hrs. on a pre-selected day. He will spend approximately 28 hours in the facility and will be released by approximately 1200 hrs. the following day. Upon entering the bedrest facility the subject will recline in a horizontal position on a tilt-table. He will be instrumented with a standard twelve lead ECG, blood pressure cuff and strain guages to measure anthropomorphic changes. The first injection of technetium will be made and baseline measurements will be taken including blood and urine samples. (see Figure 2) The subject will then be tilted 5° head-down and data collection will take place every 15 minutes for the first hour, every 30 minutes for the next two hours, and once every hour until 1700 hrs. No data will be collected during the night. The following morning at 0800 hrs. there will be another collection of data. A repeat injection of technetium will be necessary since its half-life is only 6 hours. At 0900 hrs. the subject will return to the horizontal position upon which data will also be collected. A post- maximal exercise stress test on the bicycle ergometer will be conducted and right and left heart functioning will be studied.

TIMELINE

ONE WEEK PRIOR:	(Pre) Max Exercise Stress Test (Bicycle ergometer)	1700	Data collection; radionuclide imaging and blood samples
	1st pass ^{201}Thallium injection (individual attenuations)	1200-1300	Lunch
	MUGA --- ^{99m}Technetium injection	1300-1700 (hourly)	Data collection; radionuclide imaging and blood sample every hour
	Max VO_2 measurement with radionuclide imaging	1700-1800	Dinner
TIME	ACTIVITY	1800-0800 (next day)	Rest in tilt position; no data collection during the night
0800-0900	Briefing and preparation of the subject ^{99m}Isotope (technetium) injection	0900-1000	^{99m}Isotope (technetium) injection RETURN TO HORIZONTAL Data collection; radionuclide imaging and blood sample
0900	Horizontal - baseline measurements TILT	1000-1100	(Post) Max Exercise Stress Test (Bicycle ergometer) Max VO_2 measurement with radionuclide imaging
0915 0930 0945 1000	Data collection; radionuclide imaging and blood samples		
1030 1100	Data collection; radionuclide imaging and blood samples	1200	Dismissal

FIGURE 2

The approach taken in this protocol will allow elucidation of the time course of volume changes in the right and left heart and possible correlation between these changes with other biochemical and physiological changes.

B.) Dog Study: A healthy mongrel dog(17.6 kg) was used as an experimental animal. The animal was anesthesized with thiamyl sodium (10 mg/kg i.v.) and ventilated under positive pressure through a cuffed endotracheal tube with a standard volume animal respirator (Harvard Apparatus, Waltham, MA). The animal was positioned on the right side and an incision was made on the left hind leg to expose the saphenous vein into which a catheter was inserted. The catheter was connected to a two-way stopcock so that additional thiamyl sodium could be added to effect. The dog was then placed on its side and a small neck incision was made. The right common carotid was exposed and a pulsed Doppler flow probe tied into place. Careful attention was paid to assure normal carotid artery flow velocity with no constriction. The pulsed Doppler flow probe was calibrated in terms of Doppler frequency shift. A baseline zero can be established and a linear relationship between flow and frequency shift has been established in vivo(17). A Doppler flow probe was also placed on the right femoral artery under direct vision. Precaution was taken to prevent arterial constriction. The dog was again positioned on the right side and a left flank incision made through which the left renal artery was identified and encircled with a pulsed Doppler flow probe. As with the other flow probes, careful attention was taken to avoid constriction of the artery. A catheter was inserted into the dog's urinary bladder so that urine output could be measured. Next, a catheter was inserted into the left femoral artery and threaded up into the aorta for measurement of the systemic arterial pressure (Statham p23Db pressure transducer). Electrocardiographic (ECG) lead AVF was established through a needle lead in all four limbs. A left thoracotomy was performed through the fifth intercostal space and the heart was suspended in a pericardial cradle. With the aid of a purse string suture, a solid state pressure transducer (Koingsberg P-7,

Pasadena, CA) was inserted through a stab wound into the left ventricular apex.

The instrumentation allowed for measurement of left ventricular pressure(LVP), left ventricular end diastolic pressure(LVEDP), systemic arterial systolic and diastolic pressures(SP and DP), heart rate(HR), renal, carotid and femoral artery blood flow (RF,CF,FF), urine output and calculation of the first derivative of left ventricular pressure (LVdP/dt). Recordings were made on an eight channel recorder(Gould brush mark 200).

Following instrumentation, control baseline measurements were made for five minutes of LVP, LVEDP,SP,DP,HR,CF,RF, and FF with the dog in the horizontal position. The dog was then tilted 5° head-down. Measurements were taken every five minutes for one hour. After one hour head-down the dog was then brought back up to the horizontal position and measurements were taken at five and seven minutes post-tilt. After the last measurements were made, the dog was euthanized with a saturated solution of potassium chloride.

IV. RESULTS:

The results obtained from this experiment show the expected changes in the cardiovascular parameters. The contractility of the heart increased, left ventricular pressure increased, left ventricular end diastolic pressure increased, heart rate increased, RF and CF increased, and SP and DP increased after the dog was tilted 5° for one hour. Most of the changes occurred during the first 30 minutes head-down and then leveled out for the rest of the hour. When the dog was brought back up to the horizontal position, no change in the readings were evident within five and seven minutes.

V. RECOMMENDATIONS:

Despite the many studies on man's adaptation to simulated zero gravity, no conclusive evidence exists as to the effect of physical fitness on this response to zero gravity. Although objective comparisons of athletes and non-athletes give variable results, it has been suggested

that the more physically fit person (high VO_2) is less tolerant to orthostatic stress when volume depleted as a result of diuresis during head-down rest. This loss of blood volume is analogous to heat-related volume loss(6,18,19,33).

The investigators in this study would eventually like to find out if, in fact, there are relative physiological and biochemical changes between trained and untrained subjects, dog and human, and if so, to what degree these parameters vary when the subjects go from the horizontal position to a 5° head-down position. Understanding the mechanisms involved requires elucidation of changes in plasma renin-angiotensin levels, plasma catecholamine levels and body fluid shifts. As a result, further significance could be drawn from this study regarding the use of the head-down tilt method as a means of pre-flight preparation for the pilot.

Studies have shown that increased activity of the carotid sinus nerve due to hypervolemia (23,35) or manual stimulation with electrodes (42) results in a reflex decrease in the force of atrial systole by a decrease in sympathetic activity to the heart and an increase in efferent vagal activity to the heart. A suggestion for follow-up research would be to isolate the carotid sinus nerve in an animal model. A small neck incision could be made to expose the left internal carotid artery at its origin where there is a bulbous enlargement, the carotid sinus, which is about 3 mm in diameter and 4 mm in length (25). It contains an afferent fiber called either the carotid sinus nerve or the Hering nerve which is a branch of the glossopharyngeal nerve (IX). A microelectrode could then be placed on the carotid sinus nerve to measure its electrical activity. Then one could determine if it was stimulation of the carotid sinus nerve or the reduction in circulating catecholamines that decreases sympathetic activity to the heart resulting in an attenuation of the force of atrial systole.

REFERENCES

1. Abel, F.L., J.H. Pierce, and W.G. Guntheroth, "Baroreceptors influence on postural changes in blood pressure and carotid blood flow," Am. J. Physiol. Vol. 205(2), pp. 360-364, 1963.

2. Abel, F.L., and J.A. Waldhausen, "Influence of posture and passive tilting on venous return and cardiac output. Am. J. Physiol. Vol. 215(5), pp. 1058-1066, 1968.

3. Auger, R.G. J.E. Zehr, R.G. Siekert, and W.E. Segar, "Position effect on antidiuretic hormone: blood levels in bedrest patients," Arch. Neurol. Vol. 23, pp. 513-517, 1970.

4. Bevegard, S., J. Castendors, and L. Lindblad, "Effect of changes in blood bolume distribution on circulatory variables and plasma renin activity in man," Acta Physiol. Scand. Vol. 99, pp. 237-245, 1977.

5. Blomqvist, C.G., J.V. Nixon, R.L. Johnson, J.H. Mitchell, "Early cardiovascular adaptation to zero gravity simulated by head-down tilt," Acta Astronautica Vol. 7, pp. 543-553, 1980.

6. Blomqvist, C.G., and H.L. Stone, "Cardiovascular adjustments to gravitational stress," Handbook of Physiology-The Cardiovascular Sysem IV, in publication.

7. Bonde-Peterson, F., Y. Suzki, and T. Sadamoto, "Cardiovascular responses to isometric exercise during simulated zero gravity," The Physiologist pp.37-38.

8. Brennan, L.A., R.L. Malvin, K.E. Jochim, and D.E. Roberts, "Influence of right and left atrial receptors on plasma concentrations of ADH and renin," Am. J. Physiol. Vol. 221(1), pp. 273-278, 1971.

9. Buderer, M.C., J.A. Rummel, E.L. Michel, D.G. Mauldin, and C.F. Sawin, "Exercise cardiac output following Skylab mission," Aviat. Space Environ. Vol. 47(4), pp. 365-372, 1976.

10. Epstein, M., "Cardiovascular and renal effects of head-out water immersion in man," Circ. Res. Vol. 39(5), pp. 619-628, 1976.

11. Gazenko, O.G., et. al., "Effects of various countermeasures against the adverse effects of weightlessness on central circulation in the healthy man," Aviat. Space Environ. Med. Vol. 53(6), pp. 523-530, 1982.

12. Gomez-Sanchez, C., D.C. Kim, N.M. Kaplan," A radioimmunoassay for plasma aldosterone by immunologic purification," J. Clin. Endocrin. Metab. Vol. 36, pp. 795-798, 1973.

13. Goodall. McC., M. McCally, and J.D. Graveline,"Urinary

adrenaline and noradrenaline response to simulated weightless state," Am. J. Physiol. Vol. 206(2), pp. 431-436, 1964.

14. Graveline, D.E., and M. McCally, "Body fluid distribution: implications for zero gravity,: Aerosp. Med. Vol. 33(11), pp. 1281-1290, 1962.

15. Greenleaf, J.E., D. Sciaraffa, E. Shwartz, L.C. Keil, and P,J. Brock, "Exercise training hypotension: implications for plasma volume, renin, and vasopressin," J. Appl. Physiol.: Respirat. Environ., Exercise Physiol. Vol. 51(2), pp. 298-305, 1981.

16. Haber, E., T. Koerner, L.B. Page, B. Klenian, and A. Purnode, "Application of a radioimmunoassay for angiotensin I to the physiological measurements of plasma renin activity in normal human subjects," J. Clin. Endocrinol. Metab. Vol. 29, pp. 1349-1355, 1969.

17. Hartley, C.J., H.G, Hanley, R.M. Lewis, and F.S. Cole, "Synchronized pulsed Doppler blood flow and ultrasonic dimension measurement in concious dogs," Ultrasound in Med. and Biol. pp. 99-110, 1978.

18. Katkov, V.E., V.V. Chestukkin, R.I. Lapteva, V. A. Yakovleva, V.M. Mikhailov, O. Kh. Zybin, and V.N. Utkin, "Central andcerebral hemodynamics and metabolism of the healthy man during head-down tilting," Aviat. Space Environ. Med. Vol. 50 (2), pp. 147-153, 1979.

19. Klein, K.E., H.M. Wegmann, and P. Kuklinski, "Athletic endurance training - advantage," Aviat. Space Environ, Med. Vol. 48(3), pp. 215-222, 1977.

20. Klein, K.E., H.M. Wegmann, H. Bruner, and L. Voght, "Physical fitness and tolerance to environmental extremes," Aerosp. Med. Vol. 40(9), pp. 998-1001, 1969.

21. Lamb, L.E., R.L. Johnson, P.M. Stevens, and B.E. Welch, "Cardiovascular deconditioning from space cabin simulator confinement, " Aerosp. Med. Vol. 35, pp. 420-428, 1964.

22. Liu, C.T., H.E. Hoff, and R.A. Huggins, "Circulatory and respiratory responses to plstural changes in the hemorrahagic dog," J. Appl. Physiol. Vol. 27(4), pp. 460-464, 1969.

23. Mancia, G., J.T. Shepherd, and D.E. Donald, "Interplay among carotid sinus, cardiopulmonary, and carotid body reflexes in dogs," Am. J. Physiol. Vol. 230(1), pp. 19-24, 1976.

24. Michel, E.L., J.A. Rummel, and C.F. Sawin, "Skylab experiment M-171 "metabolic activity" - results of the first manned mission," In Biomedical Results from Skylab. (Edited by R.S. Johnston and L.F. Dietlein), pp. 284-312, NASA, Washington, D.C.

25. Miller, M.E., "The head," In Guide to the Dissection of the Dog, p. 289, Ithaca, New York.

26. Miller, P.B., R.L. Johnson, and L.E. Lamb, "Effects of four weeks of absolute bed rest on circulatory functions in man," Aerosp. Med. Vol. 42, pp. 1194-1200, 1964.

27. Myhre,L.G., D.K. Brown, F.G. Hall, and D.B. Dill, "The use of carbon monoxide and T-1824 for determining blood volume," Clin. Chem. Vol. 14, pp. 1197-1205, 1969.

28. Nickel, J.F., L. Levine, and J.A. Gagnon, "Effects of acute passive tilting on arterial pressure, renal hemodynamics, and urinary electrolyte excretion in the dog." J. Appl. Physiol. Vol. 9, pp. 176-184, 1956.

29. Nixon, J.V., R. Gordon Murray, C. Bryant, R.L. Johnson, J.H. Mitchell, O. Bryan Holland, C. Gomez-Sanchez, P. Vergne-Marini, and C.G. Blomqvist, "Early cardiovascular adaptation to simulated zero gravity," J. Appl. Physiol. Respirat. Environ. Exercise Physiol. Vol. 46(3), pp.541-548, 1979.

30. Rummel, J.A., E.L. Michel, and C.A. Berry, "Physiological response to exercise after space flight - Apollo 7 to Apollo 11," Aerosp. Med. Vol. 44(3), pp. 235-238, 1973.

31. Rummel, J.A., C.F. Sawin, M.C. Buderer, D.G. Mauldin, and E.L. Michel, "Physiological response to exercise after space flight - Apollo 14 through Apollo 17," Aviat. Space Environ. Med. Vol. 46(5), pp. 679-683, 1975.

32. Rummel, J.A., E.L. Michel, C.F. Sawin, and M.C. Buderer, "Medical experiment M-171: results from the second manned Skylab mission," Aviat. Space Environ. Med. Vol. 47(10), pp. 1056-1060, 1976.

33. Saiki, H., N. Masayuki, M. Sudoh, M. Abe, Y. Taketomi, and M. Naruse, "Effect of athletic training on physical fitness under hypodynamics," The Physiologist , pp. 39-40.

34. Saltin, B., G. Blomqvist, J.H. Mitchell, R.L. Johnson Jr., K. Wildenthal, and C.B. Chapman, "Response to exercise after bed rest and after training, A longitudinal study of adaptive changes in oxygen transport and body composition," Circ. Suppl. Vol. 7, pp. 1-78, 1968.

35. Sarnoff, S.J., J.P. Gilmore, S.K. Brockman, J.H. Mitchell, R.J. Linden, "Regulation of ventricular contraction by the carotid sinus: its effect on atrial and ventricular dynamics," Circ. Res. Vol. 8, pp. 1123-1136, 1960.

36. Siggers, D.C., C. Salter, and P.A. Toseland, "A double isotope dilution method for differential determination of adrenaline and noradrenaline in plasma," Clinica Chemica Acta. Vol. 30, pp. 373-376, 1970.

37. Stegemann, J., U. Meier, W. Skipka, W. Hartleif, B. Hemmer, and U. Tibes, "Effects of a multi-hour immersion with intermittent exercise on urinary excretion and tilt-table tolerance in athletes and nonathletes," Aviat. Space Environ. Med. Vol. 46(1), pp. 26-29, 1975.

38. Stone, H.L., "Cardiac function and exercise training in conscious dogs," J. Appl. Physiol.: Respirat. Environ. Exercise Physiol. Vol. 42(6), pp. 824-832, 1977.

39. Tipton, C.M., R.A. Carey, W.C. Eastim, and H.H. Erickson, "A submaximal test for dogs: evaluation of effects of training, detraining, and cage confinement," J. Appl. Physiol. Vol. 37(2), pp. 27]-275, 1974.

40. Tipton, C.M., A.R. Hargens, P.D. Gollnick, S.J. Mubarad, M.R. Gonsalves, and B.J. Tucker, "Influence of head-down tilt on muscle function and Starling forces," Med. Sci. Sports Exercise, Vol. 13(2), p..102, 1981.

41. Torphy, D.E., "Effects of short-term bed rest and water immersion on plasma volume and catecholamine response to tilting," Aerosp. Med. Vol. 37, pp. 383-387, 1966.

42. Vatner, S.F., D. Franklin, and E. Braunwald, "Effects of anesthesia and sleep on circulatory response to carotid sinus nerve stimulation," Am. J. Physiol. Vol. 230(1), pp. 19-24, 1976.

43. Vogt, F.B., and P.C, Johnson, "Study of effect of water immersion on healthy adult male subjects: plasma volume and fluid-electrolyte changes," Aerosp. Med. Vol. 38, pp. 447-452, 1965.

44. Wilkens, R.W., S.E. Bradley, and C.K. Friedland, "The acute circulatory effects of the head-down position (negative G) in normal man, with a note on some measures designed to relieve cranial congestion in this position," J. Clin. Invest. Vol. 29, pp. 940-949, 1950.

45. Williams, J.A., and J. Fine, "Measurement of blood volume with a mass apparatus," N. Eng. J. Med. Vol. 264, pp. 842-848, 1961.

1983 USAF-SCEEE SUMMER RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

PINHOLE BEAM SENSORS II

Prepared by:	James Eldon Steelman, Ph.D., P.E.
Academic Rank:	Associate Professor
Department and University:	Electrical & Computer Engineering New Mexico State University
Research Location:	Advanced Concepts Branch, Applied Physics Division, Nuclear Technology Office, Air Force Weapons Laboratory
USAF Research:	Lt. Col. Ray Zazworsky
Date:	September 29, 1983
Contract No.:	F49620-82-C-0035

PINHOLE BEAM SENSORS II

by

James Eldon Steelman

ABSTRACT

Several different mechanizations of pinhole sensing of neutral particle beams were presented. Possible scintillators are examined and a scintillator followed by a "bent microscope" and a CCD detector is analyzed in detail. This system includes a laser alignment system and the assumption is made that the quantity to be measured is the distance between the pinhole beam spot and the laser spot. A five lens system is presented for detecting beam rotation and parallel displacement. The analysis indicates that beam centroid location within less than 1 μm is possible.

Two experiments are suggested. One uses a UV laser to create scintillation and provides operational experience with the system. The other experiment uses the 800 MeV LAMPF beam to determine the response versus dose for various scintillators.

The conclusion is presented that a pinhole system can satisfactorily detect the low energy "skirts" of the beam. Unfortunately, the relationship between the skirts of the beam and the centroid of the beam is uncertain.

Acknowledgements

The author would like to thank Professor Warren Peele and the Southeastern Center for Electrical Engineering Education for providing him with the opportunity to spend a very worthwhile and interesting summer at the Air Force Weapons Laboratory, Kirtland AFB, NM. He would like to acknowledge the laboratory, in particular the Advanced Concepts Branch, for its hospitality and excellent working conditions.

The author would like to thank Lt. Col. Ray Zazworsky for suggesting this area of research and for his participation in productive brainstorming sessions. He would also like to thank Capt. Mike Norris of Kirtland Air Force Base and SCEEE Fellow Dr. David Pegg for helpful and congenial discussions. The author also appreciates drafting services provided by David Aranda and computer assistance from Norman Schwers. Finally, he wishes to thank Dr. Jack Aldridge and Dr. David Holtkamp of the Los Alamos National Laboratory for helpful suggestions.

I. INTRODUCTION

The Air Force is currently performing system studies of neutral particle beam weapons. Among the problems under study is neutral particle beam sensing. Two primary sensing mechanisms are receiving consideration. One of these mechanisms is the laser resonant backscatter effect and the other is pinhole beam sensing near the beam edges.

The laser resonant backscatter technique senses the centroid of the complete beam and, for this reason, has been considered as the primary beam sensing mechanism (Reference 1). However, this technique requires laser excitation and a complex electro-optical system and may not be able to sample rapidly.

The pinhole beam sensing technique requires simple hardware and offers the possibility of providing beam shape or quality information. However, it does not measure the beam centroid but rather infers centroid behavior from measurements at the beam edges. This report provides an examination of several different pinhole sensing mechanisations, a proposed design, and two proposed experiments. The general pinhole system is shown in Figure 1.

In a 1982 SCEEE research report (Reference 2), Steelman examined two types of solid state detectors and concluded that solid state detectors would not withstand even a relatively low beam intensity for an adequate period of time.

In the same 1982 report, Steelman examined three different fluorescent detector systems. The system proposed by Lockheed is shown in Figure 2.

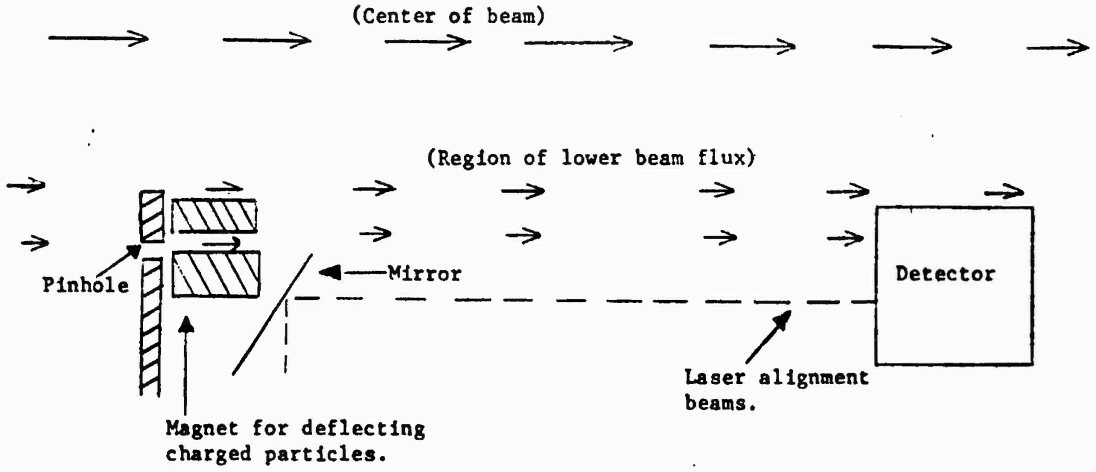


Figure 1. General Pinhole System

The Lockheed system suffers from an aspect problem, i.e. when the fluorescing cylinder is viewed off-axis, it no longer appears as a circular spot and the centroid calculation becomes difficult.

Another system which offers some possibilities uses fiber optics to image the fluorescent spot on an optical detector. Unfortunately, this system requires that the critical optical fibers be exposed to the particle beam.

Finally, Steelman selected the fluorescent detector-imaging system shown in Figure 3.

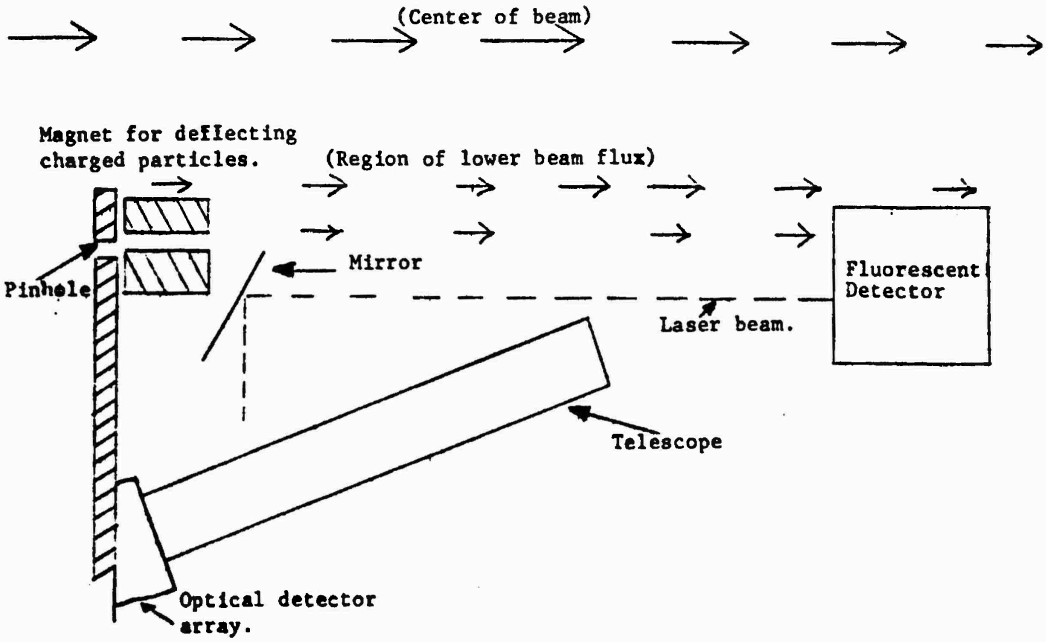


Figure 2. Lockheed Pinhole Mechanization

II. OBJECTIVES

The objectives of this research are to perform the detailed design of the pinhole beam sensing system of Figure 3 and to suggest experiments for resolving any uncertainties and for demonstrating the feasibility of the design.

III. OPTICAL SENSORS

In the 1982 SCEEE research report (Reference 2), Steelman examined a quad cell sensor proposed by Draper (Reference 1) and an array detector proposed by Stanton and Hill (Reference 3). The Draper proposed quad cell (Figure 4) will be treated first.

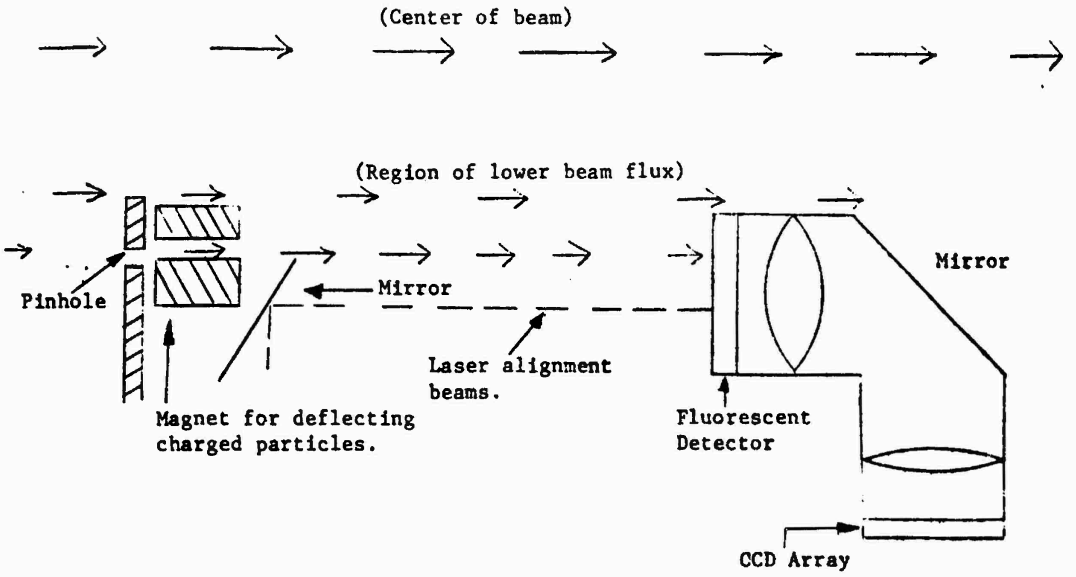


Figure 3. Proposed Pinhole System

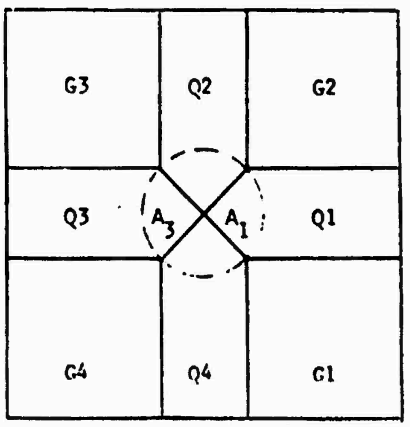


Figure 4. Quad Cell Detector (with Guard cells)

The displacement formula is

$$\Delta x_c = K D \frac{(A_1 - A_3)}{(A_1 + A_3)} \qquad (1)$$

where D = diameter of pinhole spot

A = area of spot on cell i.

For this displacement formula with the spot at the center of quad cell detector K = 0.443 and the standard deviation of the centroid is

$$\sigma_c = \frac{0.443}{\sqrt{F\Delta t}} \qquad (2)$$

Or in terms of the signal-to-noise ratio in each cell,

$$\sigma_c = \frac{0.196}{(S/N)} \qquad (3)$$

The quad cell detector was rejected for two reasons. First, the quad cell must be located with absolute positional accuracy. Second, the array detector offers superior performance.

The Texas Instruments TC-201 virtual phase (Reference 4) CCD array was selected for an optical area detector. Key parameters for the TC-201 follow (Reference 5).

Quantum efficiency = 45 to 50%.

(for wavelengths in the range .5 to .75μm)

328 pixels (horizontally) by 245 pixels (vertically)

Noise level of 150 electrons per pixel

Clock rate 6.3 MHz (typical), 15 MHz (maximum)

Full well capacity of 700,000 electrons.

Pixels are 24.4 μm square.

Maximum frame time of 1 second.

A pinhole beam image centered on the center pixel of a 3 by 3 sub-array is shown in Figure 5. If x is a small displacement of the beam image center with respect to the pixel center, the areas A_1 and A_3 are given by

$$A_1 = R^2 \arccos \frac{s}{R} - s\sqrt{R^2 - s^2} \qquad (4)$$

where $s = 0.5P - x$

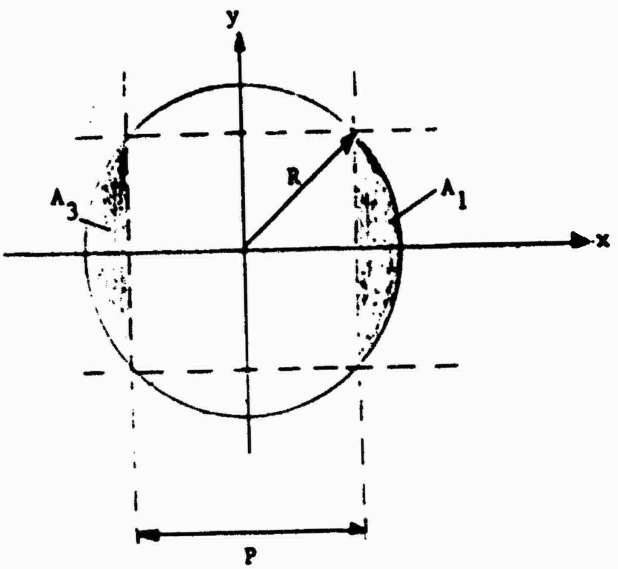


Figure 5. Pinhole Spot on Center Cell of 3 by 3 Array

$$A_3 = R^2 \arccos \frac{t}{R} - t \sqrt{R^2 - t^2} \qquad (5)$$

where $t = 0.5P + x$

The centroid formula is then

$$x = K P \frac{A_1 - A_3}{A_t} \qquad (6)$$

where P is the pixel width

A_t is total image area

K is a constant

For this formula, the standard deviation of the centroid is

$$\sigma_c = \frac{K P}{A_t} \sqrt{\frac{A_1 + A_3}{F \Delta t}} \qquad (7)$$

where Δt is sample time

F is flux in consistent units.

For the optimum configuration of Figure 5 and for small x,

$$K = \frac{\pi}{8} \frac{D}{P} \frac{1}{1-(P/D)^2} \qquad (8)$$

If the center of he pinhole image is shifted to fall on a pixel boundary,

$$K = 1 \qquad (9)$$

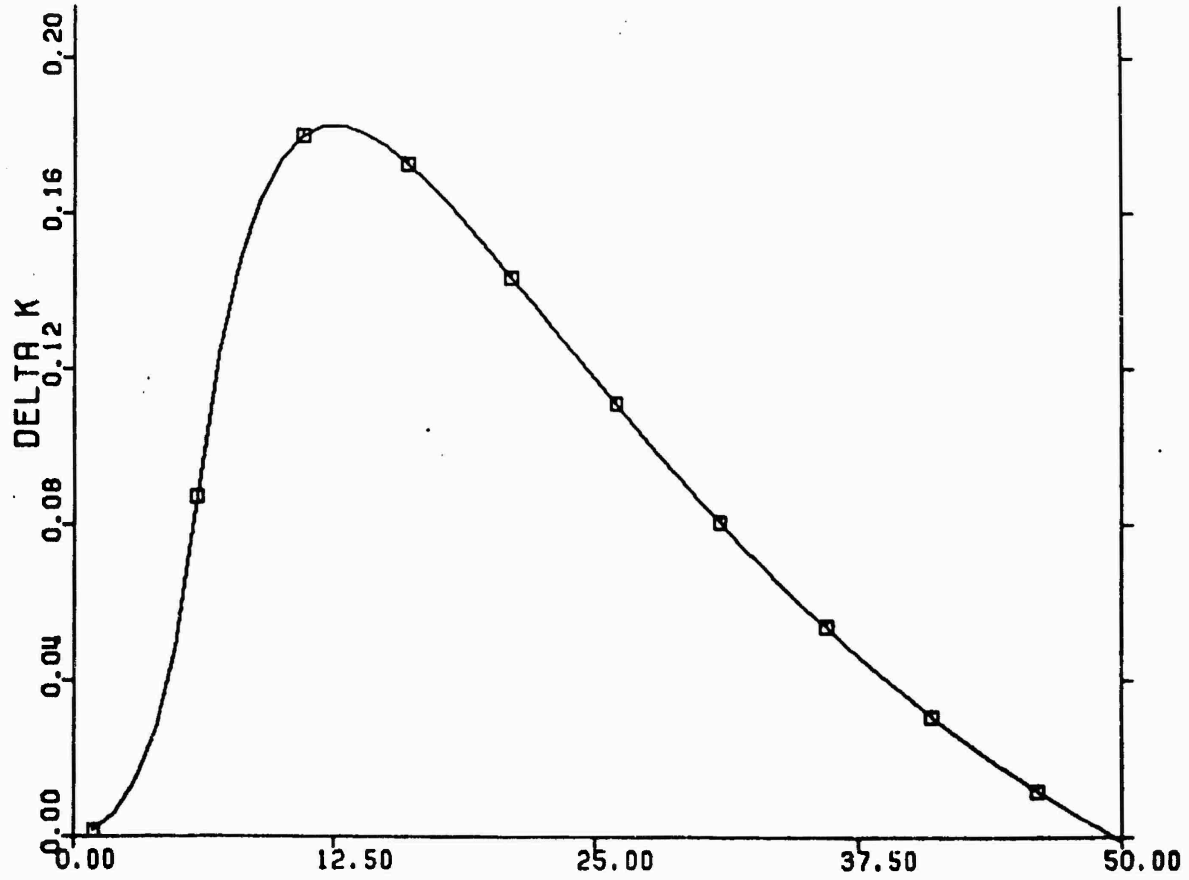


Figure 6 Error in K vs
X/D (D = Pinhole dia)

Figure 6 shows the difference between the exact K from eq. (6) and the approximate K from eq. (8) for P/D = 0.9 as a function of x. For this case the approximate K is 1.0, which matches the exact K for small x and for x = P/2. This figure reveals that, if x deviates much from the desired value of 0, then K will have to be found from a table look up procedure. Figure 7 shows the standard deviation in μm for FAt = 1.0 (from eq. 7) for the exact K and the approximate K. Note that the standard deviation is less than .75 μm for x = D/2.

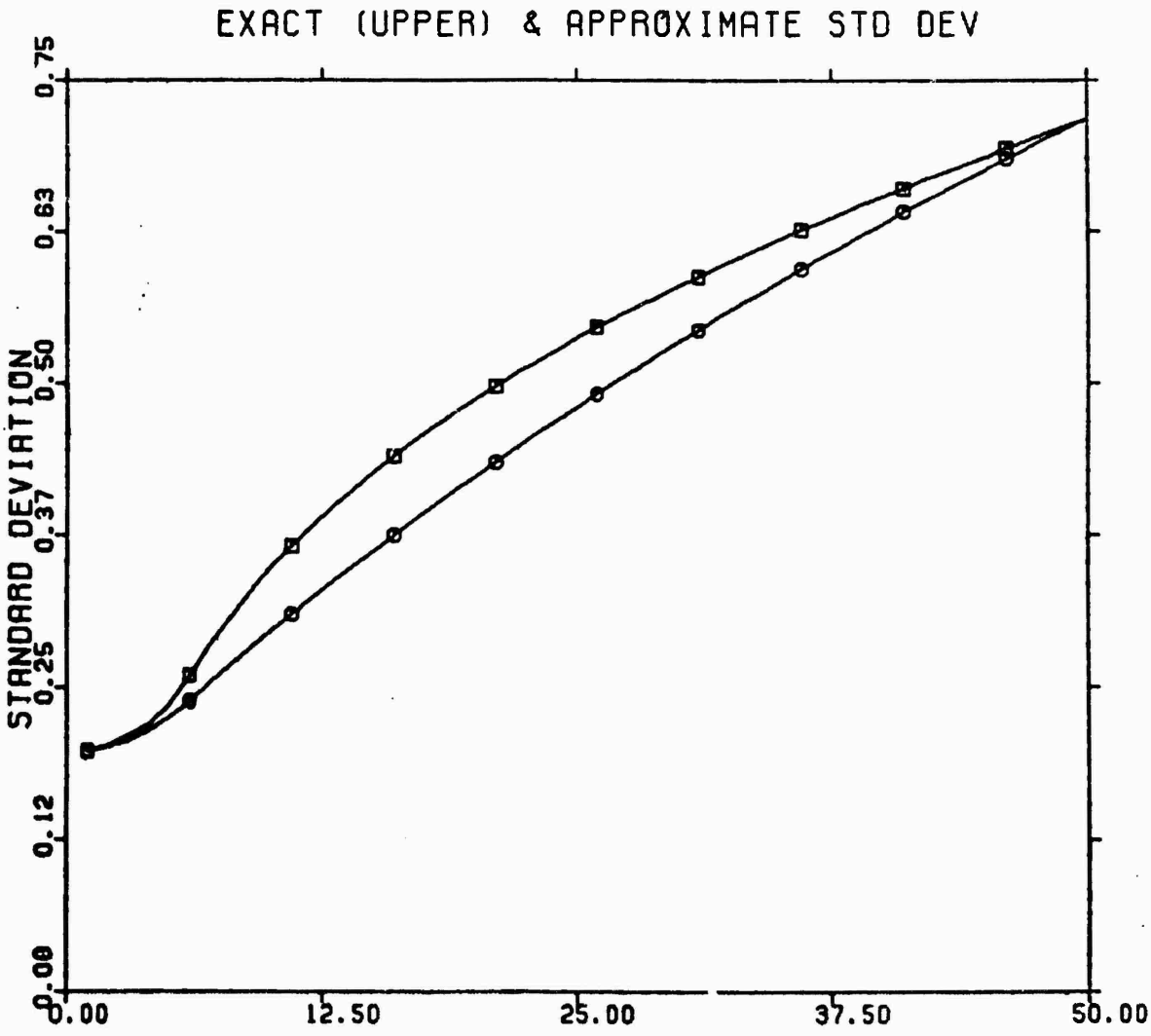


Figure 7. Standard Deviation vs X/D
for Pât = 1

IV. SCINTILLATION DETECTORS

In scintillation detectors the passage of radiation creates an ionized channel in the material (References 6, 7, 8, and 9). As the ionized electrons recombine part of the ionization energy reappears as photons. Three possible scintillators deserve consideration. (See Table I).

Material	dE/dx (MeV/cm)	Optical Efficiency	Optical Energy(eV/μm)	Response Peak(μm)
CsI(Tl)	18.26	5.85%	107	.565
NE216(Liquid)	4.0	4.4%	17.6	.425
Ruby(Al_2O_3)	5.06	.12%	.61	Red

Table 1. Possible Scintilators for 200 MeV Protons

The energy deposition rate for CsI(Tl) (cesium iodide with a thallium dopant) was obtained directly from Janni (Reference 10). The energy rate for NE216 with an H/C ratio of 1.171 was obtained by scaling Janni's rate for toulene (H/C ratio of 1.143) by the ratios of their densities. The energy rate for ruby was produced by a KAFB computer code. The optical efficiencies for CsI(Tl) and NE216 were taken from Knoll (Reference 8). He gives an absolute efficiency of 13% for NaI(Tl). His efficiency for CsI(Tl) is 45% of the value for NaI(Tl) and his value for NE216 is 78% of

the value for anthracene. Anthracene, in turn has an efficiency 43.5% that of NaI(Tl). Finally, the efficiency of ruby is an unpublished result given by Jack Aldridge of LANL.

The optical energy in eV/μm reveals that CsI(Tl) produces more optical energy per μm and is the obvious first choice. If NE216 is used the flux level must be increased a factor of 6 to produce the same optical energy - but because the color does not match the CCD array detector a flux increase of 10 is more realistic. A liquid scintillator can flow through the scintillation region and, theoretically, withstand any flux level. (Of course, creating a uniform thin flowing sheet of liquid is not easy).

Finally, ruby would require that the flux be increased by a factor of 175 to yield the same light output. Ruby was included because of its hardness.

V. BEAM INTENSITY AND GRADIENTS

Two problems concerning the beam must be solved. All these solutions assume that the beam has a circular normality distribution. The problems are:

1. For reference purposes, what is the maximum flux for each beam?

2. How far from the center of the beam to the region of desired flux intensity?

3. How large can the pinhole be before the beam gradient introduces a bias into the centroid calculation?

Table 2 shows the maximum flux for the beams of interest.

σ(cm)

I(mA)	1	2	5	8	10
1	9.93	2.48	3.97	1.55	9.93
10	99.3	24.8	39.7	15.5	99.3
100	993.	248.	397.	155.	993
500	4967.	1242.	1987.	776.	4967.
Column Multiplier	10^6	10^6	10^5	10^5	10^4

Table 2. Maximum Beam Flux (Particles/μm^2 sec)

To find regions of desired flux observe that, for a circular normal beam,

$$F(r) = F_{max} \exp(-r^2/2\sigma^2) \qquad (10)$$

Thus, for a desired flux,

$$r = \sqrt{-2\sigma^2 \ln F/F_{max}} \qquad (11)$$

For fluxes in the range of 10 to 100 particles per μm^2 second should provide small enough σ_c. Table 3 shows the distance in cm from the beam center to a region of 10 particles/μm^2 second.

σ(cm)

I(mA)	1	2	5	8	10
1	5.26	9.97	23.0	35.2	42.9
10	5.68	10.86	25.4	39.1	48.0
100	6.07	11.67	27.6	42.7	52.6
500	6.33	12.21	29.0	45.1	55.5

Note: All entries in cm

Table 3. Distance from Beam Center to Pinhole Center (F=10/μm^2 sec)

Table 4 tabulates the distance from the beam center to a region of 100 particles/μm^2 sec.

	σ(cm)				
I(mA)	1	2	5	8	10
1	4.80	9.00	20.4	30.7	37.2
10	5.26	9.97	23.0	35.2	42.9
100	5.68	10.86	25.4	39.1	48.0
500	5.95	11.43	26.9	41.7	51.2

Note: All entries in cm.

Table 4. Distance From Beam Center to Pinhole Center (F=100/μm^2 sec)

The pinhole cannot be too small because the optics will introduce a minimum blur diameter. The pinhole cannot be too large because the non-uniform statistics will introduce a bias. As a basis for pinhole size selection, compare the bias count with the count due to a 0.1μm offset. For a 0.1μm offset

$$\text{count} = 0.2 \sqrt{D^2 - P^2}\ (F\Delta t) \qquad (12)$$

To develop the bias count, consider the approximate count on A_3.

$$\text{Count } (A_3) = F(r - \Delta)\ A_3\ \Delta t$$

$$= A_3\ F_{max} \exp \frac{-1}{2} \left(\frac{r - \Delta}{\sigma}\right)^2$$

$$\approx A_3\ F\Delta t \exp \frac{r\Delta}{\sigma^2}$$

Now $r\Delta$ is small; thus

$$\text{Count } (A_3) \approx A_3\ (F\Delta t) \left(1 + \frac{r\Delta}{\sigma^2}\right) \qquad (13)$$

The expression for the count on A_1 can be obtained by changing the sign on Δ.

Finally, bias = count (A_3) - count (A_1)

$$= (2A_3(F\Delta t)(r\Delta/\sigma^2)) \qquad (14)$$

where Δ = distance from pinhole center to A_3 centroid.

Rather than calculating a Δ, the worst case value of D/2 was used. For design purposes, the ratio of the bias (eq. 14) to Δ count (eq. 12) is a more convenient measure. This ratio is tabulated in Table 5.

	P=24	p=50	p=100	
	D=20	d=55.55	D=111.1	
I	σ=1	σ=1	σ=1	σ=2
1	.158	.27	1.08	.52
10	.172	.29	1.18	.56
100	.186	.32	1.27	.62
500	.194	.33	1.32	.65
$\frac{\text{Count}}{F\Delta t}$	3.6	4.8	9.6	9.6

Table 5. Bias/Count for Δx = 0.1 μm

Table 5 indicates a 111.1 μm pinhole is satisfactory for any beam sigma of 2 cm or greater. Thus, the design for the centroid finding mechanization will be based on P=97.2 μm, and D = 108 μm (D=1.11P). (P of 97.2 μm is four pixels of the TI TC-201 CCD array).

VI. OPTICAL SYSTEM ANALYSIS

The proposed pinhole system of Figure 3 will now be analyzed in detail.

The particle beam will pass through a thin CsI(Tl) crystal and then on through the mirror. In the CsI(Tl) the particle beam will form a cylinder of isotropic radiators. The plane of best focus of the optical system will be assumed to fall in the center of the cylinder. The situation is diagrammed in Figure 8.

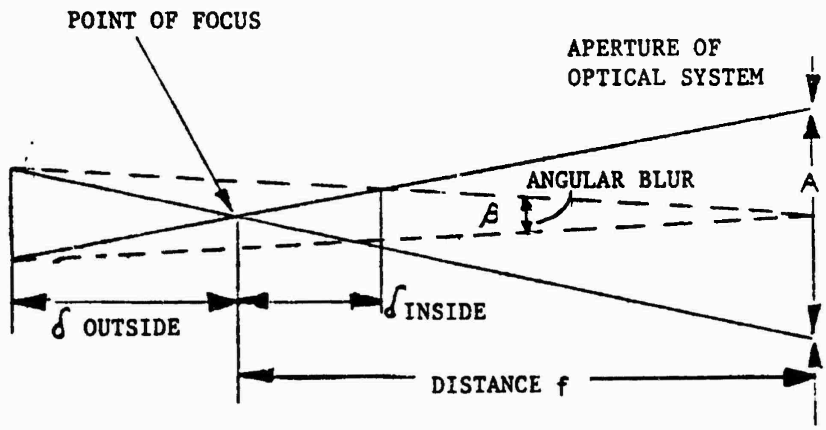


Figure 8. Blur due to cylindrical object

For this case (Reference 11),

B(blur diameter) 1 = βF

and

$$\delta = \frac{BF}{A} \tag{15}$$

If $\delta = L/2$, (half the length of the cylinder)

$$B = \frac{A}{F}\frac{L}{2} \tag{16}$$

Now assume that a blur diameter the same size as the pinhole diameter is acceptable.

Thus B = 108 μm = D (the pinhole diameter)

and

$$L = 2B\left(\frac{f}{A}\right) \qquad (17)$$

Now if the optical system has a blur diameter 10% of D

and

f/A = 2.222,

L = 2(2.22)(0.9)(108)

= 432 μm

= .432 mm

(Note that L increases directly with f/A while the energy collected by a lens is proportional to the square of A/f. Thus, a small f/A is desirable except that too small an f/A introduces too much blur in the optical system.) The normalized intensity for an image of a cylinder with this length is shown in Figure 9. Figure 9 represents a spatial frequency of about 9 cycles per mm. If the optical system can use a sharper image the intensity profile for L = 240 μm is shown in Figure 10.

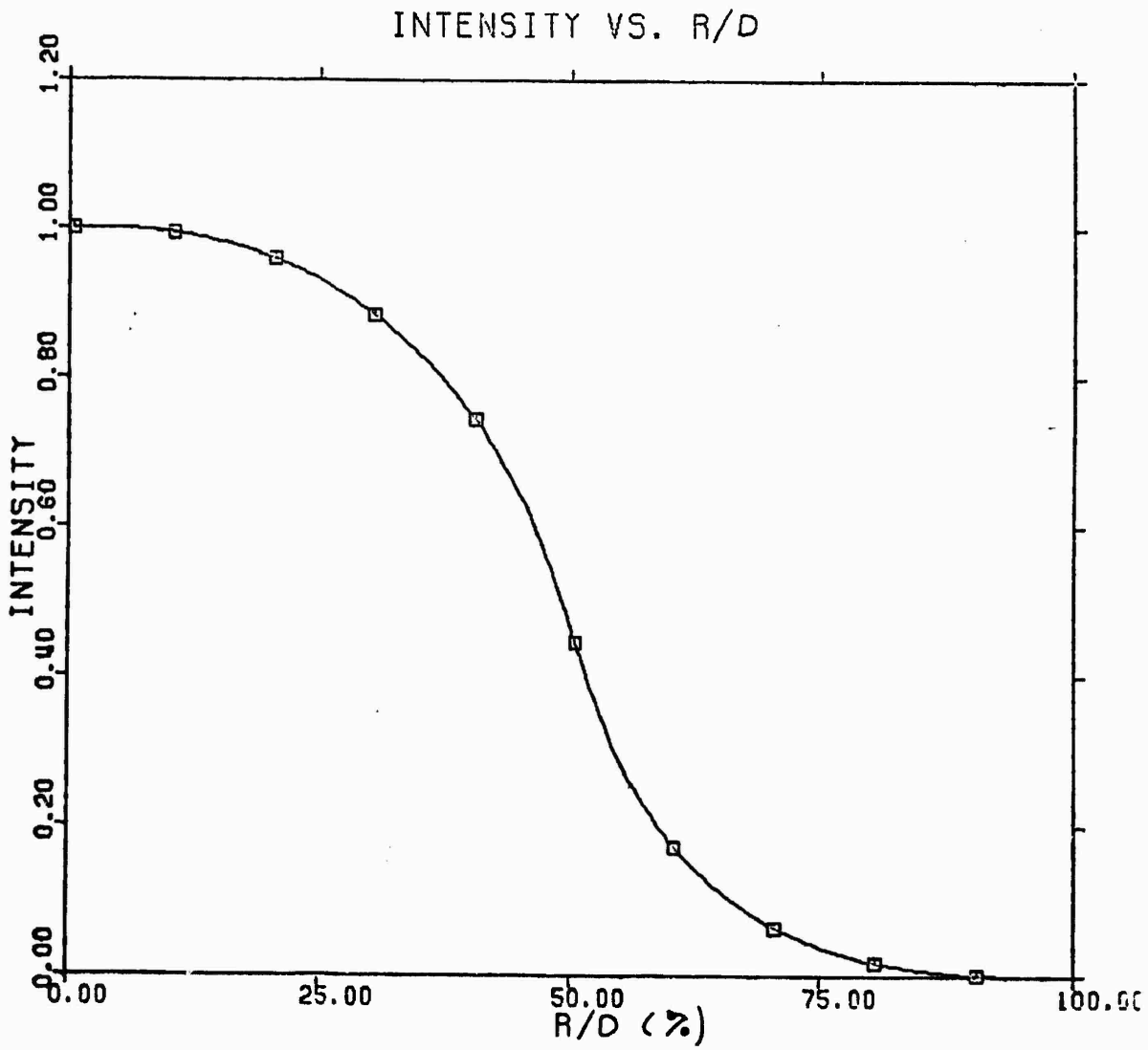


Figure 9. Intensity vs. R/D for L=432 μm

Note that Figure 10 represents a spatial frequency of about 20 cycles per mm. The optical energy created in CsI(Tl) by a single 200

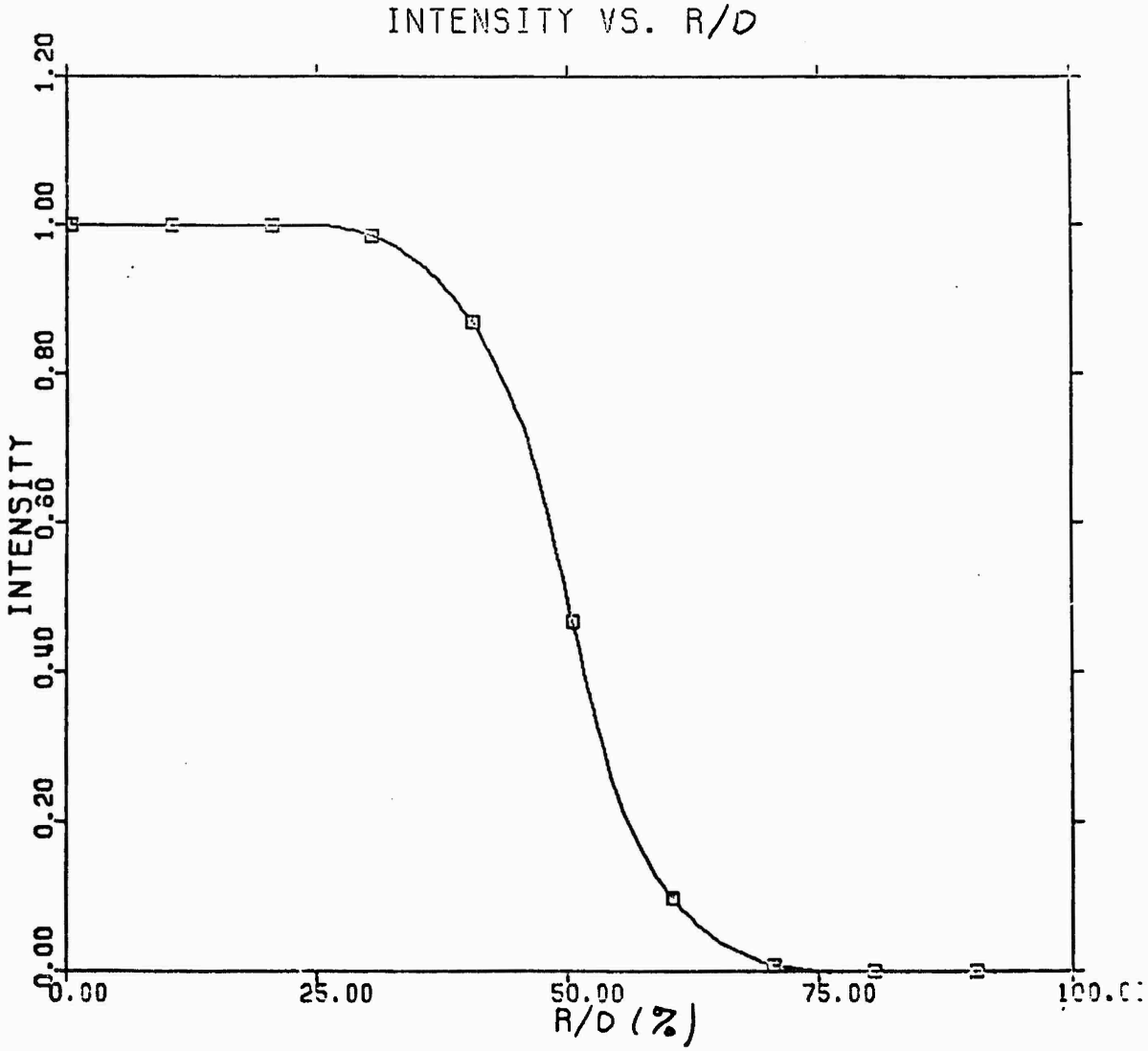


Figure 10. Intensity vs R/D for L=240 μm

MeV proton is 107 eV/μ. Thus, the optical energy created in a detector of length L by a pinhole of diameter D is

$$E = 107L\,\pi\,D^2\,F\Delta t(eV) \tag{19}$$

The response of CsI(Tl) peaks at $\lambda = 0.565$ μm. A photon of this wavelength has an energy

$$h\nu = 2.2\ eV \tag{20}$$

Thus, the total photon production is

$$N_{\gamma T} = 48.6\,\pi\,D^2\,F\Delta t \tag{21}$$

If these photons are assumed to be isotropically distributed (not exactly true), the photons intercepted by a lens of aperture A at distance F is

$$N_{\gamma A} = N_{\gamma T}\,\frac{\pi(A/2)^2}{4\pi F^2}$$

$$= \frac{1}{16}\left(\frac{A}{F}\right)^2 N_{\gamma T}$$

$$N_{\gamma A} = 3.04L\pi\,D^2\left(\frac{A}{f}\right)^2(F\Delta t) \tag{22}$$

Finally, the quantum efficiency of the TC-201 is in the 45 to 50% range. Thus, the number of electrons collected is

$$N_e = 1.4\,\pi\,D^2\,L\left(\frac{A}{f}\right)^2 F\Delta t \tag{23}$$

If the optical system were perfect (i.e., the blur radius = 0) with a magnificaion M, the electron density in the detector would be

$$F_e = \frac{5.6}{M^2} L \left(\frac{A}{f}\right)^2 F\Delta t \qquad (24)$$

For this same perfect optical system with μ=1, f/A =2.22 and L = .432 μm,

$$F_e = 490\ F\Delta t\ e/\mu m^2 \qquad (25)$$

If the blur diameter is less than or equal to the pinhole diameter then this is the electron density at the center of the detector for M = 1.

The exact performance of the optical system has not been calculated but the signal-to-noise ratio in the difference signal can be estimated (See Figure 11).

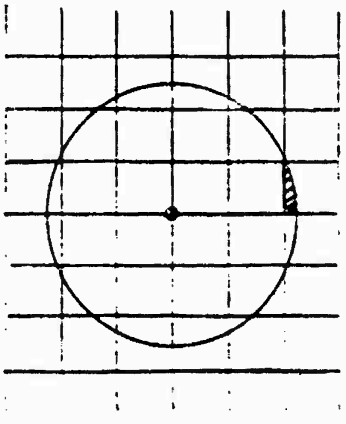


Figure 11. Critical Pixel

For a 1 μm shift of the pinhole image with respect to the CCD array the difference in area is

$$\Delta A = 94.34\ \mu m^2 \qquad (26)$$

The count $F\Delta t\ \Delta A$ is the volume of a strip from a cylinder of rotation formed by the curve of Figure 9. This strip stretches from 47.8 μm to 49.8 μm. The normalized distance to this strip varies from 45% of R/D to 63% of R/D. As a rough estimate the average normalized intensity over this area is 0.3. Thus, the signal difference is approximately.

$$S_1 - S_3 = 147\ (94.34)\ F\Delta t \qquad (27)$$

For $F\Delta t$ of 1,

$$S_1 - S_3 = 13{,}868 \text{ electrons} \qquad (28)$$

Now this signal is obtained from four pixels. Thus, the noise level for the signal is 2 times the pixel noise. The electronic noise in a pixel is

$$N_e = 150 \text{ electrons} \qquad (29)$$

If the electron count is quantized in 10 levels, the quantization noise is about

$$N_q = \frac{700{,}000}{1024} \times 0.707$$

$$N_q = 483 \text{ electrons} \qquad (30)$$

The Poisson noise is

$$N_p = \sqrt{S} = 240 \text{ e} \tag{31}$$

Thus, the total pixel noise is

$$N_{pT} = 560 \text{ e} \tag{32}$$

The signal-to-noise ratio for the difference signal is

$$S/N = 12.4 \tag{33}$$

A more detailed analysis of this system will be performed.

For beam quality information, the shape of the pinhole image is required (instead of just the centroid). To determine the shape, the same basic system can be used but a magnification will be needed. The magnification will lower the signal-to-noise ratio but this decrease can be overcome by increasing the sample time.

VII. LASER ALIGNMENT SYSTEM

The laser alignment system will require the 5 lens system shown in Figure 12.

Lens L_o, used to obtain beam angle information will have additional neutral density filters in series with it. Further, lens L_o will be located at such a distance from the CsI(Tl) that after passing through it and the optical system L_o forms a spot about the same size as the pinhole image. That such a location is possible can be

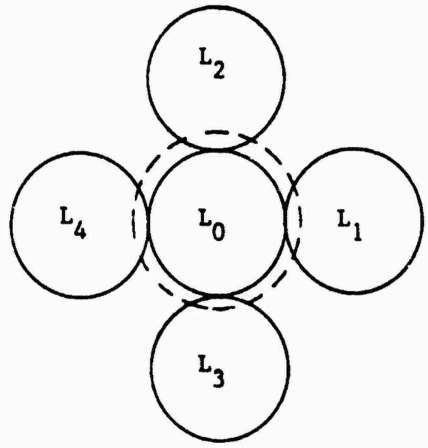


Figure 12. Laser Alignment Lenses.

established by considering an actual spot forming lens Melles Griot 01 LFS 043 (Reference 12). This diffraction limited lens forms a HeNe spot with a diameter of 37.2 μm (distance to the first dark ring of the airy disk). Even with the added blur due to the additional optics, the centroid of the spot from the central lens can probably be calculated within 0.5 μm by the techniques of Section III. (Additional research needs to be performed on this topic.) If the angle of the laser beam with respect to the lens system changes, this change will cause a shift in the centroid of the center spot.

Unfortunately, if the laser beam is translated parallel to its original location, the center spot will not move. The lenses L_1, L_2, L_3, and L_4 provide information on the translation of the laser beam. These lenses must be chosen in such a way that they form a spot with a

diameter of about 900 μm on the CCD array. The distance from the centroid of the center lens spot to the beam centroid is found by

$$\Delta C = K_c \frac{E_3 - E_1}{E_3 + E_1} \tag{34}$$

To analyze the sensitivity of eq. (34) begin by selecting 50 mm for the diameters of L_o, L_1, L_2, L_3, and L_4. (50 mm is the actual diameter of the LFS 043). Then choose 55 mm as the diameter of the laser beam. Next, assume that the spots formed by L_1, L_2, L_3, and L_4 on the CCD array are 1 mm in diameter and that the average pixel has a signal half of the saturation level or 350,000 e. This implies that 700,000 photons were incident on each pixel. Thus, the optical energy incident on each pixel is

$$E_5 = 2.45 \times 10^{-13} \text{ J} \tag{35}$$

The number of pixels covered by this spot is

$$\# \text{ pixels} = 1319 \tag{36}$$

Thus, the total optical energy incident on an outer lens is

$$E_1 = 3.245 \times 10^{-10} \text{ J} \tag{37}$$

The average energy density in that portion of the beam which intercepts an outer lens is

$$E/mm^2 = 1.212 \times 10^{-11} \ J/mm^2 \qquad (38)$$

The average number of photons in that portion of the beam which intercepts the outer lenses is

$$\text{Photons}/\mu m^2 = 34.5 \ \text{photons}/\mu m^2 \qquad (39)$$

(Of course, the laser beam will not be uniform and these calculations need to be repeated using a better model of the laser beam, probably a Gaussian model.) Now a 1 μm shift in the centroid of the laser beam will produce an increase in the lens area covered on one side and a decrease on the other side. The net change in area is

$$2\Delta A = 3.197 \times 10^4 \ \mu m^2 \qquad (40)$$

Thus, the difference in photons on the two spots is

$$\Delta \ \text{photons} = 1.10 \times 10^6 \qquad (41)$$

The 50% quantum efficiency leads to

$$S = \Delta \ \text{electrons} = 5.51 \times 10^5 \qquad (42)$$

The Poisson noise in each pixel is

$$N_p = \sqrt{s} = 592 \ e \qquad (43)$$

The quantization noise and the electronic noise were found in Section VII. The total noise per pixel is

$$N_{PT} = 732\ e \tag{44}$$

The total noise in the 1310 pixels is

$$N_T = \sqrt{1319}\ (732)$$

$$N_T = 2.66 \times 10^4 \tag{45}$$

Thus, the signal to noise ratio for the difference is

$$S/N = \frac{5.51 \times 10^5}{\sqrt{2}\ (2.66 \times 10^4)}$$

$$S/N = 14.6 \tag{46}$$

Finally, the laser power on L_o and in the beam needs to be calculated. If the L_o spot has a diameter of 108 μm, then it covers about 16 pixels. If each pixel is at half saturation,

$$E_o = 3.94 \times 10^{-12}\ J \tag{47}$$

Thus the power incident on L_o for a sample time of 10 ms is

$$P_o = 3.94 \times 10^{-10}\ W \tag{48}$$

The total laser power is

$$P_{LASER} = 2.88\ \mu w \tag{49}$$

Thus, lens L_o needs attenuating by

$$\text{Attn} = .0166 \tag{50}$$

VIII. PROPOSED EXPERIMENTS

LAMPF EXPERIMENT

One uncertainty which must be resolved is the response versus dose for CsI(Tl). An experiment at LAMPF is proposed to determine this response. The system of Figure 3 will be modified to remove the first lens from the beam and to condense the image by 1/2 to 1/5. If CsI(Tl) is unsatisfactory, the experiment will be repeated with ruby and NE216 (or another liquid scintillator).

The experiment will be performed in the "Blue Room" with an 800 MeV proton beam about 1.1 cm in diameter. In the Blue Room, 10 mA current pulses are available within a limit of 100 nA DC.

The objectives of this experiment are to

1. Determine the response vs dose of CsI(Tl). (The response will be determined for several different thallium doping levels.)
2. Determine the effect of beam caused radiation on the CCD array.

NMSU EXPERIMENT

Thallium doped cesium iodide scintillates in the presence of UV. This phenomena will be used to check out the system of Figure 3 without a 200 MeV beam. For this experiment two pinholes will be placed in contact with a thin CsI(Tl) crystal. These pinholes will be illuminated by a UV laser and the resulting fluorescence will be used to align the optical system and to determine the system performance.

The objectives of this experiment are to

1. Align and test the system (including the data recording system).
2. Gain operational experience with the system.
3. Test different optical designs.
4. Test different centroid finding algorithms.

IX. CONCLUSIONS

At this stage of the research effort, the pinhole system remains a viable option for beam sensing and alignment. However, the pinhole experiment will have to be run concurrently with a laser resonant scattering experiment to resolve two questions.

1. Is a pinhole sample on the outskirts of the beam an adequate indicator of the beam centroid?

2. What is the accuracy of the pinhole system? (No present techniques provide an independent measurement of the beam of similar accuracy.)

Finally, even if the laser resonant scattering technique is the technique of choice for final pointing, pinhole sensors could be useful for coarser pointing and for accelerator protection in case of a failure which would cause the beam to self destruct.

REFERENCE

1. T.T. Chien, et al, "Neutral Particle Beam Pointing and Tracking", AFWL-TR 80-64, pps. 50-59.

2. J.E. Steelman, "Pinhole Beam Sensors", Final Report of Contract F33615-78-D-0617, Dated October 14, 1982.

3. R.M. Stanton and R.E. Hill, "A CCD Star Sensor for Fine Pointing Control of Spaceborne Telescopes", AIAA paper 79-0394.

4. J. Hynecek, "Virtual Phase CCD Technology", Texas Instruments Reprint.

5. TI-201 CCD Area Image Sensor, Preliminary Data Sheet Dated January 1983.

6. R.L. Heath, et al, "Inorganic Scintillators: A Review of Techniques and Applications", Nuclear Instruments and Methods, Vol. 162 (1979), pps. 477-506.

7. F.D. Brooks, "Development of Organic Scintillators", Nuclear Instruments and Methods, Vol. 162 (1979), pps. 477-506.

8. G.F. Knoll, Radiation Detection and Measurement, Wiley, 1979.

9. J.B. Birks, The Theory and Practice of Scintillation Counting, Pergamon Press (distributed by MacMillan), 1964.

10. J.F. Janni, "Proton Range-Energy Tables, Part 1", Atomic Data and Nuclear Data Tables, Volume 27, Numbers 2/3, March/May 1982, Academic Press.

11. W.J. Smith, Modern Optical Engineering, McGraw-Hill, 1966.

12. Melles Griot Optics Guide 2; Melles Griot, 1770 Kettering St. Irvine, CA 92714.

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

DYNAMIC STALL: A STUDY OF THE CONSTANT PITCHING RATE CASE

Prepared by:	James H. Strickland
Academic Rank:	Professor
Department and University:	Mechanical Engineering Department Texas Tech University
Research Location:	Frank J. Seiler Research Laboratory, USAF Academy, Colorado Springs, Colorado 80840
USAF Research Personnel:	Major John Walker and Lt Hank Helin
Date:	15 July 1983
Contract No:	F49620-82-C-0035

DYNAMIC STALL: A STUDY OF THE CONSTANT PITCHING RATE CASE

by

James H. Strickland

ABSTRACT

Data were collected for an airfoil undergoing a pitching motion in which the pitching rate $\dot{\alpha}$ remained constant. These data included surface pressure data which were obtained by a previous investigator, smoke-wire flow visualization data, and airfoil surface mounted hot-wire velocity data. Large scale vortical structures on the suction side of the airfoil were seen to form during the pitching motion. As indicated from the surface pressure data, as well as the surface velocity data, these structures are highly energetic and have relatively long residence times near the airfoil. Significant increases in lift occur due to the presence of these vortical structures. These data were also compared to the predictions of an analytical model (USTAR2). Reasonable agreement between the experiment and analysis were obtained up to moderate angles of attack.

ACKNOWLEDGEMENT

The author would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research, and the Southeastern Center for Electrical Engineering Education for providing him with the opportunity to spend a very worthwhile ten weeks at the Frank J. Seiler Research Laboratory, USAF Academy, Colorado Springs, Colorado. In particular, he would like to acknowledge the helpful staff associated with the laboratory who made his stay more fruitful as well as pleasant. Finally, he would like to thank Major John Walker and Lt Hank Helin for their collaboration and hard work during this period.

I. INTRODUCTION:

Unsteady aerodynamics is an important phenomenom which has been studied more intensively in recent years. These studies have been made in connection with applications pertaining to helicopters, axial flow turbines and compressors with inlet distortions, vertical axis wind turbines, and missles and fixed wing aircraft undergoing rapid maneuvers. Lifting surfaces subjected to time dependent freestream velocity or time dependent body motions may, in some cases, also have significant stalled regions on their surfaces.

A number of approaches have been taken with regard to the prediction of unsteady stalled airfoils. Most of these approaches have been reviewed by McCroskey[1-2]. In general, these approaches range from empirical models[2-6] to models based on the Navier-Stokes equations[7-8]. The empirical models are generally applicable to small sinusoidal pitch oscillations about some relatively low angle of attack. The Navier-Stokes solutions consume enormous amounts of computer time and are usually limited to low Reynolds number solutions. Unfortunately, there are few models which can be considered as representing something in between the extreme cases of total empiricism and brute force solutions of the Navier-Stokes equations. There have been several boundary layer codes developed which can be used to predict some of the behavior associated with unsteady stall[9-11]. There have also been models of the potential flow behavior related to the shedding of leading edge vorticity as typified by the work of Ham[12]. More recently, Katz[13] simulated the unsteady separated flow over a thin cambered airfoil. The models of both Ham and Katz required empirical information regarding the appearance and position of the separation point.

The need for a general purpose unsteady airfoil model is apparent when one considers the potential applications. The number of applications along with the parametric ranges within each application legislates against the expense of either a purely experimental or brute force analytical approach to solutions.

Recently, the present author, along with other co-workers at Texas Tech University, formulated and began development on an analytical model[14] which is potentially capable of predicting dynamic effects for stalled and unstalled airfoils undergoing arbitrary airfoil motions. This model does not require input of airfoil section data and may thus be used to examine arbitary airfoil shapes. An UnSTeady AiRfoil model in 2 dimensions based on this analysis has been implemented via a computer code (USTAR2). Execution times for this code are short when compared to Navier-Stokes solutions and little empiricism is required. In order to validate this analysis, however, comparison between USTAR2 predictions and experimental data must be made for a number of cases.

Recently, workers at the USAF Academy, Frank J. Seiler Research Laboratory (FJSRL), began an experimental investigation of airfoils undergoing large amplitude pitching motion. Results for several constant pitching rate cases were obtained by Francis, Keesee, and Retelle[15]. These results consisted of pressure coefficient data which were in turn used to obtain lift and drag data. Work at FJSRL is presently continuing to obtain flow visualization data as well as hot-wire anemometer data for the constant pitching rate case. The author's AFOSR/SCEEE Summer Faculty research assignment to FJSRL can, therefore, be seen to be a logical one in that data obtained at FJSRL is useful in validating and/or developing the USTAR2 computer code while on the other hand, the understanding of experimental trends can be enhanced based on the physics contained in the USTAR2 analysis.

II. OBJECTIVES OF THE RESEARCH EFFORT:

The major research goal is to correlate unsteady airfoil data taken at FJSRL with the unsteady airfoil computer code USTAR2. These data, which will be compared with USTAR2 predictions, will include instantaneous pressure distributions, lift and drag coefficiencts, and flow visualization data from 6-inch chord NACA 0012 and NACA 0015 airfoils undergoing unsteady pitching motions. In addition, a qualitative examination of hot-wire data taken near the surface of a

pitching airfoil will be made. The motion of the airfoil in each case will be initiated from a zero angle of attack and will consist of a pitching motion where the pitching rate $\dot{\alpha}$ remains constant. The object of this work is the eventual validation of USTAR2. This work will be helpful in confirming the present level of validity of USTAR2 and in providing insight into needed revision and development of USTAR2. An additional objective is to provide some understanding of the trends evidenced in the experimental data using insight gained from physical principles.

III. RESEARCH METHODOLOGY:

In this section a brief description of analytical and experimental methods used to obtain data will be given. The analytical methods consist of the afore mentioned USTAR2 analysis as well as a simple analysis which is valid for substall angles of attack. For the experimental work, methods used in obtaining surface pressure data, flow visualization data, and hot-wire velocity data will be described briefly.

A. Analytical.

A very brief description of the method of analysis found in USTAR2 will be presented. Details of this analysis can be found in reference[14]. In addition, a simple analysis will be presented which is only valid for the constant $\dot{\alpha}$ case at substall angles of attack. This simple analysis is presented since it can be used to explain some of the data trends seen during the initial phase of motion.

1. USTAR2 Analysis.

This computer code is based upon an analysis which utilizes a doublet panel method to model the airfoil surface, an integral unsteady boundary layer scheme to model the viscous attached flow, and discrete vortices to model the detached boundary layers which form the airfoil wake region. This model has been used to successfully predict steady lift and drag coefficients as well as pressure distributions for several airfoils with both attached and detached boundary layers. In addition, calculations have been made for a limited number of cases for both attached and detached unsteady flow situations.

The potential flow model is based upon the Laplace equation for the streamfunction $\phi(\vec{r},t)$

$$\nabla^2\phi = 0 \tag{1}$$

which is valid for both steady and unsteady flow. By Green's theorem, a solution to (1) may be represented by integrals over the "boundaries" of the flow where those boundaries are replaced by surfaces across which potential jumps occur. These surfaces, as depicted in Figure 1, are represented by the airfoil surfaces and the wake sheets which spring from the trailing edge and any separation point. With Green's theorem the disturbance potential at any field point $\vec{r}$, due to the airfoil and wake surfaces, may be written as

$$\phi(\vec{r}) = \int_S \sigma \frac{\partial}{\partial \nu}\left[\frac{1}{4\pi R}\right] dS + \int_W \Delta\phi \frac{\partial}{\partial \nu}\left[\frac{1}{4\pi R}\right] dS \tag{2}$$

where σ and $\Delta\phi$ are doublet distributions on the airfoil and wake surfaces S and W respectively, ν is the surface normal at the source point, and R is the distance separating the field point and source point. Boundary conditions include a kinematic surface tangency conditions given by

$$\frac{\partial\phi}{\partial n} + (\vec{u}_\infty - \vec{u}_S)\cdot\vec{n} = 0 \qquad \text{on S} \tag{3}$$

where $\vec{u}_\infty$ and $\vec{u}_s$ are the freestream and airfoil surface velocities respectively and n is the outward normal to the airfoil surface. An additional boundary condition is the "trailing edge flow condition" which, in the present model, requires that the flow direction at the trailing edge be along the trailing edge bisector. Equations (2) and (3) are solved for σ and ϕ by first eliminating ϕ to form a single equation in σ. The airfoil and wake surfaces are then discretized to form a set of linear algebraic equations in terms of the unknown σ. The potential ϕ is then obtained from (2). The pressure distribution around the airfoil is obtained from the unsteady Bernoulli equation. Details of this analysis are presented in reference[14].

The primary function of the boundary layer analysis is to predict the presence and location of any boundary separation point on the airfoil surface. The pressure gradient and edge velocity distributions, which are used in boundary layer calculations, are obtained from the

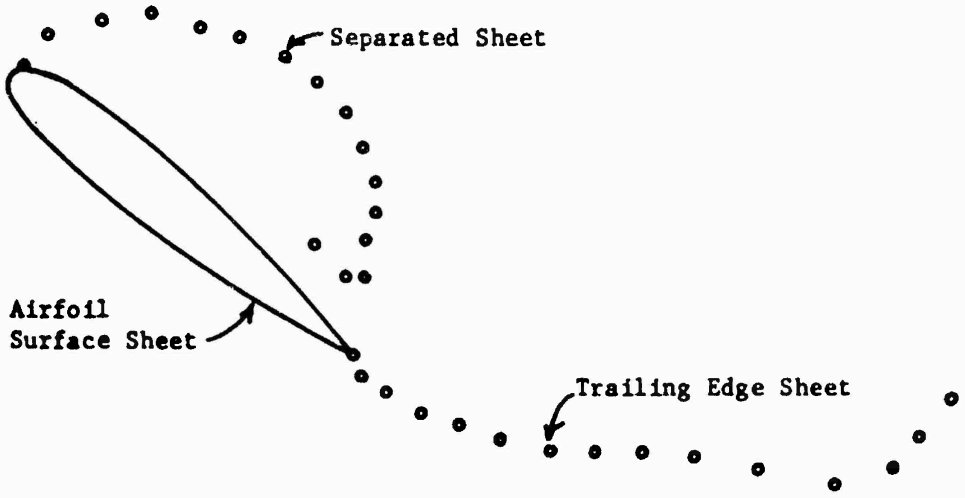


Figure 1. Schematic of Boundary (Vortex) Sheets for Unsteady Pitching Motion with Separation

potential flow model which may include sheets of vorticity shed from the boundary layer separation point. Therefore, a strong coupling between the boundary layer analysis and potential flow analysis exists for separated flow situations. The turbulent boundary layer analysis used in the present work is essentially that due to Lyrio, Ferziger, and Kline[16]. This unsteady integral technique gives excellent results for the steady flows of Tillman, Herring, and Norbury; Stratford, Samuel, and Joubert (see Coles and Hirst, reference 17); Kim[18]; Simpson and Strickland[19]; and Wieghardt (see Kim, reference 18). More importantly, this method predicts the unsteady boundary layer data of Karlson[20] and Houdeville, et. al.[21] and compares well with the finite difference methods of McCroskey and Philippe[22] and Singleton and Nash[23], while being an order of magnitude faster. The formulation

for the laminar portion of the boundary layer is based upon the author's unsteady extension of Thwaites method (see Cebecci and Bradshaw[24]). Transition is assumed to occur either due to laminar boundary layer separation or according to a natural transition criteria due to Cebecci and Smith[25].

2. Simple analysis for substall angles.

An approximate solution which will yield the lift coefficient for low substall angles of attack with the pitching rate equal to a constant can be obtained by considering Figure 2.

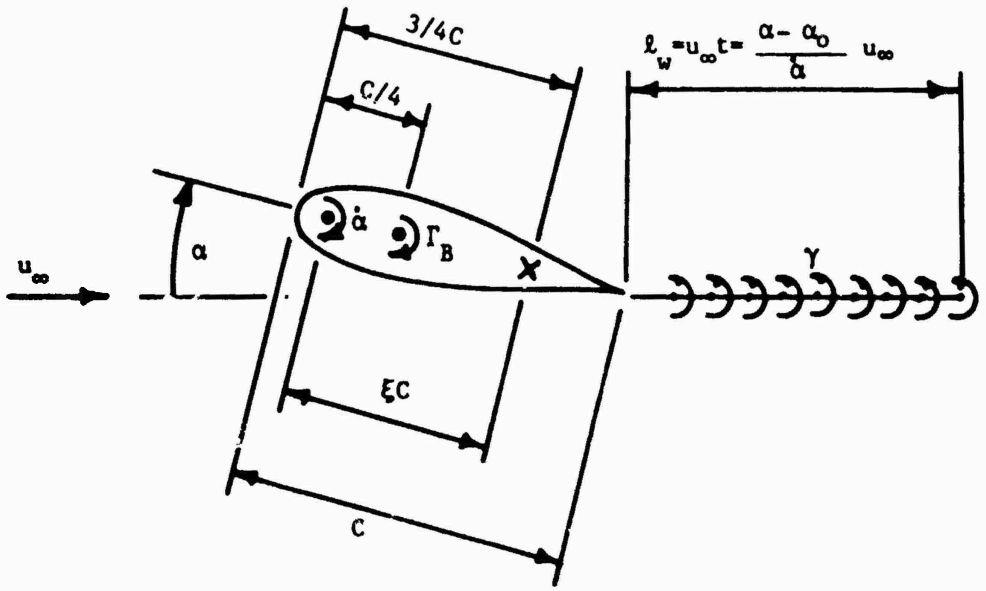


Figure 2. Airfoil and Wake Geometry for Low Substall Angles of Attack

In this simple model, the velocity at a "control point" located at the 3/4 chord point is made to be tangent to the chord line of the airfoil. In other words, the velocity normal to the airfoil at the 3/4 chord point will be made zero by adjusting the bound circulation Γ_B. The use of the 3/4 chord point as a control point location can be justified based on potential flow analysis of a pitching flat plate[26]. The "surface tangency condition" can therefore be written as:

$$\frac{\Gamma_B}{\pi C} + \frac{\gamma}{2\pi} \ln\left(\frac{4\ell_w}{C} + 1\right) = u_\infty \alpha + \xi C\dot{\alpha} \qquad (4)$$

The first term is the downwash velocity due to the bound vorticity, the second term is the downwash due to the wake, the third term is the upwash due to the freestream, and the fourth term is the velocity of the airfoil normal to the chord line due to the pitching motion. The wake circulation per unit length γ is given by:

$$\gamma = \frac{1}{u_\infty} \frac{d\Gamma_B}{dt} = \frac{\dot{\alpha}}{u_\infty} \frac{d\Gamma_B}{d\alpha} \tag{5}$$

Equations 4 and 5 form the basis for determining the bound vorticity Γ_B. The bound vorticity is related to the lift coefficient by

$$C_L = \frac{2\Gamma_B}{u_\infty C} \tag{6}$$

An approximate solution was obtained which can be given in terms of the following equations:

$$C_L = 2\pi \left(KC_L^* + \alpha_o + 2K\xi\right)$$

$$K = \frac{\dot{\alpha}C}{2u_\infty}$$

$$C_L^* = \frac{\alpha^{*2}}{\alpha^* + \ell n\ (2\alpha^* + 1)} \tag{7}$$

$$\alpha^* = \frac{\alpha - \alpha_o}{K}$$

where α_o is the angle of attack before the start of the pitching motion. It should be stressed that the approximate solution given by equation 7 is valid only for low substall angles of attack and only for constant $\dot{\alpha}$ motion whereas the USTAR2 analysis is potentially valid for all angles of attack and all airfoil motions.

B. EXPERIMENTAL

Experimental data which are used in this report consist of airfoil surface pressure data, flow visualization data, and surface hot-wire data. All data were obtained in the USAF Academy low speed 2 ft x 3 ft wind tunnel. The experimental arrangements will be briefly described.

1. Pressure data.

Airfoil surface pressure data were obtained by Francis, Keesee, and Retelle[15]. Details of the test setup can be found in that reference. A computer controlled pitch oscillator was used to impart constant $\dot{\alpha}$ pitching to a 6-inch chord NACA 0012 airfoil. Pressure taps were located at 19 positions along the surface of the airfoil. Approximately 25 repetitions of each case were run so as to obtain ensemble averages of the surface pressure coefficients at the pressure ports. These data were then used to obtain lift and drag coefficients for the airfoil as a function of time.

2. Flow visualization.

The basic flow visualization scheme was originally set up by Walker and Helin and was used in the present work to examine flow around a 6-inch chord NACA 0015 airfoil. This scheme makes use of a smoke wire placed across the tunnel upstream of the pitching airfoil. The wire was placed in a plane normal to the axis of rotation of the airfoil. A smoke producing oil (Roscoline) was coated on a 0.005 inch diameter tungsten wire which was in turn heated electrically to produce a large number of smoke streaks in the flow. These streaks tend to have a rather uniform spacing due to the regular spacing of smoke material droplets which are formed when the wire is coated. The smoke was illuminated by a high intensity strobe light placed downstream of the airfoil. The proper sequencing of airfoil pitch commands with strobe light and smoke wire triggering was accomplished by computer control. A 35 mm camera looking along the pitch axis was used to record the visual data.

3. Surface mounted hot-wire.

A NACA 0015 airfoil with a 6-inch chord was instrumented in the present work with an array of hot-wires. As indicated in Figure 3, seven hot-wires were mounted on the upper surface of the airfoil (suction side). The hot-wire sensing elements (TSI-10 hot films) were soldered to pairs of number 9 sewing needles which protruded above the airfoil surface approximately 0.20 inches. The needle supports in turn were mounted in electrically insulated plugs which were machined flush

with the airfoil surface. A TSI model 1050 hot-wire anemometer system, along with an in-house linearizer, were used to obtain velocity signals. Approximately 25 repetitions of each case were run to obtain ensemble averages of the velocity signal from each probe.

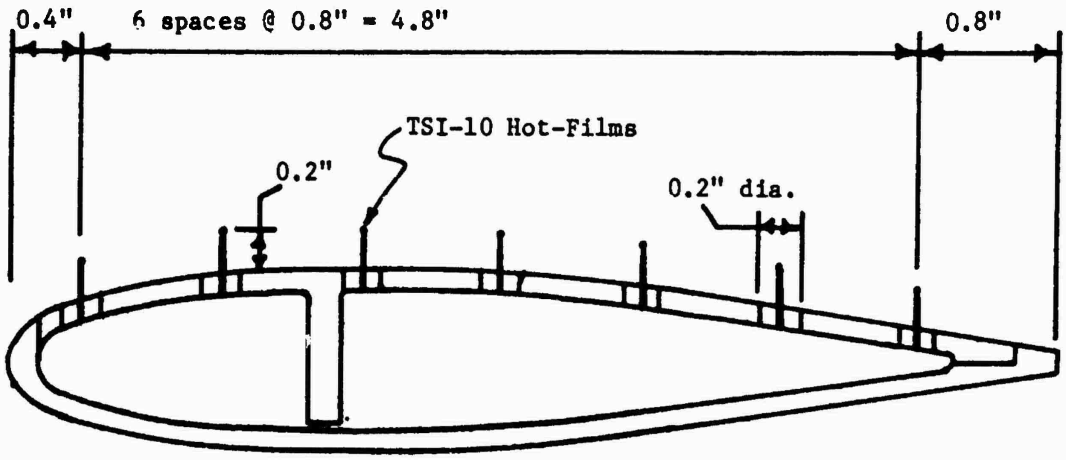


Figure 3. Surface Hot-Wire Configuration

IV. RESULTS:

In this section typical experimental data are presented and compared with analytical results. These data include lift and drag data, wake geometry, and airfoil surface velocities. It should be noted that a large body of data was obtained as part of this work and that the data presented herein are only representative samples.

A. Lift and Drag.

Lift and drag data are shown in Figure 4 for a NACA 0012 airfoil pitching up from a zero angle of attack at a non-dimensional pitching rate K of 0.089. The results from the simple analysis given by Equation 7 are seen to agree with the experimental lift data very well up to and a little beyond the stall angle which is approximately 10 degrees for the present case. The non-zero lift coefficient at zero angle of attack is due to the so-called "pitch circulation." The slope of the lift curve in this substall region can be seen to be considerably

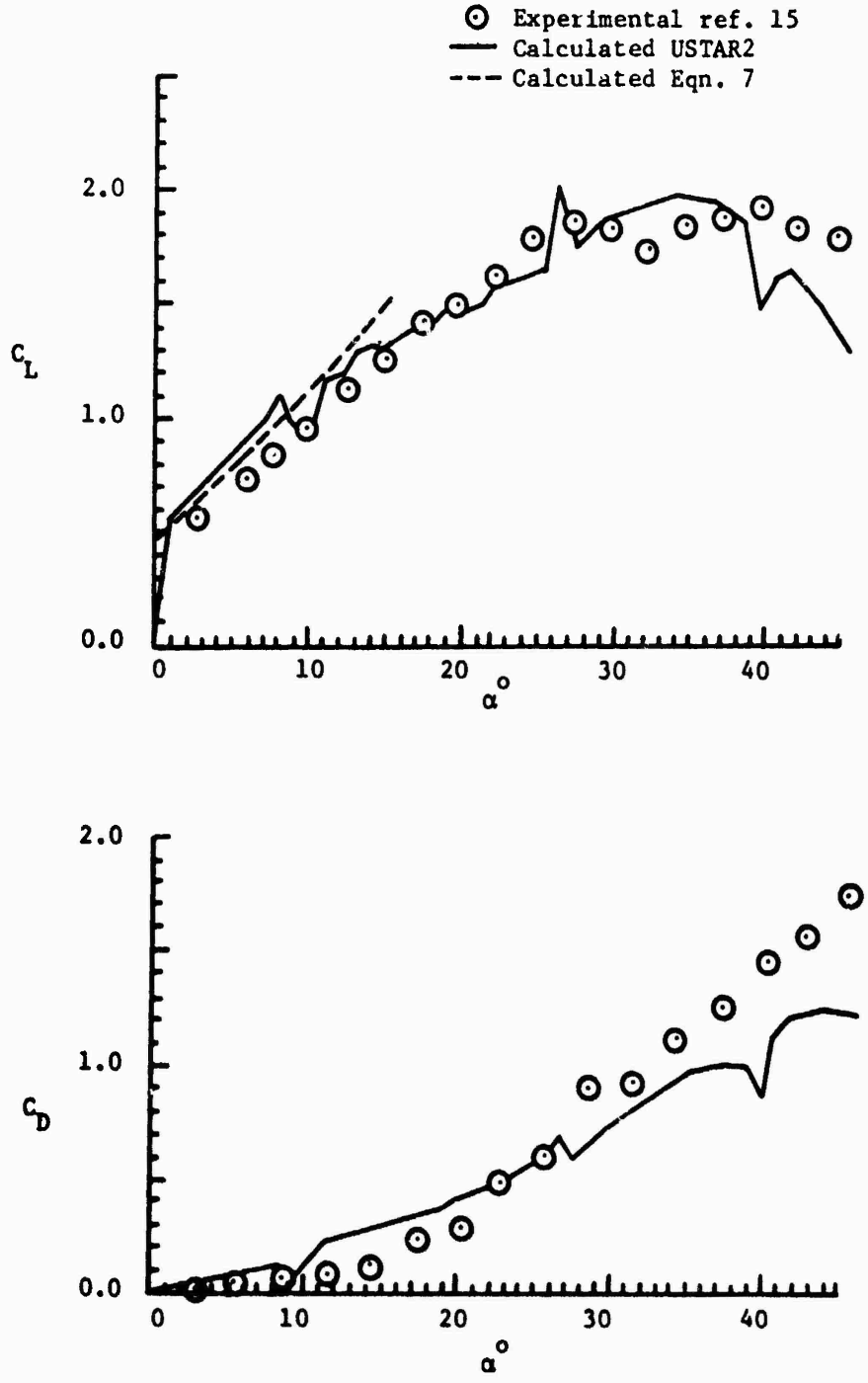


Figure 4. Lift and Drag Coefficient for a NACA 0012 Airfoil Pitching up at Constant $\dot{\alpha}$ About the 31.7 Percent Chord for Re=77,700

less than for the steady case due to the downwash on the airfoil produced by the vortex sheet springing from the trailing edge. The USTAR2 analysis is seen to predict both lift and drag coefficients reasonably well below an angle of attack equal to about 30 degrees. The small, abrupt jumps observed in the USTAR2 results can be attributed to the tendency for the boundary layer separation point to "lock in" on surface panel edges. This problem can be corrected in subsequent versions of USTAR2 by making minor changes in the procedures for introducing "edge velocities" into the boundary layer subroutine. The lack of agreement at high angles of attack is thought to be due to discrete wake vortex core growths which are excessively large in the USTAR2 analysis. This reduces the effect of the large scale vortex which grows on the suction side of the airfoil. The core growth parameter was arbitrarily selected in the past and should be reduced to represent more realistic growth rates. Examination of pressure coefficient data bears out the fact that the influence of the vortex moving over the airfoil is too weak in the USTAR2 analysis.

B. Wake Geometry.

The wake geometry obtained from the USTAR2 analysis is compared in Figure 5 to flow visualization data. The vortex sheets obtained in the USTAR2 analysis basically represent "streaklines" in that most of the fluid particles which make up these sheets were either injected into the flow at the leading or trailing edge. While exact comparisons between visual and calculated results are difficult to make, it does appear that the predicted large scale vortex is not as tightly rolled up as that indicated by flow visualization. This again indicates that the discrete wake vortex cores are growing at excessive rates in the USTAR2 analysis.

C. Airfoil Surface Velocities.

Velocities obtained from three of the seven surface mounted hot-wire probes are shown for a particular case in Figure 6. Near the nose of the airfoil (7% chord) the velocity measured by the probe rises until the probe becomes immersed in the separated boundary layer. A significant region of reversed flow first appears at an angle of attack of about 25 to 30 degrees somewhat downstream of the nose (20% chord).

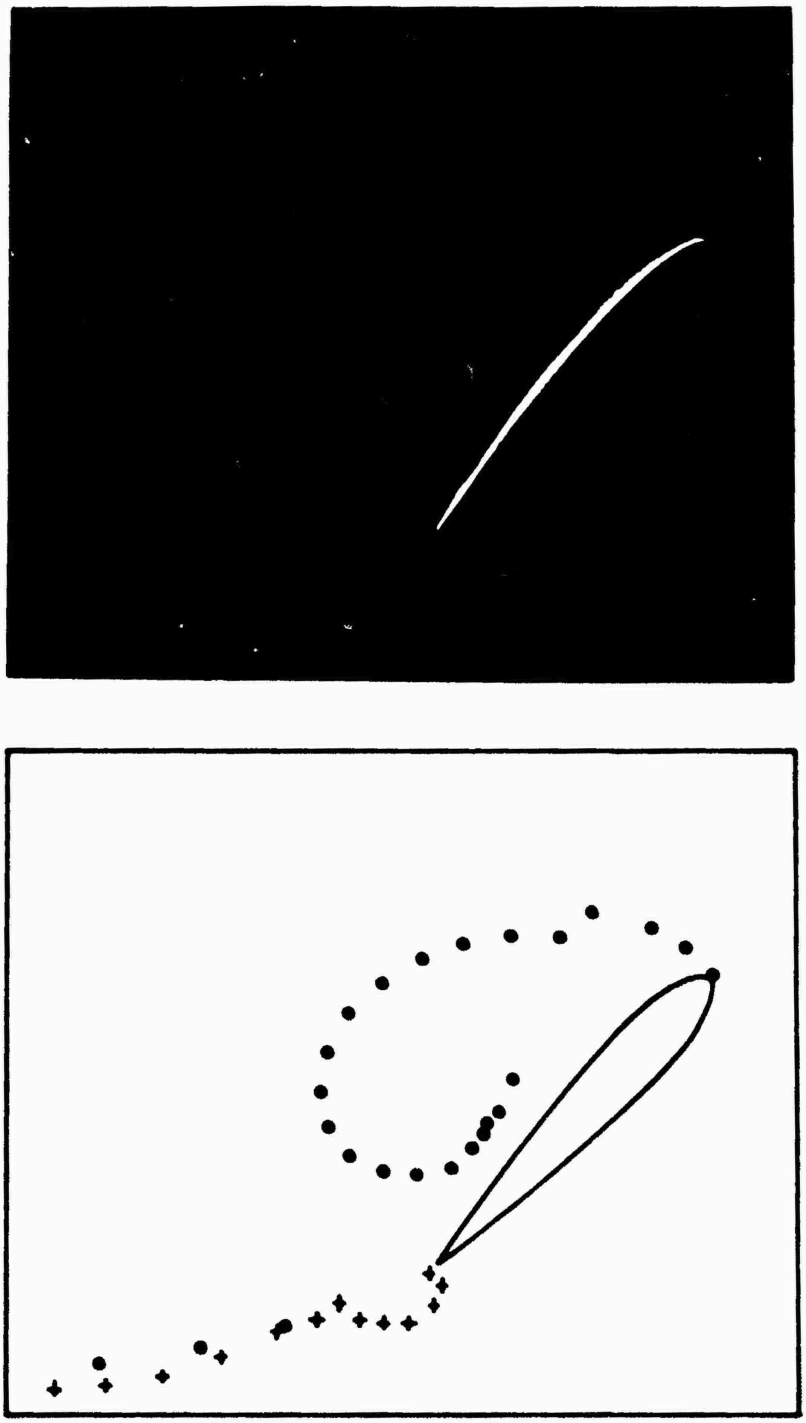


Figure 5. Wake Geometries Obtained From Flow Visualization and USTAR2 (α = 45°, K = 0.109, Re = 62,500, Pivot Point = 0.25 Chord, NACA 0015)

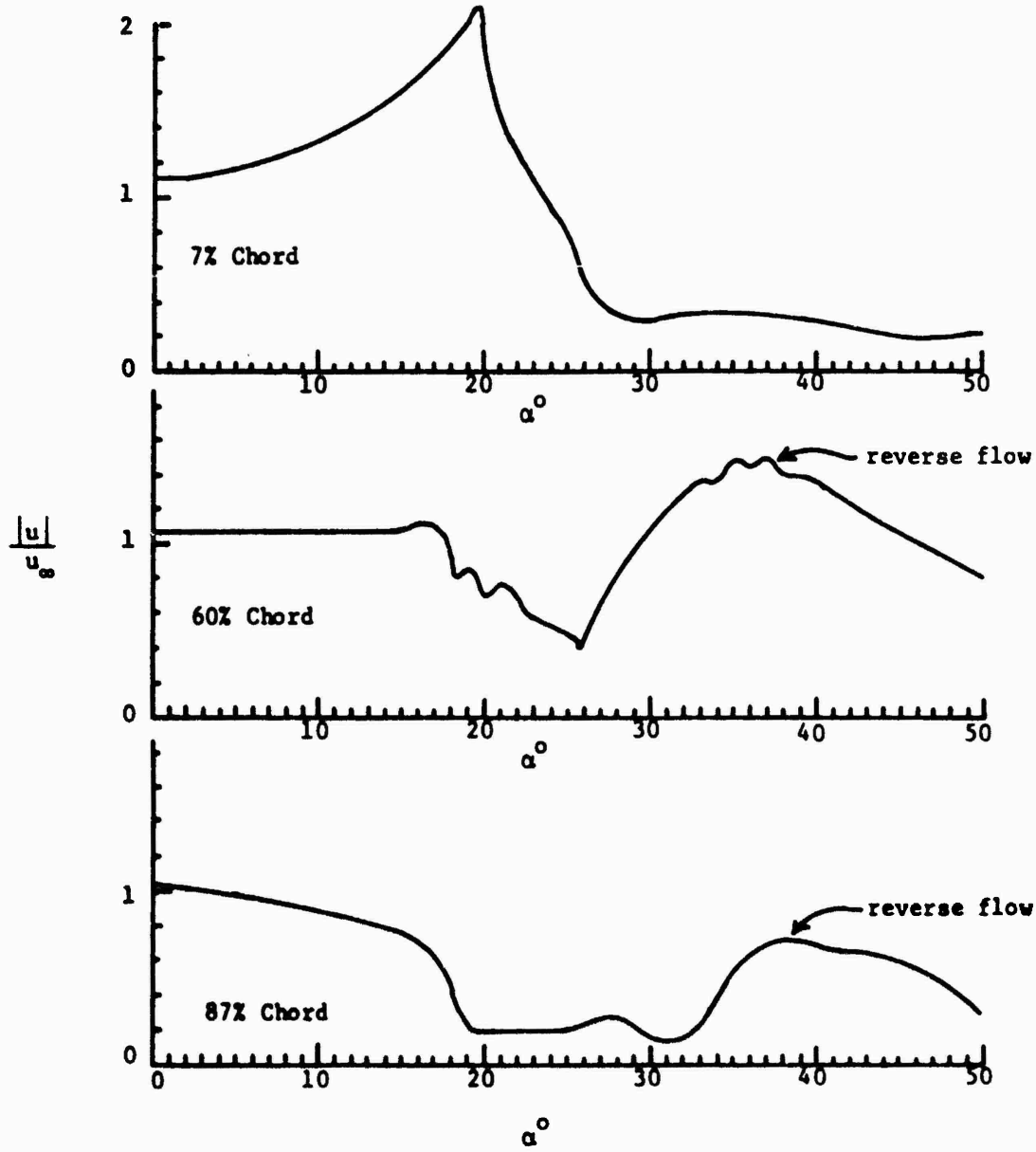


Figure 6. Velocities Obtained from Surface Mounted Hot-Wire Anemometer (K = 0.089, Re = 62,500, Pivot Point = 0.317 Chord, NACA 0015)

For this particular case the magnitude of the reversed flow reaches a maximum at about the 60% chord and has a value of 1.5 times the freestream velocity. Further downstream (87% chord), the reverse flow due to the initial large scale vortex passage is reduced and occurs at a higher angle of attack.

While no direct comparisons have been made between the hot-wire results and USTAR2, it is apparent from examination of pressure coefficient data obtained from USTAR2 that the relatively large magnitude of the reverse flow are not being predicted. This again emphasizes the need to reduce the discrete vortex core growth rate in the USTAR2 model to produce a more energetic large scale vortical structure.

V. RECOMMENDATIONS:

In general, the greatest need is to continue to expand the experimental data base. Some minor changes in experimental methodology may also be appropriate in order to improve the quality and ease of taking data. It is also apparant that the analytical model USTAR2 may be improved on the basis of the experimental data already obtained.

A. Experimental.

In order to complete the data base for the current set of runs (constant $\dot{\alpha}$ starting from zero angle of attack) surface pressure data should be obtained in addition to those obtained by Francis, et. al.[15]. After those data are taken the effect of a non-zero initial angle of attack should be studied. Other pitching schedules such as constant $\ddot{\alpha}$ should also be run. In order to improve the flow visualization methodology, a high speed movie camera should be used in place of the present 35 mm camera. In addition, a more powerful pitching mechanism may be required to run some of the desired cases. Finally, it is recommended that a series of tests be run with much larger values of the non-dimensional pitching rate K. These high values of K can be obtained in the tow tank facility at Texas Tech University. This recommended work will be the subject of a mini-grant proposal by the author.

B. Analytical.

The USTAR2 analysis which is presently being developed at Texas Tech University under a contract from Sandia Laboratories, Albuquerque, New Mexico, should be improved on the basis of the experimental data already obtained. In particular, a systematic study of the discrete vortex core growth parameter should be undertaken to yield better agreement with data. The boundary layer prediction scheme should also be improved with regard to improved numerical procedures.

REFERENCES

1. McCroskey, W.J., "Recent Developments in Dynamic Stall," Proceedings of a Symposium on Unsteady Aerodynamics, University of Arizona, Vol. 1, pp 1-33, 1975.

2. McCroskey, W.J., "Some Current Research in Unsteady Fluid Dynamics - The 1976 Freeman Scholar Lecture," J. Fluids Engineering, Vol. , pp 8-38, 1977.

3. Ericsson, L.E. and J.P. Reding, "Dynamic Stall Analysis in Light of Recent Numerical and Experimental Results," J. Aircraft, Vol. 1, pp 248-255, 1976.

4. Gormont, R.E., "A Mathematical Model of Unsteady Aerodynamics and Radial Flow for Application to Helicopter Rotors," U.S. Army AMRDL Technical Report 72-67, 1973.

5. Carta, F.O., "Unsteady Normal Force on an Airfoil in a Periodically Stalled Inlet Flow," J. Aircraft, Vol. 4, pp 416-421, 1967.

6. Johnson, W., "The Effect of Dynamic Stall on the Response and Airloading of Helicopter Rotor Blades," J. Am. Hel. Soc., Vol. 14, pp 68-79, 1969.

7. Mehta, U., "Dynamic Stall of an Oscillating Airfoil," AGARD Fluid Dynamics Panel Symposium, Ottawa, Paper No. 23, September 1977.

8. Hodge, J.K., A.L. Stone, and T.E. Miller, "Numerical Solution for Airfoils Near Stall in Optimized Boundary-Fitted Curvilinear Coordinates," AIAA Journal, Vol. 17, pp 458-464, 1979.

9. Nash, J.F., L.W. Carr, and R.E. Singleton, "Unsteady Turbulent Boundary Layers in Two-Dimensional Incompressible Flow, AIAA Journal, Vol. 13, pp 167-173, 1975.

10. Dwyer, H.A. and W.J. McCroskey, "Crossflow and Unsteady Boundary Layer Effects on Rotating Blades," AIAA Journal, Vol. 9, pp 1498-1505, 1971.

11. Telionis, D.P., "Calculations of Time-Dependent Boundary Layers," Unsteady Aerodynamics, R.B. Kinney (Ed.), Vol. 1, p 155, 1975.

12. Ham, N.D., "Aerodynamic Loading on a Two-Dimensional Airfoil During Dynamic Stall, AIAA Journal, Vol. 6, pp 1927-1934, 1968.

13. Katz, J., "A Discrete Vortex Method for the Non-Steady Separated Flow Over an Airfoil," J. Fluid Mech., Vol. 102 pp 315-328, 1981.

14. Oler, J.W., J.H. Strickland, B.J. Im, and G.H. Graham, "Dynamic Stall Regulation of the Darrieus Turbine," Sandia National Laboratory Report SAND83-7209, June 1983.

15. Francis, M.S., J.E. Keesee, and J.P. Retelle, "An Investigation of Airfoil Dynamic Stall With Large Amplitude Motions," USAF/FJSRL Report (in press).

16. Lyrio, A.A. and J. Ferziger, "A Method of Predicting Unsteady Turbulent Flows and Its Application to Diffusers with Unsteady Inlet Conditions," AIAA Journal, Vol. 21, No. 4, pp 534-540, 1983. See also:
Lyrio, A.A., J.H. Ferziger, and S.J. Kline, "An Integral Method for the Computation of Steady and Unsteady Turbulent Boundary Layer Flows, Including the Transitory Stall Regime in Diffusers," Stanford University Report PD-23, 1981.

17. Coles, D. and E.A. Hirst, Computation of Turbulent Boundary Layers, 1968 AFOSR-RFP, Vol. 2, Stanford University, 1969.

18. Kim, J., S.J. Kline, and J.P. Johnston, "Investigation of Separation and Reattachment of a Turbulent Shear Layer: Flow Over a Backward-Facing Step," Rpt. MD-37, Thermosciences Div., Mech. Engrg. Dept., Stanford University, 1978. See also TASME, J. Fluids Eng., 102, pp 302-308, 1980.

19. Simpson, R.L., J.H. Strickland, and P.W. Barr, "Features of a Separating Turbulent Boundary Layer in the Vicinity of Separation," J. Fluid Mech., Vol. 79, pp 553-594, 1977.

20. Karlson, S.K.F., "An Unsteady Turbulent Boundary Layer," J. Fluid Mech., Vol. 5, 1959.

21. Houdeville, R. and J. Cousteix, "Premiers Resultats D'une Etude sur les Couches Limites Turbulentes en Encoulement Pulse Avec Gradient de Pression Moyen Defavorable," 15E Collogue D;Aerodynamique, AAAF NT-79-05, 1978. Translated into English in NASA TN-75799, 1979.

22. McCroskey, W.J. and J.J. Phillippe, "Unsteady Viscous Flow on Oscillating Airfoils," AIAA Journal, Vol. 13, pp 71-79, 1975.

23. Singleton, R.E. and J.F. Nash, "Method for Calculating Unsteady Turbulent Boundary Layers in Two- and Three-Dimensional Flows," AIAA Journal, Vol. 12, pp 590-595, 1974.

24. Cebecci, T. and P. Bradshaw, Momentum Transfer in Boundary Layers, McGraw-Hill, 1977.

25. Cebecci, T. and A.M.O. Smith, Analysis of Turbulent Boundary Layers, Academic Press, 1974.

26. Strickland, J.H., T. Smith, and K. Sun, "A Vortex Model of the Darrieus Turbine: An Analytical and Experimental Study," Sandia Laboratory Report #SAND81-7017, 1981.

27. Sheldahl, R.E. and P.C. Klimas, "Aerodynamic Characteristics of Seven Symmetrical Airfoil Sections Through 180 Degree Angle of Attack for Use in Aerodynamic Analysis of Vertical Axis Wind Turbines," Sandia Laboratory Report, SAND80-2114, 1981.

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

EXPERIMENTAL AND THEORETICAL INVESTIGATIONS OF NEGATIVE ION-POLAR MOLECULE REACTIONS

Prepared by:	Dr. Timothy C. K. Su
Academic Rank:	Associate Professor
Department and University:	Department of Chemistry Southeastern Massachusetts University
Research Location:	Air Force Geophysics Laboratory (LKB) Hanscom AFB, Massachusetts
USAF Research:	Dr. John F. Paulson

Date:	September 6, 1983
Contract No.:	F49620-82-C-0035

EXPERIMENTAL AND THEORETICAL INVESTIGATIONS OF NEGATIVE ION-POLAR MOLECULE REACTIONS

by

Timothy C. K. Su

ABSTRACT

The rate constants for proton-transfer reactions of OH^-, NH_2^- and O^- with HCN and $CH_3CN$and for the displacement reactions of OH^- and O^- with methyl halides have been measured in the gas phase in the temperature range 200K-500K using a selected ion flow tube. Rate constants of the fastest reactions are very close to the collision rate constants predicted by the trajectory method. It was observed that the reaction efficiency decreases as exothermicity decreases and the reaction probability of S_N2 reactions decreases as temperature increases. A trajectory calculation is also developed for the kinetic energy dependence of ion-polar molecule collision rate constants.

ACKNOWLEDGEMENT

The author would like to thank the Air Force System Command, the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for providing him with the opportunity to spend a very worthwhile and interesting summer at the Air Force Geophysics Laboratory (LKB), Hanscom AFB, MA. He would like to acknowledge the laboratory for its hospitality and excellent working conditions.

Finally, he would like to thank Dr. John F. Paulson for his valuable suggestions and his guidance, and he would like to thank Dr. Albert A. Viggiano for his help in many ways.

I. INTRODUCTION:

Ion-molecule reactions are important in atmosphere chemistry, combustion systems and gas discharges. Ion-chemistry has been a fast growing field in the past couple of decades.[1] Numerous reactions have been studied involving both positive and negative ions. Much theoretical work has also been done.[2] Although many negative ion molecule reactions have been studied, temperature dependence experiments have not been performed for most of the reactions. Temperature dependence investigations are of importance in studying reaction mechanisms and the properties of reaction complexes.[3]

The present project is concerned with the temperature dependence of the reactions of the following ion-polar molecule systems: the reactions of OH^-, O^- and NH_2^- with HCN and CH_3CN, and the reactions of OH^- and O^- with methyl halides. These reactions, except that of O^- with CH_3CN, have been studied by Bohme and coworkers[4,5,6] using the flowing afterglow technique at room temperature. Here we report the reaction rate constants and branching ratios in the temperature range 200K-500K. Reactions of OH^- and O^- with HCN and CH_3CN are chosen because these ions and neutrals were found in the earth's atmosphere[7,8] and because the neutrals have very high dipole moments. These reactions are suitable for studying the dipole effects. Reactions of NH_2^- with these neutrals were also studied because their large rate constants can add further test on the temperature dependence trajectory treatment of ion-polar molecule interactions[9,10] which is used in this project as a diagnostic tool to study reaction mechanisms. Reactions of O^- and OH^- with methyl halides were chosen because they involve S_N2 reactions which require a specific site of attack for the reaction to occur. We are interested in studying the effect of restricted reaction site on the reaction efficiency.

Another aspect of this project is to develop trajectory calculations of kinetic energy dependence of ion-polar molecule collision rate constants. Temperature dependence trajectory calculations mentioned above have been developed[9] and parameterized.[10] In these

calculations, the ion is treated as a point charge and the polar molecule as a two-dimensional rigid rotor. The system Hamiltonian is given by

$$\mathcal{H} = \frac{P_r^2}{2\mu} + \frac{L^2}{2\mu r^2} + \frac{J^2}{2I} + V(r, \theta), \qquad (1)$$

where P_r is the radial momentum of the collision partners, μ is the reduced mass and L is their orbital angular momentum. The angular momentum and moment of inertia of the rotor are given by J and I, respectively. $V(r, \theta)$ is the ion-dipole potential which is a function of the ion-molecule separation, r, and the angle formed by r and the dipole, θ. The capture rate constant, $k_{cap}(E)$, at a total energy, E, is given by

$$k_{cap}(E) = F_{cap}(E)/\rho(E) \qquad (2)$$

where $F_{cap}(E)$ is the microcanonical capture flux and $\rho(E)$ is the total density of states per unit volume of the collision partners at infinite separation. The trajectories were begun at the capture radius, $r_o = (\alpha q^2/2E)^{1/4}$, and were integrated to a large ion-molecule separation. The thermal capture rate constant was obtained by averaging $k_{cap}(E)$ over a five dimensional Boltzmann energy distribution. During recent years, due to the advance of experimental techniques, numerous experimental works were performed to examine the kinetic energy dependence of ion-molecule reactions.[11] In these experiments, while the ion-reactant kinetic energy may be varied, the rotational energy of the reactant neutral remains at room temperature or at a specific temperature. Consequently, there is a growing need for a collision theory that describes the kinetic energy dependence of rate constants when the neutral remains at thermal energy.

II. OBJECTIVES

The objectives of this project were:

(1) To determine experimentally the reaction rate constants and branching ratios (if any) of the negative ion-polar molecule reaction systems mentioned above at different temperatures.

(2) To study the kinetics of these reactions by comparing experimental

results with theoretical models at different temperatures.

(3) To develop trajectory calculations for the kinetic energy dependence of ion-polar molecule collision rate constants.

III. EXPERIMENTAL APPARATUS

The rate constant measurements were carried out using a selected ion flow tube (SIFT). The SIFT is similar to that described by Adams and Smith.[12] Briefly, the reactant ions are generated in an appropriate source, mass-selected, and injected into the flow tube through a venturi-type aspirator inlet, whence they are carried downstream past the reactant-molecule inlet and are detected by a mass spectrometer sampling system. The total rate coefficient for the selected ion-molecule reaction is determined from the decline of the ion signal as a function of neutral gas addition.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

A. Reactions of NH_2^-, OH^- and O^- with HCN and CH_3CN

Figure 1 compares experimental rate constants of reactions (a), (b) and (c) with theoretical collision rate constants calculated by the trajectory method[9,10] in the temperature range 300K-500K. Figure 2 gives the similar comparison for reactions (d), (e) and (f).

	ΔH(kcal/mole)	
$NH_2^- + HCN \rightarrow CN^- + NH_3$	-54	(a)
$OH^- + HCN \rightarrow CN^- + H_2O$	-41	(b)
$O^- + HCN \rightarrow CN^- + OH$	-32	(c)
$NH_2^- + CH_3CN \rightarrow CH_2CN^- + NH_3$	-29	(d)
$OH^- + CH_3CN \rightarrow CH_2CN^- + H_2O$	-16	(e)
$O^- + CH_3CN \rightarrow CH_2CN^- + OH$	-9	(f)

The ordinate of these figures is the ratio of rate constant to the Langevin rate constant,[13] k_L. Experimental values measured by Bohme[4,5] at room temperature are also included in the figures. The rate constants of reactions (a), (b) and (d) are very close to the collision rate constants throughout the temperature range studied. Since these reactions are some of those very fast ion-polar molecule proton transfer

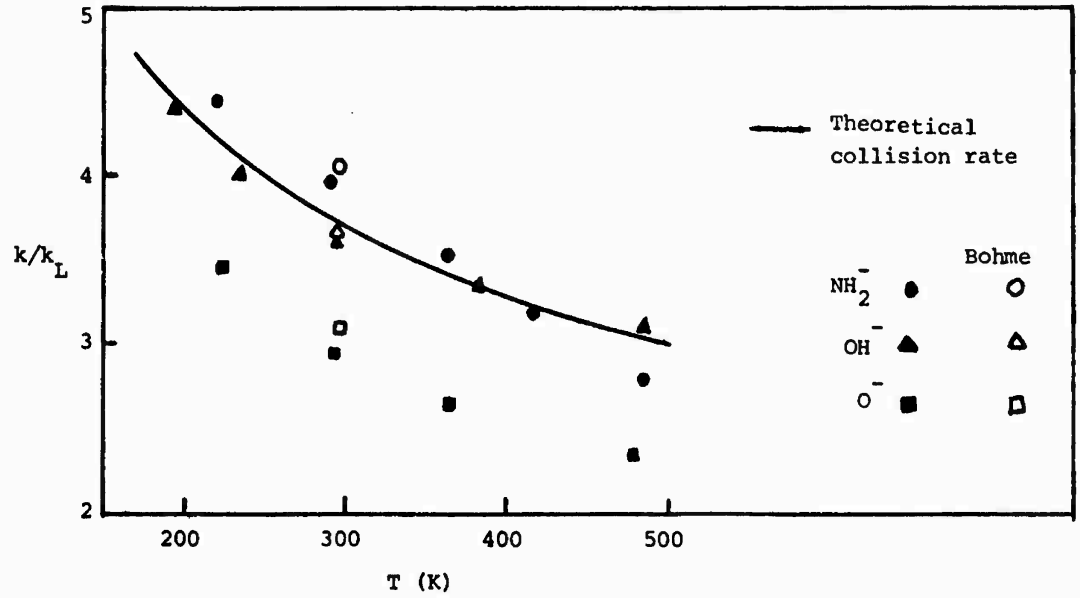


Figure 1. X^- + HCN

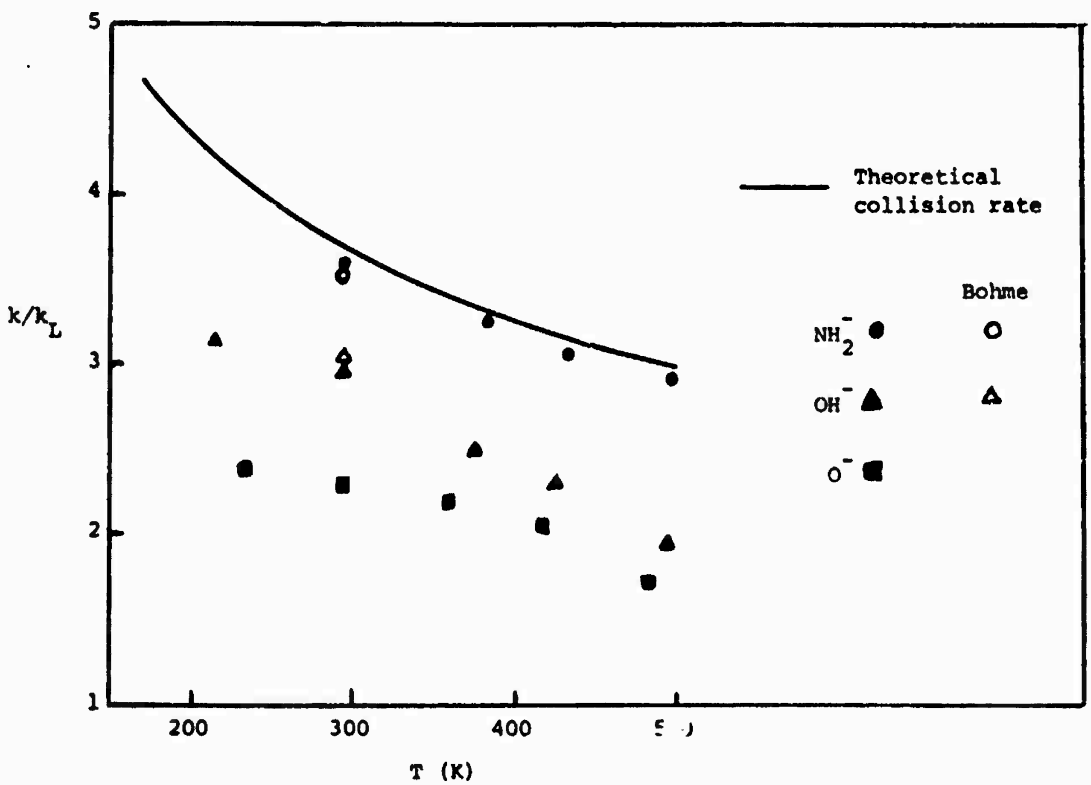


Figure 2. X^- + CH_3CN

reactions compared to the Langevin theory, it is reasonable to believe that their reaction efficiencies are very close to unity. The results here give a positive support to the theoretical model. Rate constants of reactions (c), (e) and (f) are somewhat less than the collision rate constants. It appears that in both sets of reactions, the reaction efficiency decreases as the exothermicity decreases. A second product ion, CNO^-, is also observed in reaction (c). This corresponds to the reaction $O^- + HCN \rightarrow CNO^- + H$. The amount of CNO^- formation was about 10% of the total product formation throughout the temperature range studied. For those reactions with reaction efficiency less than unity, the shapes of the reaction rate constant vs. temperature curves are not much different from those of the collision rate constants. The below unity reaction efficiency is probably a result of the backward dissociation of the ion-molecule complex to the reactants. The relative probability of this backward dissociation is expected to increase as the exothermicity decreases.

B. Reactions of OH^- and O^- with methyl halides

Figure 3 gives the comparison of the total rate constants of reactions (g), (h) and (i) with theoretical collision rate constants.

	ΔH(kcal/mole)	
$OH^- + CH_3Br \rightarrow Br^- + CH_3OH$	-57	(g)
$OH^- + CH_3Cl \rightarrow Cl^- + CH_3OH$	-50	(h)
$OH^- + CH_3F \rightarrow F^- + CH_3OH$	-18	(i)

Experimental rate constants measured by Bohme[6] at room temperature are also included in the figure. Except for reaction (g), the results of this work agree well with Bohme's results at room temperature. Note that the experimental data of reaction (i) shown in Figure 3 are the rate constants multiplied by 100. Although the major reaction channel (the S_N2 channel) of reactions (g) and (h) are quite exothermic, their reaction rate constants are significantly lower than the collision rate constants and they also appear to decrease at higher temperatures. The below unity reaction efficiencies are probably a result of the restricted reaction site in S_N2 reactions. The exothermicity of

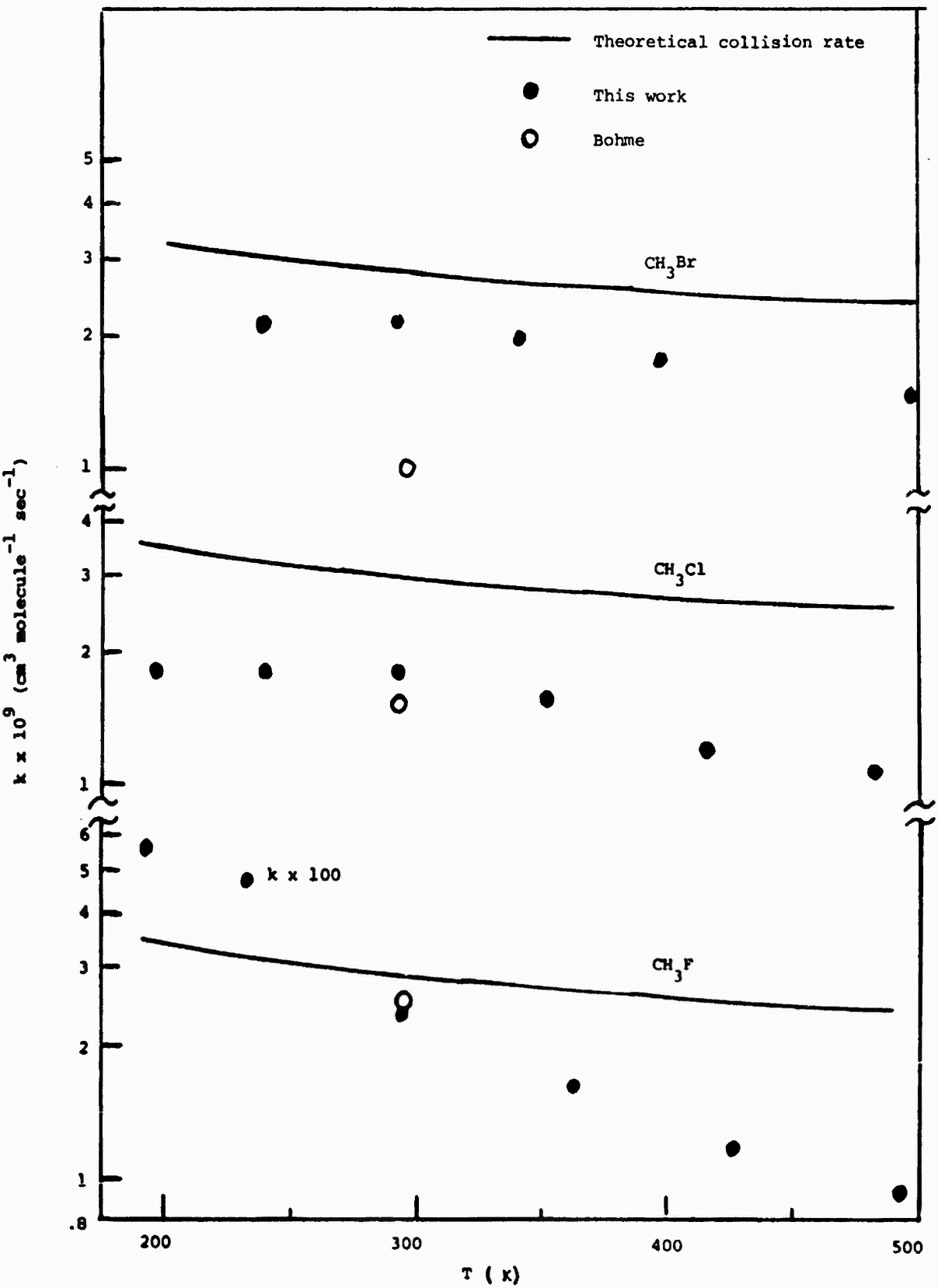


Figure 3. OH^- + CH_3X

reaction (i) is much lower. The reaction efficiency is also much lower and decreases drastically as temperature increases. It is reasonable to believe that the reaction efficiency of S_N2 reactions would increase as the lifetime of the ion-molecule complex increases. As temperature increases, both the relative kinetic energy of the ion-molecule partners and the rotational energy of the polar molecule increases. The complex lifetime would probably decrease leading to a lower S_N2 reaction efficiency.

A small amount of CH_2Br^- and CH_2Cl^- were observed as product ions in the reactions of OH^- with CH_3Br and CH_3Cl, respectively. Since the reactions $OH^- + CH_3Br \rightarrow CH_2Br^- + H_2O$ and $OH^- + CH_3Cl \rightarrow CH_2Cl^- + H_2O$ are predicted to be quite endothermic, they are not expected to be observed in the temperature range studied. A possibility of the formation of CH_2Y^- (where Y = Br or Cl) in these experiments is that some impurities, maybe CH_3CH_2Y, are present in the CH_3Y gas samples leading the reactions of the type $OH^- + CH_3CH_2Y \rightarrow CH_2Y^- + CH_3OH$ to happen.

Figure 4 gives the comparison of the total reaction rate constants of reactions (j), (k) and (ℓ) with collision rate constants.

		ΔH(kcal/mole)	
$O^- + CH_3Br$	$\rightarrow Br^- + CH_3O$	-66	(j1)
	$\rightarrow OH^- + CH_2Br$	-8	(j2)
	$\rightarrow CHBr^- + H_2O$	-20 (estimate)	(j3)
$O^- + CH_3Cl$	$\rightarrow Cl^- + CH_3O$	-59	(k1)
	$\rightarrow OH^- + CH_2Cl$	-10	(k2)
	$\rightarrow CHCl^- + CH_3O$	-20 (estimate)	(k3)
$O^- + CH_3F$	$\rightarrow F^- + CH_3O$	-27	(ℓ1)
	$\rightarrow OH^- + CH_2F$	-10	(ℓ2)
	$\rightarrow CHF^- + H_2O$	-10 (estimate)	(ℓ3)

Results from Bohme[6] are also included in the figure. Rate constants measured in this work at room temperature agree well with those obtained by Bohme for reactions (k) and (ℓ). However, similar to reaction (g), the

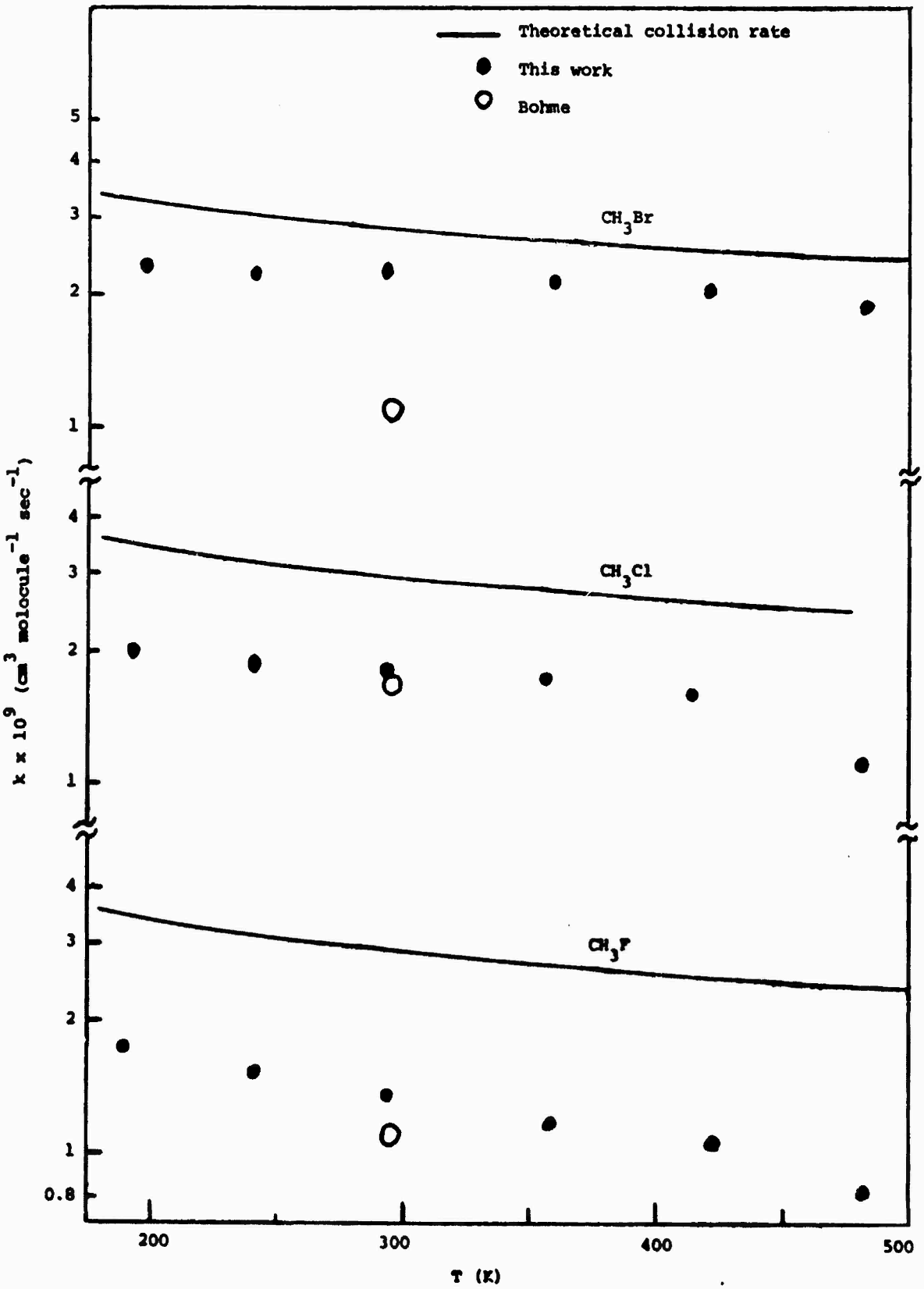


Figure 4. $O^- + CH_3X$

rate constant of reaction (j) measured in this work at room temperature is substantially higher than that measured by Bohme. This discrepancy is not understood.

Figures 5, 6 and 7 show the branching ratios (in product percentages) of reactions (j), (k) and (ℓ) respectively. Similar to the reactions of OH^-, the total reaction efficiencies of reactions of O^- with methyl halides are: $CH_3Br > CH_3Cl > CH_3F$. This has the same trend as that of the exothermicity of the S_N2 channel, the most exothermic channel of each system. The shapes of the temperature dependence curves of the total reaction rate constants are similar to those of the collision rate constants. However, the branching ratios shown in Figures 5 to 7 deserve more discussion. Figures 5 to 7 are obtained by measuring the relative ion signals at a constant neutral gas addition as a function of temperature. As pointed out by Bohme,[6] there are uncertainties in determining the branching ratios of these reactions because of the uncertainties associated with sampling and detection sensitivities and the further reactions of CHX^- (where X = Br, Cl or F) and OH^- with CH_3X. However, these results are at least semiquantitative especially the change of the branching ratios with temperature. Let's first take a look at Figure 7. Although the most exothermic channel is the formation of F^- (S_N2 reaction), the formation of OH^- is clearly the dominant channel. This is probably why the total reaction efficiency of reaction (ℓ) is much higher than that of reaction (i) where F^- is the only observable product. This also implies that the S_N2 reaction is relatively less likely to occur despite the higher exothermicity. Reaction (ℓ3) also has low reaction probability because of the low exothermicity and the relatively more complex process of transferring two H atoms.

Figures 5 and 6 show similar branching ratios as a function of temperature for reactions (j) and (k). In each case the S_N2 reaction channel is the dominant one at low temperatures but its production probability decreases as temperature increases, indicating that the S_N2 process becomes less likely to occur at higher temperatures as

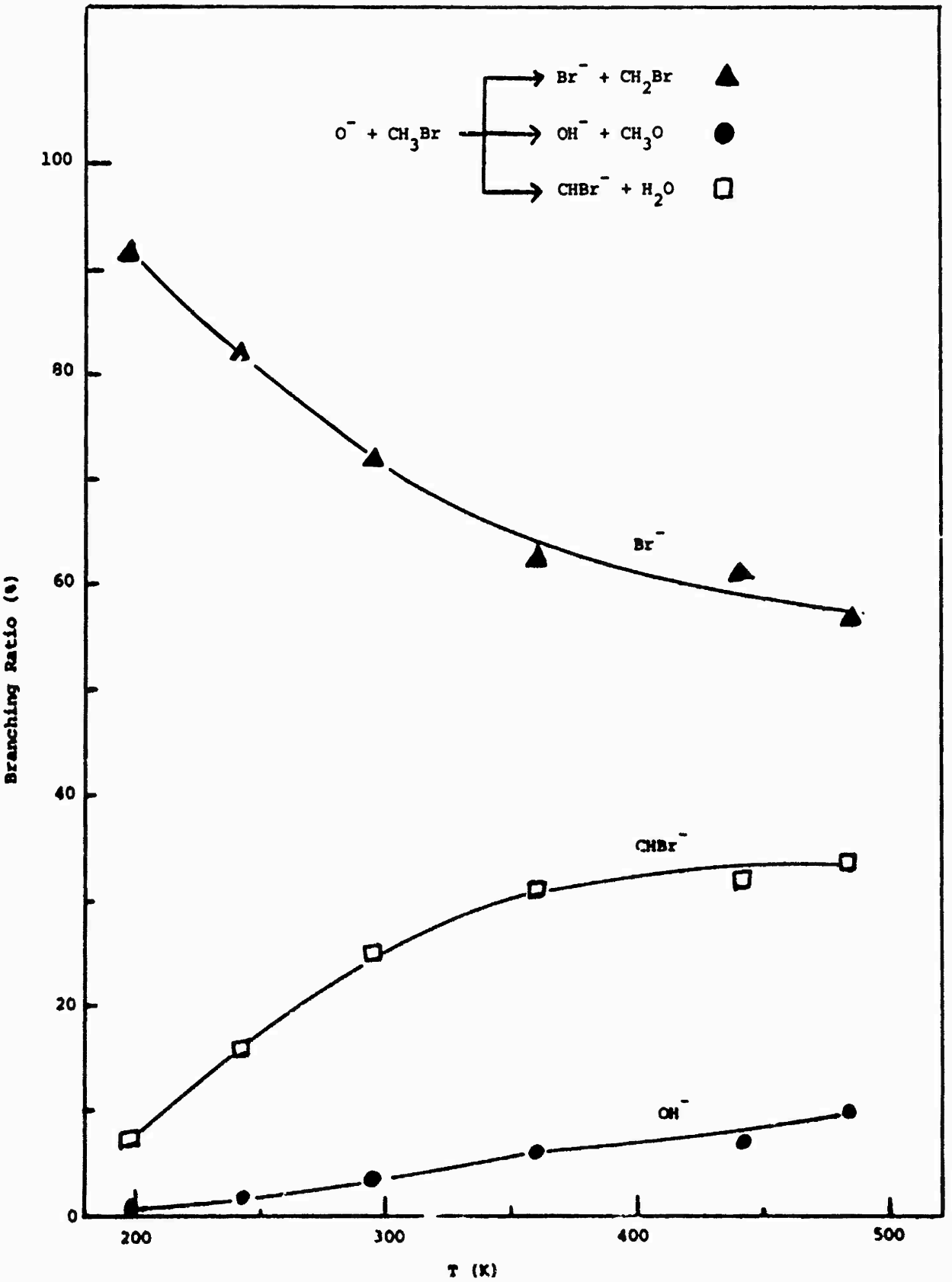


Figure 5. Branching Ratio of Reaction (1)

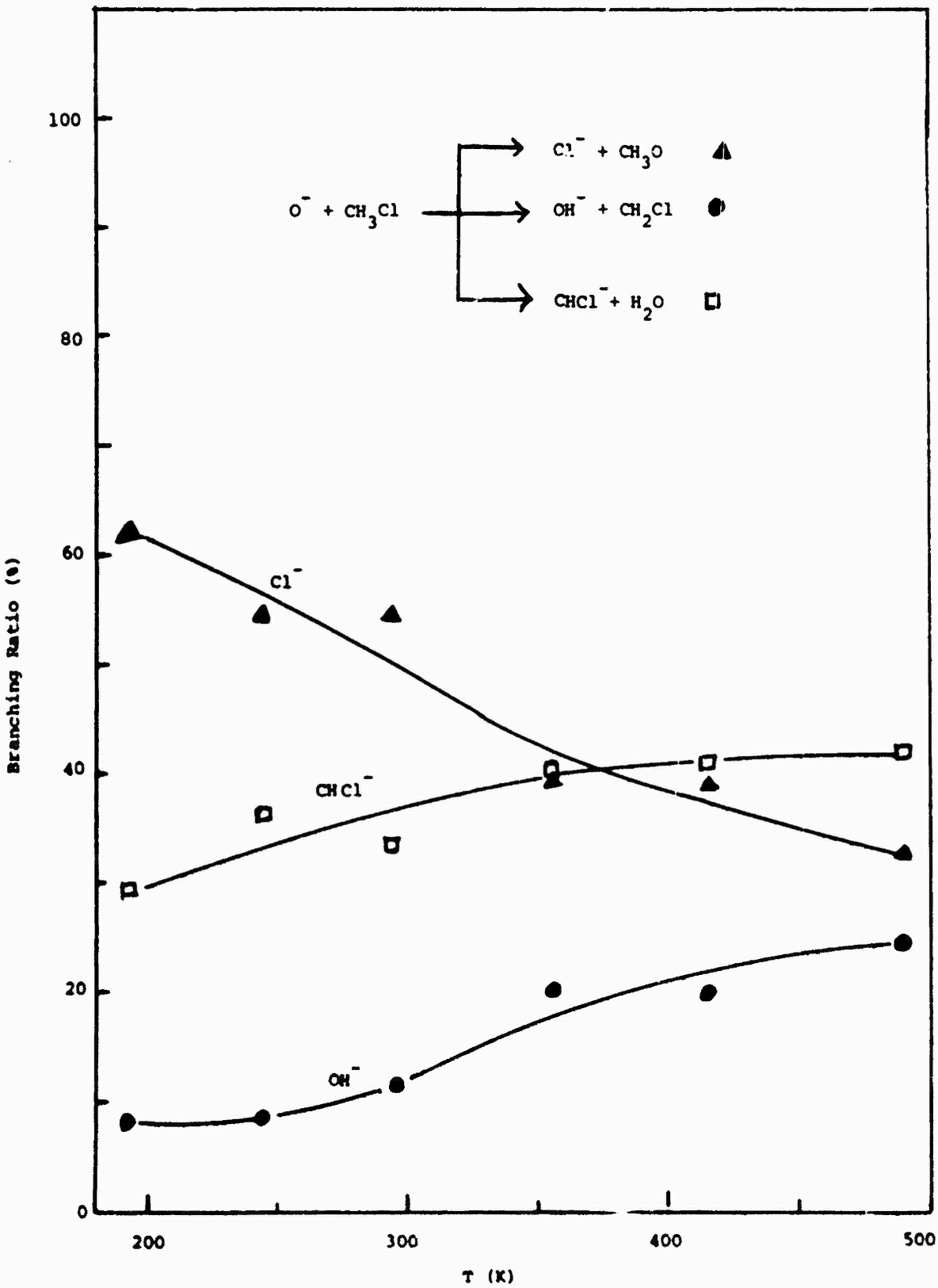


Figure 6. Branching Ratio of Reaction (k).

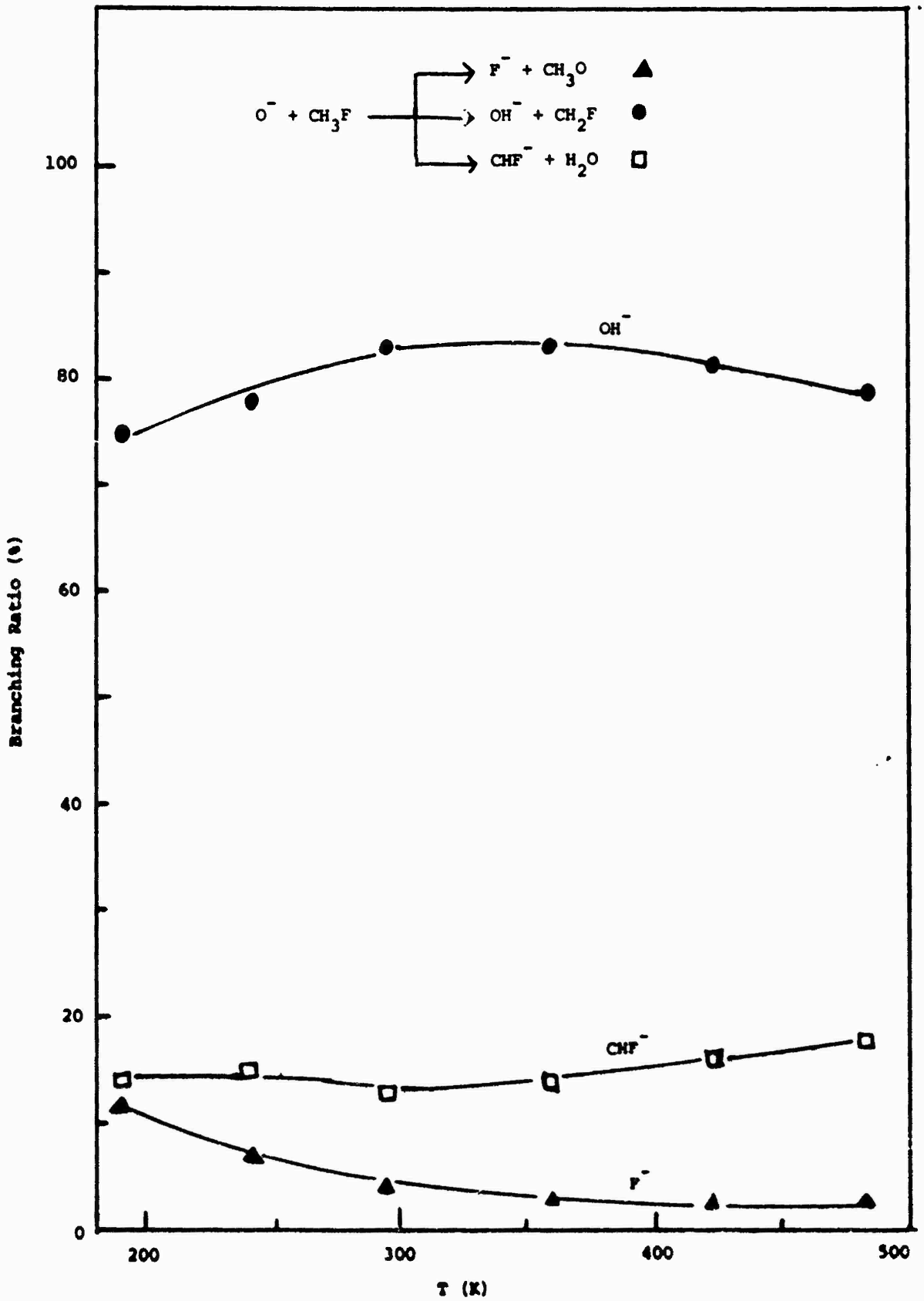


Figure 7. Branching Ratio of Reaction (1)

suggested previously. The increase in reaction probability of reactions (j2), (j3), (k2) and (k3) suggests that there may be an energy barrier in the exit channel of each of these reactions.

V. TRAJECTORY TREATMENT OF KINETIC ENERGY DEPENDENCE OF ION-POLAR MOLECULE COLLISION RATE CONSTANTS

The approach in this calculation is similar to that for thermal energy ion-polar molecule collisions.[9,10] The trajectories, however, are integrated inwardly from ca. 65 Å ion-molecule separation toward the capture separation, r_o. The capture rate constant, $k_{cap}(E_t,T)$ at a given relative kinetic energy, E_t, and temperature, T, is given by

$$k_{cap}(E_t,T) = F_{cap}(E_t,T)/\rho(E_t), \tag{3}$$

where $F_{cap}(E_t)$ is the total capture flux considering the thermal distribution of the two dimensional rotational energy, E_r, of the neutral and $\rho(E_t)$ is the translational density of states per unit volume of the colliding partners at infinite separation. Thus Equation (3) becomes

$$k_{cap}(E_t T) = \frac{1}{2\pi(2\mu)^{3/2}E_t^{1/2}k_B T}\int \chi_r\{p,q\}d\gamma_1 d\gamma_2 2LdL\exp(-E_r/k_B T)dE_r \tag{4}$$

where k_B is the Boltzmann constant. γ_1 and γ_2 are Euler angles. $\chi_r\{p,q\}$ is the characteristic function which labels the trajectory with initial condition {p,q} on r as reactive or nonreactive. If the trajectory returns to 65 Å, then $\chi_r = 0$ for that trajectory. If the trajectory reached r_o, then $\chi_r = 1$. Figure 8 shows the calculated collision rate constant as a function of E_t for the ion-polar molecule system OH^- + HCN at 300K using 1000 trajectories for each given E_t. This leads to an error of <10%. Collision rate constants of other systems can be calculated in the same way. As pointed out in the previous work,[9,10] to get more accurate results, about 3000 trajectories are needed for a given kinetic energy.

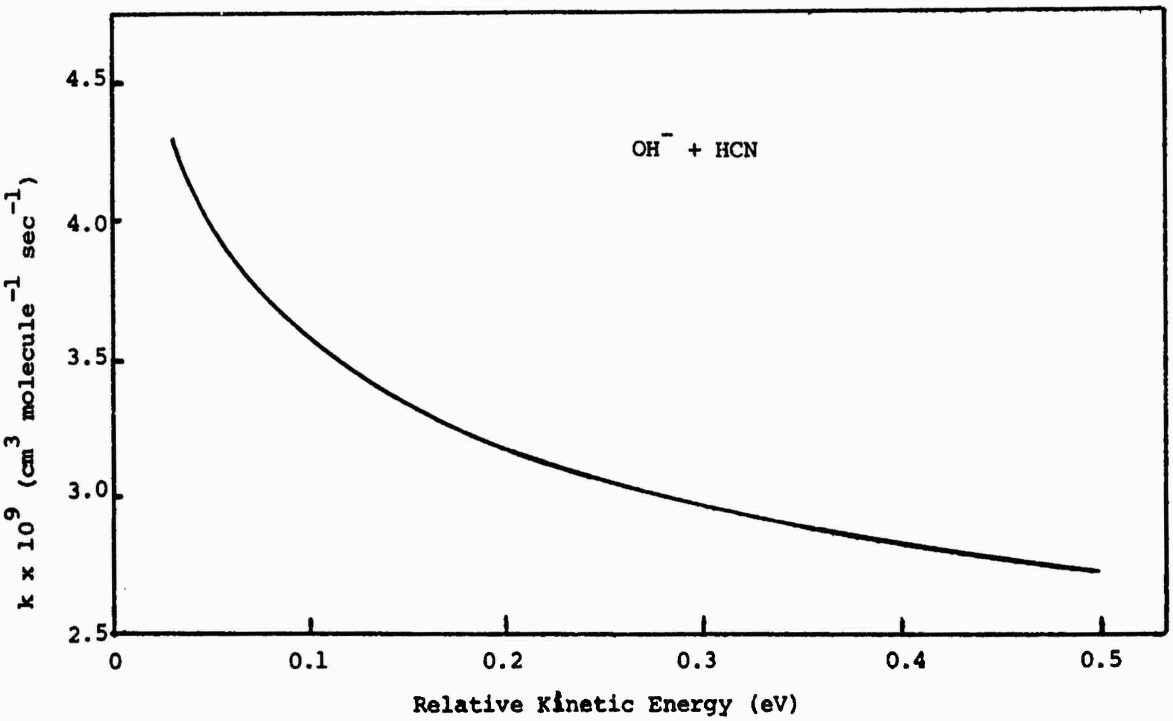


Figure 8. Kinetic Energy Dependence of Collision Rate Constant.

VI. RECOMMENDATIONS

There are at least a couple of theoretical treatments that can be performed to provide better understanding of the mechanisms of the ion-polar molecule reactions studied in this project.

(1) As mentioned previously, the reaction probability of S_N2 reactions may depend on the lifetime of the ion-molecule collision complex. The lifetime of ion-molecule complexes can be estimated by using the method of the thermal energy trajectory calculations with the trajectories integrated inwardly from large ion-molecule separations. The time that each trajectory spends within the capture radius or an assigned ion-molecule separation will be recorded. The relative lifetime of the ion-molecule complex at a given temperature can be approximated by the average time of all captured trajectories spend within this assigned separation. This can be done for different ion-molecule systems at different temperatures to gain more insight into ion-polar molecule interactions.

(2) A statistical theoretical treatment can be performed to study the branching ratios of the reactions of OH^- and O^- with methyl halides.

We suggest that the transition state switching model[14,15] be employed. This model recognizes that there are at least two transition states, a tight transition state close to the unimolecular reactant (or ion-molecule complex) and an orbiting transition state close to the separated products. The transition state switching model requires that the best transition state to use at each energy and angular momentum is the one of minimum flux. By comparing theoretical model with experimental results, information of the structure of the complex and potential energy surface of the reaction paths can be deduced.

Finally, further development in trajectory calculations of ion-polar molecule interactions are desirable in both the temperature dependence and the kinetic energy dependence of collision rate constants. For example, other potential energies such as the induced dipole-induced dipole potential and the ion-quadrupole potential can be added. The size of the particles and the three dimensional rotational motion of non-linear molecules can also be considered in the calculations.

REFERENCES

1. For recent reviews, see (a) Kinetics of Ion-Molecule Reactions, edited by P. Ausloos, Plenum, New York, 1979; (b) Gas Phase Ion Chemistry, edited by M.T. Bowers, Academic, New York, 1979.

2. See, for example, W.J. Chesnavich, T. Su and M.T. Bowers, "Ion-Dipole Collisions" in reference 1(a); A.W. astleman, Jr., "Studies of Ion Clusters" in reference 1(a); and Chapters 3, 4 and 6 in reference 1(b).

3. See, for example, Chapter 6 of reference 1(b).

4. G.I. Mackay, L.D. Belowski, J.D. Payzant, H.I. Schiff and D.K. Bohme, "ate Constants at 297K for Proton-Transfer Reactions with HCN and CH_3CN. Comparisons with Classical Theories and Exothermicity," J. Chem. Phys., Vol. 80, pp 2919-2922, 1976.

5. D.K. Bohme, "The kinetics and energetics of proton transfer," in Interactions Between Ions and Molecules, edited by P. Ausloos, Plenum, New York, 1975.

6. K. Tanaka, G.I. Mackay, J.D. Payzant and D.K. Bohme, "Gas-Phase Reactions of Anions with Halogenated Methanes at 297 ± 2K," Can. J. Chem., Vol. 54, pp 1643-1659, 1976.

7. See E.E. Ferguson, "Ion-Molecule Reactions in the Atmosphere" in reference 1(a) and Chapter 2 of reference 1(b).

8. F. Arnold, "Physics and Chemistry of Atmospheric Ions," in Atmosphere Chemistry, Physical and Chemical Science Research Report 4, edited by E.D. Goldberg,Dahlem Konferenzen, Springer-Verlag, 1982.

9. W.J. Chesnavich, T. Su and M.T. Bowers, "Collisions in a Non-central Field: A variational and Trajectory Investigation of Ion-Dipole Capture," J. Chem. Phys., Vol. 72, pp 2641-2655, 1980.

10. T. Su and W.J. Chesnavich, "Parameterization of the ion-polar molecule collision rate constant by trajectory calculations," J. Chem. Phys., Vol. 76, pp 5183-5185, 1982.

11. See D.L. Albriton, "Energy Dependence of Ion-Neutral Reactions Studied in Drift Tubes," in reference 1(a) and Chapter 1 of reference 1(b).

12. N.G. Adams and D. Smith, "The Selected Ion Flow Tube (SIFT); A Technique for Studying Ion-Neutral Reactions," Int. J. Mass Spectrom. Ion Phys., Vol. 21, pp 349-359, 1976.

13. G. Gioumousis and D.P. Stevenson, "Reactions of Gaseous Molecule Ions with Gaseous Molecules. V. Theory," J. Chem. Phys., Vol. 29, pp 294-299, 1958.

14. W.J. Chesnavich, L. Bass, T. Su and M.T. Bowers, "A multiple transition state in unimolecular reactions: A transition state switching model. Application to the $C_4H_8^+$ system," J. Chem. Phys., Vol. 74, pp 2228-2246, 1981.

15. M.F. Jarrold, L.M. Bass, P.R. Kemper, P.A.M. van Koppen and M.T. Bowers, "Unimolecular and bimolecular reactions in the $C_4H_6^+$ system: Experiment and theory," J. Chem. Phys., Vol. 78, pp 3756-3766, 1983.

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

A DYNAMIC MINI-MODEL FOR SPACE TECHNOLOGY RESOURCE ALLOCATION

Prepared by:	Patrick J. Sweeney Stephen C. Cooley
Academic Rank:	Associate Professor Graduate Student
Department and University:	Engineering Management and Systems, University of Dayton
Research Location:	Air Force Wright Aeronautical Laboratory, Material Laboratory (AFWAL/ML)
USAF Research:	Dr. Robert Barthelmy and Mr. Gary Waggoner
Date:	July 22, 1983
Contract No:	F49620-82-C-0035

A DYNAMIC MINI-MODEL

FOR SPACE TECHNOLOGY RESOURCE ALLOCATION

by

Patrick J. Sweeney
Stephen C. Cooley

ABSTRACT

The report outlines a rationale and methodology that can be used to allocate resources in the Air Force Research and Developement Space Program. The results of allocations are shown both graphically and in tabular form. The report addresses mission priorities and work unit (research projects) selection and shows how each can affect the overall mission accomplishment. That work unit subjective probability of success estimates have traditionally been overly optimistic is also addressed and a potential solution is discussed. The dynamic model developed can easily be utilized to search out best overall solutions and for exercising "what if" opportunities. An extensive bibliography of resource allocation model evaluations is included.

Acknowledgement

The authors wish to thank Dr. Robert Barthelemy and Mr. Gary Waggoner of AFWAL. Their personal involvement in this effort was singularly responsible for its success. The insights provided by the Aerospace Corporation and the Air Force Space Divisions are also greatly appreciated. A special thanks is in order for Ms. Mary Ann Penrod and her staff, who from rather poorly prepared text, prepared this report.

A special thanks to Prof. Warren Peele and his professional staff who make this summer faculty program such a gratifying experience.

I. INTRODUCTION

Knowing which technology to fund when, and at what level of effort, has plagued research organizations for years. The risk associated with pushing technology at the leading edge, at the expense of more convential technologies, is high indeed. Similarly, if the leading edge research program is successful it results in great payoff.

The drug companies must expend vast sums of money on basic research that they hope will result in a marketable product. In a like manner the Defense Department must fund a multitude of research efforts in hopes that this basic research will assist in mission accomplishment fifteen to fifty years hence. The Air Force Space Mission has been defined for the 1990's and today's research must answer those questions necessary to put the best hardware into space.

Limited budgets are a fact of life with the Air Force Space Program. Therefore, it is absolutely essential that the Air Force support those research efforts that lead most directly to mission accomplishment. Anything less than this is an inefficient expenditure of funds and may potentially lead to non-mission accomplishement. When the cost of losing is to jeopardize national security, the pressure to assure an optimum research program is invaluable.

This report addresses the key risks associated with research funding and suggests a dynamic simulation as a methodology for selecting the most favorable research project portfolio. All allocations are based upon the accomplishment of the Air Force Space Mission in the nineties. the risks associated with new technologies and the priorities of various missions are all included in the modeling effort.

II. OBJECTIVE

The objective of this report was to determine the feasibility of designing an improved research and development work unit selection model for use by the Space Applications Major Thrust (SAMT) group.

The model should address the six identified space technology goals, include the 34 major tasks, be relatively simple but inclusive of all significant variables, be interactive if possible, provide a visual result, and select the appropriate work units to "optimize" an overall SAMT time varying measure of effectiveness.

The authors have a combined total of over 24 years of military service and therefore, began with a general understanding of the interworkings of the current R & D project (work units in this case) selection methodology. A computerized literature search and comprehensive review were used as background for the modeling that followed. It was agreed upon early in this work that the potential value of a computer simulation model was very high. This modeling technique can include probabilities, risk, large numbers of work units, sensitivity analysis, and excellent (easy to understand) graphics.

III. LITERATURE REVIEW

An initial search of the current literature addressing Research and Development (R and D) investment strategies yielded an abundance of individual R and D "project selection" models. It is estimated by one prominent author in this field that literally hundreds of models have been developed to accomplish this task[1]. In response to a profusion of previously developed models; the decision was made to seek out model comparison articles which might provide some insights into the

strengths and weaknesses of currently available R and D selection models and identify which approaches have previously been explored. The following is a short history of the comparative literature in the field of R and D project selection modeling.

The article which appears to be the baseline for most of the current comparative R and D project selection work done was published in 1964 by N. R. Baker and W. E. Pound. Appropriately entitled "R and D project Selection: Where We Stand,"[2] Baker and Pound discuss in fair detail three separate models which they believed were representative of the models in existence at the time. These classifications and their brief descriptions are provided below:

a) Decision Theory Approach: Represented by the Mattley-Newton model. This approach involves rating each project subjectively with respect to a number of important criteria such as: 1) promise of technical success, 2) duration of project, 3) estimated cost etc. A projects decision score is then calculated as the product of these individual scores. Those with the highest product are considered most desirable.

b) Economic Analysis Approach: Represented by the Disman model. This approach involves the calculations of a maximum expenditure justified (MEJ) for each project. This project MEJ is equal to the product of the discounted net value of the project and the project probability of success. The MEJ divided by the cost of the project then provides a relative "benefits to cost" ratio to which other projects can be compared.

c) Operations Research Approach: Represented by the unique Hess model. This approach involves the application of sophisticated mathematics to the project selection problem. Estimates are required of the benefit value of each project as of the nth period of time prior to success, and the probability of success in each period. Selection is accomplished by maximizing the expected value mathematically. Data requirements for this model are restrictive.

These three basic model types were an initial attempt to model what Baker and Pound realized was "a relatively unstructured problem." The approaches were for the most part derivatives of previous Operation Research-Management Science techniques and the authors further realized that the problem required more study before the validity of any of these approaches could be ascertained.

Following the approach of the Baker/Pound article was a model comparison article published by Cetron, Martino, and Roepske[3]. This article expanded the number of models considered to 30 with the bulk of the new models falling into one of the three categories previously established by Baker/Pound. Each of these models is then judged against a fixed set of 26 requirements such as sensitivity to small input changes, time constraints, etc. The end result is a series of charts which provide at least a rudimentory overview of the then current modeling capabilities.

An interesting side-note in the Cetron/Martino/Pound article is their mention of several models intended for Defense Department use. Representative of these models is Esch's PATTERN model which attempts to

predict technology requirements necessary to support future national security objectives and then direct the R & D optimization efforts into filling the predicted shortfalls. This appears to be a significant refinement over previous approaches.

In 1975 N. R. Baker with co-author J. Freeland published a very significant follow-on article to Baker's initial 1964 report[3]. This article summarized some of the significant work which had been accomplished during the previous decade, including some very sorely needed quantitative model evaluation techniques which allowed for more accurate comparisons of the expanding number of R & D project selection models. This topic will be discussed in greater detail later in this paper.

In a departure from previous model classification procedures, Baker and Freeland grouped competing models under two major categories; namely benefit measurement models and project selection/resource allocation models. Benefit measurement models are those which attempt to measure the contribution which individual projects or a combination of projects make in satisfying established research goals. The authors point out that the use of a benefit measurement type model allows for an interactive portfolio analysis procedure which may have inherently more meaning than the evaluation of isolated projects. They also point out that an interactive benefit contribution model can also supply its user with a valuable form of sensitivity analysis which allows the operator to "test" economic and portfolio changes.

On the topic of benefit contribution models, it is worthwhile to note that in their 1976 assessment of military R & D project selection models for Project RAND[4] E.S. Ojdana and J.P. Weyant strongly favored

benefit contribution models over other types for this application. The primary reason cited by these authors is apparently organizational compatability. "The use of benefit contribution models appears better suited to the needs of the Air Force (and the DOD) because these departments now justify R & D projects on the basis of mission requirements or goals."

Finally, the most current model comparison articles provide a final update on current modeling techniques. In his 1978 article, W. E. Souder discusses several modeling concepts which could have utility in future R & D project selection models[1]. One of these techniques involves risk analysis which allows the decision maker to more accurately input his/her chosen investment strategies into the project selection process. Unfortunately risk analysis requires a substantial amount of additional subjective estimating of future project characteristics which may negate some of its benefits. A second technique known as frontier modeling may function as a screening device to eliminate unacceptable projects prior to entry into the general model. Frontier modeling establishes a graphical acceptance region based on estimated project risk and return criterion which allows for quick indentification of unsuitable projects. Also discussed in Souder's article are project profiling procedures and Q-sort technique which may also be of some value.

Attempting to model a problem as unstructured and complex as R and D project selection has generated an abundance of models each with different capabilities and limitations. W. E. Souder has developed a systematic methodology for analytically scoring perspective models based

upon a fixed set of criteria[5] which seems to have gained some acceptance by other authors in the field[3,6]. This model may provide a much needed comparison methodology by which the utility of new models may be measured. Before discussing this system in more detail a basic modeling assumption shall also be addressed.

Inherent in almost any of the modeling forms in current existance is the requirement for some sort of project "probability of technical success" estimate, usually provided by the project manager or sponsor. This subjective probability estimate has been pointed out by a number of authors as a source of error. In his 1979 review of the literature A. Schnarch concluded "Some of the optimism in forecasting costs and completion times may be -- according to the literature -- due to deliberate manipulation on part of R & D management with the aim of making projects look attractive."[6] W.E. Souder also accomplished a preliminary study in this field and concluded that in the absence of organizational pressure project managers are capable of providing meaningful project probability estimates even for preliminary research projects[7]. Since these subjective probability estimates usually provide the "underpinnings" upon which the subsequent project selection model is built, it is essential that some efforts should be made to eliminate organizational bias.

This organizational and individual bias or over optimism can be controlled by using the following technique. Advise each work unit manager that he will be making probability estimates concerning his work units from start to finish on each project and that the relative accuracy of those estimates will be factored into future funding

exercises. In other words the better his estimates the higher the probability that his future estimates will be accepted at face value. The better he estimates now, the higher his future credibility will be and thus the better his chances of future funding. Figure 1 shows a typical work unit that was managed to success. The vertical axis is the probability of success. In this case the manager's optimism was a continually growing (higher probability of success) estimate with passing time. The summation of the area beneth the curve is equal to 80%. Figure 2 shows a similar optimistic view, however this time the project failed. The area above the curve was 20%. In the first case the manager was "80%: correct and in the second only "20%". These percentages would then be used to off-set over optimism in future work-unit forecasts.

IV. THE MODEL

Modeling of the SAMT resource allocation (project selection) problem must be attacked realistically from both the top (SAMT goals) and the bottom (work units) simultaneously. The SAMT goals are established by high level DOD personel, whereas the work unit resources required for success must be determined at the project engineer level, see Figure 3.

The dynamic computer simulation techniques applied for this effort is Systems Dynamics and uses the DYNAMO Compiler. In order to demonstrate the capabilities of dynamic simulation a small mini-model was designed and executed. Figure 4 is a schematic of this model. Dr. Robert Barthelemy, AFWAL/ML provided the data for this test edition of

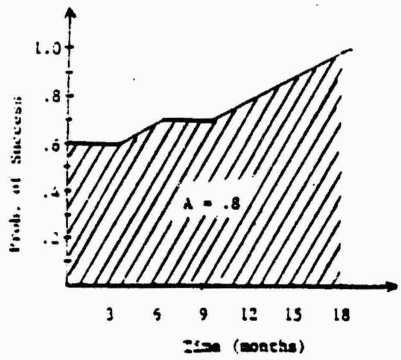


Figure 1 - Accuracy of Probability Estimates over Time

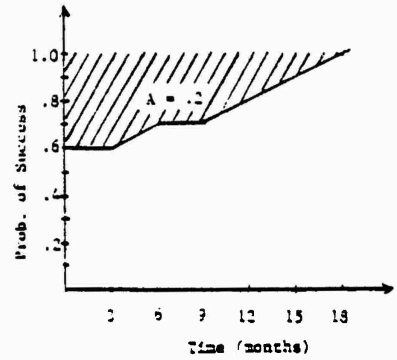


Figure 2 - Accuracy of Probability Estimates over Time

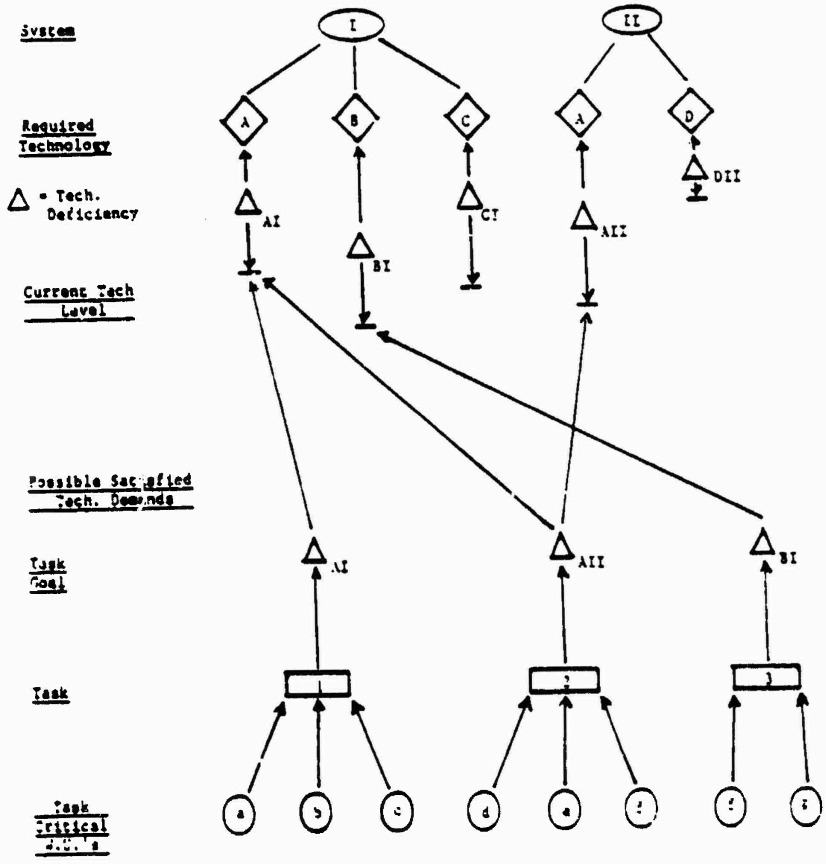


Figure 3 - SAMT Resource Allocation

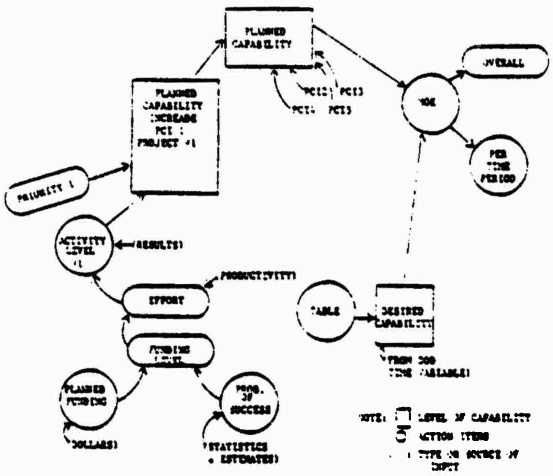


Figure 4 - Schematic of the Model

```
 1.0000 NOTE<*****************************************************>
 2.0000 NOTE<*                                                   *>
 3.0000 NOTE<*            RESOURCE ALLOCATION MODEL:             *>
 4.0000 NOTE<*                                                   *>
 5.0000 NOTE<*     THIS EARLY EDITION OF THE MODEL ENABLES THE   *>
 6.0000 NOTE<* USER TO TEST THE AFFECTS OF PRIORITY, FUNDING, THE *>
 7.0000 NOTE<* TIMELINESS OF THE FUNDING UPON THE CONTRIBUTION OF *>
 8.0000 NOTE<* EACH PROGRAM AND THE OVERALL ARRAY CONTRIBUTION.  *>
 9.0000 NOTE<*                                                   *>
10.0000 NOTE<*****************************************************>
11.0000 NOTE
12.0000 NOTE
13.0000 NOTE<<<PARAMETERS
14.0000 NOTE
15.0000 NOTE  PROGRAM                PRIORITY FUNDING DURATION(START)
16.0000 NOTE<< 1)GA-AS SOLAR CELLS      3        9         7(80)
17.0000 NOTE<< 2)NI-MH BATTERIES        2        3         5(83)
18.0000 NOTE<< 3)HVPS                   1       14         6(83)
19.0000 NOTE<< 4)SOLAR CONCENTRATOR     5        4         5(84)
20.0000 NOTE<< 5)25 KW P.S.             4       19         5(85)
21.0000 NOTE
22.0000 L PC.K=PC.J+DT(PCI1.JK+PCI2.JK+PCI3.JK+PCI4.JK+PCI5.JK)
23.0000 NOTE.....PC=PLANNED CAPABILITY
24.0000 N PC=0
25.0000 A PCI1.KL=CLIP(0,PC1.K,PC1.K,5)*PRIOR1.K*ACT1.K
26.0000 NOTE.....PCI1=THE RATE INCREASE OF PC CAUSED BY PROGRAM 1.
27.0000 NOTE.....PC1 AND PCI2 ARE THE FACTORS THAT DEFINE THE CONTRIBUTION
28.0000 NOTE.....OF PROGRAM 1.   PCI1= THE PLANNED CAPABILITY INCREASE
29.0000 NOTE.....CAUSED BY PROGRAM 1.  PCI2 AND TPCI2 CONSTITUTE THE TABHL
30.0000 NOTE.....FUNCTION.
31.0000 A PC1.K=STEP*PCI2.K,SID)
32.0000 A PCI2.K=TABHL(TPCI2,(TIME.K-ACT1.K)-SID,0,0,0)
33.0000 T TPCI2=0/9
34.0000 C SID=0
35.0000 NOTE.....SID=THE DELAY IN STARTING PROGRAM 1 FROM START TIME.
36.0000 A ACT1.K=TABHL(TACT1,EFRT1.K,0,200,100)
37.0000 NOTE.....ACT1=ACTIVITY LEVEL OF PROGRAM PROGRESS.  THIS IS TIED
38.0000 NOTE.....TO TACT1 AND IS CONTROLLED BY EFRT1.
39.0000 T TACT1=0/1/1.9
40.0000 A PRIOR1.K=TEN/PR1
41.0000 C TEN=10
42.0000 C PR1=3
43.0000 NOTE.....PRIOR1.K=THE PRIORITY OF THE PROGRAM (H(6= IS GOOD).
44.0000 A EFRT1.K=TABHL(TEFRT1,FUND1.K,0,10,10)
45.0000 NOTE.....EFRT1=EFFORT IN PERCENTAGE OF PLANNED EFFORT, NORMAL
46.0000 NOTE.....LEVEL OF EFFORT IS 100.
47.0000 NOTE.....THIS IS A TABHL FUNCTION AND IS CONTROLLED BY FUND1.
48.0000 T TEFRT1=0/200
49.0000 A FUND1.K=TABHL(TFUND1,PROD1.K,-.5,.5,.1)
50.0000 T TFUND1=0/.5/.5/.5/.5/.5/.5/.5/.5/1
```

Figure 5 - The SAMTMOD Computer Program

```
51.0000 A PROB1.K=SAMPLE(NSN1.K,INTV,ITSAM1)
52.0000 C INTV=1
53.0000 C ITSAM1=4
54.0000 A NSN1.K=NOISE()
55.0000 NOTE.....FUNB1=THE FUNDS ALLOCATED FOR THE PROJECT.
56.0000 P PC12.KL=CLIP(0,PC2.K,PC2.K,3)+PRIOR2.K*ACT2.K
57.0000 A PC2.K=STEP(PC2E.K,E2B)
58.0000 A PC2E.K=TABHL(TPC2E,(TIME.K*ACT2.K)-E2B,0,60,60)
59.0000 T TPC2E=0/3
60.0000 C E2B=36
61.0000 A ACT2.K=TABHL(TACT2,EFRT2.K,0,200,100)
62.0000 T TACT2=0/1/1.5
63.0000 A PRIOR2.K=TEN/PR2
64.0000 C PR2=2
65.0000 A EFRT2.K=TABHL(TERFT2,FUNB2.K,0,6,6)
66.0000 T TERFT2=0/200
67.0000 A FUNB2.K=TABHL(TFUN22,PROB2.K,-.5,.5,.1)
68.0000 T TFUNB2=3/3/3/3/3/3/3/3/3/4
69.0000 A PROB2.K=SAMPLE(NSN2.K,INTV,ITSAM2)
70.0000 C ITSAM2=8
71.0000 A NSN2.K=NOISE()
72.0000 P PC13.KL=CLIP(0,PC3.K,PC3.K,14)+PRIOR3.K*ACT3.K
73.0000 A PC3.K=STEP(PC3E.K,E3B)
74.0000 A PC3E.K=TABHL(TPC3E,(TIME.K*ACT3.K)-E3B,0,72,72)
75.0000 T TPC3E=0/14
76.0000 C E3B=36
77.0000 A ACT3.K=TABHL(TACT3,EFRT3.K,0,200,100)
78.0000 T TACT3=0/1/1.5
79.0000 A PRIOR3.K=TEN/PR3
80.0000 C PR3=1
81.0000 A EFRT3.K=TABHL(TERFT3,FUNB3.K,0,20,20)
82.0000 T TERFT3=0/200
83.0000 A FUNB3.K=TABHL(TFUNB3,PROB3.K,-.5,.5,.5)
84.0000 T TFUNB3=14/14/20
85.0000 A PROB3.K=SAMPLE(NSN3.K,INTV,ITSAM3)
86.0000 C ITSAM3=8
87.0000 A NSN3.K=NOISE()
88.0000 P PC14.KL=CLIP(0,PC4.K,PC4.K,4)+PRIOR4.K*ACT4.K
89.0000 A PC4.K=STEP(PC4E.K,E4B)
90.0000 A PC4E.K=TABHL(TPC4E,(TIME.K*ACT4.K)-E4B,0,60,60)
91.0000 T TPC4E=0/4
92.0000 C E4B=40
93.0000 A ACT4.K=TABHL(TACT4,EFRT4.K,0,200,100)
94.0000 T TACT4=0/1/1.5
95.0000 A PRIOR4.K=TEN/PR4
96.0000 C PR4=5
97.0000 A EFRT4.K=TABHL(TERFT4,FUNB4.K,0,0,0)
98.0000 T TERFT4=0/200
99.0000 A FUNB4.K=TABHL(TFUNB4,PROB4.K,-.5,.5,.5)
100.0000 T TFUNB4=4/4/10
101.0000 A PROB4.K=SAMPLE(NSN4.K,INTV,ITSAM4)
102.0000 C ITSAM4=8
103.0000 A NSN4.K=NOISE()
104.0000 P PC15.KL=CLIP(0,PC5.K,PC5.K,15)+PRIOR5.K*ACT5.K
105.0000 A PC5.K=STEP(PC5E.K,E5B)
106.0000 A PC5E.K=TABHL(TPC5E,(TIME.K*ACT5.K)-E5B,0,60,60)
107.0000 T TPC5E=0/15
108.0000 C E5B=60
109.0000 A ACT5.K=TABHL(TACT5,EFRT5.K,0,200,100)
110.0000 T TACT5=0/1/1.5
111.0000 A PRIOR5.K=TEN/PR5
112.0000 C PR5=4
113.0000 A EFRT5.K=TABHL(TERFT5,FUNB5.K,0,30,30)
114.0000 T TERFT5=0/200
115.0000 A FUNB5.K=TABHL(TFUNB5,PROB5.K,-.5,.5,.1)
116.0000 T TFUNB5=15/15/15/15/17/19/21/23/25/27/30
117.0000 A PROB5.K=SAMPLE(NSN5.K,INTV,ITSAM5)
118.0000 C ITSAM5=.2
119.0000 A NSN5.K=NOISE()
120.0000 A BC.K=TABHL(TBC,TIME.K,0,110,110)
121.0000 T TBC=1000/0000
122.0000 A SHTFL.K=BC.K-PC.K
123.0000 NOTE*****SHTFL=THE PERIOD DIFFERENCE BETWEEN DESIRED
124.0000 NOTE*****CAPABILITY AND PLANNED CAPABILITY.
125.0000 R SHTFLR.KL=SHTFL.K/TIMTOT
126.0000 NOTE*****SHTFLR=RATE EQUATION USED TO COMPUTE THE TOTAL DIFFERENCE
127.0000 NOTE*****BETWEEN DESIRED CAPABILITY AND PLANNED CAPABILITY.
128.0000 C TIMTOT=1100
129.0000 NOTE*****TIMTOT=A CONSTANT DETERMINED BY MULTIPLYING DT (.1)
130.0000 NOTE*****BY THE TIME (110).  WITH THIS CONSTANT THE AREA BETWEEN
131.0000 NOTE*****DESIRED AND PLANNED CAPABILITY IS CALCULATED IN SHTFLL.
132.0000 L SHTFLL.K=SHTFLL.J+DT(SHTFLR.JK)
133.0000 N SHTFLL=0
134.0000 A FUNTOT.K=PC1E.K+PC2E.K+PC3E.K+PC4E.K+PC5E.K
135.0000 SPEC DT=.1
136.0000 SPEC LENGTH=130
137.0000 SPEC PLTPER=8
138.0000 SPEC PRTPER=8
139.0000 PLOT PC/PC11/PC12/PC13/PC14/PC15/BC/SHTFL/SHTFLL/FUNTOT
140.0000 PRINT PC,PC11,PC12,PC13,PC14,PC15,BC
141.0000 RUN
142.0000 C E5B=24
143.0000 RUN
144.0000 T TFUNB5=7/7/7/7/8/10/11/12/13/14/15
145.0000 RUN
146.0000 C PR5=2
147.0000 RUN
148.0000 SPEC PLTPER=8
149.0000 SPEC PRTPER=10
150.0000 RUN
```

Figure 5 - Concluded

the model, see lines 16-20 of Figure 5, The Model Program. This test simulation can show the effects of varying priorities, funding levels and timeliness upon the overall (five projects) effectiveness in reducing the projected overall technology deficiency.

V. RESULTS AND CONCLUSIONS

Figure 6 is the mini-model basic run. The difference between the desired capability and the planned, is shown by the cross hatched area. The period shortfall ranges from 3700 to 100 units. The accumulative difference is 230 units.

Figure 7 shows the effects of beginning program #5 early. The total differences vary from 3600 to 300. This shows an obvious improvement in overall measure of effectiveness (MOE). Starting this program early helps improve MOE.

Figure 8 shows a reduced funding on program #5. The total difference increases to 280 and the period difference from 3700 to 1100. Reducing funds on program #5 reduces the MOE.

Figure 9 shows the affects of increased priority. The total difference is 195 and the period difference varies from 3700 to minus 1500 (more than required). The MOE is significantly altered by priority change. Tabular data can be produced by the model, but was omitted, due to space limitations.

The model responds appropriately. The results are easily seen and understood. This small mini-model demonstrates the capabilities of dynamic modeling.

This modeling technique given good mission priorities and work unit probabilities can be a significant aid to research fund allocation.

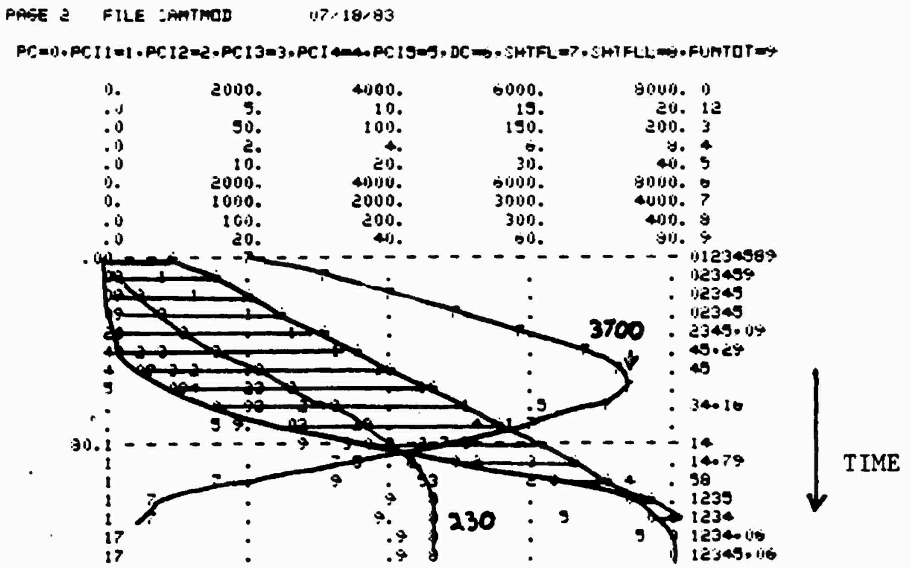


Figure 6 - Basic Run

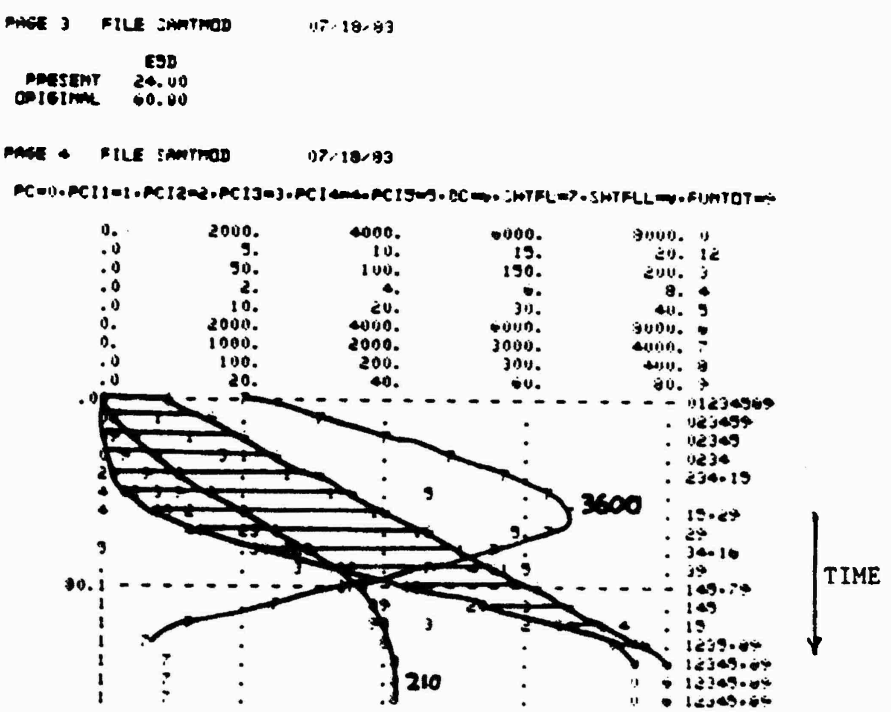


Figure 7 - Early Start Time on #5 Program

PAGE 5 FILE SAMTMOD 07/18/83

	TFUND5							
PRESENT	7.000	7.000	7.000	7.000	8.000	10.00	11.00	12.00
ORIGINAL	15.00	15.00	15.00	15.00	17.00	19.00	21.00	23.00
	. . .	. . .	. . .					
PRESENT	13.00	14.00	15.00					
ORIGINAL	25.00	27.00	30.00					

PAGE 6 FILE SAMTMOD 07/18/83

PC=0,PC11=1,PC12=2,PC13=3,PC14=4,PC15=5,DC=6,SWTFL=7,SWTFLL=8,FUNTOT=9

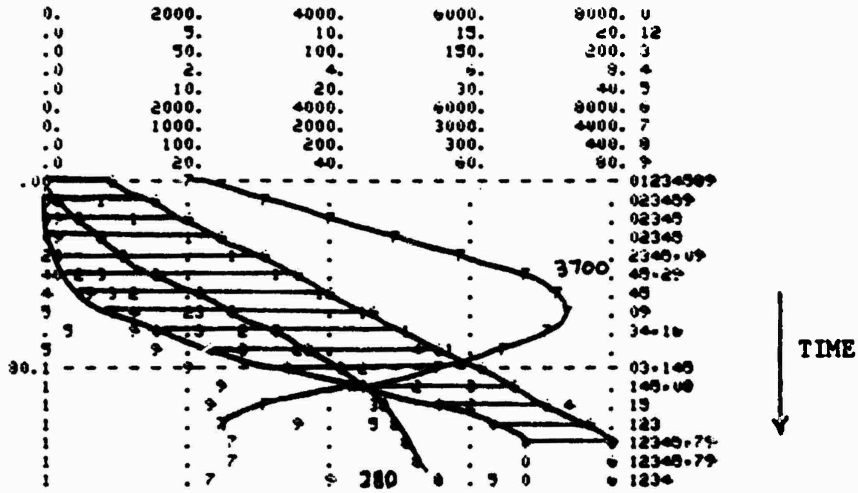


Figure 8 - Reduced Funding on #5 Program

PAGE 7 FILE SAMTMOD 07/18/83

	PRI
PRESENT	2.000
ORIGINAL	4.000

PAGE 8 FILE SAMTMOD 07/18/83

PC=0,PC11=1,PC12=2,PC13=3,PC14=4,PC15=5,DC=6,SWTFL=7,SWTFLL=8,FUNTOT=9

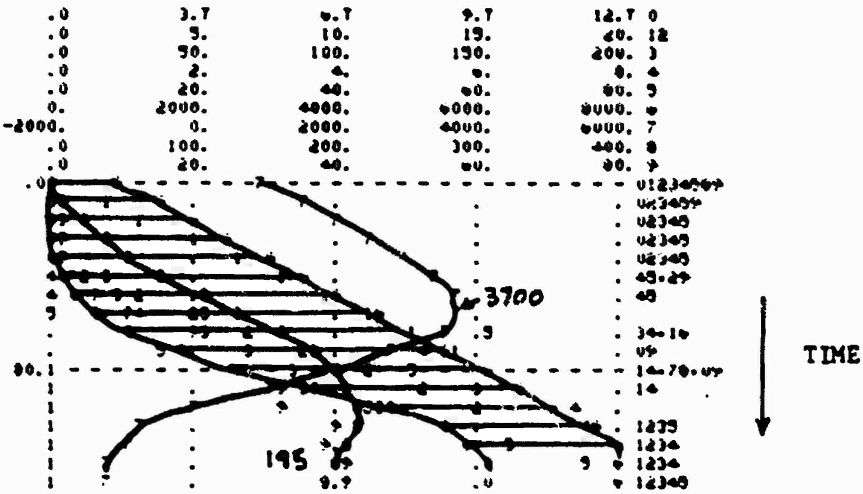


Figure 9 - Improved Priority on #5 Program

VI. RECOMMENDATIONS

The importance of funding the research that best accomplishes the Air Force Space Mission is self-evident. This report outlines a methodology that will accomplish this funding allocation. The methodology is straight forward, simple, and easy to understand. All of this enhances the probability that the results will be used by Air Force decision makers.

This work should continue. The authors recommend that an actual Space Mission Model be completed as soon as possible. This should then be validated and evaluated. There appears to be no reason why this methodology could not be expanded to the AFWAL Missions the Air Force Missions, and subsequently to the Defense Department Mission.

Research in the area of resource allocation is critical if the Defense Department is to get the "most bang for the buck" possible.

REFERENCES

1. Souder, W.E., "A System for Using R & D Project Evaluation Methods," Research Management, pp. 29-37, September 1978.

2. Baker, N.R. and W.H. Pound, "R and D project selection: Where we stand," IEEE Trans. on Engineering Management, pp. 124-134, December 1964.

3. Baker, N.R. and J. Freeland, "Recent Advances in R & D Benefit Measurement and Project Selection Methods," Management Science, pp. 1164-1175, June 1975.

4. Ojdana, E.S. and J.P. Weyant, "An Assessment of Selected Models Used for Evaluating Military R & D Projects," The Rand Corporation, R-1847-PR, September 1976.

5. Souder, W.E., "Comparitive Analysis of R & D Investment Models," AIIE Transactions, pp. 57-64, March 1972.

6. Schnarch, A., "The Budgeting of Industrial R & D Projects: A Review of the Literature," Society of Research Administrators Journal, pp. 43-46, Fall 1979.

7. Souder, W.E., "The Validity of Subjective Probability of Success Forcasts by R & D Project Managers," IEEE Trans. on Engineering Management, pp. 35-49, February 1969.

ADDITIONAL REFERENCES

8. Cetron, M.J., J. Martino, and L. Roepcke, "The selection of R & D Program Contine-Survey of Quantitative Methods," IEEE Trans. on Engineering Management," pp. 4-13, March 1967.

9. Liberatore, M.J., "An Incremental Approach for R & D Project Planning and Budgeting," Research Management, pp 17-20, March 1981.

10. Cooper, M.J., "An Evaluation System for Project Selection," Research Management, pp. 29-33, July 1978.

11. Aaker, D.A. and T.T. Tyebjee, "A Model for the Selection of Interdependent R & D Projects," IEEE Trans. on Engineering Management, pp. 30-36, May 1978.

12. Cochran, M.A., E.B. Pyle, III, L.C. Greene, H.A. Clymer, and A.D. Bender, "Investment Model for R & D Project Evaluation and Selection," IEEE Trans. on Engineering Management, pp. 89-100, August 1971.

13. Goodwin, P.G., "A Method for Evaluation of Subsystem Alternate Designs," IEEE Trans. On Engineering Management, pp. 12-21, February 1972.

14. Borison, A.B., B.R. Judd, P.A. Morriss, and S.S. Sussman, "Applying Analytic Methods to R & D Planning: Two Case Studies," Research Management, pp. 17-24, November 1982.

15. Maher, P.M. and A.H. Rubenstein, Factors Affecting Adoption of A Quantitative Method for R & D Project Selection Models," Management Science, pp. 119-129, October 1974.

16. Souder, W.E., "Utility and Perceived Acceptability of R & D Project Selection Models," Management Science, pp. 1384-1394, August 1973.

17. Ebert, R.J., "Methodology for Improving Subjective R & D Estimates," IEEE Trans. on Engineering Management, pp. 108-116, August 1970.

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

EARLY PERFORMANCE OF THE GALLIUM ARSENIDE PHOTOVOLTAIC ARRAY ON THE LIVING PLUME SHIELD (LIPS) SATELLITE

Prepared by:	**Dr. Paul P. Szydlik**
Academic Rank:	**Professor**
Department and University:	**Department of Physics State University of New York College at Plattsburgh**
Research Location:	**Air Force Wright Aeronautical Laboratories, Aero Propulsion Laboratory, Aerospace Power Division, Energy Conversion Branch**
USAF Research:	**Mr. Joseph F. Wise**
Date:	**August 24, 1983**
Contract No:	**F49620-82-C-0035**

EARLY PERFORMANCE OF THE GALLIUM ARSENIDE PHOTOVOLTAIC ARRAY ON THE LIVING PLUME SHIELD (LIPS) SATELLITE

by

PAUL P. SZYDLIK

ABSTRACT

The electrical characteristics of an array of gallium arsenide photovoltaic cells in earth orbit on the Living Plume Shield (LIPS) satellite are studied as a function of time in orbit. A theoretical solar cell model and least-squares methods are used to extract those characteristics from current vs. voltage data corrected to one sun and 28°C. The decrease in electrical output of the array is related to the equivalent 1 MeV electron fluence required to produce the observed decrease. Results for the first 172 days in orbit show either that deterioration of the array based on the anticipated radiation environment is greater than expected or that the equivalent 1 MeV electron fluence for the orbit is about an order of magnitude greater than predicted.

ACKNOWLEDGEMENT

The author thanks the Air Force Systems Command, the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for providing this research opportunity. He expresses his gratitude to the Aero Propulsion Laboratory of the Air Force Wright Aeronautical Laboratories for the facilities and support in carrying out that research. He also expresses his appreciation to Mr. Joseph Wise for suggesting the problem and for his guidance and collaboration. Helpful discussions with Dr. Pat Rahilly, Mr. Terry Trumble, Dr. Jim Holt and Lt. Ken Masloski are gratefully acknowledged.

I. INTRODUCTION:

As a result of its use in a wide variety of semi-conductor devices, the technology of silicon is well developed and solar cells of this material have proven to be effective and reliable. Photovoltaic arrays of silicon solar cells have been a primary energy source for satellites for years. However, if cells of higher efficiency and greater proton and electron radiation resistance could be developed, the weight of a photovoltaic array required to supply a given power at the end of mission life would be reduced with a decrease in required payload and an associated cost savings.

Gallium arsenide (GaAs) solar cells are known to have theoretical and attainable efficiencies substantially greater than those of silicon. Recent terrestrial radiation experiments with 1 MeV electrons[1] indicate that at least up to fluences approaching 10^{16}, the resistance of GaAs cells to radiation damage is better than that of planar silicon.

Gallium arsenide solar cell technology is virtually untested in space. A limited study[2] on the Navigational Technology Satellite-2 (NTS-2) of a single passive 5-cell module of GaAs cells (with an AlGaAs window) bore out these expectations to some extent. However, the initial efficiency of these cells was in itself low (13.6%), in the same range as that of silicon and considerably below the efficiencies which have been attained in limited laboratory production of GaAs cells. Additionally, the lifetime characteristics and ability to operate in a particular radiation environment in space did not match those of certain textured silicon solar cells, for example.

An array of 300 GaAs solar cells was launched with the Living Plume Shield (LIPS) satellite under the auspices of the Naval Research Laboratory. This array can provide 1/6 of the electrical energy to the power bus

of this satellite. Sufficient data are acquired to generate a reasonably accurate current-voltage (I-V) curve. The active nature of the array coupled with the detailed knowledge of the voltage and current characteristics should provide the most comprehensive data to date on the performance of GaAs solar cells in space.

II. OBJECTIVES

The broad purpose of this study is to determine the performance in space of a GaAs photovoltaic array and ultimately to compare this performance to similar arrays of silicon cells. The performance of these arrays is measured through the variation of their electrical characteristics with time in orbit. Specifically the objectives are:

1) From current vs. voltage data for the LIPS GaAs array supplied by the Naval Research Laboratory to compute the open-circuit voltage (V_{oc}), the short circuit current (I_{sc}), maximum power (P_{max}), voltage (V_{mp}) and current (I_{mp}) at maximum power and the fill factor;
2) to develop computational tools (specifically interactive computer programs) to facilitate computation of the array's electrical characteristics in view of the large volume of data which must be analyzed;
3) from variation of the array's electrical characteristics to assess the effects of electron and proton radiation damage; to compare these results to expectations based on terrestrial experiments.

Since the radiation environment to which the array is subjected is uncertain, estimates will be made of the equivalent 1 MeV electron fluence which produces the observed degradation in performance.

These estimates will be compared to the expected fluence for the given orbit.

III. CHARACTERISTICS OF THE LIPS GaAs PHOTOVOLTAIC ARRAY

The GaAs photovoltaic array consists of 300 2 cm x 2 cm (nominal) cells. Electrically these are connected in 12 parallel banks of 25 cells in series. Open circuit voltage of the array approaches 25 volts, short circuit current at air mass zero is 1.2 amperes and maximum power at the beginning of life is about 23 watts. Cells are Hughes Liquid Phase Epitaxy with a junction depth $\sim$ 0.5 micron. The cells have a single-layer Ta_2O_5 anti-reflective coating and a UV filter. They are bonded to a 10 mil fused silica coverglass using Dow-Corning DC 93-500 adhesive.

IV. TELEMETRY DATA AND CORRECTIONS

The raw telemetry data consist of information from which the sun angle with respect to the normal to the plane of the array can be determined, the array's temperature, plus current and voltage data. Provision is made for measuring the projections (α and β) of the sun angle on two perpendicular planes which in turn are perpendicular to the plane of the array. The sun angle can then be computed from these projections. Temperature is measured using a thermistor attached to the back of the aluminum honeycomb skin on which the GaAs cells are mounted. Sufficient current-voltage pairs are measured, particularly near the "knee" of the I-V curve, from which a complete I-V curve can be generated reliably using a suitable solar cell model.

Corrections are made to the intensity for sun angle and (solar) orbital position to reduce all data to air mass zero. Appropriate temperature coefficients (-2.04×10^{-3} volts/C°/cell and $+ 3.01 \times 10^{-5}$ amperes/C°/cell) are used to correct the voltage and current data to 28°C.

The data were scrutinized to determine under what sun angle conditions shading of the array occurs. There is serious shading whenever angle $\beta > 30°$ for positive α. Minor shadowing effects occur whenever $\alpha \lesssim -40°$. There may be shading whenever α and β are both negative, but the effect is not pronounced. There are a few cases where the short-circuit current is anomalously high. It is not clear whether this is the result of an enhancement of the solar intensity because of reflections or just an unreliable measurement. Such results were isolated and ignored in the data analysis as were the cases where shading of the array was probable.

Recommendations for modification of the data corrections will be made in the final section of this report.

V. DATA ANALYSIS-PRIMARY ELECTRICAL CHARACTERISTICS OF THE GaAs ARRAY

The quality of the performance of the LIPS GaAs array is determined from the variation of the electrical characteristics with time in orbit. The characteristics to be monitored include the open circuit voltage, short circuit current, maximum power, fill factor and the voltage and current at maximum power. While the open-circuit voltage is measured directly and an excellent measure of the short circuit current can be obtained from the corrected telemetry data, further analysis is necessary to determine the other quantities. Even for a series of current and

voltage measurements during the same orbit there are variations and some averaging or fitting procedure is required. The analysis is performed using all the corrected data acquired in a particular orbit. Since data can only be obtained in real time when the satellite is in range of the receiving station a series of measurements is taken in only a few minutes. Using a simple solar cell model, a least-squares fit is made to all these current-voltage data. From this least-squares fit, the required electrical parameters are easily computed.

VI. SOLAR CELL MODEL

An equation which describes realistically the relation between current and voltage for a solar cell is [3]

$$I = I_L - I_0 (\exp [q (V + IR_S)/nKT] - 1) - (V + IR_S)/R_{SH} \quad (1)$$

where I is the current,

V is the voltage,

q is the charge of the current carriers,

I_L is the light-generated current,

I_0 is an effective saturation current,

n is a constant sometimes referred to as the "curve" or "quality" factor,

k is Boltzmann's constant,

T is the Kelvin temperature,

R_S is the series resistance and

R_{SH} is the shunt resistance.

For our purposes the shunt resistance and series resistance are neglected. The actual equation used to fit the LIPS data then has the form

$$I = P_1 - P_2 (\exp (P_3 V) - 1) \quad (2)$$

where P_1 and P_2 are the same as I_L and I_0, respectively and $P_3 = q/nKT$.

Here the P_i are made free parameters. These are adjusted to obtain the best fit to the current-voltage data as will be described in more detail in the following section.

Note that P_1 is the short-circuit current. The open-circuit voltage is obtained from equation (2) for I = 0

$$V_{oc} = \frac{1}{P_3} \ln\left[\frac{P_1}{P_2} + 1\right] \tag{3}$$

The power is obtained as the product of equation (2) and the voltage. Maximum power is determined by taking the derivative of the power equation with respect to V setting the result equal to zero and solving for the voltage at maximum power. The result is:

$$V_{mp} = \frac{1}{P_3} \ln\left[\frac{\left(\frac{P_1}{P_2} + 1\right)}{P_3 V_{mp} + 1}\right] \tag{4}$$

which is easily solved by iteration. Equation (2) is used to find the current at maximum power and the maximum power itself is simply $V_{mp} I_{mp}$.

VII. LEAST-SQUARES FIT TO THE EXPERIMENTAL DATA

Equation (2) is non-linear in parameter P_3. As a result, special fitting techniques using an iterative approach with initial parameter estimates are required. It is especially important to have a good initial estimate of P_3 because of the exponential dependence. Gradient or steepest descent procedures move rapidly toward a best fit, but do poorly in "fine-tuning" the parameters near optimum. An optimum fit is defined as a chi-squared minimum, where chi-squared is a measure of the deviations between the experimental and computed currents. Linearization methods involving a Taylor expansion of the fitting function work well close to a chi-squared minimum, but they can fail completely if initial parameter estimates are poor.

The Levenberg-Marquardt algorithm[4] attempts to combine the best features of both methods. The changes in the parameters to move to a chi-squared minimum are determined from both the gradient and linearization procedures. A larger proportion of the former is used when far from a chi-squared minimum. Of course the reverse is true as an optimum fit is approached.

Subroutine packages are available for performing the required fits, but they tend to be cumbersome in their application. Simpler procedures and programs were adapted from Bevington[5].

A typical fit to experimental data from the LIPS array is shown in Figure 1. The quality of the fit is quite good. There is one minor flaw in the analysis. Since the procedure minimized deviations in the current, it tends to weight more heavily the data closer to V_{oc} where a small error in voltage will produce a relatively large current change. As a result, the short circuit current is often lower than any measured value by 5-10 mA.

The tic on the figure marks the maximum power point. This and all other required electrical characteristics of the array are computed from the fitted curve and displayed in the output.

VIII. ELECTRICAL PROPERTIES OF THE LIPS GaAs ARRAY VS. TIME IN ORBIT

Figure 2 illustrates the variation of the short-circuit current and maximum power with time in orbit for the LIPS GaAs array. The open circuit voltage is not plotted since fluctuations mask any observable trend. The same is true for the fill factor. These values are 24.4 ± 0.1 volts and 0.773 ± 0.006, respectively. The quoted uncertainties are standard deviations from the mean. The variations in the voltage

and current at maximum power tend to exhibit the same behavior as the open circuit voltage and short-circuit current, respectively. Unfortunately the satellite was not in range of the receiving station while in the sunlight early after launch. Hence there are no data until 40 days in orbit. For similar reasons, there are no data from about day 80 to day 120. (Actually data were acquired from day 100 to day 120, but the array always experienced some shading during this period.)

While it may be hazardous to extrapolate the data to zero time, this is necessary to assess the effects of radiation damage. This extrapolation gives a zero time short-circuit current of about 1.24 amperes and a maximum power of 23.6 watts. Beginning of life (BOL) laboratory measured values are 1.28 amperes and more than 24.5 watts. On day 172, the short circuit current is 1.134 amperes and the maximum power is 21.3 watts. The drop in short circuit current is about 8.5% and that in the maximum power is about 9.7%. The estimated effect of adhesive darkening[6] for the time in orbit (about 4000 hours) is 2.5% for the short-circuit current. (The effect on the maximum power should be approximately the same.) It will be assumed that the residual deterioration in the electrical properties (about 6% for the short-circuit current and 7% for the maximum power) is caused by direct radiation damage to the cell. The 1 MeV electron fluence to produce this damage as determined in terrestrial experiments is[1] $\sim 1.5 \times 10^{14}$ electrons/cm^2. While cover glasses on the GaAs cells used in these experiments were 12 mils thick, whereas those on LIPS were 10 mils, this difference should not alter substantially the interpretation of the results.

The LIPS orbit is 600 nm circular inclined at 63°. The equivalent 1 MeV electron fluence for 172 days in this orbit assuming infinite backshielding and a 10 mil cover glass is 1.9×10^{13} electrons/cm^2 for P_{max} and 1.2×10^{13} electrons/cm^2 for I_{sc}[7]. This is about an order of magni-

tude smaller than that determined from the deterioration in P_{max} and I_{sc}. Several factors should be kept in mind in comparing these fluence estimates. Virtually all the damage in orbit is the result of proton bombardment. The equivalent 1 MeV electron fluences given in Reference 7 are for silicon, not gallium arsenide. The terrestrial measurements of radiation damage[1] used 1 MeV electrons and fluences of 10^{14}, 10^{15} and 10^{16} electrons/cm^2. Experiments were performed at each fluence using temperatures ranging from -20°C to 200°C. There is evidence[8] that annealing of the radiation damage begins to occur at 200°C. Thus, these terrestrial experiments may underestimate the damage occurring at a particular electron fluence.

From the deterioration of silicon cell properties Reference 2 estimates the annual fluence in a 20,000 km circular orbit inclined at 63° to be 2×10^{14} electrons/cm^2 a factor of five higher than that given in Reference 7. While this is a vastly higher orbit than that for LIPS and the radiation consists mainly of electrons, possibly similar discrepancies may exist in the effects of proton radiation for the lower orbit.

The data trends in Figure 2, particularly for the short-circuit current are of some concern. The data drop rapidly with time until day 80 when the satellite is no longer in a position to transmit reliable daytime data. There is an apparent leveling off between days 80 and 120 with the decrease continuing at virtually the same rate after day 120. There is no reason to believe that between days 80 and 120 any known damage mechanism has been suppressed nor that inordinately high temperatures and associated annealing occurred.

It may be argued that the scale on which the data are plotted is quite fine (each major division is only about 2% of the values) and that the apparent trends may still result from scatter in the data. This does

not appear likely for the short-circuit current. Additional data later in satellite life may resolve this difficulty.

IX. RECOMMENDATIONS

Some improvement in the reliability of the analysis presented here could be obtained by improving the analytical tools and the ingredients of those tools. A detailed study of the BOL current data suggests that the temperature coefficient for the current is low. This study indicates that a coefficient of $3.9\ (\pm 0.4) \times 10^{-5}$ amperes/cell/C^{o} is more consistent with BOL data than the value presently in use. Intensity corrections to the voltage are not often made since the logarithmic dependence is deemed not to be significant. However, on the high side of the knee of the I-V curve small voltage variations produce large current changes. The correction can make a difference of a few tenths of volts in the open circuit (and other) voltages similar in magnitude to the observed fluctuations. The suggested intensity correction to voltage derived from our solar cell model and BOL data is

$$\Delta V = -2.05 \times 10^{-4}\ T \ln (H)$$

where ΔV is the voltage correction in volts/cell,

T is the Kelvin temperature and

H is the solar intensity in suns, resulting from the effects of the sun angle and the earth's position in solar orbit.

This correction is in good agreement with BOL experimentally measured values in Reference 1. These corrections are to be made to raw data acquired at the Naval Research Laboratory.

The series resistance has been neglected in the solar cell model used here. This can be included without difficulty if voltage rather than current is made the independent variable. The series resistance is expected to increase with increasing radiation damage.

Little is known of the effects of proton radiation in GaAs cells. Experiments to determine the variation of macroscopic cell properties and microscopic damage which occurs as a result of proton irradiation would be extremely helpful in assessing the performance of GaAs cells in the radiation environment of space.

The LIPS satellite uses 5 silicon photovoltaic arrays. If data are available, the variation in the properties of these Si arrays should also be studied. Radiation damage effects for Si are better known than for GaAs. By comparison the effects in Si could shed light on the character of the radiation to which the GaAs array is subjected and on the reasons for the degradation in the characteristics of the GaAs cells.

REFERENCES

1. Anspaugh, B. E., Downing, R. G., Miyahira, T. F., and Weiss, R. S., "Characterization of Solar Cells for Space Applications: Volume XIV. Electrical Characteristics of Hughes Liquid Phase Epitaxy Gallium Arsenide Solar Cells as a Function of Intensity, Temperature and Irradiation", JPL 78-15, November 15, 1981.

2. Walker, D. H., "Performance of the Solar Cell Experiments Aboard the NTS-2 Satellite After Three Years in Orbit", NRL Memorandum Report 4580, July 30, 1981.

3. Tada, H. Y., Carter, Jr., J. R., Anspaugh, B. E. and Downing, R. G., "Solar Cell Radiation Handbook", Third Edition, JPL 82-69, November 1, 1982, p. 1-23.

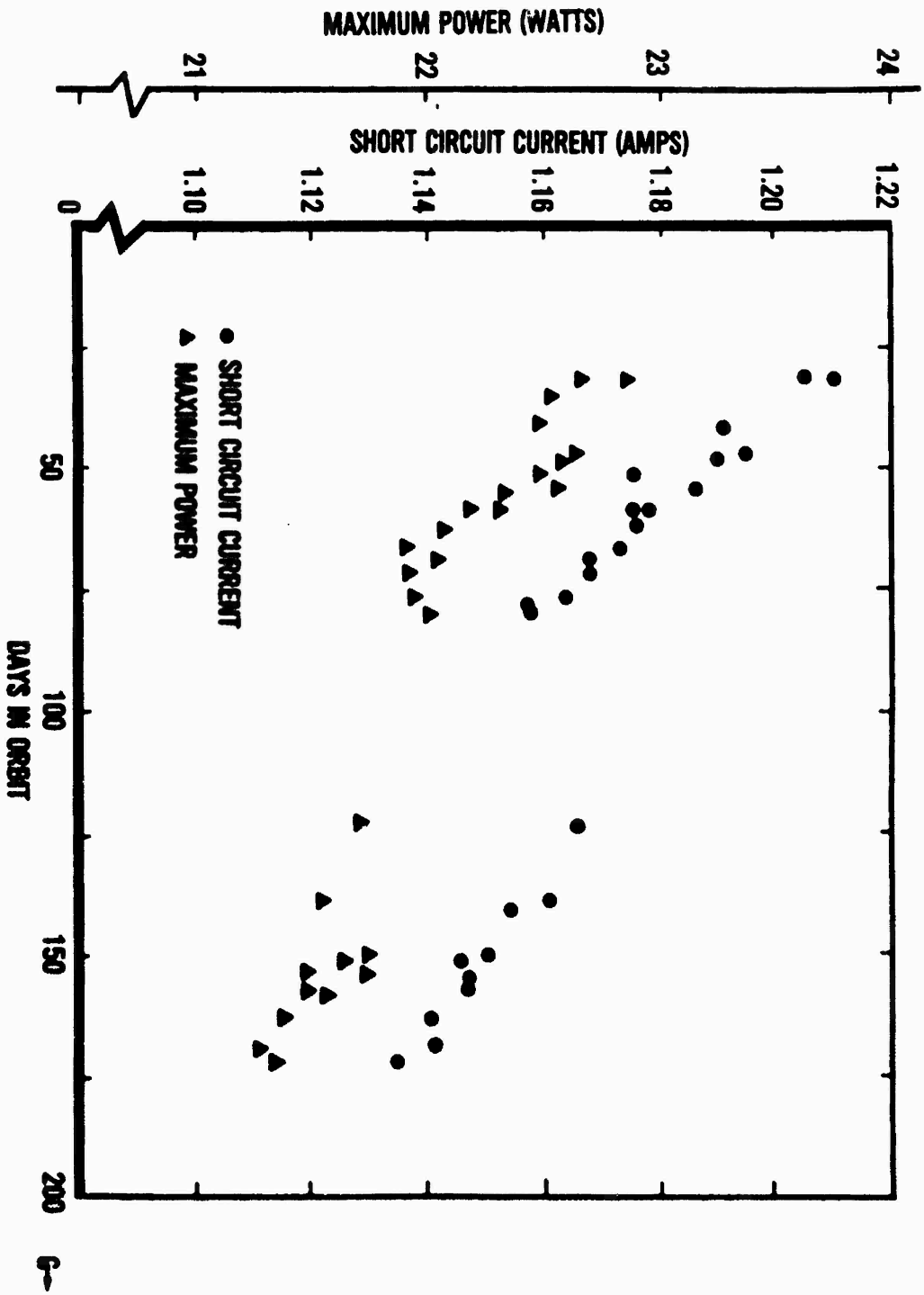
4. Marquardt, D. W., "An Algorithm for Least-Squares Estimation of Non-Linear Parameters", J. Soc. Indust. Appl. Math., 11, 431 (1963).

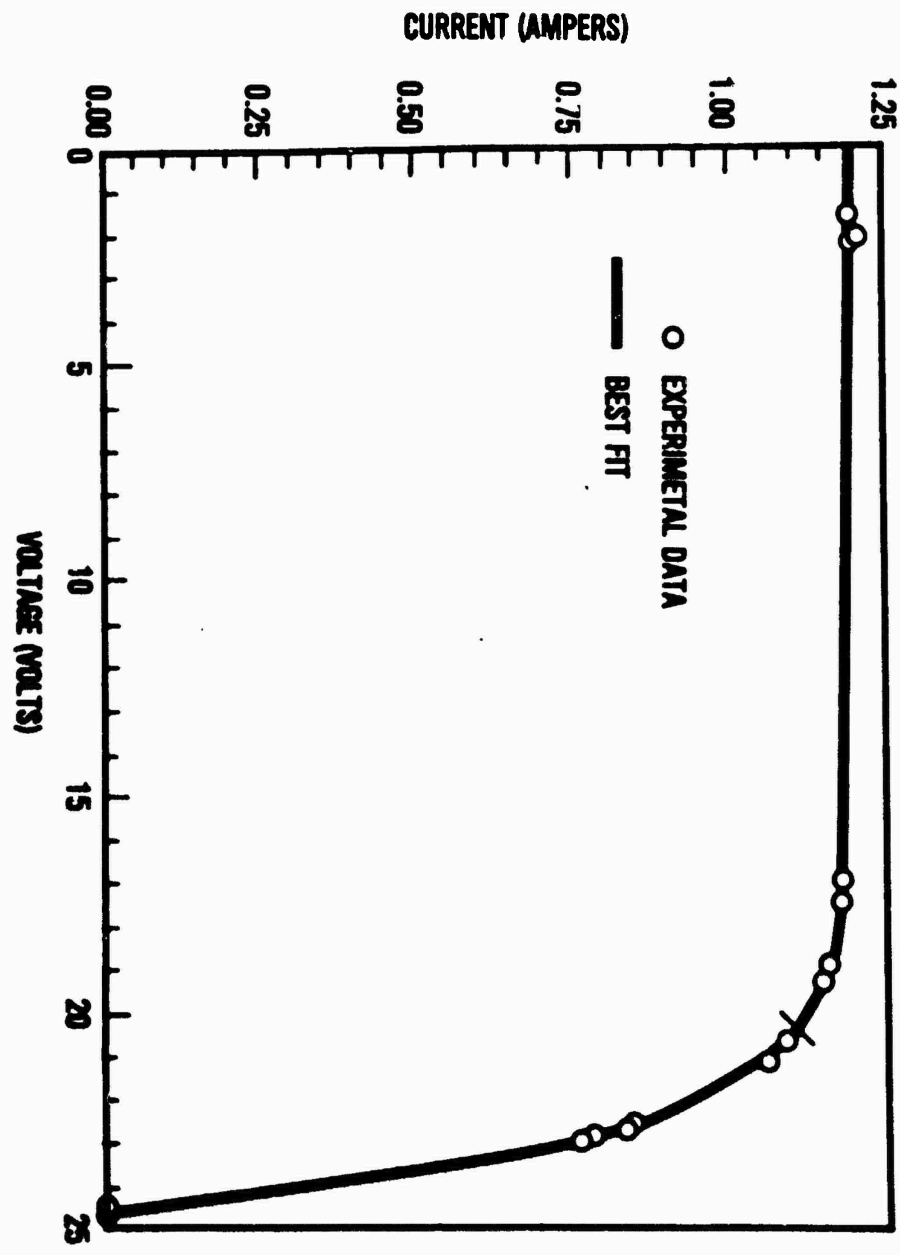
5. Bevington, P. R., "Data Reduction and Error Analysis for the Physical Sciences", McGraw-Hill, New York, 1969.

6. Meulenberg, Jr., A., "Space Environment Damage to Solar Cell Coverslide Assemblies", Proceedings of the Thirteenth Photovoltaic Specialists Conference, pp. 107-115 (1978).

7. Reference 3, pp. 6-37 to 6-42.

8. Loo, R., Kamath, G. S. and Knechtli, R., "Electron Radiation Damage of (AlGa)As-GaAs Solar Cells", Final Report for Contract 955062, Hughes Research Laboratories, October 1979.


MAXIMUM POWER (WATTS)
21
22
23
24
SHORT CIRCUIT CURRENT (AMPS)
1.10
1.12
1.14
1.16
1.18
1.20
1.22
0
50
100
150
200
DAYS IN ORBIT
● SHORT CIRCUIT CURRENT
▲ MAXIMUM POWER


CURRENT (AMPERS)
0.00
0.25
0.50
0.75
1.00
1.25
0
5
10
15
20
25
VOLTAGE (VOLTS)
EXPERIMETAL DATA
BEST FIT

FIGURES

1. Figure 1 - Representative current vs. voltage data for the LIPS gallium arsenide photovoltaic array and least-squares fit.

2. Figure 2 - Short circuit current and maximum power vs. time in orbit for the LIPS gallium arsenide photovoltaic array.

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

Building a Multiple Regression Equation When Many Variables Are Available

Prepared by:	Enoch C. Temple, Ph.D.
Academic Rank:	Associate Professor
Department and University:	Department of Mathematics Alabama A&M University
Research Location:	Leadership and Management Development Center Directorate of Research and Applications Maxwell Air Force Base, AL 36112
USAF Research Colleague:	Capt. Jeffrey S. Austin, Ph.D.
Date:	August 8, 1983
Contract No.	F49620-82-C-0035

Building a Multiple Regression Equation When Many Variables Are Available

by

Enoch C. Temple

ABSTRACT

This paper summarizes some of the problems encountered when a linear model is constructed from a large set of potential regressor variables. Limitations imposed on the multiple regression coefficient by survey data are discussed. The problem of multicol linearity is reviewed. A method of selecting a set of regressor variables and simultaneously identifying multicollinearities is presented through an example.

ACKNOWLEDGEMENTS

I would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for providing this excellent research opportunity. Thanks also are extended to the staff of the Leadership and Management Development Center for the help willingly given during this research period. Special thanks to Captain Jeff Austin and Major Larry Short for suggesting and discussing the topic addressed in this report. Further appreciation is extended to Captains Mark Cochran and Dan Salvino who provided computer assistance and to Ms. Maria Crownover, who typed this manuscript.

I. INTRODUCTION

The research applications staff of the Leadership and Management Development Center (LMDC) at Maxwell Air Force Base, Alabama, is charged with the task of statistical analysis which identifies trends, relations and differences among many of the variables that are connected to the ultimate goals of the United States Air Force (USAF). This task is in direct support of the ultimate goal of the Air University which is to enhance USAF leadership and management effectiveness.

For the purpose of achieving its objectives, LMDC has developed several instruments (survey tools). Among these are the Organization Assessment Package (OAP), Supplemental Survey #1 (Combat Motivation) and Supplemental Survey #2 (Spouse Views). The OAP contains 109 items, Supplemental Survey #1 contains 56 items and Supplemental Survey #2 contains 73 items. For each survey tool, each nondemographic response item is limited to a choice of one out of seven possible. LMDC has a large data base on each survey tool, the largest being on the OAP.

In order for the research staff to make sound recommendations through an appropriate analysis of the available data base, it is necessary to identify and analyze relationships among the variables on which data has been collected. Specifically, there is an interest in the development of a best possible linear model. Such a model should contain the variables that have the greatest influence on a particular target variable. Also, when assessing this influence, comments in this paper are expected to serve as a reminder of limitations imposed on the statistical evaluation of the model by the data set. Additionally, one should be aware of precautions that should be taken in order to ensure that standard statistical procedures used to evaluate the model are not misused or misinterpreted. For example, one of the criteria used to evaluate the quality of a prediction equation is the value of the multiple regression coefficient R or its square, R^2. It is known that the R^2 value is inflated when one attempts to apply a stepwise variable selection procedure to a regression problem where the sample size is

smaller than the number of potential variables. Rencher and Pun (1980) observed this property and also evaluated some distribution properties of R^2.

It is also known that data collected by survey tools tend to have a large number of repeats. (i.e., one variable constant while the other variable varies.) These repeats impose restrictions on the value of R^2. Thus any large R^2 value obtained from a data set with a large number of repeats should be suspect.

At this stage of the linear model construction at LMDC, researchers have identified several dependent variables for which they would like to identify good predictor variables. For each of the dependent variables, they also have a good idea as to which variables may be good predictors. However, when the number of variables is large, one always wonders if perhaps some important variable has been left out. Furthermore, often times a set of regressor variables may be highly intercorrelated (a situation known as multicollinearity). Even if pairwise correlations are not large, multicollinearity may still exist among the variables.

When multicollinearity is present, the classical method of multiple regression is unreliable. That is, regression coefficients are unstable from sample to sample. For example, one sample may generate a regression coefficient whose value is negative and yet another sample may produce a positive value for that same coefficient. Any model exhibiting this behavior is useless. In fact, it is highly possible that the presence and ill-effects of multicollinearity (if not checked) may go completely unnoticed. Hence, as an intermediate step for the construction of a linear model from a large set of potential variables, it is recommended that exploratory tools be applied to a preliminary sample. The application of such tools allows the researcher to select predictor variables and simultaneously identify multicollinearities.

Finally, a few words of caution is in order here. First, exploratory tools should not be allowed to replace a researcher's insight nor any well established statistical tools. Instead, they are to serve as an auxiliary to existing tools and the researcher's

knowledge. As pointed out by Hawkins (1973) and Baskerville and Toogood (1982), an exploratory method will expose the underlying structure of all variables involved in an investigation, thus, permitting the construction of a best possible linear model. Secondly, since survey research never completely satisfies the assumptions of the linear model, statistical inference should be interpreted with care. Nonetheless, the study of linear relationships, even for survey data, is a very effective way of allowing the researcher to understand patterns and trends in the data.

II. OBJECTIVES

Procedures for estimating the regression coefficients in a multiple regression problem are well established. As pointed out in the introduction, sample size, repeats in the data and multicollinearity may inhibit the effectiveness of the application of established regression procedures. Therefore, specific objectives for this project are to:

(1) Identify the best technique, based on the available data, for constructing a multiple regression model.

(2) Provide theoretical support for the technique recommended in objective (1).

(3) Recommend a variable selection technique.

III. THE GENERAL LINEAR MODEL

Basic Properties of a Linear Model

For the purpose of conciseness, completeness and convenience, the basic properties of the linear model will be presented through the use of vectors and matrices. Furthermore, a global picture of multicollinearity can be better presented through the matrix representation than through any alternate approach.

To develop the essential properties of the linear model, we let y be the dependent variable and x_1, x_2, ..., x_p be the independent variables. (Independent variables, regressor variables and predictors will be used interchangeably.) It is assumed that a linear relationship exists between the variable y and the variables x_1, x_2 ...., x_p. The mathematical form of this assumption is

$$y = b_o + b_1x_1 + \ldots + b_px_p \qquad (1)$$

where b_o is the constant term and b_i $(i>o)$ is the numerical coefficient of x_i. At the outset of the construction, the values of b_o, b_1, ..., b_p are unknown and must be estimated from sample data.

Let us also assume that there are n observations on the variables y, x_1, x_2, ...,x_p. Substituting these n observations into the assumed model in (1), one gets

$$\begin{array}{ll} \text{observation 1} & y_1 = b_o + b_1x_{11} + \ldots + b_px_{1p} \\ \text{observation 2} & y_2 = b_o + b_1x_{21} + \ldots + b_px_{1p} \\ \cdot & \cdot \quad \cdot \quad \cdot \qquad\qquad \cdot \\ \text{observation n} & y_n = b_o + b_1x_{n1} + \ldots + b_px_{np} \end{array} \qquad (2)$$

An alternate matrix/vector representation of the above observations is

$$\begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} \cdots x_{1p} & b_o \\ 1 & x_{21} \cdots x_{2p} & b_1 \\ & & \cdot \\ & \cdots & \cdot \\ & & \cdot \\ 1 & x_{np} \cdots x_{np} & b_p \end{bmatrix} \qquad (3)$$

Where $\underline{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{bmatrix}$ is the nx1 vector of observations

on the dependent variable y,

$X = \begin{bmatrix} 1 & x_{11} \cdots x_{1p} \\ 1 & x_{21} \cdots x_{2p} \\ & \cdots \\ 1 & x_{n1} \cdots x_{np} \end{bmatrix}$ is the nxp matrix of

observations on the independent variables x_1, x_2, ..., x_p and

$$\underline{B} = \begin{bmatrix} b_o \\ b_1 \\ . \\ . \\ . \\ b_p \end{bmatrix}$$ is the vector of

coefficients to be estimated from the observed data. When $\underline{Y}$, X and $\underline{B}$ are substituted into equation (3), it becomes

$$\underline{Y} = X\underline{B} \tag{4}$$

For the purpose of any inferential statistical analysis, the set-up in (1) is represented by

$$y_i = b_o + b_1 x_{1i} + \cdots b_p x_{pi} + E_i \tag{5}$$

where E_i represents the random error in the ith observation and hence the corresponding set-up in (4) becomes

$$\underline{Y} = X\underline{B} + \underline{E} \tag{6}$$

where $\underline{E} = \begin{bmatrix} E_1 \\ E_2 \\ . \\ . \\ . \\ E_n \end{bmatrix}$ is an nx1 error

vector. When (6) is used, it is also often assumed that the variance of $\underline{E}$ is

$$\text{var}(\underline{E}) = \sigma^2 I \tag{7}$$

where I is an nxn identity matrix, $\underline{E}$ has the multivariate normal distribution and σ^2 is positive. However, these additional assumptions are necessary only if we are interested in the development of statistical distributions. This paper does not focus in that direction.

It is also worthwhile to remark here that many users will standardize the values in both vector $\underline{Y}$ and matrix X. Under standardization, the column of ones in matrix X will disappear. The column of ones disappears because under standardization, the constant term b_o in (1) disappears. Nevertheless, whether the observations are standardized or not, the basic set-up in (4) or (6)

remains the same. That is, if the observations are not standardized, $\underline{Y}$ and matrix X contain the original observations. If the observations are standardized, the vector $\underline{Y}$ and matrix X contain the standardized observations with vector $\underline{B}$ being $\underline{B}^1 = [b_1, b_2, \ldots, b_p]$ in case of standardization. ($\underline{B}^1$ represents the transpose of vector $\underline{B}$).

For the actual computation of the values in vector $\underline{B}$, we refer to equation (5) or (6). From (6), we see that $\underline{E} = \underline{Y} - X\underline{B}$ and

$$\underline{E}^1\underline{E} = (\underline{Y} - X\underline{B})^1 (\underline{Y} - X\underline{B}) \tag{8}$$

which is the sum of the squared error in the dependent variable y. The computed values of $\underline{B}$ are the values that will minimize (8). The application of partial differentiation to (8) [See Graybill, (1976) for theoretical details] shows that (8) is minimized when $\underline{B} = (X^1X)^{-1} X^1Y$. For the purpose of clarity, we call $\underline{B}$ the unknown values of the coefficients and $\hat{\underline{B}}$ the estimate of $\underline{B}$. Therefore, $\hat{\underline{B}}$ is the actual vector of values computed from the observed values in vector $\underline{Y}$ and matrix X.

Hence,
$$\hat{\underline{B}} = (X^1X)^{-1} X^1Y \tag{9}$$

Additional expressions in the matrix notation are as follows: [See Graybill (1976) for details.]

Variance of the components of $\underline{B} = (X^1X)^{-1}\sigma^2$ (10)

Vector of estimated values of $y = \hat{\underline{Y}} = X\hat{\underline{B}}$ (11)

Sum of squares due to regression $= \hat{\underline{B}}^1X^1\underline{Y}$ (12)

Total sum of squares $= \underline{Y}^1\underline{Y}$

Residual sum of squares $= \underline{Y}^1\underline{Y} - \hat{\underline{B}}^1X^1\underline{Y}$ (13)

Expressions (9) throughout (13) are included for easy reference in later developments.

Objectives of the Linear Model

The general linear model and its theory are used to justify a large class of statistical tools. For example, tools of analysis of variance, covariance analysis, path analysis, among others are justified through the theory of the linear model. Although the linear model can be applied to a large class of problems, Draper and Smith (1981, page 412) have identified three major types of mathema-

tical models. They are:

(1) The functional model

(2) The control model

(3) The predictor model

Of these three, Draper and Smith raised doubt as to whether models (1) and (2) can be achieved through regression alone. Nevertheless, we will identify the above as three major objectives of the linear model. In fact, researchers at LMDC have expressed interest in the best possible model so that testable hypotheses can be formulated on the degree of control one may exercise over the dependent variable by manipulating the independent variable(s).

Influence of Repeat Runs on R^2.

To evaluate the influence of repeat observations on R^2, we use an approach presented by Draper and Smith (1981). Suppose we are interested in establishing a simple regression equation of y on x. Let there be n observations on variables y and x. Of the n observations on x, suppose that only m (<n) of them are distinct. That is, for the distinct x value

x_1, there are n_1 observations on y;

x_2, there are n_2 observations on y;

x_j, there are n_j observations on y;

x_m, there are n_m observations on y.

In general, for each distinct x_j, there are n_j different observations on y and hence we say that the data set contains repeats at x_j and $n_j = n$.

Let y_{jk} = repeat observations at x_j where $k = 1, 2, \ldots, n_j$. The residual at any x_j is $y_{jk} - \hat{y}_j$ where $\hat{y}_j$ is the estimated value of y at x_j and y_{jk} is the observed value. By adding $\bar{y}_j - \bar{y}_j$ the residual at x_j becomes

$$y_{jk} - \hat{y}_j = (y_{jk} - \bar{y}_j) - (\hat{y}_j - \bar{y}_j), \qquad (14)$$

where $\bar{y}_j$ is the mean of the repeats at x_j.

Squaring both sides of (14) and applying the appropriate summations, we get

$$\sum\sum (y_{jk} - \hat{y}_j)^2 = \sum\sum(y_{jk} - \bar{y}_j)^2 + \sum n_j(\hat{y}_j - \bar{y}_j)^2 \qquad (15)$$

The left side of (15) is the residual sum of squares and the two terms on the right side partition the residual sum of squares into pure error sum of squares (PSSE) and a lack of fit sum of squares (LFSS). That is, PSSE = $\sum\sum(y_{ik} - \bar{y}_j)^2$ and LFSS = $\sum n_j (\hat{y}_j - \bar{y}_j)^2$ is the error for which no model, be it linear or otherwise, can account. Therefore, when repeats exist in a data set, the maximum value of $R^2 = 1$ cannot be attained. That is,

$$R^2 = 1 - \frac{PSSE + LFSS}{TSS} \tag{16}$$

where TSS = total sum of squares.

If PSSE $\neq 0$, no matter the value of LFSS, $R^2 < 1$. Hence an R^2 of 1 cannot be achieved if the data set contains repeats.

In a regression analysis where repeats appear, it may be helpful to know the value of the maximum attainable R^2 value, denoted max R^2, for a particular data set.

To compute this max R^2 value, recall that PSSE remains unchanged no matter what model is fitted to the data. Therefore,

$$\text{Max } R^2 = \frac{TSS - PSSE}{TSS} . \tag{17}$$

It should be kept in mind that the max R^2 value is data dependent. If there are no repeats, PSSE = 0 and max $R^2 = 1$. However, for the LMDC data, repeats are almost assured after the seventh observation on any pair of variables.

Since the R^2 value of (16) was computed for a simple regression, the absolute value of the correlation for the x and y variables is $(R^2)^{1/2}$. Hence the absolute value of the maximum achievable correlation between two variables is $(\text{max } R^2)^{1/2}$ where max R^2 is given in (17).

The reader may have observed that the partition in (15) is somewhat similar to the partition for the total sum of squares in a one-way analysis of variance. Actually, if one is interested in computing max R^2 for a simple regression of y on x where x has m distinct values and repeats are available, variable x can be thought of as a single factor with m levels. (For LMDC data $m \leq 7$.) Thus the analysis of variance set-up gives PSSE as the within cell sum of

squares. Therefore, from a one-way ANOVA table, max R^2 is easily computed.

In a similar manner, max R^2 for the multiple regression problem can be computed. For example, suppose one is interested in computing a max R^2 for the multiple linear regression of y on variables x_1, x_2 and x_3.

The values TSS and PSSE are easily computed by a computer statistics package, say SPSS, by using a three-way ANOVA where x_1, x_2 and x_3 are the factors and each of the factors has m levels.

Here max $R^2 = \frac{TSS - PSSE}{TSS}$ where PSSE is the pooled within cell sum of squares.

It is worthwhile to remark here that the ratio (actual value of R^2)/(max R^2) gives some measure of how well the model fits when compared to its potential. Max R^2 is its max potential.

The Problem of Multicollinearity.

The symptoms of multicollinearity may take many forms. Some are listed in Nie et.al (1975). However, one may use a systematic approach to determine the existence of collinearity in a set of variables.

The first look at the concept of multicollinearity will be presented through the use of a four independent and one dependent variable regression problem. That is, consider the linear equation:

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4. \qquad (18)$$

Furthermore, suppose it is known that $x_2 + 2x_3 + .5x_4 = 0$. In this case we say that there is an exact collinearity involving the independent variables x_2, x_3 and x_4 and hence one variable can be expressed as a linear combination of the other two. This means that at least one of variables x_2, x_3 or x_4 can be eliminated.

When exact collinearity exists, from a mathematical point of view, it makes no difference which variable is removed. However, the researcher may have a preference as to which variable should be removed. For example, since it is known that $x_2 + 2x_3 + .5x_4 = 0$, we get $x_2 = -2x_3 - .5x_4$.

Hence the equal for x_2 can replace x_2 in equation (18). After

substitution and simplification, equation (18) becomes $y = c_0 + c_1X_1 + c_3X_3 + c_4X_4$ where c_0, c_1, c_3 and c_4 are the new coefficients obtained by the substitution process. The reader should observe that either variable x_3 or x_4 also could have been eliminated.

In most applied problems, it is unusual for exact multicollinearity to exist. That is, instead of exact collinearity, a case of "near collinearity" exists. From the example in (18), we may say that a near collinearity exists between variables x_2, x_3 and x_4 if, say, $x_2 + 2x_3 + .5x_4$ is near zero. When it is determined that this linear combination is near zero, we may say that collinearity exists and at least one variable can be eliminated by equating the linear combination to zero and proceeding as though an actual exact collinearity exists. Here again, mathematically, the elimination process is straight forward but the actual identification of a collinearity may not be clear-cut.

The mechanics of identifying variables involved in a multicollinearity can be more compactly presented through matrices. This is particularly true when a large number of potential regressor variables are involved. This treatment is provided in the next section.

How to Detect Multicollinearity

A researcher using regression analysis needs methods of determining when multicollinearity is present. By using the matrix notation given in III.1, we recall that $(X^1X)\underline{B} = X^1\underline{Y}$ or $B = (X^1X)^{-1}X^1\underline{Y}$. Also recall that if the original data set is standardized, X^1X is the correlation matrix for the independent variables. For the sequel that follows, it is assumed that the data is standardized.

To determine if multicollinearity is present, one may examine the diagonal of $(X^1X)^{-1}$. If there is a large element on the diagonal of $(X^1X)^{-1}$ then this is an indication that the associated variable is involved in a multicollinearity. Alternately, matrix X^1X can be factored such that

$$X^1X = T^1 \Lambda T \tag{19}$$

where T is a matrix of eigenvectors, Λ is a diagonal matrix con-

taining the eigenvalues of X^1X along its diagonal and T^1 denotes the transpose of matrix T. Let vector $\underline{t}_i$ be the ith eigenvector of X^1X (ith column of T) and let λ_i be the associated eigenvalue. According to Massey (1965), the linear combination $\underline{t}_i\underline{x}$ where $\underline{x}$ is some vector of the independent variables, represents a multicollinearity if λ_i is small. (Criteria for judging smallness will be given in section IV).

Another alternate indicator of the presence of multicollinearity is the determinant of matrix X^1X, denoted det (X^1X). If det (X^1X) is small, then there is a possibility of collinearity. However, the reader should be reminded that in equation (9), it is implicitly assumed that $n > p$, where n is the number of observations and p is number of variables. If $n < p$, then det (X^1X) is exactly zero and hence $(X^1X)^{-1}$ does not exist. Thus, when attempting to estimate regression coefficients, we should be certain that the number of observations is sufficiently large. Anderson (1958) has recommended 3 to 4 observations per variable when attempting an initial investigation of a linear relationship. For example, if ten variables are involved in the investigation, then 30 to 40 observations should be used.

Solutions for Multicollinearity

Several solutions for the problem of multicollinearity have been proposed. Among them are ridge regression introduced by Hoerl and Kennard (1970), principal components regression discussed by Massey (1965) and latent root regression proposed by Webster et.al (1973). All three of the above methods rely on an alteration of the eigenvalues of the matrix (X^1X) to achieve their results. These methods carry the general label of biased regression. For specifics, the reader may consult the above refrences. Also a paper by Hocking (1976) gives an excellent overview of the methods mentioned here as well as of two additional methods. The theoretical justification for the recommended approach suggested by this paper is given in the next section. This approach allows the researcher to select good regressor variables and simultaneously identify collinearities.

IV. RECOMMENDED METHOD

The recommended method here follows the method proposed by Hawkins (1973). The set-up consists of a vector $\underline{x}^1 = (x_1, x_2, \ldots, x_p)$ of independent variables and the dependent variable y. Augment vector $\underline{x}$ by adding variable y to obtain vector $Z^1 = (y, x_1, \ldots, x_p)$ and assume that all variables in vector $\underline{z}$ are standardized. Let R be the (p + 1) by (p + 1) correlation matrix of $\underline{Z}$ where R is assumed to be nonsingular.

The problem of multiple regression consists of finding a vector so that the hyperplane

$$y - b_1x_1 - b_2x_2 - \ldots - b_px_p \tag{20}$$

is close to the data in the y-norm. That is, so that the mean squared deviation along the y-axis is minimized. Denote the y-norm by s^2 and denote the vertical norm (mean squared deviation along a normal to the hyperplane) by λ . From vector analysis it is known that if $\underline{Z} = (y, x_1, \ldots, x_p)$ is any vector in the (p + 1) dimensional space, then

$$\frac{y - b_1x_1 - \ldots - b_px_p}{1} = \lambda^{1/2} \tag{21}$$

where $1 = (1 + b_1^2 + \ldots + b_p^2)^{1/2}$ is the vertical distance between $\underline{Z}$ and the hyperplane in (20). An application of trigonometry gives the equation

$$s^2 = 1^2 \quad . \tag{22}$$

Rewriting the left hand side of equation (21) we get

$$\frac{y}{1} - \frac{b_1x_1}{1} - \ldots - \frac{b_p}{1}x_p. \tag{23}$$

Now for each hyperplane $\underline{t_iZ}$ determined by the eigenvector $\underline{t_i}$ of matrix R, $_i$ is the corresponding vertical norm. Simply stated, $_i$ is the ith eigenvalue of R.

An expansion of $\underline{t^1{}_i Z}$ gives

$$\underline{t_iZ} = t_oy + t_1x_1 + \ldots + t_px_p \tag{24}$$

Equating coefficients of y in (23) and (24) and solving for 1 we get

$$1 = \frac{1}{t_o}.$$

Hence we now have a situation where 1 and λ are known for equation (22). By substituting $\frac{1}{t_o}$ for 1 and λi for λ in equation (22), we get

$$s^2 = \frac{\lambda i}{t_o^2}$$

Since the original objective was to minimize s^2, it is now clear that s^2 is minimized when λi is small and t_o is large. The reader should note that $|t_o| \leq 1$ because t_o is the first component of an eigenvector of R. Also note that S^2 is small if $\frac{t_o^2}{\lambda i}$ is large.

A convenient representation of the ratio $\frac{t_o^2}{\lambda i}$ can be obtained by rescaling the matrix of eigenvectors. That is, by rescaling matrix T where $R = T^1 \Lambda T$, T contains the eigenvector of R in its columns and is the eigenvalue diagonal matrix. Let matrix $D = \Lambda^{-1/2} T^1$. Here $d_{ij} = \frac{c_{ij}}{\lambda 1^{1/2}}$ where t_{ij} is the jth element of the ith eigenvector. The D matrix provides guidelines for determining which variables are likely to be good predictors of the dependent vairable y. First note two results from Hawkins. Let $\underline{d}^1_i = (d_{io}, d_{i1}, \ldots, d_{ip})$ to be the ith row of matrix D. Then

(1) $\frac{1}{\sum d_{io}^2}$ is the mean squared error for the regression of all independent variables of y.

(2) $\frac{1}{d_i^2}$ is the mean squared error of a regression using the variables provided by row i of matrix D.

(Some elements in row i may be zeros. Thus, this row shows a subset of variables that are suitable regressors.)

By using result (2), it can be seen that the best subset of variables for predicting y is indicated by the row of matrix D having

the largest value in column zero (the 1st column). The best subset of regressor variables will be those independent variables corresponding to the largest d-values in that row.

If there are two or more such rows in matrix D, then this is an indication that there are two or more clusters of variables that are good predictors of y. Furthermore, if d_{io} is near zero, then the independent variables loading most heavily on this row are involved in a multicollinearity. The following example is illustrative of the above description.

V. EXAMPLE

A sample run of the subprogram named FACTOR from the SPSS package is shown below. The objective of the run was to obtain the D matrix for the 93 nondemographic variables in the OAP. The variables are denoted by V_1, V_2, ... V_{93} and variable V_{93} is the dependent variable. A sample of size 116 was used to determine the correlation matrix. (This sample size is small but sufficient for illustrative purposes.) Because of a limitation on space, the 93 by 93 correlation matrix is not printed here. Also, only an abbreviation of certain rows of the rotated D matrix is printed in table 1. Row 1 contains a partial list of the 93 variables with variable 93 being the dependent variable. Rows 2, 3 and 4 provide some information about the relationship between the variables. Row 2 is printed because, of all 93 rows of matrix D, it contains the largest value in the column associated with variable 93 and row 3 contains the second largest value in that column. Row 4 contains a small value (rounded to zero) in the first column.

TABLE 1

93	3	11	32	38	42	59	61	64	66	72	80	85	86	91
3.9	-.3	-.1	0	.2	-.8	-.3	-.7	.8	-.4	.2	.9	-.4	1.0	.9
-1.3	-1.3	4.3	-1.4	-3.0	1.2	.6	.8	-.3	-.2	-.6	-.6	1.4	.2	0
0	-.1	-.2	.5	-.1	.1	-1.1	1.0	0	-1.2	3.7	.8	-.5	.4	.3

From row 2, variables 42, 64, 80, 86 and 91 have the largest d-values for this row. Hence they represent a cluster of five good

predictors for V_{93}. Row 3 indicates that variables 3, 11, 32, 38 and 85 is a second cluster of good predictors for V_{93}. Since row 4 contains a very small value in the first column, the regressor variables corresponding to the largest d-values in this row are involved in a collinearity. That is, variables 59, 61, 66, 72 and 80 are near collinear. However, the collinearity as defined by row 4 does not affect the regressor variables identified in rows 2 and 3 because the identified regressor variables there do not contain any of the collinear variables. If the collinear variables were involved with the identified regressor variables, it would be possible to delete the collinearity by using the deletion method given in section II.

In general, from a set of 93 variables, it is expected that more than five Regressor variable will be identified. No matter what number of variables are identified from matrix D, this paper recommends that a stepwise regression then applied to the identified variables to get the actual regression equation.

VI. COMMENTS AND CONCLUSIONS

In this paper we have presented a method for constructing a linear prediction equation when a large number of variables is available. Emphasis was placed on the limitations imposed on simple correlation coefficients and the R^2 value because of the likelihood for repeat observations. Also, it was suggested that regressor variables be identified through an inspection of a D matrix. This inspection allows the user to identify one or more subsets of regressor variables as well as any collinearities that may exist.

It is worthwhile to point out that if the researcher is interested in treating any variable of $\underline{Z} = (y, x_1, \ldots, x_p)$ as a dependent variable, a set of good regressor variables can be obtained by inspecting matrix D. For example, suppose we are interested in treating x_k as a dependent variable to be regressed on variables $y, x_1, \ldots x_{k-1}, x_{k+1}, \ldots x_p$. We first locate the column in D that is associated with x_k and the largest value in this column identifies the row of D that will identify the variables that are

good predictors of x_k. Thus, matrix D may allow the researcher the versatility of detecting good predictors of any variable in $\underline{Z}$ through a single preliminary analysis.

Matrix D can be obtained in its rotated form by using the principal factoring method followed by a varimax rotation. These methods are available in the SPSS subprogram FACTOR.

Further applications of exploratory tools to the various types of linear models are possible. Also, the influence of an oblique rotation, with respect to regression, on the D matrix should be investigated. Because of the necessary shortness of this research period, these problems were not addressed. These are good candidates for a follow-on research project.

References

Anderson, T.W. (1958, An Introduction to Multivariate Statistical Analysis, New York: John Wiley and Sons, Inc.

Baskerville, J.C. and Toogood, J.H. (1982) "Guided Regression Modeling for Prediction and Exploration of Structure With Many Explanatory Variables," Technometrics, 24, pp. 9-17.

Darling, J.R. and Tamura, H. (1970) "Use of Orthogonal Factors for Selection of Variables in a Regression Equation - An Illustration," Applied Statistics, 19, pp. 260-268.

Graybill, F.A. (1976), Theory and Application of the Linear Model, Massachusetts: Duxbury Press.

Hawkins, D.M. (1973), "On the Investigation of Alternative Regressions by Principal Component Analysis," Applied Statistics, 22, 275-286.

Hocking, R.R. (1976), "The Analysis and Selection of Variables in Linear Regression," Biometrics, 32, 1-47.

Hoerl, A.E., and Kennard, R.W. (1970), "Ridge Regression: Biased Estimation for Non-orthogonal Problem," Technometrics, 12, 55-68.

Massey, W.F. (1965), "Principal Components Regression in Exploratory Statistical Research," Journal of the American Statistical Association, 60, 234-257.

Nie, N.H. et al (1975), Statistical Package for the Social Sciences. New York: McGraw Hill.

Rencher, A.C. and Pun, F.C. (1980) "Inflation of R^2 in Best Subset Regression," Technometrics, 22, pp. 49-53.

Webster, J.T., Gunst, R.R., and Mason, R.L. (1973), "Recent Developments in Stepwise Regression Procedure," Proceeding of the University of Kentucky Conference on Regression with a Large Number of Predictor Variables, 24-53.

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

A PRELIMINARY INVESTIGATION OF THE UTILITY OF LINEAR DIGITAL FILTERS FOR ANALYZING ECONOMIC SYSTEM PERFORMANCE DATA

Prepared by:	William Robert Terry, PhD
Academic Rank:	Associate Professor
Department and University:	Department of Industrial Engineering University of Toledo
Research Location:	Business Research Management Center, Wright Patterson AFB
USAF Research	Capt. Michael Tankersley
Date:	September 9, 1983
Contract No:	F49620-82-C-0035

A PRELIMINARY INVESTIGATION OF THE UTILITY OF LINEAR DIGITAL FILTERS FOR ANALYZING ECONOMIC SYSTEM PERFORMANCE DATA

by

William Robert Terry, Ph.D.

ABSTRACT

The utility of linear digital filter (LDF) methods for analyzing performance data for economic systems was investigated. LDF methods typically used for analyzing physical systems were found to be not well suited for analyzing economic systems. Alternative methods for analyzing economic systems were identified. Data series from a number of economic systems of interest to the Air Force were identified and LDF models were fitted to each series and selected pairs of series. Analysis of the models fitted to the individual series indicated that the assumption of independent errors was not usually appropriate and that LDF methods could be used to account for such dependencies. Analysis of the simultaneous behavior of selected pairs of series indicated that feedback relationships could exist between certain pairs of series and that multivariate LDF models could be appropriate. The analysis of attempts to model certain pairs of series also indicated that a non-linear digital filter method could be needed.

ACKNOWLEDGMENTS

The author would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for providing the author with the opportunity to spend a very worthwhile and intellectually stimulating summer at the Air Force Business Research Management Center, Wright Patterson AFB, Ohio. He would also like to acknowledge the special assistance that he received from a number of persons: Major Tom Fitzgerald for invaluable assistance in helping the author gain access to the General Dynamic F-16 TECH MOD data; Major Lyle Lockwood for discussions which helped the author gain a clearer and more comprehensive understanding of the acquisition research process; Dr. Dave Lee for his helpful suggestions for locating archival performance data; Messers Al Reese, Joe Gertge and Tom Spitler for their assistance in helping the author gain a better understanding of the AFLC productivity measurement problem and for helping him locate historical productivity data; Capt. Don Brechtel, PhD and Lt. Col. Tom Clark, PhD for insightful discussions and references in the area of Systems Dynamics; and Lt. Col. Dan Robinson for stimulating discussion regarding problems associated with measuring the impact of a productivity improvement program; and Col. Ron Deep, PhD for his help in obtaining computer resources and locating data resources and for his support and encouragement. Special thanks are due to Capt. Mike Tankersley for his help in locating data resources, obtaining reference materials, and gaining access to computer resources and especially for the numerous discussions which helped the author gain a clearer understanding of the acquisition research problems. Finally the author would like to thank his graduate assistant, Mr. Norman Zaenglein, for his loyalty and untiring efforts on behalf of the project.

I. INTRODUCTION

In this report any organization which contributes to the production of goods or services will be referred to as an economic system. Two different types of feedback can be used in controlling such a system. Negative feedback (NF) assumes that a desired level of performance has been established and that the goal is to maintain performance at this level. NF involves comparing the actual performance with that desired and then adjusting the inputs to the system so as to eliminate the difference between actual and desired performance. In contrast positive feedback (PF) assumes that the goal is to improve the performance of the system. PF involves changing the inputs to the system on an experimental basis, observing the impact on performance, retaining those changes which improve performance and discarding the remainder.

Economic systems are almost invariably found to be operating in environments which complicate the process of exercising either NF or PF control. In such environments it is not usually possible to observe and control all of the causal factors which can influence the performance of the system. These unobservable or uncontrollable causal factors can cause the actual performance of the system to fluctuate in an unexplained fashion. In such environments there are usually a number of factors which introduce errors into system performance measurements. Both unexplained fluctuations in actual system performance and errors made in measuring the performance of a system are referred to as noise.

The presence of such noise can increase the risk that an inappropriate control action will be taken. When the system is operating in the NF mode, noise can cause a system which in actuality does not require corrective action to look like one which does. It can also cause a

system which in actuality needs corrective action to look like one which does not. Also when the system is operating in a FF mode, noise can cause a change which in actuality had a positive impact on performance to look like one which had a negative or neutral impact. Conversely, it can also cause a change which in actuality produced a negative impact to look like one which produced a positive impact.

In many situations existing knowledge of system behavior is not sufficient for specifying how changes in system inputs will affect system performance. In such situations it is often necessary to use historical performance measurement data to infer a relationship between input and performance. Unfortunately, the presence of noise can obscure the relationship between inputs and performance. In particular noise can create the illusion of a relationship between inputs and performance which is quite different from the one which exists in reality. Using such an illusionary relationship as the basis for control actions could produce results which are opposite to those which are desired and expected.

The risk of taking an inappropriate control action could be reduced if there were a method for extracting the signal which describes the true behavior of a system from the noise which tends to obscure it. This problem is similar in some respects to one which arises in the control of physical systems. In such situations linear digital filter (LDF) methods are often used to separate the signal from the noise.

However, LDF methods typically used in controlling physical systems are not directly applicable to economic systems. In applying LDF methods to physical systems, the first step is to utilize existing theoretical knowledge to specify the general form of a mathematical model which describes the behavior of the physical system. Unfortunately

the equivalent theoretical knowledge for an economic system is not generally available.

Conventional statistical methods such as linear regression analysis are based on the assumption of independent errors. Unfortunately, there are a number of unobservable and uncontrollable factors which could cause an economic system to behave in such a manner that the errors which would result from fitting a linear regression model would not be independent. Examples of such factors include the learning effect, the deterioration of equipment, seasonal effects, and alternating relationships in which a high value in one period tends to be followed by a low value in the next and vice versa.

II. OBJECTIVES OF THE RESEARCH EFFORT

The primary objective of this research was to determine the utility of LDF methods for analyzing performance data for economic systems. This objective was decomposed into three subobjectives which represented specific tasks necessary for the accomplishment of the primary objective. The first subobjective was to identify a LDF method which could be suitable for analyzing performance data for an economic system. The second subobjective was to identify potential applications of LDF's which could be of value to the Air Force. The third subobjective was to actually use LDF methods to analyze performance data for selected economic systems.

III. IDENTIFICATION OF SUITABLE LDF METHOD

The task of identifying a suitable LDF method which could be used for analyzing performance data for an economic system began with a review of the engineering literature in which LDF methods had been used to analyze the behavior of physical systems. This review revealed that the basic approach for developing an LDF model of a physical system involved two phases. The purpose of the

first phase was to specify the general form of a mathematical model for describing the behavior of the physical system. In the second phase the purpose was to estimate the unknown values of the parameters of the mathematical model specified in the first phase. The basic process for identifying the general form of the mathematical model consisted of using a knowledge of the basic relationships between the various components in the system and the laws of physics to specify a system of equations for describing the behavior of the physical system. In the second phase the basic process was to use historical data on the systems performance to estimate the values of the unknown parameters in the mathematical model specified in the first phase.

The utility of an LDF model of any system will be crucially dependent on the adequacy of the general form of the mathematical model of the system's behavior. Failure to specify an adequate form for this model could result in an LDF model which is not capable of adequately fitting historical data. Alternatively, it could result in a model which fits past data well, but which forecasts the future poorly.

Economic systems tend to be more complex than most physical systems. As a result the processes governing the behavior of such systems are not understood as well as those which govern the behavior of physical systems. In particular verifiable laws which govern the behavior of economic systems have yet to be developed, whereas well established laws of physics are available for describing the behavior of physical systems. Lack of such laws suggested the need for an alternative approach for determining the general form of a mathematical model for describing the behavior of an economic system. This prompted a search of the engineering and statistics

literature for approaches which could be used for determining the general form of such a model.

This search revealed two approaches which were considered to be potentially useful. One was developed by Pandit (1973) and the other by Akaike (1976). Both of these approaches are similar in two important ways. First, each utilize a highly general family of models each of which possesses sufficient flexibility to adequately model the behavior of many systems found in practice. Second, each has a theoretically sensible system for determining which members of its family of flexible models could be appropriate for modeling the behavior of a particular system.

In spite of the similarities, these approaches also differ in several important ways. First, they differ with respect to the structure of the family of flexible models which is assumed. The Pandit approach employs the auto-regressive moving average (ARMA) representation advocated by Box and Jenkins (1976), while the Akaike approach employs the Markovian state vector (MSV) representation traditionally used by systems engineers. The ARMA and MSV representations differ primarily with respect to the information which they store and utilize for forecasting the future. Basically the ARMA representation stores as historical values of the observed data for a number of prior periods, whereas the MSV approach stores forecasts for future periods made during the preceding period. Thus the ARMA representation stores "look back" information whereas the MSV representation stores "look ahead" information.

The second way in which the Pandit and Akaike approaches differ is in the process used to determine which member of the flexible family of models is most appropriate for modeling the behavior of a particular system. In

particular the Pandit approach employs a statistically based criterion for selecting an appropriate model, whereas the Akaike approach employs a criterion based on the concept of entropy.

The availability of two alternative approaches which could be used for developing LDF models for economic systems prompted the question as to which approach might be most effective. A search of both the engineering and statistics literature for investigations which could be useful in answering this question revealed that none had been reported. Given this situation the Akaike approach was selected on the basis that software for implementing it was immediately available, whereas software for implementing the Pandit approach would have required further development.

IV. IDENTIFICATION OF POTENTIAL LDF APPLICATIONS

Personnel in a number of organizational units within AFLC and ASD were contacted and interviewed for the purpose of identifying potential LDF applications. The particular AFLC units contacted were: (1) Directorate Organization Resources and Management-Financial Management and Productivity Analysis Division (MAJA); (2) Directorate Industrial Maintenance Process Control (MAQ), and (3) Directorate Facilities Production Engineering (MAX). The ASD units contacted were: (1) Directorate Cost Analysis-Research Cost Division (ACCR) and (2) Directorate Manufacturing Quality Assurance-Manufacturing Support Air Vehicle Planning Division (YPMG). The potential projects which were identified as a result of these interviews are briefly described in the next subsection.

AFLC Maintenance Productivity Measurement. The MAJA unit is currently in the process of developing a system for measuring the productivity of AFLC maintenance operations. Their basic goal is to develop a measure which: (1) is

comprehensive in that it considers all input and outputs; (2) is not affected by factors not related to the efficiency and effectiveness with which maintenance operations are performed; and (3) describes how well a unit is doing in relation to how well they should be doing.

It is almost inevitable that any measure developed will yield productivity measurements in which the true value of productivity is contaminated by noise. To illustrate how such noise might corrupt the data it will be helpful to consider an example. Suppose the measure of productivity is defined to be the ratio of the average cost of units serviced during a base period divided by the average cost of units serviced during the current period. There are a number of factors which could change and thereby cause the above ratio to yield an inaccurate measure of productivity. First, inflation or deflation in the cost of inputs could give a false signal of a change in productivity. In particular an increase in the cost of inputs could give the illusion of a productivity decrease, whereas a decrease in input costs could give the illusion of an increase. Second, a change in the workload mix could also give a false signal of a change in productivity. A change in the workload mix in which units with high service costs are replaced by those with low service costs could give the illusion of a productivity increase. Conversely, a change in which low service cost units are replaced by high service cost units could give the illusion of a productivity decrease.

Part of the distortions created by inflation/deflation and changes in the workload mix can be removed by relatively simple means. The effects of inflation/deflation might be partially accounted for by expressing the cost of all inputs in constant dollars. Also the effects of workload mix changes might be partially

accounted for by including in the ratio only those units which were serviced in both the base period and the current period. However, the process of making these corrections for inflation and changes in the workload mix could cause the resulting measurements to be contaminated with noise. Noise could be introduced by the inflation adjustment process if the rate of inflation for the inputs differ from the rate of inflation in the price index used to convert current cost to constant dollar equivalents. Noise could also be introduced by the process of correcting for the workload mix if the average productivities for the units serviced in the base period and the current period differ.

There are a number of additional factors which could cause the productivity measure to be distorted. (1) Overcharging for the depreciation for a new facility could cause the productivity measure to be falsely depressed, whereas undercharging could cause it to be falsely inflated. (2) The amount of service required is likely to increase as the average age of units in the field increases. Failure to correctly adjust for this phenomena could give a false picture of productivity. (3) Changes in the quality of materials used in the repair process could influence the amount of labor required and could thereby cause the measure of productivity to be distorted. (4) An increase in the workload which requires the use of excessive amounts of overtime could cause the efficiency of labor to fall which could in turn cause productivity measures to be depressed. In addition a change in the workload which creates excessive amounts of idle time could also cause productivity measure to be depressed. (5) A failure to correctly account for the cost of disruptions in customer organizations, caused by low quality service or late deliveries could cause the productivity measure to be distorted. True productivity would be overstated if such

costs were underestimated and vice versa.

From the above it can be seen that there are a large number of factors which could distort productivity measurements. Furthermore it will be virtually impossible to identify and correctly adjust for all of these factors. Thus it is highly likely that any series of productivity measures are likely to be contaminated by noise.

Stimulating productivity growth is one of the potential benefits which could result from a system for measuring productivity. In order for this potential to be fully realized it will be necessary for those responsible for productivity to be able to discriminate between real changes in productivity and distortions in the productivity measure caused by noise.

ASD-F-16 Tech Mod Program Impact. In order to stimulate productivity growth the Air Force has established a program to encourage firms in the aerospace industry to modernize their facilities. This program which is known as the technology modernization (TECH MOD) program utilizes various types of financial incentives to stimulate investment in new technologies which will hopefully increase productivity.

The TECH MOD program requires firms which participate to collect data to verify that the projected cost reductions have been achieved. Unfortunately, this performance data is typically contaminated by noise. In order to correctly assess the impact of a TECH MOD project it will be necessary to separate the signal, which represents the true productivity gain, from the noise which almost invariably contaminates production cost data. LDF methods could be used to accomplish the above separation.

V. LDF ANALYSIS OF PERFORMANCE DATA

Performance Data Selected for Analysis. Two sources of performance data were identified as a result of the

interviews described in the preceding section. First the interviews with the YPMG unit in ASD resulted in receiving permission to use data collected internally by General Dynamics, Fort Worth, for the purpose of assessing the impact of F-16 TECH MOD projects on direct labor cost. This data consists of the direct labor hours charged to each production run of a number of products which had been produced both before and after the implementation of a F-16 TECH MOD project. Second, the interviews with the MAJA and the MAQ units in AFLC were successful in locating a number of data series which were currently being used to monitor the productivity of maintenance activities at each Air Logistics Center.

The AFLC data was immediately available, whereas the F-16 TECH MOD data, which had to be obtained from General Dynamics Fort Worth, was not. The lack of immediate availability of the F-16 TECH MOD data combined with the limited time available for completing this project made it necessary to give priority to the analysis of the AFLC productivity data.

Results of the LDF Analysis. The Akaike method for developing LDF models was used to analyze the AFLC productivity data. There were two key questions which motivated the analysis of this data. The first was whether or not standard statistical methods, which assume statistically independent errors, would be appropriate for analyzing such data. The second was to determine if non-linear digital filters (NLDF) might be needed to obtain an appropriate model for the system which generated the data.

The strategy for answering the first question involved a two phase approach. In the first phase Akaike's approach was used to determine the most appropriate LDF model. In the second phase the basic structure of the fitted LDF model was analyzed to determine whether or not

it was consistent with the assumption of statistically independent errors.

The strategy for obtaining a preliminary indication of the answer to the second question also involved two phases. The first phase consisted of applying Akaike's method for developing a LDF model to the data. The second phases consisted of analyzing the results of the first phase model building effort to determine if there is evidence which suggest that a NLDF model might be appropriate.

The results obtained from utilizing Akakie's approach to developing LDF models for each of the AFLC productivity data series are summarized in Table I. This table also indicates whether or not the fitted model is consistent with the assumption of independent errors made in standard statistical methods. If a model is not consistent with the assumption, then the model is classified according to the nature of the dependency which it implies. This classification of the dependence which the model implies is based on the recognition that there are two classical ways in which a system can recover from an exogoneous shock. One way is for the system to gradually recover from the shock over an extended period of time. The other is for the system to suffer from the full effect of the shock for a finite period of time and to then abruptly recover. Systems which gradually recover are said to possess a gradual recovery mechanism, while those that abruptly recover are said to possess an abrupt recovery mechanism. In practice there is nothing to prevent a system from utilizing both the gradual and the abrupt recovery mechanisms. Such systems are said to possess a hybrid recovery mechanism.

Table I also describes the fitted models in terms of the number of steps into the future with which it is

TABLE I
UNIVARIATE ANALYSIS OF AFLC PRODUCTIVITY DATA

Location	Variable Name	# Step Ahead	Inde-pend.	Type Depend.
Ogden	Direct Labor Effic.	0	No	Gradual
	Indirect Labor Effic.	0	No	Gradual
	Direct Labor Hours	11	No	Hybrid
	Indirect Labor Hours	1	No	Hybrid
	Overtime Direct Labor	2	No	Hybrid
	Overtime Indirect Lab.	1	No	Hybrid
	Output Index	0	No	Gradual
	Input Index	0	No	Gradual
	Productivity Index	0	No	Gradual
Okla. City	Direct Labor Effic.	0	No	Gradual
	Indirect Labor Effic.	0	No	Gradual
	Direct Labor Hours	5	No	Hybrid
	Indirect Labor Hours	8	No	Hybrid
	Overtime Direct Labor	1	No	Hybrid
	Overtime Indirect Lab.	0	No	Gradual
	Output Index	0	No	Gradual
	Input Index	1	No	Hybrid
	Productivity Index	0	No	Gradual
Sacramento	Direct Labor Effic.	1	No	Hybrid
	Indirect Labor Effic.	1	No	Hybrid
	Direct Labor Hours	9	No	Hybrid
	Indirect Labor Hours	7	No	Hybrid
	Overtime Direct Labor	2	No	Hybrid
	Overtime Indirect Lab.	1	No	Hybrid
	Output Index	0	No	Gradual
	Input Index	0	No	Gradual
	Productivity Index	0	No	Gradual
San Antonio	Direct Labor Effic.	4	No	Hybrid
	Indirect Labor Effic.	0	No	Gradual
	Direct Labor Hours	11	No	Hybrid
	Indirect Labor Hours	12	No	Hybrid
	Overtime Direct Labor	2	No	Hybrid
	Overtime Indirect Lab.	0	No	Gradual
	Output Index	0	No	Gradual
	Input Index	0	No	Gradual
	Productivity Index	0	No	Gradual
Warner Robins	Direct Labor Effic.	3	No	Hybrid
	Indirect Labor Effic.	0	Yes	NA
	Direct Labor Hours	9	No	Hybrid
	Indirect Labor Hours	7	No	Hybrid
	Overtime Direct Labor	3	No	Hybrid
	Overtime Indirect Lab.	0	No	Gradual
	Output Index	0	No	Gradual
	Input Index	1	No	Hybrid
	Productivity Index	0	No	Gradual

necessary to forecast in order to adequately characterize the dynamic behavior of the system. The forecast for one step into the future can be used to account for the velocity of the system, whereas the forecast for two steps into the future can be used to account for the acceleration of the system. Forecasts for more distant steps can be used to account for seasonal changes in the behavior of the system.

Table II contains results obtained from analyzing the behavior of two variables simultaneously. These tables indicate whether or not the model of the variables behave in a pairwise independent fashion. If they are not independent, it describes the nature of the pairwise dependency which exists between the variables.

An examination of Table I indicated that the productivity data are likely to be pairwise dependent and that LDF methods should be used to analyze such data.

An examination of Table II revealed that feedback relationships could exist between the output and the input indices and also between the indirect labor and direct labor efficiencies. It also indicated that a nonlinear relationship could exist between (1) direct labor overtime hours and direct labor efficiency and also between (2) indirect labor overtime hours and indirect labor efficiency. If true this could suggest that nonlinear digital filter methods might be needed to develop adequate models of such relationships.

VI. RECOMMENDATIONS

The above results suggest that the performance data for most economic systems are not likely to satisfy the assumption of independence made by standard statistical techniques such as regression analysis, analysis of variance and Shewhart control charts. The consequences of incorrectly assuming independence can cause the estimated variance to be badly distorted. The potential magnitude

TABLE II
BIVARIATE ANALYSIS OF AFLC PRODUCTIVITY DATA

Location	Variables	Pairwise Indep.	Type Depend.	Evid. Nonlin.
Ogden	Direct Labor Effic. & Indir. Labor Effic.	No	Feedback indir.to dir.	No
	Overtime Direct Labor & Dir. Labor Effic.	?	?	Yes
	Overtime Indirect Lab. & Indir. Lab. Effic.	?	?	Yes
	Output & Input Index	Yes	NA	No
Okla. City	Direct Labor Effic. & Indir. Labor Effic.	No	Feedback indir.to dir.	No
	Overtime Direct Labor & Dir. Labor Effic.	?	?	Yes
	Overtime Indirect Lab. & Indir. Lab. Effic.	?	?	Yes
	Output & Input Index	No	Feedback out to in	No
Sacramento	Direct Labor Effic. & Indir. Labor Effic.	?	?	Yes
	Overtime Direct Labor & Dir. Labor Effic.	?	?	Yes
	Overtime Indirect Lab. & Indir. Lab. Effic.	?	?	Yes
	Output & Input Index	No	Feedback out to in	No
San Antonio	Direct Labor Effic. & Indir. Labor Effic.	?	?	Yes
	Overtime Direct Labor & Dir. Labor Effic.	?	?	Yes
	Overtime Indirect Labor & Indir. Labor Effic.	?	?	Yes
	Output & Input Index	Yes	NA	No
Warner Robins	Direct Labor Effic. & Indir. Labor Effic.	?	?	Yes
	Overtime Direct Labor & Dir. Labor Effic.	?	?	Yes
	Overtime Indirect Labor & Indir. Labor Effic.	?	?	Yes
	Output & Input Index	Yes	NA	No

of such distortions is demonstrated by an example in Box, Hunter, and Hunter (1978) which demonstrates that failure to satisfy the assumption of independence may cause the variance of the mean in a sample of size 10 to change by a factor of 19.

In the past the high cost of storing the data or of performing the necessary calculations would have made the use of LDF methods prohibitively expensive for many applications. Fortunately, the advent of microcomputers have reduced costs to the point where this is no longer the case. However, the lack of the availability of software which could be utilized by non-experts on a microcomputer is a factor which could limit the extent to which LDF methods are applied.

This suggests that the next task in this research area should be to develop user friendly LDF software which could be used by non-experts on a microcomputer. One danger with this approach is that such software could be used in situations in which it is ill-suited. The risk of such misuse could be greatly reduced by performing simulation experiments to gain a better understanding of the characteristics of data for which LDF methods should not be used. In particular such simulation experiments are needed to determine how well LDF methods can cope with systems which behave in a non-linear fashion.

The results in the preceding section also indicate that some of the economic systems analyzed could behave in a non-linear fashion. In addition the System Dynamics paradigm developed by Forrester (1961) suggests that this could be the case. These considerations suggest that the user friendly software should be capable of identifying when LDF methods are not appropriate. To be complete the software should be capable of fitting non-linear filter models when LDF methods are not deemed to be appropriate.

REFERENCES

1. Akaike, H., "Markovian Representation of Stochastic Processes and its Application to the Analysis of Autoregressive Moving Average Processes," Annals Institute of Statistical Mathematics, 1976, pp. 363-387.
2. Akaike, H., "Canonical Correlation of Time Series and the Use of an Information Criterion." In System Identification: Advances and Case Studies, Eds: R.K. Mehre and D.G. Lainiotis, New York: Academic Press, 1976, pp. 27-96.
3. Box, G.E.P.; Hunter, W.G.; and Hunter, J.S., Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building, New York: John Wiley & Sons, 1978.
4. Box, G.E.P. and Jenkins, G.M., Time Series Analysis: Forecasting and Control, San Francisco: Holden-Day, 1970.
5. Forrester, J.W., Industrial Dynamics, Cambridge, Massachusetts: The M.I.T. Press, 1961.
6. Pandit, S.M., Data Dependent Systems: Modeling Analysis and Optimal Control via Time Series, Doctoral Dissertation, University of Wisconsin-Madison, 1973.

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

Shallow Donor Impurity Binding Energies in Asymmetric Quantum Wells

Prepared by:	Dr. John Thomchick
Academic Rank:	Assistant Professor
Department and University:	Department of Engineering Science and Mechanics The Pennsylvania State University
Research Location:	Air Force Avionics Laboratory Electronic Research Branch
USAF Research Colleague:	Dr. Krishan K. Bajaj
Date:	August 9, 1983
Contract No.:	F49620-82-C-0035

Shallow Donor Impurity Binding Energies in Asymmetric Quantum Wells

by

John Thomchick

ABSTRACT

The energy of an electron bound by a donor impurity center inside a GaAs/AlGaAs layered structure has been investigated. In the present study the donor impurity atom was assumed to be located at an arbitrary site inside a GaAs layer sandwiched between two GaAlAs layers with differing Al concentrations. The different Al concentrations cause different potential energy barrier heights on the two sides of the GaAs layer resulting in an asymmetric quantum well. The calculation was formulated so that the binding energies could be obtained as a function of the ratio of the barrier heights at the two interfaces, as a function of the location of the impurity atom inside the GaAs layer, and as a function of the width of the GaAs layer.

ACKNOWLEDGEMENTS

The author is indebted to Dr. K. K. Bajaj of the Avionics Laboratory for suggesting the research problem considered in this report and for many helpful discussions. The author is also indebted to Captain George Norris of the Avionics Laboratory, Dr. Sid Chadhuri of the University Research Center of Wright State University, and Dr. Ronald Greene of the Department of Physics of the University of New Orleans for many stimulating discussions concerning the present problem and closely related problems in the area of GaAs research.

Finally, the author would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research, the Electronics Research Branch of the Air Force Avionics Laboratory, and the Southeastern Center for Electrical Engineering Education for the opportunity of working on the research discussed in this report.

I. INTRODUCTION

The Electronic Research Branch of the Avionics Laboratory has for a number of years been interested in the use of GaAs for application to device technology.[1] The specific interest in this material has resulted because GaAs possesses certain advantages over other semiconductors such as Si or Ge in the areas of opto-electronics, high frequency microwave devices, and high speed logic.[2] A closely related semiconductor $Al_xGa_{1-x}As$ in which a certain fraction x of Al atoms (usually $x \lesssim 0.4$) replace Ga atoms in the crystal lattice also has application in the above areas and has been studied extensively in other laboratories.

More recently, crystal growth techniques such as molecular beam epitaxy[3] (MBE) and metal-organic chemical vapor deposition[4] (MOVCD) have held promise for vast improvements in semiconductor technology in the area of man-made tailoring of structures for specific device applications.[5,6] With MBE it is possible to grow single-crystal semiconductor films having smooth surfaces and finely defined film thicknesses. To date one of the smoothest and most highly controlled MBE grown structures has been the $GaAs/Al_xGa_{1-x}As$ multilayers in which alternating layers of GaAs and $Al_xGa_{1-x}As$ are grown into a "superlattice" configuration. Because of these developments and because of its on-going interest in GaAs research, it has been natural for the Electronic Research Branch to undertake a program to study the growth and characterization of the new multilayer structures.

Included in the characterization program has been a systematic theoretical study[7-9] of the binding energies due to shallow donor

impurity atoms and Wannier excitons inside quantum wells formed when a GaAs layer is sandwiched between layers of $Al_xGa_{1-x}As$. It has been customary both to grow and to investigate theoretically systems in which $Al_xGa_{1-x}As$ layers have had the same fraction x of Al, and thus the same physical properties, on both sides of the GaAs layer. This leads to a quantum well structure known as a "square well." However, it has become apparent that for certain applications structures where the Al concentration on the two sides of the GaAs layer is different may be useful[10]; such a structure would lead to an "asymmetric quantum well." It is expected that the properties of various physical systems such as shallow impurity atoms and Wannier excitons in asymmetric wells could differ significantly from those in a square well. It is the study of the binding energy of shallow donor impurity atoms in an asymmetric quantum well with which the present report is concerned.

II. OBJECTIVES

The objectives for the summer program were to investigate electronic properties associated with $GaAs/Al_xGa_{1-x}As$ quantum well structures. Specific objectives that developed with regard to the asymmetric quantum well problem are as follows:

(1) obtain an expression for the subband level energies in an asymmetric quantum well;

(2) obtain a criterion for the existance of bound solutions in an asymmetric quantum well;

(3) obtain an expression for the binding energy of a shallow donor impurity atom in an asymmetric quantum well as a function of

the ratio of the potential barriers on the two sides of the well and as a function of the total width of the well, and incorporate into the calculation the possibility of locating the donor impurity atom at an arbitrary site inside the well.

III. THE ASYMMETRIC QUANTUM WELL PROBLEM

The one-dimensional asymmetric quantum well potential energy function $V_A(z)$ is defined by the following mathematical relation:

$$V_A(z) = \begin{cases} V_1 , & z > L/2 \\ 0 , & -L/2 < z < L/2 \\ V_2 , & z < -L/2 \end{cases} \qquad (1)$$

In the present work the origin of coordinates is taken to be at the center of the well and the behavior of the electron in the x-y plane is assumed to follow that for plane wave motion. In Eq. (1) V_1 and V_2 are uniform potential energy values such that $V_1 \neq V_2$ and L is the total width of the well. The interval $z > L/2$ will be referred to as region I, $-L/2 < z < L/2$ as region II and $z < -L/2$ as region III. The above potential energy function would exist physically in a layered structure if region I consisted of a layer of $Al_xGa_{1-x}As$, region II of GaAs, and region III of $Al_yGa_{1-y}As$, such that $x \neq y$. For a large majority of applications one keeps $x,y \lesssim 0.4$ because AlGaAs remains a direct band gap semiconductor for these values. It should be clear that if x=y,

then $V_1=V_2$, and the asymmetric quantum well potential reduces to the special limiting case of the square quantum well potential.

To find the subband energy levels corresponding to the potential $V_A(z)$ one needs to obtain the solutions to the Schrodinger equation[11]

$$-\frac{\hbar^2}{2m}\frac{d^2f(z)}{dz^2} + V_A(z)f(z) = E_n f(z) \tag{2}$$

where $f(z)$ is the wavefunction for an electron in the potential $V_A(z)$, m is the effective mass associated with the conduction band in GaAs (for GaAs, $m = 0.067m_e$; this same value is assumed for the AlGaAs layers) and E_n refers to the energy of the n^{th} subband. In the present work we are interested only in the energy of the lowest subband E_1.

In what follows it will be convenient to describe all lengths in terms of an effective Bohr radius given by

$$b_o = \frac{\hbar^2 \epsilon_o}{m e^2} \tag{3}$$

where ϵ_o is the dielectric constant for GaAs (for GaAs, $\epsilon_o = 12.5$; this same value is assumed for the AlGaAs layers) and e is the electronic charge, and to describe all energies in terms of the effective Rydberg

$$R = \frac{e^2}{2\epsilon_o b_o} \tag{4}$$

One finds for GaAs that $b_o = 98.7$A and $R = 5.83$ meV. In these units the Schrodinger equation becomes

$$-\frac{d^2}{dz^2}f(z) + V_A(z)f(z) = E_n f(z) \tag{5}$$

with $V_A(z)$ and E_n now dimensionless.

The solution of Eq. (5) differs in two significant respects from that for the square well problem. First because of the asymmetry of the potential $V_A(z)$ the solutions f(z) cannot be divided into even and odd parity solutions. Secondly, although at least one bound solution exists for all non-zero values of the potential barriers in the square well problem, bound solutions do not always exist for the asymmetric quantum well problem; the condition for the existence of solutions in this latter case will be given below.

The solution of Eq. (5) proceeds by taking wavefunctions appropriate to the various regions of the one-dimensional coordinate space and imposing matching conditions on these functions and their first derivatives at the sides of the well ($z = \pm L/2$). The wavefunction chosen in the present situation is

$$f(z) = \begin{cases} A_1 e^{-K_1 z} & , \quad z > L/2 \\ A_0 \sin(kz + \phi) & , \quad -L/2 < z < L/2 \\ A_2 e^{K_2 z} & , \quad z < -L/2 \end{cases} \tag{6}$$

where A_0, A_1 and A_2 are normalization factors,

$$K_1 = (V_1 - E_n)^{1/2} ,$$

$$K_2 = (V_2 - E_n)^{1/2} , \tag{7}$$

and

$$k = E_n^{1/2}$$

and where ϕ is a phase factor required due to the asymmetry of $V_A(z)$. The matching conditions at $z = \pm L/2$ lead to the condition on ϕ ,

$$\phi = \frac{1}{2} kL + \tan^{-1}\left(\frac{k}{K_2}\right), \tag{8}$$

and to the energy eigenvalue equation

$$n\pi - kL = \tan^{-1}\left(\frac{k}{K_1}\right) + \tan^{-1}\left(\frac{k}{K_2}\right). \tag{9}$$

In Eqs. (8) and (9) the principle values of $\tan^{-1}$ are to be taken and the integer n takes on the values, n = 1, 2, 3, ... Equation (9) is a transcendental equation for the energy eigenvalues E_n and is most conveniently solved either graphically or numerically.

The condition for the existence of bound state solutions is most easily obtained by rearranging Eq. (9) to read

$$n\pi - KL\xi = \sin^{-1}\xi + \sin^{-1}(\xi \cos\gamma) \tag{10}$$

where

$$\xi = k/K$$

$$K = (k^2 + K_1^2)^{1/2} = V_1^{1/2} \tag{11}$$

and $$\cos\gamma = (V_1/V_2)^{1/2}$$

and where, as for $\tan^{-1}$, the principle values for $\sin^{-1}$ are to be taken.

The left hand side of Eq. (10) is a series of straight lines with respect to the variable ξ (one for each value of n) and for a solution to exist the straight lines must intersect the function of ξ on the right hand side of Eq. (10). Since the principle values of $\sin^{-1}$ are involved, one must have $|\xi| \leq 1$; physically this condition implies that for a bound state to exist the energy E_n must be smaller than or equal to the smaller of the potential barriers V_1 or V_2 (in deriving Eq. (10) it was assumed that $V_1 < V_2$). Substituting $\xi = 1$ into Eq. (10) leads to the condition for the number n of bound state eigenvalues

$$KL \geq (n-1)\pi + \gamma . \tag{12}$$

If this condition is <u>not</u> obeyed for n=1, then there are <u>no</u> bound state solutions; thus, there are no bound state solutions if $KL < \gamma$, or

$$V^{1/2} L < \cos^{-1}\left(V_1/V_2\right)^{1/2} . \tag{13}$$

IV <u>SHALLOW DONOR IMPURITY ATOM IN AN ASYMMETRIC QUANTUM WELL</u>

To obtain the energy E of an electron bound to a hydrogenic donor atom associated with the first subband in a GaAs asymmetric quantum well one must consider the Schrodinger equation $H\Psi = E\Psi$ with

$$H = \frac{p^2}{2m} - \frac{e^2}{\epsilon_0 r} + V_A(z) \tag{14}$$

The first term in the Hamiltonian operator H is the kinetic energy of the electron in the effective mass approximation where, as before, m is the electron effective mass associated with the lowest subband. The last two terms in H are, respectively, the attractive Coulomb potential due to the donor ion and the asymmetric quantum well potential discussed in Section III. As indicated in Section III, it will be assumed in the present work that the mass m and the dielectric constant ϵ_o are constant across the interfaces between the GaAs and AlGaAs layers. This approximation is probably good as long as the donor atom is near the center of the quantum well and less good when the donor atom is near the interface regions. Expressing the momentum p in terms of its equivalent operator form (in cylindrical coordinates) and using the dimensionless units b_o and R introduced in Eqs. (3) and (4), one finds that H can be written in the form

$$H = -\left[\frac{1}{\rho}\frac{\partial}{\partial\rho}\rho\frac{\partial}{\partial\rho} + \frac{1}{\rho^2}\frac{\partial^2}{\partial\phi^2}\right]$$

$$-\frac{\partial^2}{\partial z^2} - \frac{2}{\left[\rho^2 + (z-z_i)^2\right]^{1/2}} + V_A(z) \qquad (15)$$

In this last expression z_i is the arbitrary position of the donor atom inside the well.

Since it appears unlikely that an exact solution of the Schrodinger equation is possible with H given by Eq. (15), a variational approach was taken. The trial wavefunction Ψ used is of the product form

$$\Psi = f(z)\, g(\rho, \phi, z - z_i) , \tag{16}$$

where f(z) is the wavefunction corresponding to the solutions of the asymmetric quantum well problem and where

$$g(\rho, \phi, z - z_i) = \sum_n a_n g_n(\rho, \phi, z - z_i) \tag{17}$$

with the g_n having the form

$$g_n = (z - z_i)^n \exp\left\{-\alpha\left[\rho^2 + (z - z_i)^2\right]^{1/2}\right\} . \tag{18}$$

In Eqs. (17) and (18) the a_n and α are variational parameters to be determined. Note from Eq. (18) that the exponential factor is that appropriate for the ground state of the hydrogen atom problem.

In terms of the product wavefunction Eq. (16), the Schrodinger equation can be written as

$$\left[T_\perp + T_z + V_c + V_A\right](fg) = E_n(fg) , \tag{19}$$

where

$$T_\perp = -\left[\frac{1}{\rho}\frac{\partial}{\partial\rho}\rho\frac{\partial}{\partial\rho} + \frac{1}{\rho^2}\frac{\partial^2}{\partial\phi^2}\right] , \tag{19a}$$

$$T_z = -\frac{\partial^2}{\partial z^2} \quad , \tag{19b}$$

$$V_c = -\frac{2}{[\rho^2 + (z-z_1)^2]^{1/2}} \quad , \tag{19c}$$

and where V_A is given by Eq. (1). One finds that

$$T_z(fg) = g(T_z f) - 2\frac{\partial f}{\partial z}\frac{\partial g}{\partial z} + f(T_z g) . \tag{20}$$

However, $(T_z + V_A)f = E_A f$ since f is the solution of the asymmetric quantum well problem. This leads to the result

$$f(T_\perp + V_c)g - f(T_z g) - 2\frac{\partial f}{\partial z}\frac{\partial g}{\partial z} = (E-E_A) fg . \tag{21}$$

It is really $(E-E_A)$ that we wish to calculate. The quantity $(E-E_A)$ is the total energy minus the energy of the first subband and so this term represents the binding energy of an electron to the donor ion; we will denote this binding energy by E_B.

To evaluate this energy it is necessary to take the quantum mechanical expectation values of the terms in Eq. (21). This leads to the result

$$\int_{-\infty}^{\infty} dz \int_0^{2\pi} d\phi \int_0^{\infty} \rho d\rho \left[|f|^2 (T_\perp + V_c) |g|^2 \right.$$

$$\left. - |f|^2 g^* (T_z g) - 2 \left(f^* \frac{\partial f}{\partial z} \right) \left(g^* \frac{\partial g}{\partial z} \right) \right] \qquad (22)$$

$$= E_B \int_{-\infty}^{\infty} dz \int_0^{2\pi} d\phi \int_0^{\infty} \rho d\rho \, |f|^2 |g|^2 .$$

The last two terms on the left hand side of this expression can be combined into a single term through a Green's theorem type transformation and by invoking the physical condition that the wavefunction must vanish at large distances from the center of the quantum well. The integral on the right hand side of Eq. (22) is just a normalization integral which will be denoted by N. Taking these considerations into account one obtains for the binding energy the expression

$$\int_{-\infty}^{\infty} dz \int_0^{2\pi} d\phi \int_0^{\infty} \rho d\rho \left[|f|^2 (T_\perp + V_c) |g|^2 + |f^2| \left(\frac{\partial g}{\partial z} \right)^2 \right] \qquad (23)$$

$$= \bar{E}_B N$$

Analytic work on these integrals which was fairly tedious has been completed during the course of the summer program; due to space

limitations the results of these calculations cannot be given in the present report.

The final step in finding values for E_B is to vary the results obtained from Eq. (23) with respect to the variational parameters a_n and α for each value of L, z_i and the ratio V_1/V_2 that is of interest. This is an exceedingly formidable task to perform analytically and instead will be carried out numerically. This part of the calculation will be completed after the summer program and will be discussed briefly in the section on recommendations. The numerical evaluation required may be simplified somewhat by setting up the calculation in the following determinantal form as illustrated for the case where $g = a_0 g_0 + a_1 g_1$:

$$\begin{vmatrix} E_{00} - E_B N_{00} & E_{01} - E_B N_{01} \\ E_{10} - E_B N_{10} & E_{11} - E_B N_{11} \end{vmatrix} = 0 \quad , \tag{24}$$

where E_{00} is a shorthand notation for the result of the integration of the left hand side of Eq. (23) for only the terms involving the wavefunction g_0, E_{01} involves only g_0^* and g_1, and so forth; in a similar way N_{00} is a normalization term involving only g_0, etc. For our calculation $E_{01} = E_{10}$ and $N_{01} = N_{10}$. By performing the calculation in this manner (taking the lowest value for E_B from Eq. (24)) the variation

with respect to a_o and a_1 is automatically taken into account with only the variation with respect to α remaining to be performed numerically.

V RECOMMENDATIONS

In completing the numerical calculations associated with the work described in this report, it is anticipated that plots will be obtained for the binding energy E_B versus the width L of the GaAs quantum well layer for various constant values of the ratio V_1/V_2 and for various constant positions z_i of the donor ion inside the well. In addition, a more detailed investigation of the relationship between E_B and V_1/V_2 for constant L and z_i, and between E_B and z_i for constant V_1/V_2 and L will be made. Each of these relationships will aid in understanding and characterizing the subband structure in GaAs multilayers for the case of an asymmetric quantum well configuration.

As a consequence of this calculation one can also obtain, by simply forcing $V_1 = V_2$, the binding energy for a shallow donor in a square quantum well as a function of arbitrary position z_i of the impurity ion. Although the binding energy as a function of position z_i has been calculated for the case where both potential barriers V_1 and V_2 are infinite[12], and also for the case of finite square barriers when the impurity ion is located at the edge of the well[13] ($z_i = \pm L/2$), to our knowledge E_B has not been calculated for a square quantum well for arbitrary z_i.

A number of offshoots stemming from the calculation presented in this report seem feasible. Perhaps the most immediate in terms of further characterization of these layered structures is the

determination of the binding energy of a Wannier exciton in the asymmetric quantum well. An exciton is an electron and hole bound by their mutual Coulombic attraction and such entities play an important role in the understanding of absorption spectra from these systems. This calculation would proceed in a fashion similar to that for the donor impurity case except that a wavefunction incorporating the presence of both an electron and a hole must be utilized.

Other aspects that can be studied in relation to the present calculation are the effects on the binding energy due to the non-parabolicity of the conduction band or due to the inclusion of a magnetic field; also the calculation of excited state energies could be attempted. Results from calculations such as these would provide a more detailed and more comprehensive characterization of the multilayer structure.

REFERENCES

1. G. B. Norris, "Survey of GaAs Research in the Electronic Research Branch," Internal Report (unpublished).
2. For an excellent review on GaAs, see J. S. Blakemore "Semiconducting and Other Major Properties of Gallium Arsenide," J. Appl. Phys. 53, R123 (1982).
3. For an excellent review on MBE, see A. C. Gossard, "Molecular Beam Epitaxy of Superlattices in Thin Films," in Thin Films: Preparation and Properties, edited by K. N. Tu and R. Rosenberg (Academic Press, New York, 1982).
4. L. Hollan, J. P. Hallais and J. C. Brice, "The Preparation of Gallium Arsenide," in Current Topics in Materials Science, edited by E. Kaldis (North-Holland, Amsterdam, 1980).
5. T. Mimura, S. Hiyamizu, K. Joshin, and K. Hikosaka, "Enhancement-mode High Electron Mobility Transistors for Logic Applications," Japan. Jour. Appl. Phys. 20, 4714 (1981).
6. B. A. Joyce, "Present Status and Future Directions for MBE," Surf. Sci. 86, 92 (1979).
7. R. L. Greene and K. K. Bajaj, "Energy Levels of Hydrogenic Impurity States in GaAs - $Ga_{1-x}Al_xAs$ Quantum Well Structures," Solid State Commun. 45, 825 (1983).
8. R. L. Greene and K. K. Bajaj, "Binding Energies of Wannier Excitons in GaAs - $Ga_{1-x}Al_xAs$ Quantum Well Structures," Solid State Commun. 45, 831 (1983).

9. R. L. Greene and K. K. Bajaj, "Energy Levels of Hydrogenic Impurities and Wannier Excitons in Quantum Well Structures," Jour. Vac. Sci. Technol. B1, 391 (1983).

10. B. A. Vojak, W. D. Laidig, N. Holonyak, Jr., M. D. Camras, J. J. Coleman and P. D. Daphus, "High Energy (visible-red) Stimulated Emission in GaAs," J. Appl. Phys 52, 621 (1981).

11. A. Messiah, Quantum Mechanics, Vol. I (Amsterdam, North-Holland, 1961), Chapter 3.

12. G. Bastard, "Hydrogenic Impurity States in a Quantum Well: A Simple Model," Phys. Rev B24, 4714 (1981).

13. C. Mailhiot, Y. C. Chang, and T. C. McGill, "Energy Spectra of Donors in GaAs - $Ga_{1-x}Al_xAs$ Quantum Well Structures in the Effective-mass Approximation," Phys. Rev. B26, 4449 (1982).

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

STATISTICAL SIMULATION AND CORRELATION STUDIES OF GaAs MESFETS

Prepared by:	Dr. Arthur R. Thorbjornsen
Academic Rank:	Professor
Department and University:	Department of Electrical Engineering The University of Toledo
Research Location:	Air Force Wright Aeronautical Laboratories, Avionics Laboratory, Electronics Technology Division, Microelectronics Branch, Device Technology Group
USAF Research Colleague:	Dr. Chern I. Huang
Date:	August 19, 1983
Contract No:	F49620-82-C-0035

STATISTICAL SIMULATION AND CORRELATION STUDIES OF GAAS MESFETS

by

Arthur R. Thorbjornsen

ABSTRACT

A method for the statistical simulation of GaAs MESFETs has been tested by comparing simulated device data with measured device data. The Kolmogorov-Smirnov two-sample goodness of fit test indicated that the simulated and measured parameter values have the same distribution in about 70% of the cases tested. The MESFET model used produces correlation coefficients between parameter values that have realistic values for all except one device parameter. Overall, the simulation method seems to produce results that are adequate for design purposes. A statistical test on measured device data indicates that the distance separating two GaAs MESFETs has no effect on correlation.

ACKNOWLEDGEMENTS

The author would like to express his sincere gratitude to the U.S. Air Force Systems Command, the Air Force Office of Scientific Research, and the Avionics Laboratory at Wright-Patterson Air Force Base for the opportunity to participate in the 1983 USAF-SCEEE Summer Faculty Research Program. Special thanks are due to my Effort Focal Point, Dr. Chern I. Huang, for his cooperation and assistance. Thanks are also due to Mr. Roy Newman for providing experimental device data and to Mr. Rick Stormont for his help with the computer system.

I. INTRODUCTION

The USAF Avionics Laboratory at Wright-Patterson Air Force Base has been engaged in research on GaAs (gallium arsenide) MESFETs (metal-semiconductor field-effect transistors) for a number of years. The reason for this interest is the potential that GaAs devices have for high speed operation. GaAs has an electron mobility that is about 5.5 times higher than that of silicon at relatively low levels of electric field strength [1]. A GaAs device should then be theoretically capable of operating at speeds 5.5 times greater than an equivalent silicon device. GaAs MESFET logic gates have been constructed which have gate delay times on the order of 100 picoseconds or less.

GaAs MESFETs are constructed in a manner similar to silicon JFETs, with the main difference being the method of formation of the gate junction. The silicon JFET gate junction is formed by diffusion while the GaAs MESFET gate junction is actually a Schottky barrier diode formed by depositing a metal film (Cr-Pt-Au) directly on the GaAs material [2].

In order to increase the operating speed of the GaAs MESFET the gate is made with a very small length and a very great width. For example, some gate structures are 1.0 micrometers long and 50 micrometers wide. Producing patterns 1.0 micrometers in size is near the limit of the capabilities of optical lithography, and as a result, one would expect a large variability in the values of electrical device parameters which are dependent on gate length. Measured device data obtained at the Avionics Laboratory show a very large variability for practically all parameters. Other causes of large parameter variability lie with the relatively immature status of GaAs device fabrication technology.

Because of the large variabilities present in GaAs devices, a statistical approach to GaAs circuit design should be used. Of the methods used in statistical design, Monte Carlo analysis produces the most accurate and complete information.

One requirement of statistical analysis and design of electronic circuits is the capability to statistically simulate electronic devices.

In other words, one must be able to generate sets of device parameter values that have realistic joint probability density functions. This means that each set of generated parameter values should have a properly shaped distribution and that the correlation coefficients between the various parameters should be in the proper ranges.

II. OBJECTIVES OF THE RESEARCH EFFORT

The author has previously developed a method for the statistical simulation of GaAs MESFETs [3]. One object of this SFRP research effort, therefore, was to assess the adequacy of the above method of simulation. This assessment was to be done by generating simulated device data and comparing it with measured device data using statistical tests. If required, the mathematical device model used in the simulation would be revised in order to produce more satisfactory results.

Another research objective was to try to determine if the correlation coefficients between the parameters of different GaAs MESFETs on a single chip was dependent on the distance between the devices. When simulating groups of devices it is necessary to know whether or not relative device location must be an input parameter.

III. THE GAAS MESFET MODEL

Because of the way that electron mobility varies with applied electric field intensity in GaAs, a two-region equation has been developed by Pucel, Haus, and Statz to model the electrical behavior of an intrinsic GaAs MESFET [4,5]. In one region of the device, where the electric field is relatively weak, the mobility is assumed to be constant. In the other region the velocity of electrons is assumed to be at the constant saturation value. The dividing line between the two regions will depend on the applied terminal voltages. The two-region mathematical equations describing intrinsic GaAs MESFET behavior are:

$$V_{13} = V_p\left\{\left(u_2^2 - u_1^2\right) + \frac{2t}{\pi\beta L}\sinh\left[\frac{\pi(L - L_1)}{2t}\right]\right\}, \qquad (1)$$

$$L_1 = \beta L\left[(u_c^2 - u_i^2) - \frac{2}{3}(u_c^3 - u_i^3)\right] / (1 - u_c) \quad , \tag{2}$$

and $$I_D = q N_d v_s t W (1 - u_c) \quad , \tag{3}$$

where $$V_p = q N_d t^2 / 2\epsilon_s \quad ,$$

$$u_i = \left[(V_{32} + V_{bi}) / V_p\right]^{1/2} \quad , \tag{4}$$

and $$\beta = \mu_n V_p / v_s L \quad .$$

Also, v_s is the saturated velocity in region 2, L is the total channel length, W is the gate width, u_c is the normalized depletion layer width at the boundary between regions 1 and 2, t is the thickness of the active GaAs layer, N_d is the concentration of donor electrons in the GaAs layer, ϵ_s is the permitivity of GaAs, V_{bi} is the built-in potential of the gate junction, μ_n is the electron mobility in GaAs, L_1 is the channel length of region 1, and q is the magnitude of the charge of an electron.

The intrinsic device model does not account for parasitic source and drain resistances which are included in the extrinsic device model shown in Figure 1. The parasitic source resistance is given by

$$R_s = \frac{L_{as}}{q N_d \mu_n t W} + \frac{4 \times 10^{-5}}{L_{cs} W \sqrt{N_d}} \quad .$$

The second term in the expression for R is the contact resistance [6]. L_{as} is the distance from the channel to the source contact and L_{cs} is the length of the source contact. A similar equation is used for R_D. The extrinsic equations are

$$V_{DS} = I_D (R_s + R_D) + V_{13} \tag{5}$$

and

$$V_{GS} = -V_{32} + I_D R_s \quad . \tag{6}$$

Given a pair of values for V_{DS} and V_{GS} and initial estimate for u_c, the

standard Newton-Raphson method is used to solve equations (1) through (6) for the drain current I_D [7].

IV. STATISTICAL SIMULATION

Equations (1) through (6) are in terms of physical constants or material or dimensional variables. These variables will be referred to as the input physical parameters. A good assumption to make is that the input physical parameters are normally distributed random variables. With the exception of dimensional pairs L_{as}, L_{aD} and L_{cs}, L_{cD} all of the input physical parameters are assumed to be independent, or noncorrelated. Sets of values for the input physical parameters with any desired means, standard deviations, and correlation coefficients can easily be computed in seconds using a standard random number generating subroutine.

Once sets of values for the input physical parameters are obtained, then the nonlinear equations can be solved using each set of input parameters for I_D provided that V_{GS} and V_{DS} are specified. An alternate solution is for V_{GS}, given values of I_D and V_{DS}.

It is desired to simulate the measurement of standard electrical device parameters for the GaAs MESFET. A set of measured parameters that is routinely used in the Avionics Lab consists of the following parameters:

I_{DSS} = the saturation drain current with V_{DS} = 2.5v and V_{GS} =0v,
V_T = the pinchoff voltage, or the value of V_{GS} at which I_D =0.01xI_{DSS}.
R_{LIN} = the drain-to-source resistance in the linear region,
R_{ON} = the drain-to-source resistance in saturation,
R_D = the parasitic drain resistance,
and R_S = the parasitic source resistance.

The above parameters are illustrated in Figure 2. Other parameters are measured, such as the transconductance and the knee voltage, but the six listed above are sufficient to compute the parameters of the Curtice model of the GaAs MESFET [8].

If the model equations accurately represent the device electrical

behavior, then the computed electrical parameter values should have the same distributions and correlation coefficients as measured parameter values.

V. COMPARISON OF MEASURED AND SIMULATED RESULTS

Measured data were obtained for about 50 samples of each of five GaAS MESFETs that were fabricated and tested in the Integrated Circuits facility at the Avionics Lab. A program was written that computes the means, standard deviations, correlation coefficients, and other statistics for the measured parameters. The program also plots frequency histograms, which allow one to see the shape of the parameter distributions, and the scatter diagrams, which show if the correlation between two parameters is linear or nonlinear. All of the statistics and plots were generated for the measured data.

Previously, another computer program had been written to perform the statistical simulation referred to earlier. A trial and error procedure was used in order to determine the values of the input physical parameters that gave the best fit for each of the five devices. These values are listed in Table 1.

Using the best fit input parameters the statistical simulation program was run and produced the same type of data, statistics, and plots as for the measured data. The Kolmogorov-Smirnov two-sample goodness-of-fit test [9,10] was performed in order to compare the simulated and measured parameter distributions. The results of the K-S test are listed in Table 2, where it is seen that 46.7% of the test statistics are less than the critical value at a level of significance of 0.95. This means that for those cases the hypothesis that the measured and simulated distributions are the same cannot be rejected.

Figure 3 shows two superimposed frequency histograms for which the K-S statistic is 0.33. From a practical point of view, and with a small sample size, the two distributions seem to have an acceptable match. Using 0.33 as a critical value we see that 76.7% of the parameters in Table 2 then have an acceptable match between measured and simulated distributions.

In Table 3 are listed the computed correlation coefficients for the simulated and measured device parameters. Although no statistical test was performed, one can see that the correlation coefficients match quite well. There is one exception in the case of parameter R_{ON}, for which the measured and simulated correlation coefficients have the opposite signs. It is believed that this discrepancy is due to the two-region GaAs MESFET model not being perfectly descriptive of device behavior. There are many assumptions and simplifications in the derivations of equations (1) through (6).

VI. TESTING FOR THE DISTANCE EFFECT ON PARAMETER CORRELATIONS

When simulating groups of devices, as when performing a Monte Carlo analysis of an integrated circuit, we will have to account for the correlation between parameters of different devices as well as for the correlation between parameters within a single device. Previous work has indicated that the closer together two devices are placed on a chip or substrate, the more highly correlated their parameters will be [11,12].

The correlation coefficients between like parameters on the various devices (e.g., between I_{DSSA} and I_{DSSB}) were computed and the linear distances separating the device centers were measured. These values are shown in Table 4(a). Since it is difficult to determine behavioral trends with such large variations as are shown, a linear regression line was computed for each correlation coefficient. A statistical test [13] was then performed to see if the slope of the line is not significantly different from zero, or in other words, that there is no distance effect. The results of the "t" test are also listed Table 4(a). The test results show that there is no test statistic greater than the critical value of 1.86 at a level of significance of 0.95. Therefore, it is concluded that we cannot reject the hypothesis that the slope of the regression line is equal to zero.

This test was repeated for the identical devices on a second wafer. These results are listed in Table 4(b). It is seen that in only one case can we reject the hypothesis that the slope is equal to zero.

VII. CONCLUSIONS

The results of statistical tests on simulated and measured parameter data for GaAs MESFETs indicate that the method of statistical simulation employing the two-region model produces an acceptable simulation for practical use. The simulation is not an exact one and one parameter, R_{ON}, does not have the correct correlation with respect to the other parameters.

A test of measured data from two wafers indicates that the correlation between parameters of two devices is not dependent on how closely the two devices are placed together. This result may be due to the very large variabilities present in GaAs MESFET parameters. If this result is true, then the task of simulating groups of devices will be simplified.

VIII. RECOMMENDATIONS

Although the method of statistical simulation of GaAs MESFETs appears to be acceptable from a practical point of view, there are some deficiencies in the method, particularly the correlation of R_{ON} with the other parameters. Follow-on research should include a search for and trial of other MESFET mathematical models in the statistical simulation method.

The simulation method used in this effort simulates individual GaAs MESFET devices. When performing Monte Carlo analyses of integrated circuits it is necessary to simulate groups of devices with correlation between the parameters of different devices. The statistical simulation method should be extended to simulate groups of devices and then should be coupled with a circuit analysis program, such as SPICE2, to produce Monte Carlo analyses.

Further research can make use of the above mentioned Monte Carlo analysis program to study GaAs circuit designs and to develop Monte Carlo analysis as a circuit design tool.

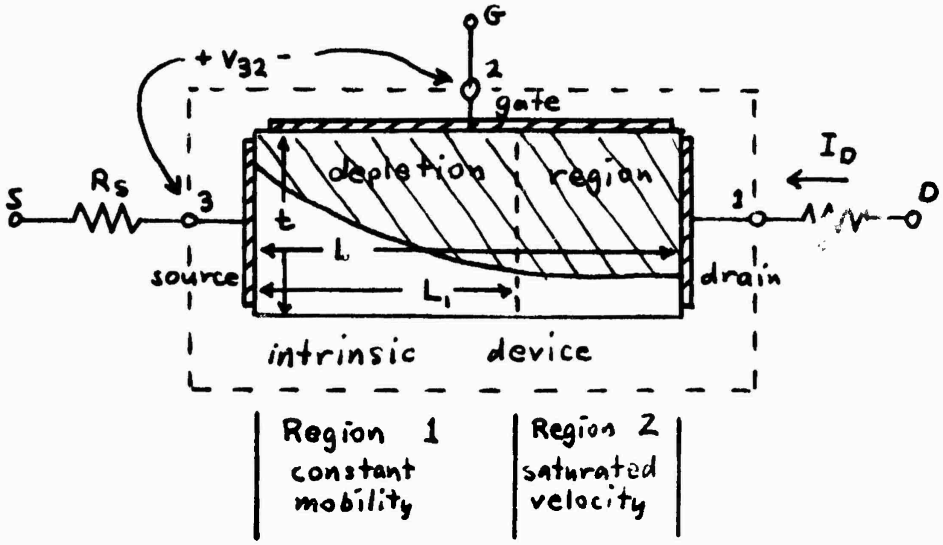


Figure 1. The extrinsic GaAs MESFET model.

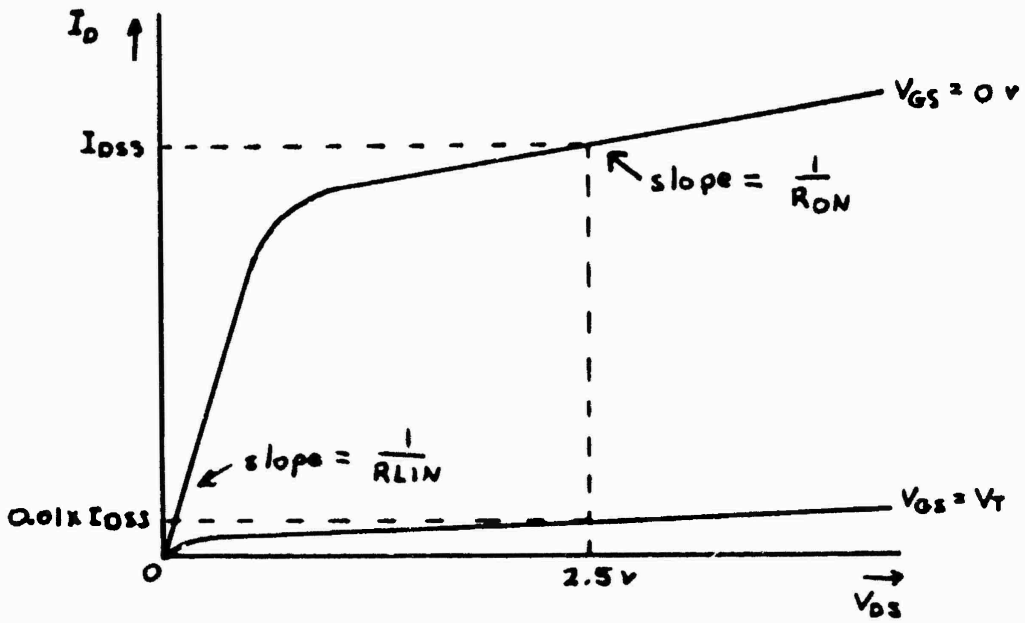


Figure 2. Measured MESFET parameters.

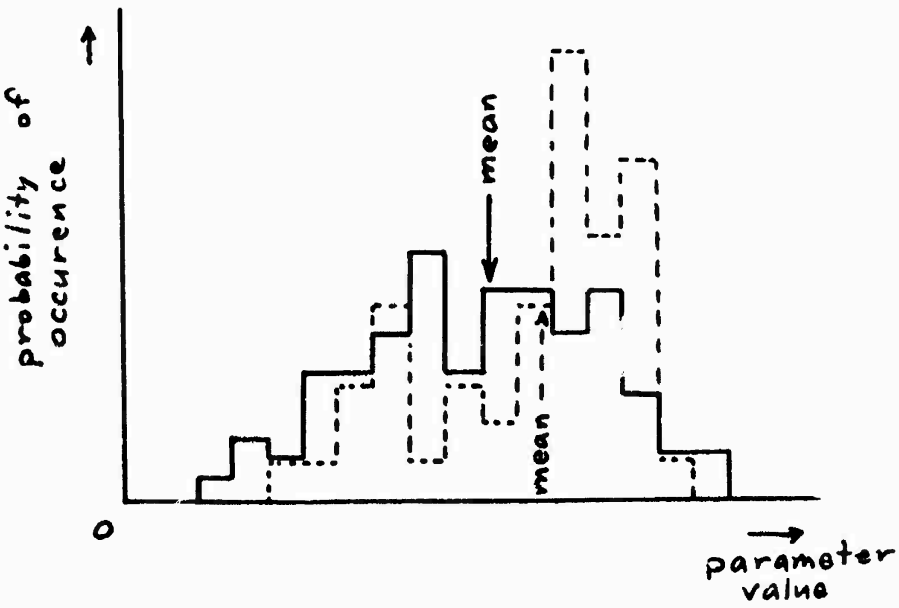


Figure 3. Frequency histograms for two distributions which have a test statistic of 0.33 .

Table 1. Best fit input parameters for statistical simulation

Device	N_d (cm^{-3})	v_s (cm/sec)	t (cm)	V_{bi} (volts)
A	0.78×10^{17}	1.3×10^{7}	0.25×10^{-4}	0.8
B	0.73×10^{17}	1.6×10^{7}	"	"
C	0.65×10^{17}	1.7×10^{7}	"	"
D	0.78×10^{17}	1.3×10^{7}	"	"
E	0.78×10^{17}	1.3×10^{7}	"	"

Table 2. Results of the Kolmogorov-Smirnov test on simulated and measured device parameter data.

	K-S test statistic value				
Parameter	A	B	C	D	E
I_{DSS}	0.272**	0.128*	0.225*	0.199*	0.169*
R_{ON}	0.530	0.211*	0.110*	0.220*	0.272**
R_{LIN}	0.287**	0.198*	0.437	0.514	0.546
R_D	0.296**	0.237**	0.232	0.204*	0.589
R_S	0.309**	0.229**	0.211	0.361	0.384
V_T	0.275**	0.128*	0.175*	0.249**	0.196*
Critical Value	0.233	0.225	0.241	0.232	0.225

* Less than critical value.

** Less than the acceptable value of 0.33 .

Table 3. Computed parameter correlation coeffients. The upper number in each box was computed from measured data and the lower from simulated data.

	I_{DSS}	R_{LIN}	R_{ON}	R_D	R_S
R_{LIN}	-0.956 -0.967				
R_{ON}	0.629 -0.969				
R_D	-0.455 -0.818	0.535 0.809	-0.142 0.797		
R_S	-0.921 -0.895	0.913 0.877	-0.549 0.893	0.256 0.556	
V_T	0.993 0.993			-0.498 -0.845	-0.892 -0.869

Table 4. Correlation coefficients vs. distance for interdevice parameters. (a) Results for wafer I. (b) Summary of results for wafer II.

Device Pair	Distance (mils)	Correlation Coefficient r_{IDSS}	r_{RLIN}	r_{RON}	r_{RD}	r_{RS}	r_{VT}	r_{GM}
D-E	11.3	0.895	0.803	0.519	0.386	0.544	0.901	0.532
A-B	14.3	0.913	0.909	0.542	0.781	0.906	0.906	0.759
C-E	32.0	0.886	0.840	0.420	0.459	0.492	0.895	0.443
A-C	37.3	0.927	0.923	0.595	0.720	0.933	0.907	0.807
B-C	37.7	0.914	0.915	0.435	0.881	0.931	0.908	0.819
C-D	42.0	0.900	0.855	0.605	0.788	0.932	0.900	0.682
B-E	58.8	0.910	0.859	0.378	0.422	0.510	0.911	0.540
A-E	66.7	0.889	0.820	0.481	0.404	0.520	0.907	0.554
B-D	70.3	0.893	0.843	0.289	0.825	0.907	0.905	0.667
A-D	75.7	0.944	0.946	0.600	0.783	0.928	0.942	0.780
Slope ($mils^{-1}$)		1.72×10^{-4}	2.83×10^{-4}	-1.30×10^{-3}	1.16×10^{-3}	8.58×10^{-4}	3.30×10^{-4}	4.08×10^{-4}
"t" statistic		0.076	0.130	-0.99	0.72	0.44	0.15	0.25

t(critical)=1.86 @ 0.95 level of significance

(a)

Slope ($mils^{-1}$)	-1.19×10^{-3}	1.13×10^{-2}	-3.81×10^{-3}	-2.33×10^{-3}	-5.43×10^{-4}	-3.66×10^{-4}	-2.67×10^{-3}
"t" statistic	-0.57	0.64	-3.96*	-1.51	-0.30	-0.19	-1.70

* critical value exceeded

(b)

REFERENCES

1. Lawrence Livermore National Laboratory, "Predicting Integrated Circuit Performance," Energy and Technology Review, August 1981, pp. 1-11.

2. R.L. Van Tuyl and C.A. Liechti, "Gallium Arsenide Digital Integrated Circuits." Air Force Avionics Lab., AFSC, WPAFB, Tech. Rept. AFAL-TR-74-40, March 1974.

3. A.R. Thorbjornsen, "Statistical Simulation of GaAs MESFETs," The University of Toledo, Final Rept. for Research Grant AFOSR-82-0119, August 1983.

4. R.A. Pucel, H.A. Haus, and H. Statz, "Signal and NoiseProperties of GaAs Microwave Field-Effect Transistors," in L. Martin, Editor, Advances in Electronics and Electron Physics, v. 38, Academic, New York, 1975, p. 195.

5. S.M. Sze, Physics of Semiconductor Devices, 2nd Ed., John Wiley & Sons, New York, 1981, p. 329.

6. H. Fukui, "Determination of the Basic Device Parameters of a GaAs MESFET," The Bell Sys. Tech. J., v. 58, n. 3, March 1979, pp. 771-797.

7. L.O. Chua and P.M. Lin, Computer-Aided Analysis of Electronic Circuits, Prentice-Hall, Englewood Cliffs, N.J., 1975, pp. 214-221.

8. W.R. Curtice, "A MESFET Model for Use in the Design of GaAs Integrated Circuits," IEEE Trans. on Microwave Theory & Tech., v. MTT-28, n. 5, May 1980, pp. 448-456.

9. A.R. Gideon and D.E. Muller, "Computation of the Two-Sample Smirnov

Statistics," The American Statistician, v. 32, n. 4, November 1978, pp. 136-137.

10. P.J. Kim and R.I. Jennrich, "Tables of the Exact Sampling Distribution of the Two-Sample Kolmogorov-Smirnov Criterion: $D_{m,n}$; $m \leq n$," in Selected Tables in Mathematical Statistics, v. 1, Cambridge Univ. Press, London, 1973, pp. 79-170.

11. M.A. Battista, et al., "Resistor Tracking on Monolithic Integrated Circuits," IBM Technical Disclosure Bulletin, July 1972, pp. 423-424.

12. A.R. Thorbjornsen, M.A. Dvorack, and A. Riad, "Correlation Between Thick-Film Resistance Values," IEEE Trans. on Parts, Hybrids, and Packaging, v. PHP-13, n. 2, June 1977, pp. 138-143.

13. M.R. Spiegel, Statistics, Schaum's Outline Series, McGraw-Hill, New York, 1961, p. 247.

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

DELIVERY ACCURACY

Prepared by:	Jon W. Tolle
Academic Rank:	Professor
Department and University:	Mathematics and Operations Research University of North Carolina, Chapel Hill
Research Location:	AFATL/DLYW, Eglin AFB, FL.
USAF Research:	Daniel A. McInnis
Date:	July 25, 1983
Contract No:	F49620-82-C-0035

DELIVERY ACCURACY

by

JON W. TOLLE

ABSTRACT

The Delivery Accuracy Working Group (DAWG) has recently begun to produce delivery accuracy figures for unguided air-to-surface bombs launched by high performance aircraft. This report contains a description and a critical study of the methodology behind the production of these figures. In addition, the predicted range errors are compared to actual errors observed in training tests for the F111 A/E aircraft. Recommendations are put forth for improvements in delivery accuracy analysis.

ACKNOWLEDGEMENTS

The author would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research, the Air Force Armament Laboratory, and SCEEE for their support. In addition the assistance of the individuals at DLYW and the Oklahoma State University Field Office at Eglin AFB, especially Dan McInnis and John Feldmann, is gratefully acknowledged.

I. INTRODUCTION:

The ability to correctly predict the accuracy with which unguided bombs can be delivered from high performance aircraft under a variety of conditions plays an important role in both the development of weapons systems and their implementation in combat situations. Clearly the accuracy of projected weapons systems should be a significant factor in making procurement decisions. In combat, the level of accuracy of various weapons systems is an essential ingredient in determining bomb loads and numbers of sorties against prescribed targets.

Although the study of bomb impact patterns is almost as old as bombing itself, technological developments have made these traditional studies inapplicable to the weapons delivery systems in modern aircraft. For instance the development of advanced computer-guided tracking and release equipment has decreased and/or changed the role of the aircrew in weapons delivery. In addition, the presence of highly sophisticated defensive firepower in target areas has made the usual constant flight path deliveries extremely hazardous, if not impossible. Consequently, there is a trend towards deliveries in which the aircraft performs complicated maneuvers relatively far from the target and the bombs are released automatically when the aircraft reaches a computer-designated state.

The analysis of delivery accuracy in such settings is a complex undertaking. Initially it involves a study of the trajectory of a given weapon to determine the appropriate launch parameters required for the weapon to strike the target (for a given aircraft maneuver). Next the components of an entire delivery process must be analyzed in order to identify those sources of error in the realization of these ideal parameters. These errors are both mechanical and pilot-induced in nature and may be affected by external factors such as the quality of the radar image of the target, the stress under which the aircrew is operating, or the wind conditions. Having identified these sources of error, the analyst must attempt to quantify them and deduce their combined effects on the trajectory of the weapon. The final step in the accuracy analysis is to compare the theoretical accuracy with data on

weapons accuracy gathered from actual combat or training records. This comparison is usually greatly complicated by the poor quality of such records.

In sum, the problem of establishing delivery accuracy is both important and difficult. Because of its difficulty there has been a tendency in the past to shy away from fully exploring the technical issues and, instead, to rely on the experimental data to establish accuracy figures. But the lack of combat experience with the newer equipment and the often prohibitively high cost of obtaining sufficient training data suggest that the technical analysis will have to be carried out if realistic delivery accuracy figures are to be obtained.

II. OBJECTIVES OF THE RESEARCH EFFORT:

An intraservice group, the Delivery Accuracy Working Group (DAWG), is charged with establishing delivery accuracy estimates for aircraft weapons systems as part of their task of updating the Joint Munitions Effectiveness Manual on Delivery Accuracy for air-to-surface weapons. DAWG is now in the final stages of producing the estimates for unguided bombs which will be included in the revised manual.

Mr Dan McInnis, a member of DAWG and my research colleague, suggested that I evaluate these proposed delivery accuracy figures. The basic objectives of this effort were:

(i) to determine if the approach being used by DAWG was valid;

(ii) to statistically compare, to the extent possible, the DAWG estimates with whatever relevant training and test data that could be found;

(iii) to make recommendations suggesting new or different avenues of approach which would improve the analysis of delivery accuracy and would be applicable to future developments in the weapons delivery field.

III. CURRENT METHODS OF COMPUTING DELIVERY ACCURACY:

In order to define delivery accuracy for unguided bombs four categories of information must be specified: The particular aircraft and its equipment, the flight path, relevant external factors, and the bomb type. Aircraft equipment includes the specific measuring devices used to locate the aircraft both absolutely and relative to the target as well as the means by which the target is designated and bomb release commanded. The flight path refers to the exact description of the aircraft trajectory and configuration prior to and at bomb release. External factors which affect accuracy include such conditions as type of target, the amount of stress on the aircrew, and wind conditions. Finally, the bomb type refers to the ballistic properties of the bomb. A specification of all of this information will hereafter be called the launch condition.

A standard theoretical approach to the accuracy problem can be formulated as follows for a given launch condition. First the trajectory is computed under the assumption that the launch condition is satisfied exactly. Then each measurement device in the system (including pilot/navigator judgements) which could cause a deviation from the desired launch condition through an incorrect reading is identified. For convenience we index these error sources by i, $i = 1, \ldots, n$ and denote the reading given by device i as μ_i and the true value of the quantity being measured as x_i. Given the lanuch parameter values $x_1, \ldots, x_n$, the bomb impact point can be theoretically computed as a two-dimensional vector-valued function

$$g(x_1, \ldots, x_n)$$

which gives the coordinates in the ground plane. These two coordinates are usually taken to be the range (along-track) and deflection (cross-track) components. When the readings are all correct i.e. $x_i = \mu_i$, $i = 1, \ldots, n$ the value

$$g(\mu_1, \ldots, \mu_n)$$

yields the nominal impact point for the given specifications (which we call the target).

The values X_i can be thought of as random variables with means μ_i and variances σ_i^2. Then the distribution of the two-dimensional random variable

$$z = g(x_1,\ldots,x_n) - g(\mu_1,\ldots,\mu_n) \tag{1}$$

represents the theoretical error distribution for the prescribed launch condition. It is important to note that the distribution of the X_i will depend on the specified launch condition. As examples, the error in the altimeter will depend on the altitude while the magnitude of error in the radar reading for range to target will depend upon the type of target. In any case the determination of the distribution of z or, more realistically, the estimation of its parameters is the core of the theoretical study of delivery accuracy.

As an example (relevant to the DAWG approach), if it is assumed that the X_i are <u>independent</u> random variables then the assumption that g is a linear function of the X_i, i.e.,

$$g(x_1,\ldots x_n) = \sum_{i=1}^{n} a_i x_i \tag{2}$$

where the a_i are two-dimensional vectors, results in

$$z = \sum_{i=1}^{n} a_i (x_i - \mu_i), \tag{3}$$

and

$$\mu_z = 0, \tag{4}$$

$$\sigma_z^2 = \sum_{i=1}^{n} a_i^2 \sigma_i^2, \tag{5}$$

where the a_i can be calculated directly by

$$a_i = \frac{1}{\sigma_i}\left\{ g(\mu_1,\ldots,\mu_i+\sigma_i,\ldots,\mu_n) - g(\mu_1,\ldots,\mu_n)\right\}. \tag{6}$$

Thus, if the σ_i are known, the means and standard deviations of z can be obtained by computing n trajectories, each corresponding to the inclusion of a single error source.

There are a variety of error sources which comprise the X_i. Omitting ballistic error due to the bomb itself, they can be divided into three classes; sensor errors, pilot errors, and post-release errors. Sensor errors represent errors in the navigational and targeting equipment such as altimeter and radar range finder. Pilot errors are judgement errors of the aircrew. For example a pilot may have the hud pipper slightly off target when the bomb is released. Post-

release errors represent errors in the mechanisms which release the bomb from the aircraft. Examples include time delays in release and ejection velocity errors.

The determination of the properties of the distribution of Z depends upon the joint distribution of the X_i. It is traditional to assume that the X_i are independently normally distributed. There is little hard evidence available to justify this assumption however. For the sensor and post-release errors, the values of the σ_i are often estimated from test data on the equipment or from manufacturer's specifications. For pilot errors the estimation is not at all simple. In some cases, such as parallax effects, errors can be estimated from the geometry. In other cases, as for instance in time-to-reaction errors, there is no clear theory available and estimates are often worst-case assumptions.

The DAWG approach to approximating the distribution parameters for Z is based on the following basic assumptions.

(i) The X_i are independently distributred normal variables, and

(ii) The function $g(x_1,\ldots,x_n)$ is well-approximated by a linear function of the X_i.

With these assumptions and given estimated values for the σ_i, the estimates of the parameters of Z important to delivery accuracy can be obtained through the use of equations (1) - (6). The evaluation of the Q_i in (6) requires obtaining the terms on the right-hand side of that equation. The vehicle for obtaining these values is the weapons delivery routine of the Avionics Evaluation Program (AEP) (see reference 2). This routine integrates the equations of motion of the aircraft and bomb under the hypothesized condition to evaluate $g(\mu_1,\ldots,\mu_n)$. Then $g(\mu_1,\ldots,\mu_i+\sigma_i,\ldots,\mu_n)$ is evaluated either by integrating the equations of motion for the perturbed system or by estimating the value of using a linear approximation of the geometry of the delivery configuration.

In order to obtain values for the σ_i, DAWG has investigated specifications for the equipment, looked at test results, and talked to aircrews and engineering personnel. Synthesizing this soft data they have come up with sets of values of the σ_i which reflect flight path

conditions and given external factors. Any confidence in the figures in the DAWG delivery accuracy tables reflects a confidence in the validity of these estimates for the σ_i .

One other special error included in the DAWG approach should be mentioned here. Because the AEP subroutine makes some simplifying assumptions in the calculations of trajectories there are some errors in the calculated values of the components of $\bar{z}$. The DAWG approach is to allow for the effect of this error by adding a general term to the value of σ_z^2. This is called the bomb fall algorithum error and is discussed further in Section VI.

IV. F111 DATA BASE:

One objective of this research effort was to compare the DAWG figurs with actual test data for different aircraft systems and launch conditions. Unfortunately there is a paucity of such data. The only usable data was a record of training runs for F111 A/E aircraft at various ranges during 1979 and 1980.

This data, in its raw form, consisted of report sheets filled out by the aircrews. In many cases the reports were incomplete and sometimes they were unintelligible. This data was compiled and tabulated by the Oklahoma State University station at Eglin AFB and placed on a computer file. The resulting file consists of the records of approximately 3000 individual bombs drops. Most of the records contains the altitude and dive angle at which the bomb was released as well as the radial miss distance and the deflection angle of the impact-point. In addition, the records included the tactic used in each run. The tactic encompassed the mode of delivery (radar or visual), the maneuver (constant flight path angle or loft), and the bomb type (high or low drag).

From this data base, a set of launch conditions were chosen such that for each condition there were enough completely reported runs to justify a statistical analysis of the results. This set is displayed in Table 1. In that table R=radar, V= visual, C=constant flight path angle, L= loft, HD=high drag, and LD=low drag. The dive angle is given

in degrees measured counterclockwise from the horizontal and the altitudes are measured in feet.

Tactic #	Mode	Maneuver	Bomb	Angle	Altitudes
11	V	C	HD	0	200,500
13	V	C	LD	-10	750
14	R	C	HD	0	200,400,500,1000
16	R	L	LD	20	500

Table 1. Launch Conditions Analyzed

In Table 2 the number of runs, the sample mean, $\bar{X}$, and the sample standard deviation, $\bar{S}$, of the range component of the miss for the various launch conditions are presented. Only the range errors are studied in this report. Generally the deflection errors are much smaller than the range errors and hence the statistical fluctuations are less. It was felt that any differences between the DAWG figures and the training data would be more easily detected in the range errors.

Tactic #	Altitude	Sample Size	$\bar{X}$	$\bar{S}$
11	200	156	-25	196
11	500	27	-100	382
13	750	130	9	171
14	200	81	-17	243
14	400	107	-38	235
14	500	101	55	320
14*	400-500	208	6	282
14	1000	59	105	332
16	500	42	-77	737

Table 2. Sample Data

It should be emphasized that this data is refined from poorly recorded aircrew observations. Thus, the recorded altitudes and dive angles at release are not to be taken as exact. In the starred entry in Table 2, the results of runs at 400 and 500 feet altitudes for tactic 14 have been combined as a further aid in analysis. A missing factor in the above data is the airspeed of the aircraft at delivery. Although not recorded on the aircrew sheets all the deliveries were reported to be executed at "about 500 knots". As will be seen, this omission is compensated for by condering theoretical deliveries in the range of 450-550 knots.

It is generally assumed that a normal distribution fits bomb data reasonably well. Accordingly, simple chi-square tests were performed for each of the launch conditions to determine if the data could be assumed to be taken from a normal distribution with mean 0. No reason was found to reject that hypothesis.

V. COMPARISON OF DAWG FIGURES AND F-111 DATA:

(This section is classified)

VI. CONCLUSIONS AND RECOMMENDATIONS:

A. Comments on the DAWG figures

As was observed in the preceding section there is little agreement between the DAWG predicted accuracy and the accuracy represented by the F111 training data. Unfortunately, it is difficult to tell which, if either, of the sets of data reflects the actual delivery accuracy of the F111 A/E aircraft.

On the one hand, the training data was culled from very sketchy reports written by the aircrews on completion of their runs. Inspection of those sheets leads one to question whether even the tabluated data can be taken seriously. As an example of the discrepancies, some crews had reported miss distances in the wrong units and although some of these mistakes were caught it is readily conceivable that many were not. This exercise illustrates the folly of trying to do delivery accuracy analysis on the basic of tests not designed for that purpose.

The more relevant question for this report is the validity of the DAWG figures. There is much that is praiseworthy in the derivation of these figures. DAWG has done a good job of tracking down the sources of error in the delivery systems and in ascertaining the dependence of these errors on the launch conditions. Moreover, they have made a start in the more complex task of isolating possible pilot errors in the system. These investigations are, to my knowledge, the first studies undertaken along these lines and they suggest a methodology for standardizing the analysis of delivery accuracy. (See the recommendations below).

It is the translation of this analysis of error sources into quantitative form that, in my opinion, puts the DAWG results on shaky ground. There are two primary reasons for my skepticism concerning the DAWG figures. First, in attempting to distill the contribution of each error source into a single number (in the absence of any data from tests designed to provide this information) they are potentially injecting a large error into the calculations. As seen in Section V, the total error is often very sensitive to a few specific errors. Second, the assumptions of normality, independence, and linearity (described in Section III) required by the use of the AEP program are probably unwarranted. In fact, the procedure implicitly recognizes the nonlinearity of one part of the process by including an absolute error, called the bomb fall algorithum error, to compensate for the fact that the linearized computations of the AEP program are not accurate.

B. Recommendations

1. The use of the AEP program for delivery accuracy analysis should be discontinued. A new model should be developed which will: (a) make full use of modern computer software technology; (b) provide an adequate simulation procedure for analyzing the effect of all errors; and (c) allow three-dimensional flight-paths.

The AEP program is at least 10 years old and therefore does not take advantage of the many imnprovements in computer technology which would enable a quicker and more accurate anlaysis of aircraft and weapons trajectories. Moreover, the limitation caused by the

one-error-at-a-time calculations of AEP precludes good statistical analysis of the delivery accuracy. Finally the AEP package is severely restricted in the type of flight paths it will accept. Loft deliveries are permitted only by ad hoc additions to the program and three-dimensional flight paths are not possible. In the future it is likely that these more complex flight paths will become common. (See, for example, references 3 and 4).

2. The effects of equipment errors and aircrew errors should be analyzed separately. That is, a detailed error analysis should be carried out for each launch condition assuming that the aircrew performs faultlessly. A separate analysis should then be done incorporating only aircrew errors.

The advantages of this approach are two-fold. First, it is quite reasonable to expect that a very accurate statistical analysis can be done where only equipment errors are included. It should be possible, say, by comparing on-board equipment readings with highly accurate ground radar or laser fixes, to obtain good sample estimates of the parameters of the distribution of the errors in the instrument readings. This would include determination of the dependence of the instrument errors on the flight path characteristics as well as the correlation, or lack thereof, among the individual errors. Given this data it would be possible, using an appropriate simulation program (as envisioned in recommendation 1), to obtain good sample estimates for the distribution of impact errors as well as confidence intervals for these estimates. A second advantage of this approach is that it would result in an optimal system accuracy standard. That is, the resulting impact error distribution would represent the best attainable given the equipment used. The skill and stress performance of the aircrew would not be a factor.

As suggested above, aircrew errors could be studied separately. Using the DAWG analysis as a starting point, further research could be done on obtaining statistical measures of these errors as, for example, a function of training. It would seem plausible to study these errors on aircraft simulators. Once these aircrew errors were quantified it would be an relatively simple matter to superimpose them on the instrument

errors and use the simulation program to obtain overall accuracy figures.

3. As a final recommendation, I would reiterate a suggestion that has been implicit in all of the foregoing. Namely, that tests of aircraft instruments and pilot skills be designed and carried out specifically for delivery accuracy analyses. The impression I have gained in my research is that delivery accuracy data is obtained only as a byproduct of other missions, e.g. aircrew training, bomb damage, etc. Since weapons effectiveness ultimately depends on delivery accuracy, it makes good sense to undertake a serious effort to obtain meaningful figures.

REFERENCE

1. Joint Munitions Effectiveness Manual, Air-to-Surface, Delivery Accuracy, USAF TH61A1-3-3, April 1970, Classified.

2. D.W. Welp, R.A. Brown, and R.W. Scott, "A Computer Program for Simulation and Effectiveness Evaluation of Aviances," AFAL-TR-73-44, Volume 1, February 1973

3. Captain M. Baker, "F-16 Loft", USAF Fighter Weapons Review, Winter 1982, pp 2-12.

4. R.R. Ropelewski, "Integrated Systems Evaluated on F-15," Aviation Week & Space Technology, April, 1983, pp 47-57.

1983 USAF SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

AN INTRODUCTORY DYNAMICAL THEORY FOR FULLY COMPRESSIBLE TURBULENCE

Prepared by: George Treviño

Academic Rank: Associate Professor

Department and University: Physics Department, Del Mar College

Research Location: Air Force Weapons Laboratory - ARLB

USAF Research: Maj Charles C. Allen

Date: July 1983

Contract No: F49620-82-C-0035

AN INTRODUCTORY DYNAMICAL THEORY FOR FULLY COMPRESSIBLE TURBULENCE

by

George Treviño

ABSTRACT

A stochastic theory for use in the study of the dynamics of fully compressible turbulence is presented for the case where the turbulence is isotropic. It is shown that the time-decay of all correlation tensors encountered in this theory is dependent upon the equations of both continuity and Navier-Stokes; the qualitative implication of this kinematic-kinetical "coupling" is that in the formulation of an appropriate "closure", kinematics (and kinematical constraints) must be explicitly included in the closure scheme. A closure is formulated by adapting the method of "modified tensor invariants".

ACKNOWLEDGEMENT

The author extends his sincere thanks to the Air Force Office of Scientific Research for its sponsorship of the Summer Faculty Research program, and to the Air Force Weapons Laboratory for its participation in same.

I. INTRODUCTION: The U.S. Air Force, through its Weapons Laboratory at Kirtland AFB, is critically interested in determining the effects of a turbulent fluid flow on the integrity of a monochromatic electromagnetic wave as the wave propagates through the flow. One of the fluid parameters which (indirectly) affects this integrity is the turbulent velocity, $\underset{\sim}{u}(\underset{\sim}{x},t)$, and the equation which governs the space-time behavior of $\underset{\sim}{u}(\underset{\sim}{x},t)$ is the Navier-Stokes equation of fluid dynamics. This equation, being a nonlinear partial differential equation in the stochastic variable $\underset{\sim}{u}(\underset{\sim}{x},t)$, can be solved in the stochastic sense only by formulating a "closure" for the appropriately statistically-averaged equation. Since the author recently had published some research results on the problem of "turbulence closure" (in the case of incompressible turbulence), and since the Air Force is particularly interested in the effects of compressibility, the author was assigned to ARLB to extend his recent closure work to the case of compressible homogeneous isotropic turbulence.

II. OBJECTIVES OF THE RESEARCH EFFORT: The research effort had two goals, and these remained unchanged throughout the ten-week tenure. The first goal was to test the appropriate theoretical predictions of a recently developed closure scheme for homogeneous isotropic incompressible turbulence by comparing these predictions with existing atmospheric turbulence data; the second was to extend that theory to the case of compressible turbulence. The first goal could not be achieved because the available data was not in a form amenable to the requirements of the theory. The approach taken toward the attainment of the latter goal was the theory of "modified tensor invariants".

III. THEORETICAL PRELIMINARIES: Various facets of the fundamental statistical theory of the dynamics of compressible turbulence have been proposed, formulated, and studied by several researchers[1-11]; the relation of compressible turbulence to the generation of sound is analyzed in such works (among others) as Refs. 12-17, and in a recent publication[18] the existence and uniqueness of a general solution to the Navier-Stokes equation for compressible fluids in general was analytically established. The purpose of this communication is to present the introductory features of a comprehensive stochastic theory of homogeneous isotropic compressible turbulence which are fundamental and essential to the formulation of a "closure" for the relevant dynamical equation. In a later communication it is the intention of the author to duly extend this theory to the nonhomogeneous case by accordingly adapting the theory developed in Ref. 19.

The most general exact equations of motion for an arbitrary compressible fluid under no external forces are [12]

$$\frac{\partial \rho}{\partial t} + \frac{\partial(\rho u_i)}{\partial x_i} = 0, \tag{1}$$

$$\frac{\partial(\rho u_i)}{\partial t} + \frac{\partial}{\partial x_j}(\rho u_i u_j + p_{ij}) = 0, \tag{2}$$

where

$$p_{ij} = p\delta_{ij} - \mu\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} - \frac{2}{3}\frac{\partial u_k}{\partial x_k}\delta_{ij}\right); \tag{3}$$

in Eqs. (1)-(3) p is the pressure, ρ the density, u_i, $i = 1,2,3$, the turbulent velocity (of zero mean-value), and μ the viscosity of the

fluid, all of these parameters characteristically being functions of both space and time, i.e. $\rho \sim \rho(\underset{\sim}{x},t)$, $p \sim p(\underset{\sim}{x},t)$, $u_i \sim u_i(\underset{\sim}{x},t)$, and $\mu = \nu\rho$. The parameter ν is the so-called "kinematic viscosity" and is here taken to be constant. Equation (1) is the equation for the condition of continuity, and is usually referred to as the equation of kinematics while Equ. (2) is the Navier-Stokes equation (Newton's Second Law), and is (usually) referred to as the equation of dynamics although, strictly speaking, is should be referred to as the equation of kinetics since dynamics is known to be composed of both kinematics and kinetics[20,21]; nevertheless, both of these equations together completely describe the total dynamics of the fluid phenomena.

IV. DYNAMICS OF TWO-POINT CORRELATIONS: In accordance with established principles for the statistical study of turbulence, the two-point correlation theory of compressible turbulence begins with Equ. (1) from which automatically follows

$$\frac{\partial \rho}{\partial t}\rho' + \frac{\partial(\rho u_i)}{\partial x_i}\rho' = 0 \ , \quad \rho\frac{\partial \rho'}{\partial t} + \rho\frac{\partial(\rho' u_i')}{\partial x_i} = 0 \ ,$$

where $\rho' = \rho(\underset{\sim}{x}',t)$, $u_i' = u_i(\underset{\sim}{x}',t)$, etc. Adding these two equations and subsequently taking mathematical expected values produces

$$\frac{\partial A_i(\underset{\sim}{r},t)}{\partial r_i} = \frac{\partial}{\partial t}\langle\rho\rho'\rangle, \tag{5}$$

where

$$A_i(\underset{\sim}{r},t) = \langle \rho u_i\rho' - \rho\rho' u_i' \rangle$$

and $\underset{\sim}{r} = \underset{\sim}{x}' - \underset{\sim}{x}$. This very same equation was originally derived by Chandrasekhar[6], and he cleverly used it to establish that under appro-

priate conditions the integral

$$\int_0^\infty r^2(\langle\rho\rho'\rangle - \langle\rho\rangle^2)dr \ , \ r = \left|\underset{\sim}{r}\right|, \tag{6}$$

is an invariant of the turbulent motion, and as such is quite analogous to "Loitsiansky's invariant" of the theory of incompressible turbulence. The "physical" meaning of this invariant is that "...the largest eddies in the density fluctuations are determined by the initial conditions of the problem and [therefore] represent permanent features of the system".[6]

In order to obtain the equation for the time-decay of $A_i(\underset{\sim}{r},t)$ we naturally begin, as is done in the theory of incompressible turbulence, with the equation of Navier-Stokes. Accordingly, we have

$$\left\langle \frac{\partial(\rho u_i)}{\partial t}\rho' \right\rangle + \left\langle \frac{\partial(\rho u_i u_j)}{\partial x_j}\rho' \right\rangle + \left\langle \frac{\partial p_{ij}}{\partial x_j}\rho' \right\rangle = 0,$$

$$\left\langle \frac{\partial(\rho' u_i')}{\partial t}\rho \right\rangle + \left\langle \frac{\partial(\rho' u_i' u_j')}{\partial x_j}\rho \right\rangle + \left\langle \frac{\partial p_{ij}'}{\partial x_j}\rho \right\rangle = 0,$$

and subtraction of these equations results in

$$\left\langle \frac{\partial(\rho u_i)}{\partial t}\rho' \right\rangle - \left\langle \frac{\partial(\rho' u_i')}{\partial t}\rho \right\rangle - \frac{\partial}{\partial r_j}\langle \rho u_i u_j \rho' + \rho' u_i' u_j' \rho \rangle + \cdots = 0, \tag{7}$$

where terms such as $\langle p_{ij}\rho'\rangle$, etc., are for present purposes ignored. Equation (7), although sound, is not yet in its proper time-decay form since the time-derivative of $A_i(\underset{\sim}{r},t)$ is

$$\frac{\partial A_i(\underset{\sim}{r},t)}{\partial t} = \left\langle \frac{\partial(\rho u_i)}{\partial t}\rho' + (\rho u_i)\frac{\partial \rho'}{\partial t} \right\rangle - \left\langle \frac{\partial \rho}{\partial t}(\rho' u_i') + \rho \frac{\partial(\rho' u_i')}{\partial t} \right\rangle,$$

and is

$$\neq \left\langle \frac{\partial(\rho u_i)}{\partial t}\rho' - \frac{\partial(\rho' u_i')}{\partial t}\rho \right\rangle,$$

the relevant time-decay term present in Equ. (7); however Equ. (7) can be put into its proper time-decay form if we make judicious use of the continuity condition to first obtain

$$<(\rho u_i)\,\frac{\partial \rho'}{\partial t}> + <(\rho u_i)\,\frac{\partial(\rho' u_j')}{\partial x_j}> = 0$$

$$<\frac{\partial \rho}{\partial t}\,(\rho' u_i')> + <\frac{\partial(\rho u_j)}{\partial x_j}\,(\rho' u_i')> = 0.$$

Subsequent subtraction of these two results then produces

$$<(\rho u_i)\,\frac{\partial \rho'}{\partial t} - \frac{\partial \rho}{\partial t}\,(\rho' u_i')> + \frac{\partial}{\partial r_j}<(\rho u_i)(\rho' u_j') + (\rho u_j)(\rho' u_i')> = 0\ ,$$

so that the time-decay term in Equ. (7) can be written as

$$<\frac{\partial(\rho u_i)}{\partial t}\,\rho' - \frac{\partial(\rho' u_i')}{\partial t}\,\rho> = \frac{\partial A_i}{\partial t} + \frac{\partial}{\partial r_j}<(\rho u_i)(\rho' u_j') + (\rho u_j)(\rho' u_i')>\ .$$

The appropriate equation for the <u>total</u> time-decay of $A_i(\underset{\sim}{r},t)$ is then

$$\frac{\partial A_i(\underset{\sim}{r},t)}{\partial t} - \frac{\partial}{\partial r_j}<\rho u_i u_j \rho' + \rho' u_i' u_j' \rho - \rho u_i \rho' u_j' - \rho u_j \rho' u_i'> + \cdots = 0, \tag{8}$$

and if we now define the second-order tensor

$$C_{ij}(\underset{\sim}{r},t) = <\rho\rho'(u_i - u_i')(u_j - u_j')>, \tag{9}$$

Equ. (8) accordingly rewrites as

$$\frac{\partial A_i(\underset{\sim}{r},t)}{\partial t} - \frac{\partial}{\partial r_j}\,C_{ij}(\underset{\sim}{r},t) + \cdots = 0. \tag{10}$$

Immediately we see that the not unexpected "closure problem" exists here too, since the time-decay of $A_i(\underset{\sim}{r},t)$ is specified in terms of a second tensor, viz. $C_{ij}(\underset{\sim}{r},t)$, whose time-structure is essentially unknown. This closure problem, however, is much more profound, and interesting, than the closure problem of incompressible isotropic turbulence since, <u>and this is extremely important</u>, in order to obtain

the above (total) time-decay equation for $A_i(\underline{r},t)$ the equation of kinematics necessarily had to be employed in the mathematical manipulations, a circumstance which is not correspondingly true for incompressible turbulence. Actually, it is not at all difficult to positively establish that the time-decay of all tensors encountered in the study of compressible turbulence unmistakably rests upon the equations of both kinematics and kinetics, a consequence of the very simple fact that all of the related tensors include ρ in their definition and the fluid density is now a distinct function of the time. If one then subscribes to the well-established notion[22] that a bonafide full solution to the classical problem of "closure" will in principle deal, as it inherently should, with all types of turbulence and not only with the homogeneous isotropic incompressible variety, then what the above dynamical results very clearly imply is that, contrary to what is (shall we say) "accepted wisdom", any closure scheme which relies primarily on kinetics and/or kinetical constraints for its physical and/or mathematical justification, and which does not explicitly incorporate kinematics as an integral part of the so-called "closure formulation", can not only not be seriously considered as a true solution to the problem, but moreso, cannot even be considered as a sufficiently valid "first approximation". For if such a "limited scheme" were to be structured, i.e. without the explicit inclusion of kinematics within the essential features of the "closure", such a scheme might approximately apply to some types of homogeneous isotropic incompressible turbulence and maybe even to some types of nonhomogeneous anisotropic incompressible turbulence, but such a scheme could never apply to any practical compressible turbulence

since, because of its limited scope, it could never effectively capture the total time-decay of such a fluid. This is not to say, however, that "kinetical closures", and/or "kinetical closure theories", for use in the study of incompressible flows cannot be, in some logical fashion, systematically created and studied; however it is important to realize that such "closures" can at best shed only half of the total light required to clearly see the exact solution to the true problem. It therefore follows that even for incompressible turbulence any closure scheme that reasonably unites both kinematics and kinetics in the closure formulation, although such a unification may not at all be necessary for that case, necessarily has to be more general and viable than any scheme which unfortunately does not. This particular result, to say the least, is very fundamental because heretofore it was generally accepted that any dynamical results derived from a closure analysis which relied even mildly on kinematics necessarily, by their very nature, had to contain a basic fallacy; indeed, it was openly thought that kinematics could be employed only to obtain "bounds" on the closure and that not much (if anything) could be said about its exact functional form without explicitly invoking the Navier-Stokes condition.[23] This misconception should be partially eradicated by the present results, and completely eradicated by the results of Ref. 24. And although it could be validly argued that compressible flows are entirely different from incompressible flows, and therefore that a "closure" for one type of flow is essentially different from a "closure" for the other type of flow, a true closure undeniably applies to all types of flows.

At this point is is possible to correspondingly develop a time-decay equation for $C_{ij}(\underset{\sim}{r},t)$, and it is obvious that such an equation will inevitably contain third-order terms of the form $\frac{\partial}{\partial r_k} S_{ijk}(\underset{\sim}{r},t)$, where

$$S_{ijk}(\underset{\sim}{r},t) = < \rho\rho'(u_i - u_i')(u_j - u_j')(u_k - u_k')> .$$

A "closure problem" exists there too but this is essentially the same problem as before— since the total time-decay of $C_{ij}(\underset{\sim}{r},t)$ is dependent upon the equations of both kinematics and kinetics, kinematics and kinematical constraints must therefore be explicitly included in the appropriate closure formulation. Such an approach has already been suggested by the author for the case of incompressible isotropic turbulence[25,26], and a more thorough and complete solution for both that case and the isotropic compressible case is presented in Ref. 24.

V. KINEMATICAL CONSTRAINTS FOR COMPRESSIBLE TURBULENCE: In order to obtain the equations of kinematics for the tensors $C_{ij}(\underset{\sim}{r},t)$ and $S_{ijk}(\underset{\sim}{r},t)$ introduced earlier, it is necessary to employ only Equ. (1) and certain elaborate variations of it; the equation of kinematics for $A_i(\underset{\sim}{r},t)$ was written in Section IV (Equ. (5)). For the tensor

$$C_{ij}(\underset{\sim}{r},t) = <\rho\rho' u_i u_j - \rho\rho' u_i u_j' - \rho\rho' u_i' u_j + \rho\rho' u_i' u_j'>$$

we begin with the continuity condition in its standard form and proceed to write

$$< \frac{\partial \rho}{\partial t} u_j \rho' > + < \frac{\partial(\rho u_i u_j)}{\partial x_i} \rho' > = <\rho\rho' u_i \frac{\partial u_j}{\partial x_i} > .$$

Since $(\rho\rho')$ is a scalar quantity, the term $\langle \rho\rho' u_i \frac{\partial u_j}{\partial x_i} \rangle$ in the above is a first-order tensor whose tensor character is determined only by the product[4] $(u_i \frac{\partial u_j}{\partial x_i})$; the expected value of this product is

$$\langle u_i \frac{\partial u_j}{\partial x_i} \rangle = -u^2 \left[\frac{\partial R_{ij}}{\partial r_i} \right]_{\underset{\sim}{r}=0}$$

where

$$\langle u_i(\underset{\sim}{x}) u_j(\underset{\sim}{x}') \rangle = u^2 R_{ij}(\underset{\sim}{r}) \ .$$

It therefore follows that

$$\langle \rho\rho' u_i \frac{\partial u_j}{\partial x_i} \rangle = c(r,t) \left[\frac{\partial R_{ij}}{\partial r_i} \right]_{\underset{\sim}{r}=0},$$

$c(r,t)$ being a sufficiently well-defined even function of r; and since $\left[\frac{\partial R_{ij}}{\partial r_i} \right]_{\underset{\sim}{r}=0} \equiv 0$, it follows that

$$\langle \rho\rho' u_i \frac{\partial u_j}{\partial x_i} \rangle \equiv 0,$$

and that

$$\langle \frac{\partial \rho}{\partial t} u_j \rho' \rangle + \frac{\partial}{\partial x_i} \langle \rho u_i u_j \rho' \rangle = 0 \ .$$

Using this "modified" continuity condition and the continuity condition in its standard form eventually produces

$$\frac{\partial}{\partial r_i} C_{ij}(\underset{\sim}{r},t) = \langle \frac{\partial(\rho\rho')}{\partial t} (u_j - u_j') \rangle \ ;$$

if we now interpret Equ. (5) as an operator equation, i.e. as

$$\frac{\partial}{\partial r_i} \langle (\cdots)(u_i - u_i') \rangle = \frac{\partial \langle (\cdots) \rangle}{\partial t} \ ,$$

we then have

$$\frac{\partial^2}{\partial r_i \partial r_j} C_{ij}(\underset{\sim}{r},t) = \frac{\partial}{\partial r_j} \langle \frac{\partial(\rho\rho')}{\partial t} (u_j - u_j') \rangle = \frac{\partial^2}{\partial t^2} \langle \rho\rho' \rangle \qquad (11)$$

as the kinematic constraint on $C_{ij}(\underset{\sim}{r},t)$. A correspondingly similar conclusion results for the case of $S_{ijk}(\underset{\sim}{r},t)$, viz. that

$$\frac{\partial^3}{\partial r_i \partial r_j \partial r_k} S_{ijk}(\underset{\sim}{r},t) = \frac{\partial^3}{\partial t^3} \langle \rho\rho' \rangle \ , \qquad (12)$$

but in order to establish this latter result it is necessary to make use of the continuity condition in the following form:

$$\rho' \frac{\partial \rho}{\partial t} u_i u_j + \rho' \frac{\partial(\rho u_k u_i u_j)}{\partial x_k} = \rho\rho' u_k \frac{\partial(u_i u_j)}{\partial x_k} \ ,$$

where

$$\langle \rho\rho' u_k \frac{\partial(u_i u_j)}{\partial x_k} \rangle = -d(r,t) \left[\frac{\partial S_{ijk}(\underset{\sim}{r},t)}{\partial r_k}\right]_{\underset{\sim}{r}=0} \equiv 0.$$

Naturally, a corresponding invariant condition can be established for the case of $C_{ij}(\underset{\sim}{r},t)$, and for $S_{ijk}(\underset{\sim}{r},t)$, since

$$\frac{\partial^2}{\partial r_i \partial r_j} C_{ij}(\underset{\sim}{r},t) = (r \frac{\partial}{\partial r} + 3)(r \frac{\partial F}{\partial r} + 4F + \frac{1}{r} \frac{\partial G}{\partial r}) = \frac{\partial^2}{\partial t^2} [\langle \rho\rho' \rangle - \langle \rho \rangle^2],$$

where

$$C_{ij}(\underset{\sim}{r},t) = F(r,t) r_i r_j + G(r,t) \delta_{ij} \ ;$$

following Chandrasekhar[6], we accordingly have

$$\frac{\partial^2}{\partial t^2} \int_0^\infty r^2 [\langle \rho\rho' \rangle - \langle \rho \rangle^2] \, dr = 0,$$

and this constraint is obviously automatically satisfied if the previously derived invariant (Equ. (6)) is satisfied.

VI. CLOSURE OF THE DYNAMICAL EQUATION: In order to write the complete dynamical equation for the time-decay of $A_i(\underset{\sim}{r},t)$ it is first necessary to evaluate such terms as $\frac{\partial}{\partial x_j} \langle \rho\rho' (\frac{\partial u_i}{\partial x_j}) \rangle$, etc.. Accordingly, we write

$$\frac{\partial}{\partial x_j} \langle \rho\rho' \left(\frac{\partial u_i}{\partial x_j}\right)\rangle = \frac{\partial}{\partial x_j} \left\langle \frac{\partial(\rho\rho' u_i)}{\partial x_j} - \frac{\partial(\rho\rho')}{\partial x_j} u_i \right\rangle ,$$

and subsequent differentiation produces

$$\frac{\partial^2 \langle \rho\rho' u_i \rangle}{\partial r_j \partial r_j} - \left\langle \rho' \frac{\partial^2 \rho}{\partial x_j \partial x_j} u_i \right\rangle - \left\langle \rho' \frac{\partial \rho}{\partial x_j} \frac{\partial u_i}{\partial x_j} \right\rangle .$$

Since

$$\left\langle \frac{\partial^2 \rho}{\partial x_j \partial x_j} u_i \right\rangle = \left[\frac{\partial^2 a_i(\underset{\sim}{r},t)}{\partial r_j \partial r_j} \right]_{\underset{\sim}{r}=0}, \text{ etc.},$$

where

$$a_i(\underset{\sim}{r},t) = \langle \rho(\underset{\sim}{x},t) u_i(\underset{\sim}{x}',t) \rangle = \alpha(r,t) r_i ,$$

we eventually have

$$\frac{\partial}{\partial x_j} \langle \rho\rho' \left(\frac{\partial u_i}{\partial x_j}\right)\rangle \equiv \frac{\partial^2 \langle \rho\rho' u_i \rangle}{\partial r_j \partial r_j} ; \tag{13}$$

the complete dynamical equation includes the term

$$\left\langle \frac{\partial p_{ij}}{\partial x_j} \rho' \right\rangle - \left\langle \frac{\partial p_{ij}'}{\partial x_j} \rho \right\rangle ,$$

which in view of Equ. (3) expresses as

$$\left\langle \frac{\partial p}{\partial x_i} \rho' \right\rangle - \left\langle \frac{\partial p'}{\partial x_i} \rho \right\rangle + \nu \frac{\partial}{\partial x_j} \left\langle \rho\left(-\frac{\partial u_i}{\partial x_j}\right)\rho' \right\rangle - \nu \frac{\partial}{\partial x_j} \left\langle \rho'\left(-\frac{\partial u_i'}{\partial x_j}\right)\rho \right\rangle$$

$$+ \nu \frac{\partial}{\partial x_j} \left\langle \rho\left(-\frac{\partial u_j}{\partial x_i}\right)\rho' \right\rangle - \nu \frac{\partial}{\partial x_j} \left\langle \rho'\left(-\frac{\partial u_j'}{\partial x_i}\right)\rho \right\rangle$$

$$+ \left(\frac{2\nu}{3}\right) \left[\frac{\partial}{\partial x_j} \left\langle \rho \frac{\partial u_k}{\partial x_k} \rho' \right\rangle - \frac{\partial}{\partial x_j} \left\langle \rho' \frac{\partial u_k'}{\partial x_k} \rho \right\rangle \right] \delta_{ij} .$$

Enforcing the result established by Equ. (13) yields a time-decay equation for $A_i(\underset{\sim}{r},t)$ (cf. Equ. (10)), viz.

$$\frac{\partial A_i(\underset{\sim}{r},t)}{\partial t} - \frac{\partial C_{ij}(\underset{\sim}{r},t)}{\partial r_j} - 2\frac{\partial <p\rho'>}{\partial r_i} - \nu\frac{\partial^2 A_i(\underset{\sim}{r},t)}{\partial r_j \partial r_j} - (\nu/3)\frac{\partial^2 A_j(\underset{\sim}{r},t)}{\partial r_i \partial r_j} = 0. \tag{14}$$

Note that this equation is identically satisfied for $\underset{\sim}{r} = 0$ so that "one-point closures" have no appropriate meaning for compressible turbulence. Also note that $<p\rho'>$ can be formulated in terms of $A_i(\underset{\sim}{r},t)$ simply by assuming that $p = RT\rho$, where T is the temperature and is here (cautiously) taken to be constant. With this assumption it follows that

$$\frac{\partial <p\rho'>}{\partial r_i} = RT\frac{\partial <\rho\rho'>}{\partial r_i} = RT\int_{t_0}^{t}\frac{\partial^2 A_j(\underset{\sim}{r},t)}{\partial r_i \partial r_j}\,dt\ .$$

A closure for the complete equation is formulated by adapting the theory of Ref. 25, and by employing said theory, it is possible to generate several "bonafide" closures, some of which are

$$C_{ij}(\underset{\sim}{r},t) = H(r,t)[A_i(\underset{\sim}{r},t)r_j + A_j(\underset{\sim}{r},t)r_i]\ ,$$

$$C_{ij}(\underset{\sim}{r},t) = H_1(r,t)A_i(\underset{\sim}{r},t)A_j(\underset{\sim}{r},t)\ ,$$

$$C_{ij}(\underset{\sim}{r},t) = H_2(r,t)[A_i(\underset{\sim}{r},t)\int A_j(\underset{\sim}{r},t)dr + A_j(\underset{\sim}{r},t)\int A_i(\underset{\sim}{r},t)dr]\ ,$$

$$C_{ij}(\underset{\sim}{r},t) = H_3(r,t)[A_i(\underset{\sim}{r},t)\frac{\partial}{\partial t}\int A_j(\underset{\sim}{r},t)dr + A_j(\underset{\sim}{r},t)\frac{\partial}{\partial t}\int A_i(\underset{\sim}{r},t)dr]\ .$$

From among these we select for use here the first, since this appears to be the "simplest" and such a correspondingly simple closure has already produced "adequate" results for the incompressible isotropic case[24,26]. Proceeding accordingly, the continuity constraint, Equ. (11), eventually provides

$$\frac{\partial^2}{\partial r^2}\{r^2 A(r,t)H(r,t)\} = \frac{1}{2}\frac{\partial^2 <\rho\rho'>}{\partial t^2} = \frac{1}{2}\frac{\partial^2 A_i(\underset{\sim}{r},t)}{\partial t \partial r_i}\ ,$$

where $A_i(\underset{\sim}{r},t) = A(r,t)r_i$; the function $H(r,t)$ is then found to be

$$H(r,t) = [2r^2A(r,t)]^{-1} \int_0^r \int_0^\xi \frac{\partial^2 A_i(\underset{\sim}{\lambda},t)}{\partial t \partial \lambda_i} d\lambda d\xi, \quad \lambda = \left|\underset{\sim}{\lambda}\right| . \tag{15}$$

The resulting dynamical (scalar) equation in its "closed" form is respectively

$$\frac{\partial A}{\partial t} - 2[\frac{\partial}{\partial r}(rAH) + 3AH] - 2RT \int_{t_0}^{t} (\frac{\partial^2 A}{\partial r^2} + \frac{4}{r}\frac{\partial A}{\partial r})dt - (4\nu/3)(\frac{\partial^2 A}{\partial r^2} + \frac{4}{r}\frac{\partial A}{\partial r}) = 0 , \tag{16}$$

where $H(r,t)$ is given by Equ. (15); noting that

$$r^4 \frac{\partial^2 A}{\partial r^2} + 4r^3 \frac{\partial A}{\partial r} \equiv \frac{\partial}{\partial r}(r^4 \frac{\partial A}{\partial r}) ,$$

Equ. (16) conveniently condenses to

$$\frac{\partial A}{\partial t} - 2[\frac{\partial}{\partial r}(rAH) + 3AH] - 2RTr^{-4} \int_{t_0}^{t} \frac{\partial}{\partial r}(r^4 \frac{\partial A}{\partial r})dt - (4\nu/3)r^{-4} \frac{\partial}{\partial r}(r^4 \frac{\partial A}{\partial r}) = 0. \tag{17}$$

VII. CONCLUDING REMARKS: In examining the overall character of Equ. (14) it is important to remember that the tensors $A_i(\underset{\sim}{r},t)$ and $C_{ij}(\underset{\sim}{r},t)$ present there are defined as

$$A_i(\underset{\sim}{r},t) = \langle \rho\rho'(u_i - u_i') \rangle ,$$

$$C_{ij}(\underset{\sim}{r},t) = \langle \rho\rho'(u_i - u_i')(u_j - u_j') \rangle ,$$

and not as defined for the study of incompressible isotropic turbulence. The source of this peculiar definition, recall, lies back in Equ. (5) where it was initially desired to obtain a constraint equation for the time-derivative of $\langle \rho\rho' \rangle$, and it is there that the essential difference between the statistical theory of compressible turbulence and the classical theory of incompressible turbulence first presents itself. The

question that naturally then arises is that since incompressible turbulence is only a "special case" of compressible turbulence, then why does not the herein developed statistical theory of compressible turbulence reduce to that of incompressible turbulence simply by setting ρ equal to a constant in the formulated compressible theory, as it is obvious from "physical considerations" that is surely should. The reason that it does not is that in the classical theory of turbulence the assumption of incompressibility is made, rather rightly or wrongly, for the sake of "simplicity" rather early in the analysis of the problem, and in so doing some of the "richness" of the problem is obviously removed from it; further simplification is then introduced when it is assumed that velocity products such as $\langle u_i u_j' \rangle$, $\langle u_i u_j u_k' \rangle$, etc., are <u>isotropic</u>, although the isotropy of "velocity differences" is perhaps the better approximation in the majority of cases of practical interest. In fact, it was the invariance of probability distribution of <u>velocity differences</u> that was suggested by Kolmogorov in his original hypothesis[27], since it was felt by him that isotropy in the sense of Taylor[28] was a rather poor approximation of reality. By reformulating the incompressible theory in terms of velocity differences as

$$\frac{\partial}{\partial x_i}(u_i - u_i') = 0, \text{ (condition of continuity)}$$

$$\frac{\partial(u_i - u_i')}{\partial t} + \frac{\partial(u_i u_k)}{\partial x_k} - \frac{\partial(u_i' u_k')}{\partial x_k} + \ldots = 0 \text{ , (Navier-Stokes equation) ,}$$

it is then a rather easy task to deduce that the theory developed herein reduces, as it necessarily should, to the (reformulated) classical theory simply by setting ρ equal to a constant. What these arguments

are collectively meant to suggest is that in the formulation of the statistical theory of compressible turbulence the concept of isotropy of "velocity differences", in accordance with the principles established by Kolmogorov, is introduced as a matter of due course, and accordingly this theory is then a more "natural" theory than the already established incompressible theory; correspondingly, the "kinematic-kinetical coupling" of the closure formulation emerges, also "naturally" (as it should), since the true closure problem is indeed a dynamical problem in toto, and not simply a kinetical one as it is customarily construed to be.

In closing it is instructive to point out that the compressible theory proposed here could just as readily be formulated from the wavenumber space approach; for such an approach we would require Fourier-Stieltjes integrals of the form

$$\rho(\underset{\sim}{x},t) = (2\pi)^{-3} \int \exp(\imath \underset{\sim}{\kappa}\cdot\underset{\sim}{x}) dG(\underset{\sim}{\kappa},t)$$

and

$$u_i(\underset{\sim}{x},t) = (2\pi)^{-3} \int \exp(\imath \underset{\sim}{\kappa}\cdot\underset{\sim}{x}) dZ_i(\underset{\sim}{\kappa},t) \ .$$

Products such as $\langle \rho\rho' u_i' \rangle$ would then be

$$\langle \rho\rho' u_i' \rangle = (2\pi)^{-9} \int \exp[\imath(\underset{\sim}{\kappa}\cdot\underset{\sim}{x} - \underset{\sim}{\kappa}''\cdot\underset{\sim}{x}' + \underset{\sim}{\kappa}'\cdot\underset{\sim}{x})] \langle dG(\underset{\sim}{\kappa},t) dG^*(\underset{\sim}{\kappa}'',t) dZ_i(\underset{\sim}{\kappa}',t) \rangle ,$$

$*$ denoting complex conjugate, and since this product must be a function of $\underset{\sim}{r} = \underset{\sim}{x}' - \underset{\sim}{x}$ it follows that $\underset{\sim}{\kappa} - \underset{\sim}{\kappa}'' + \underset{\sim}{\kappa}' = 0$, and that

$$\langle \rho\rho' u_i' \rangle = (2\pi)^{-9} \int \exp[-\imath(\underset{\sim}{\kappa} + \underset{\sim}{\kappa}')\cdot\underset{\sim}{r}] \langle dG(\underset{\sim}{\kappa},t) dG^*(\underset{\sim}{\kappa} + \underset{\sim}{\kappa}',t) dZ_i(\underset{\sim}{\kappa}',t) \rangle \ . \tag{18}$$

In this approach the corresponding "closure problem" is the so-called problem of "spectral-transfer"; it is obvious from Equ. (18) that the

relevant "spectral-transfer" will not be a simple quantity. Recalling the related discussion of Section 2, it is also obvious that the solution to the "problem of spectral-transfer" will necessarily require an explicit kinematic-kinetical "coupling" in its appropriate formulation.

VIII. RECOMMENDATIONS: In order to properly implement the product(s) of the author's summer research, its analytical results should first be validated with compressible turbulence data. The tensors

$$A_i(\underset{\sim}{r},t) = \langle \rho\rho'(u_i - u_i') \rangle$$

and

$$C_{ij}(\underset{\sim}{r},t) = \langle \rho\rho'(u_i - u_i')(u_j - u_j') \rangle$$

should be approximated from turbulence data records, and the "strength" of the closure,

$$C_{ij}(\underset{\sim}{r},t) = H(r,t)\ [A_i(\underset{\sim}{r},t)r_j + A_j(\underset{\sim}{r},t)r_i] \ ,$$

should be clearly ascertained. Related follow-on research should proceed to the case of nonhomogeneous incompressible turbulence, and untimately to nonhomogeneous compressible turbulence. It is the intention of the author to pursue these investigations at a later date.

REFERENCES

1. Yaglom, A. M., "Homogeneous and Isotropic Turbulence in a Viscous Compressible Fluid", Izvestiya Akademia Nauk SSSR, Vol. 12, No. 6, 1948, pp. 501-522.
2. Krzywoblocki, M. Z. E., "On the Two-Dimensional Steady Turbulent Flow of a Compressible Fluid Far Behind a Solid Symmetrical Body. I.(Momentum Transfer Theory).", Journal of the Franklin Institute, Vol. 247, 1949, pp. 33-61.
3. Krzywoblocki, M. Z. E., "On the Invariant Theory of Isotropic Turbulence in Compressible Fluids", unpublished Naval Ordinance Laboratory report, White Oak, MD, 1949.
4. Krzywoblocki, M. Z. E., "On the Generalized Fundamental Equations of Isotropic Turbulence in Compressible Fluids and in Hypersonics", Proceedings of the First U.S. National Conference of Applied Mechanics, 1951. pp. 827-835.
5. Krzywoblocki, M. Z. E., "On Locally Isotropic Turbulence in Compressible Fluids", Proceedings of the Second Midwestern Conference on Fluid Mechanics, 1952, pp. 35-47.
6. Chandrasekhar, S., "The Fluctuations of Density in Isotropic Turbulence", Proceedings of the Royal Society, Vol. 210A, 1951, pp. 18-25.
7. Moyal, J. E., "The Spectra of Turbulence in a Compressible Fluid; Eddy Turbulence and Random Noise", Proceedings of the Cambridge Philosophical Society, Vol. 48, 1952, pp. 329-344.
8. Kovasznay, L. S. G., "Turbulence in Supersonic Flow", Journal of the Aeronautical Sciences, Vol. 20, No. 10, 1953, pp. 657-674.

9. Sitnikov, K. A., "Invariants of Homogeneous Isotropic Turbulence in a Compressible Viscous Fluid", Doklady Akademia Nauk SSSR, Vol. 122, No. 1, 1958, pp. 29-32.

10. Gavrilin, B. L., and Zaslavskii, M. M., "Lagrangian Invariants in the Dynamics of Nonviscous Compressible Fluids", Soviet Physics - Doklady, Vol. 15, No. 4, 1970, pp. 436-938.

11. Moiseev, S. S., Tur, A. V., and Yanovskii, V. V., "Influence of Compressibility on the Vortex-Turbulence Spectrum", Soviet Physics - Doklady, Vol. 25, No. 2, 1980, pp. 110-111.

12. Lighthill, M. J., "On Sound Generated Aerodynamically, I. General Theory", Proceedings of the Royal Society of London, Vol. 211A, 1952, pp. 564-586.

13. Proudman, I., "The Generation of Noise by Isotropic Turbulence", Proceedings of the Royal Socity A, Vol. 214, 1952, pp. 119-132.

14. Lighthill, M. J., "On Sound Generated Aerodynamically, II. Turbulence as a Source of Sound", Proceedings of the Royal Society A, Vol. 222, 1954, pp. 1-32.

15. Zakharov, V. E., and Sagdeev, R. Z., "Spectrum of Acoustic Turbulence", Soviet Physics - Doklady, Vol. 15, No. 4, 1970, pp. 439-441.

16. Moiseev, S. S., Sagdeev, R. Z., Tur, A. V., and Yonovskii, V. V., "Structure of Acoustic-Vortical Turbulence", Soviet Physics - Doklady, Vol. 22, No. 10, 1977, pp. 582-584.

17. L'vov, V. S., and Mikhailov, A. V., "Sound and Hydrodynamic Turbulence in a Compressible Liquid", Soviet Physics JETP, Vol. 47, No. 4, 1978, pp. 756-762.

18. May, M., "On the Theory of Compressible Fluids", ZAMM, Vol. 62, 1982, pp. 609-613.

19. Treviño, G., "The Invariant Theory of Nonhomogeneous Turbulence", To be published.

20. Beer, F. P., and Johnston, E. R., Mechanics for Engineers, McGraw-Hill Book Company, New York, 1957.

21. Robertson, J. M., Hydrodynamics in Theory and Application, Prentice-Hall, Englewood Cliffs, NJ, 1965.

22. Leslie, D. C., Developments in the Theory of Turbulence, Clarendon Press, Oxford, 1973.

23. Kraichnan, R. H., Reviewer comments on a consequently unpublished paper by G. Treviño; Reviewer comments forwarded to G. Treviño by Prof. A. Acrivos, Editor, Physics of Fluids, 1982.

24. Treviño, G., "Tensor Invariant Theory and the Problem of Turbulence Closure", Contract Report submitted to Air Force Weapons Lab., June 1983.

25. Treviño, G., "An Invariant Theory Approach to the Problem of Closure", Journal of Boundary-Layer Meteorology, Vol. 22, No. 4, 1982, pp. 463-474.

26. Treviño, G., "Closure and Isotropic Turbulence", in Proceedings of the Fourth International Conference on Physico-Chemical Hydrodynamics, Annals of the New York Academy of Sciences, Vol. 404, 1983, pp. 83-84.

27. Kolmogorov, A., "The Local Structure of Turbulence in Incompressible Viscous Fluids for Very Large Reynolds' Numbers", Comptes Rendus (Doklady) de l'Academie des Sciences de l'URSS, Vol. 30, No. 4, 1941, pp. 301-305.

28. Taylor, G. I., "Statistical Theory of Turbulence, Parts I-IV". Proceedings of the Royal Society A, Vol. 151, 1935, pp. 421-478.

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

ELECTRONIC EXCITATION OF ATOMIC XENON BY ELECTRON IMPACT

Prepared by:	Dr. Keith G. Walker
Academic Rank:	Professor
Department and University:	Department of Physics Bethany Nazarene College Bethany, OK 73008
Research Location:	Aero Propulsion Laboratory Wright-Patterson AFB, OH
USAF Research:	Capt. Greg Schneider
Date:	September 7, 1983
Contract No:	F49620-82-C-0035

ELECTRONIC EXCITATION OF ATOMIC XENON BY ELECTRON IMPACT

by

Keith G. Walker

ABSTRACT

The 6p and 6p' optical excitation functions of atomic xenon have been measured from onset to 100 eV's. Particular interest has centered on their pressure dependence. Most of the levels betray a pressure sensitivity for electron energies greater than 25 eV's. Higher np, ns and nd levels have been measured and many of these excitation functions exhibit a pressure dependence. Resonant type structure around 20 eV has also been observed in many of the functions. The xenon excitation functions have been measured when various amounts of helium are mixed with the xenon. Some functions are changed drastically while others appear 'inert' against the helium. Possible mechanisms for the above behavior are discussed.

Acknowledgement

The author would like to express his graditude to the various agencies involved in the administration and promotion of the Summer Faculty Research Program. The efficiency of administration by the Southeastern Center for Electrical Engineering Education was very helpful and appreciated. The friendliness and helpfulness of the personnel in the Power Division of the Aero Propulsion Lab made the summer pass all too quickly.

Two people stand out for special acknowlegdement. Capt. Greg Schneider provided tremendous help in the procurement of equipment and supplies...he was always ready with a helping hand. Dr. Alan Garscadden was an outstanding catalyst in the formation of ideas and inspiration. His insights and helpful discussions proved invaluable throughout the summer.

Introduction

Recent interest in high pressure rare gas laser systems has resulted in increased need to understand the various mechanisms which determine the laser's gain and intensity. The Air Force is particularly interested in the xenon high pressure gas laser in that it has strong laser lines falling between 1.73 and 4 microns. The importance of such a wavelength region lies within the fact that this corresponds to transmission 'windows' of the atmosphere. As such, this laser could have important uses in offensive and/or defensive purposes.

One of the factors which influences the gain of the xenon laser is the electron excitation cross sections of its various excited states from the ground state. Figure (1) is an energy level diagram of xenon. The ground state of xenon is a $5p^6$ state. The excited states are hence designated as $5p^5ns$, $5p^5np$, $5p^5nd$, and so forth. The l quantum number represents the orbital angular momentum of the running electron which results in the smaller l values lying deeper in energy for a given n since it penetrates more severely the core of the atom. The core, $5p^5$, results in an inverted

doublet designated as $^2P_{1/2}$ and $^2P_{3/2}$. The ionization potential of the $^2P_{3/2}$ core is 12.1 eV's. The ionization potential of the $^2P_{1/2}$ core lies about 1.3 eV's higher. Xenon follows J_c-l coupling rules. This essentially means that the angular momentum of the core couples with the orbital angular momemtum of the electron giving the angular momentum K. K then couples with the spin of the running electron to give the total angular momentum of the atom, J. Therefore, the state designation 6p(3/2,1) represents a 6p electron coupled to the $^2P_{3/2}$ core to give a K = 3/2 and a total angular momentum of J = 1. The state 6p'(3/2,1) represents a 6p electron coupled to the $^2P_{1/2}$ core with K = 3/2 and J = 1.

The laser lines of greatest importance originate from the 5d levels and terminate on the 6p levels. The lifetimes of the 6p states are several orders of magnitude shorter than the 5d states. This investigation was primarily concerned with the reaction:

$$e + 5p^6 \rightarrow 5p^5 6p + e$$

The excited $5p^5 6p$ state will betray itself with the emission of a photon. As we vary the energy of the impacting electron, the probability of exciting one of the 6p levels will vary. As the probability varies so will the photon

flux. By monitoring the photon flux as a function of the bombarding electron's energy we obtain a curve known as the optical excitation function. All our data was obtained from threshhold (about 9.5 eV's) to 100 eV's. Figure (2) illustrates a detailed energy level diagram of the 6p and 6p' levels and the transitions monitored are denoted by a solid circle.

The investigation of the excitation functions of xenon led to some surprising results whose interpretation is still under study. The excitation functions tended to divide themselves into three catagories: a) no pressure dependence, b) slight pressure dependence and c) significant pressure dependency. In some cases we observed some current dependence. The cause of these effects must be explained to fully appreciate the data.

Not only did we investigate the 6p states but numerous higher levels and certain trends were discovered in the excitation functions. States as high as the 11p, 11d and 6f were measured in hopes of fully understanding pressure and cascade effects on the lower levels.

Very little work on electron impact excitation of xenon has been performed. Rostovikova (1) has studied the optical excitation functions of xenon in the 220 - 640 nm range.

Quantitative values of the excitation cross section for the 6p levels have been published by Zapesochnyi (2). Unpublished efforts (3) and (4) have been attempted in recent years to study the levels associated with transitions in the near IR. Williams (5) has measured the electron impact energy-loss spectra of xenon at 20 eV in hopes of determining optical selection rules. Perhaps the most work in xenon has been associated with detecting resonances in elastic scattering of electrons (6), (7). The first four references above were concerned with optical excitation functions of xenon directly but all failed to note the pressure dependencies that we have found to exist.

Experiment

The experimental setup consists mainly of an electron gun (collimated source of monoenergetic electrons) enclosed in a very high vaccuum chamber in which controlled amounts of gas can be inserted. Figure (3) diplay the schematic of the e-gun. It can produce beam currents up to 5 mA over a 1 cm^2 circular area. At 1 mA an energy resolution of .5 eV's is easily obtainable. The cathode is indirectly heated and is of the impregnated barium type. The cathode is at a negative potential which is the accelerating voltage. The first two grids control the current and focusing properties of the e-beam. The third grid is at cathode potential and

essentially stops all electrons whose energy is less than the cathode potential. The last grid is at ground potential as is the Faraday cage which serves as the collision chamber. Secondary electrons which have been created by the primary electrons hitting the back side of the Faraday cup can distort the excitation curves since they are at much lower energy. Figure (4a) illustrates this for the 6p(5/2,3) level. To help negate such effects the end of the Faraday cup is constructed of a wire mesh which is transparent to the electrons but opaque to external fields. The cup-shaped collector is operated at a positive voltage with respect to the cage and collects secondary electrons before they can enter the interaction region. Figure (4b) shows the same excitation function with the bias present. No secondary electron effect.

The cathode of the gun obtains temperatures around 1100^{o}C. Hence, large quantities of impurities are boiled off. Usually liquid nitrogen traps are utilized to pump out these impurities. Since xenon also condenses at liquid nitrogen temperatures another pump source is needed. A non-evaporable getter pump was utilized. This getter operates via chemisorption and bulk diffusion into a zirconium-aluminum alloy.

The xenon gas is dosed into the system and the

resultant pressure is measured with a MKS Baratron pressure head readable to 0.1 millitorr.

The slender retangular area denoted in Figure (3) is the region of the collision chamber which radiation is gathered by the detection system. The radiation was viewed at right angles to the electron beam. The detecting portion of the apparatus is a half-meter Jarrell-Ash monochromator coupled to a cooled RCA 31034A photomultiplier tube. To cancel amplifier drift and general drift of background sources, the electron beam was chopped by biasing grid #1 to 25 volts below the cathode voltage 200 times a second. The chopping voltage was also used to synch in the photon counter. Using this method, the photon counter counted signal + background in one channel when e-beam on then counted background in another channel with e-beam off. The photon counter then subtracted the two channels and displayed the result. The number of cycles for counting was selectable.

The entire system was controlled and monitored by a PET 8032 microcomputer with high-resolution graphics. The sequence of events that occurred were as follows:

a) Set e-beam energy

b) Initiate Photon Count

c) Sense the end of count and read results

d) Measure the e-beam energy

e) Measure the e-beam current

f) Plot on high-res graphics the photon count divided by the e-beam current versus the e-beam energy.

g) If not end of run then increment e-beam energy and begin at a) again

h) Store data run on disk

i) Hard copy print out of high-res display

Our efforts in constructing the necessary interface hardware and writing the required software to obtain the above automation proved invaluable since most of the time very low pressures were necessary (with subsequent low signals) and run times were any where from 30 minutes to 16 hours. The resulting plot in f) above is what is termed an optical excitation function.

To determine the optical cross sections absolutely, the detection system in total was moved horizontally to line up with a standardized source of photon flux (tungsten strip filament lamp). The standardized source was oriented in the same relative position to the detection system as the collision region of the electron gun. The intensity of a given transition was first monitored -- being careful that

we collected all the radiation eminating from a solid angle set by a stop on the focusing lense. Then we monitored the photon flux from a given area of the standardized source through a set solid angle. , comparing these intensities the optical cross section can be obtained at a particular electron energy. Again, utilization of the computer faciliated the ease of obtaining the necessary numbers. Appendix A gives a typical printout of results. A good discussion of standardization procedures is given by St.John (8).

Data

Apparent Excitation Functions of the np and np' States

The excitation functions exhibited in this report are what are termed 'Apparent' excitation functions. What that means is that the contribution of the levels population due to cascade from higher states has not been subtracted out. When one subtracts out cascade one obtains the 'Direct' excitation function. To perform a cascade analysis of the xenon levels requires a large effort in time. Also, the main cascade contributors to the population of the 6p and 6p' states are the 5d levels...these are the very intense laser lines which lie in the mid-IR and were not capable of detection with our present apparatus.

Figure (5) gives the apparent excitation function of the 6p(5/2,3) level from threshhold to 100 eV. This function was perhaps the 'purest' function monitored. It 'goes up and comes down' like a typical excitation function. It is impervious to current and pressure variations. It remained unperturbed by gross variations in experimental parameters. We feel fortunate that it 'existed' in that it tended to negate any thoughts that many of our other non-typical observations were e-gun or detection aberations. It is a rather sharp excitation function which one is not surprised at since the only good quantum number J, the total angular momentum, is 3 and therefore we would expect little mixing with the other nearby p-levels, all of which had J < 3.

In contrast, Figure (6a) displays the 6p(1/2,0) level. It appears badly distorted with two definite peaks present...one around 15 eV's, the other at 25 eV's. There is also the hint of another, much sharper peak, at 20 eV's. Figure (6b) displays this same level and shows the drastic reduction of the curve's tail after 30 eV when the pressure is reduced below 1 mTorr. Figures (7a), (7b), (7c) and (7d) display the 7p, 8p, 9p and 11p levels which have the same K and J, i.e., 1/2 and 0 repectively. Figure (8) shows that the 7p(1/2,0) level was also quite pressure dependent.

Compare with Figure (7a). Figure (9) displays an overlay of 7p(1/2,0) onto 6p(1/2,0) which tends to demonstrate that whatever is feeding one is feeding the other. However, the 7p(5/2,3) is quite nice (Figure (10)) and compares favorably with the 6p level with the same K and J. Hence, it appears that for any p level with quantum numbers K = 1/2 and J = 0 the 'double' hump shape and pressure dependence will exist. Also, for K = 5/2 and J = 3 the curve will be rather sharp with no noticeable pressure or current dependence.

The 8p(1/2,0) level behaved anomalously as well. Figures (7b) above. Figure (11a) shows it's pressure dependence. No current dependencies were noted for any of these p levels. The small narrow peak around 20 eV's which was just barely discernable on the other np(1/2,0) levels is quite evident on the 8p(1/2,0) excitation function. Figure (11b) examines the excitation function from threshhold to 25 eV's...the small peak at 19.3 eV's is very prominant.

The 6p(3/2,2) level is 'intermediate' in pressure effects when compared to the K = 5/2, J = 3 and K=1/2, J =0 levels. Figure (12) illustrates this level's pressure sensitivity. The 6p'(3/2,1) level is another 'intermediate' in pressure dependency as illustrated in Figure (13). Taking the 6p'(3/2,1) level to 500 microtorr (Curve D) produces an excitation function very similar to the nice

6p(5/2,3) state. Previous experiments have never gone this low and in not doing so produced results very similar to curve A in Figure (13).

The 6p'(1/2,1) is much like the 6p'(3/2,1) in that as we lower the pressure to 500 microtorr an excitation function is produced which is quite 'normal' and pleasing. See Figure (14), curve C.

The 6p(3/2,1) was quite sensitive to the pressure...pressures as low as 300 microtorr were needed to get the energy tail to begin to drop appropriately. Figure (15) illustrates this quite well.

One suprising result was the 6p'(1/2,0) curve. It appears as the other curves for K = 1/2 and J = 0 but it is not pressure dependent. Therefore, it appears that whatever is causing the pressure abnormalities prefers the $^2P_{3/2}$ core configuration when K = 1/2 and J = 0. Figure (16) gives the 6p'(1/2,0) curve.

The 6p(3/2,2) level was just within the spectral response of my PMT. Figure (17) illustrates its pressure sensitivity. Again 500 microtorr was necessary to produce a quality curve.

The 6p'(3/2,2) state behaved as many of the others, having a slight pressure dependence. Figure (18). The lone remaining 6p state, 6p(1/2,1), was not obtained since transitions originating from it lay outside of our spectral response.

Apparent Excitation Functions of s, d and f States

Present now are the excitation functions of various and sundry states that were examined. The spectra of Xenon is quite complex and hence requires the slits of the monochromator to be quite narrow in order to resolve many of the lines. Coupled to this is the requirement that the pressures be kept relatively low in order to maintain a stable system and prohibit arcing. The arcing occurs because of the high ionization cross section of xenon. Hence, because of these two factors, our signals are generally quite weak and significant times are required to obtain data. What we have tried to do is pick those levels that perhaps might offer some insight into the physics of the resultant shapes of the excitation functions.

Of particular interest is the 688.2 nm line originating from the 7d(7/2,3) level. See Figures (19a & 19b). No differences were observed with pressures ranging from 4 mTorr to .5 mTorr. We observe several features on this function. An extremely sharp onset with peak, another

structure about 20 eV, followed by a rather broad 'hump' at 40 eV's. Figure (20) provides an interesting picture. It is an overlay of Figure (19b) with the excitation function of the 6p(1/2,0) level at the same pressure of 500 microtorr. The small structure we observed at 20 eV on the 6p levels (K = 1/2, J = 0) appears very evident in this high lying 7d level. Hence, some of the contribution of that structure in the 6p levels arises from cascade from higher levels...7d in particular. The question still arises...What causes the narrow structure in the 7d level? Perhaps the source of it in the 7d level is cascade from some phenomena at a higher energy which also feeds the 6p levels.

The 7d(7/2,4), which cascades into the very clean looking 6p(5/2,3) level, is quite clean as well (Figure (21)). Then there is the 8s(3/2,1) and 6+(3/2,1) excitation functions given in Figure (22) and Figure (23) repectively. These functions appear highly distorted. The 6d(3/2,2) was an <u>exceptionally</u> clean function with no pressure or current dependence. See Figure (24). <u>Why some of the levels possess clean shapes and others are so distortable with pressure is a challenging question</u>.

Figure (25a) and (25b) give the 5d'(5/2,3) function at two different pressures. Again, it appears that those levels associated with the $^2P_{1/2}$ core are more immune from

pressure effects.

Excitation Functions of Xenon in Helium/Xenon Mixtures

We were curious as to the effects that large amounts of Helium would have upon the xenon excitation functions. If large quantities of a gas are inserted into the system, then one must be careful that the mean free path of the electrons in the beam, as measured from the last grid, is longer than the distance observations are made from the last grid. If we observe the photon flux at too great a distance from the last grid then the atoms we observe radiating are those excited by electrons which have already suffered a collision. In xenon at 1 mTorr, the mean free path of an electron (λ_e) is 16 cm. In 150 mTorr of helium, λ_e = 1 cm. Figure (26a) and (26b) illustrate this point experimentally. In Figure (26b) we are observing the flux at a distance of .5 cm in 150 mTorr helium and 2 mTorr xenon. The line in question is the 6p(5/2,3) line and its shape agrees when only pure xenon is present. Compare with Figure (5). Figure (26a) shows a hump appearing when we observe this line at a distance a little over 1 cm. The hump is the result of electrons that have already suffered one collision...then many of the resultant lower energy electrons will have energies at the peak cross section and will therefore cause a distorted excitation function.

Fortunately, the first level we tried gave dramatic effects. The 7d(7/2,3) was one of our most distorted functions (See Figure (19)). Figure (27) shows the dramatic effects on the high energy tail as the pressure of helium is increased. One observes that as helium is added, we obtain a curve much like a 'normal' excitation function. Consider the 11d(7/2,3) curve. Figures (28a) and (28b) show that the tail drops when helium is present. Figure (28c) shows the same function but with pure xenon at 1.2 mTorr. It is almost identical to Figure (28b). In other words, adding helium is equivalent to lowering the pressure of xenon. Same results were obtained for 7p(1/2,0) state. Compare Figures (29a) and (29b) with Figures (7b) and (7c). Figures (30a - 30e) show the sequencial effects of adding helium as well as reducing the xenon pressure for the prominant 6p(1/2,0) level. The first structure observed at 15 eV's becomes extremely prominant and the large peak at 25 eV's begins to drop in intensity relative to the peak at 15 eV's.

Resonant Structure in Helium

Since Helium was to be used, we spent some time looking at several of its excitation functions at various pressures. The 587.6 nm line is a good line to look at because it has a well known structure at onset which we can use to get an idea of our energy resolution. Figure (31) is the excitation

function of this line from onset to 30 eV's. One can see that the structure is well resolved and our energy resolution, even at these high pressures and currents, is easily .5 eV. Figure (32a) displays the same excitation function except to 100 eV's. Please note the structure around 57 eV's. Figure (32b) gives a 'close up' view of this region by scanning from 45 to 60 eV's. The energy separation of these two peaks is about 1.2 eV's. Scans of other lines in this same energy range revealed no such structure implying it was not an e-gun effect. I ran this line under varying conditions. Figure (33a) illustrates that the structure still exists at low pressures and low currents. The reason the general shape of the function is different at different pressures is because this function is quite pressure sensitive (9). Figure (33b) shows the function at an extremely high pressure ... structure still present. Figure (33c) when xenon is mixed in...structure still present. These curves also tell us that the e-gun characteristics are not perturbed when xenon is added and the energy resolution of beam is still quite good at 50 eV's when large quantities of gas are present.

We feel that the source of these two structures is from the creation of negative ions. Such resonances have been observed by Kuyatt (6) in electron transmission experiments. This is the first time they have been observed in optical

transitions. Kuyatt observed two resonances at 57.1 and 58.2 eV's with an uncertainty of 0.1 eV. This separation cooresponds quite well with what we observed. When the contact potential of our gun is corrected for, we are almost spot on the energy for which these resonances occur. Fano and Cooper (10) have identified the negative ion states first seen by Kuyatt (6) as the $(2s^22p)^2P$ and $(2s2p^2)^2D$ states. As such they lie in the continuum and can decay by two different channels. They can decay into excited states of Helium by the emission of a single electron. Or they can emitt two electrons and form He^+. It appears we are observing the first decay channel into our 3^3D state.

Discussion

The electron impact excitation functions are confusing. In one way this has proven fortunate in that one must then search diligency for the source of confusion...which has resulted in a much broader learning experience.

When one observes pressure effects in electron impact experiments it usually follows a quadradic dependence. We did not find this dependence. Hence, many of the classic causes of pressure effects were not sufficient. At pressures where we do observe effects, the mean free path of a xenon atom is 10 cm. Our observation region is never

greater than .4 mm. Also, the time between atom-atom collisions is close to 500 microsec...the lifetimes of the majority of the states are in the area of nanosec. To be able to obtain pressure effects that were observed due to excitation by collision transfer, one must assign collision cross sections of the order of 10^{-14} cm^2...an order of magnitude greater than measured values (11). Hence, it would appear that the population of the xenon states would only be by direct electron impact and cascade from higher lying states...effects not sensitive to pressure.

One channel we considered was excitation to those states associated with the $^2P_{1/2}$ core that lie between the ionization energies of the $^2P_{3/2}$ core and the $^2P_{1/2}$ core. Figure (34) is a schematic of this area. The levels ns' (n>7) and nd' (n>5) are the predominant levels. Huffman (12) has examined the absorption coefficients of these levels and has found that the 8d' and 6s' have extremely large absorption coefficients of 5000 cm^{-1}, (2×10^{-16} cm^2). Transitions with large dipole moments usually imply large electron excitation cross sections. Hence, one can assume that these levels are readily being populated. After being populated, the state has basically two alternatives. One...it can autoionize and form a xenon ion. Two...it can decay to a level beneath the first ionization level. If process two does occur and it is to the Rydberg states, then

we have states occupied with long lifetimes which could undergo collisional effects. However, White (13) has observed no such radiative or radiationless decay occuring (lifetime of 10^{-8}sec)...only autoionization (lifetime of 10^{-14}sec). This channel produces mainly, therefore, ions.

Much of the observed pressure effects could be explained if there were a large supply of Xe_2^+ present. Figure (35) is a diagram taken from a study by Biondi (14). He studied the dissociative recombination process in Xenon. The reaction under investigation was:

$$Xe_2^+ + e \longrightarrow (Xe_2)^* \longrightarrow Xe^* + Xe$$

After producing the Xe_2^+, they observed the intensities of the atomic transitions in the afterglow. The heavier the lines in Figure (35) the greater the intensity of the line. Or, the heavier the line, the more heavily populated that level by dissociative recombination of Xe_2^+. The dashed lines indicated transitions that they observed when they heated the electrons to 8000^{o} K, or to a little over 1 eV, thus producing higher excited states of Xe_2. It is evident from the figure that states we found pressure sensitive were monitored as recipients of the dissociative recombination process. In particular, the 828 nm line, 6p(1/2,0), which we found very distortable with pressure was one of the heaviest populated by dissociative recombination. For room temperature electrons the rate coefficient for this process

is extremely large (10^{-6} cm^3/sec). As the electron temperature increases, the rate coefficient falls off as $T^{-.7}$.

Xe_2^+ can be produced via several mechanisms. See Table I below. In each of these mechanisms, we are considering processes that require three bodies and/or several steps. Such processes are pressure dependent. Our problem then is to try and formulate some mechanism by which a copious supply of Xe_2^+ is produced in order to explain our observed pressure effects.

Table I

Process 1
$Xe + e \rightarrow Xe^+ + e$
$Xe^+ + Xe + Xe \rightarrow Xe_2^+ + Xe$

Process 2
$Xe + e \rightarrow Xe^* + e$
$Xe^* + Xe \rightarrow Xe_2^+ + e$
(Hornbeck and Molnar Process (15))

Process 3
$Xe + e \rightarrow Xe^* + e$
$Xe^* + Xe + Xe \rightarrow Xe_2^* + Xe$
$Xe_2^* + Xe_2^* \rightarrow Xe_2^{+2} + Xe + Xe + e$
(Self-Penning Ionization (16)

Process 1 in our table is a two step process which requires a three-body collision. Such a collision cross-section at our pressures is

extremely small. Process 3 is even more exotic and less likely. The Hornbeck and Molnar process has some promise if the excited state is a metastable state. There are two metastable states, 6s(3/2,2) and 6s'(1/2,0)...their calculated (17) lifetimes are 150 sec and 78 msec respectively. But even then they would only account for less than 10% of the gas density as observed in the collision region.

In Figure (36a) we have drawn an end view of the grounded Faraday cup surrounding the electron beam and from which we observe the collision phenomena. Heddle (18) has calculated the potential V as a function of 'r', the distance from the center of the beam. In Equation (1) below, R is the radius of the beam, 'i' is the current in microamps and E is the energy of the electrons in eV's. V_1 is the potential due to zero space charge and V_2 is the potential depression due to the space charge of the electron interacting upon themselves.

$$V_1 = -(3 \times 10^{-2}) i E^{-1/2} \mathrm{Ln}(R/r) \text{ eV's}$$
$$V_2 = -(1.5 \times 10^{-2}) i E^{-1/2} \text{ eV's} \qquad (1)$$

Figure (36b) displays the potential depression as a

function of 'r'. What this means is that we have a well close to 5 eV's deep at beam currents of 500 microamps. This well can therefore act as an ion trap. Hence, since xenon has such a high ionization cross section as well as heavily populated autoionizing states we can well visualize a copious supply of ions available to fall into this well. Using an ionization cross section for Xenon of $5 \times 10^{-16} cm^2$ and a velocity for the xenon ion as 5×10^4 cm/sec we calculate 'filling' of the well at a xenon pressure of 0.3 mTorr. These have been rough calculations but one can feel assured that the well would be neutralized for Xenon pressure below 1 mTorr. What this means is that we have a plasma with a larger than thought concentration of xenon ions lying in the beam. Consistant with this line of thinking is the observation that at low pressures the onset potential is several volts high. As the pressure of the gas increases then the onset lowers and even reaches a point where it is lower than the true onset. Many of the figures which are presented as overlays of one function upon another have been shifted in order that the onsets match...some I have not shifted and the true relative onsets are displayed. See Figures (12), (13) and (15). Such a large concentration of ions enables us to produce

more Xe_2^+ via process 1 in Table I than originally thought possible. Also, radiative recombination,

$$Xe^+ + e \dashrightarrow Xe^* + \text{photon},$$

can become prominant. Each of these result in a population of excited xenon atoms which is pressure sensitive.

What happens when large quantities of Helium are added? Shuker (19) in examining Penning ionization in He/Xe mixtures observed that with low partial pressures of Xenon (such as 1 mTorr) dissociative recombination is neglible and radiative recombination became the dominant mechanism. Also, our 'e-beam ion trap' becomes 'clogged' with He^+. The net result is to decrease the population of various xenon levels due to dissociative and radiative recombination. Also, when 150 mTorr of Helium is introduced, time between collisions drops from 500 microsec to several hundred nanosecs. This can affect some of the excited states whose lifetimes approach these values. The effect is to relax them to lower levels and again change the population schemes of various levels due to cascade. We are presently developing rate equations to determine if such mechanisms would produce the necessary yield of excited states.

Optical Cross-Sections

The optical cross-sections for those levels whose excitation functions are displayed in this report are presented in Table II on the next page. These cross-sections are given for an electron energy of 15 eV's at a pressure of 4 mTorr. Even though this pressure is one in which we find distortion present in many of the excitation functions, the distortion begins about 25 eV's. It would be much more preferred to actually obtain these cross sections at pressures below 0.5 mTorr but at such pressures the signals are extremely weak and long count times are required...we simply did not have enough time. To obtain <u>apparent</u> cross-sections for the levels, one can utilize branching ratios (20),(21) to obtain those optical cross sections not measured. For the most part, these cross sections are good to 30%. Below 400 nm, the standardization procedure was hampered by large amounts of scattered light with the available detection system and errors of 50% would not be unexpected.

Table II

6p(5/2,2) --> 6s(3/2,2)	904.5 nm	9.8×10^{-18} cm^2
6p(5/2,3) --> 6s(3/2,2)	881.9	43.0
6d(3/2,2) --> 6p(1/2,1)	873.9	6.3
7p(1/2,0) --> 6s'(1/2,1)	857.6	1.5
6p(3/2,1) --> 6s(3/2,2)	840.9	6.4
6p(1/2,0) --> 6s(3/2,1)	828.0	13.0
6p'(1/2,1) --> 6s'(1/2,1)	826.6	1.0
6p(3/2,2) --> 6s(3/2,2)	823.2	19.7
6p'(3/2,1) --> 6s'(1/2,1)	820.6	1.7
6p'(1/2,0) --> 6s'(1/2,1)	788.7	2.2
6p'(1/2,1) --> 6s'(1/2,0)	764.2	0.9
5d'(5/2,3) --> 6p(5/2,2)	733.6	1.5
7d(7/2,4) --> 6p(5/2,3)	711.9	1.0
7d(7/2,3) --> 6p(5/2,2)	688.2	0.7
11d(7/2,3) --> 6p(5/2,2)	548.8	0.05
6f(3/2,1) --> 6s'(1/2,0)	539.3	0.05
6p'(3/2,2) --> 6s(3/2,1)	473.4	0.12
7p(5/2,3) --> 6s(3/2,2)	467.1	0.6
6p'(1/2,0) --> 6s(3/2,1)	458.3	0.14
8p(1/2,0) --> 6s(3/2,1)	407.9	0.11
9p(1/2,0) --> 6s(3/2,1)	380.9	0.04
11p(1/2,0) --> 6s(3/2,1)	358.7	0.01

Conclusion

The electron excitation functions of atomic xenon have, in general, been found to be quite sensitive to pressure. A fact hitherto undetected by previous workers. The source of this sensitivity has offered to us a very challenging problem. Lending to the confusion is that those states which are pressure sensitive appear to be randomly selected. In general, we can say that if one wants to monitor the excitation functions of atomic xenon, he should work at pressures at or below 500 microTorr.

After reviewing many possibilities for the pressure distortions, we are led to believe that what we are seeing in the collision region is a plasma effect. Due to the high ionization cross section of xenon and its very poor mobility, the xenon ions are becoming trapped in the beam. This high concentration of Xe^+ enables the production of Xe_2^+ with the subsequent dissociative recombination producing excited xenon atoms.

Further work should be concerned with two areas:

1) Investigation of the trapped ion theory and
2) Electron impact cross section of the 5d levels which give rise to the mid-infrared laser lines.

The second area above is very important in that to fully appreciate the effects of electronic excitation by electron impact on laser gain one must know these cross sections. The testing of the trapped ion theory is going to require much experimental creativity. In general, the xenon atom is a complex atom and great finesse is needed in dealing with it.

References

1. G.S. Rostovikova, V.P. Samoilov, and Yu.M. Smirnov, Opt. Spectrosc. 34, 3 (1973).

2. I.P. Zapesochnyi and P.V. Feltsan, Opt. Spectrosc. 20, 521 (1966).

3. Mark D. Stephen, Master's Thesis, Air Force Institute of Technology, WPAFB, Ohio, (1979).

4. J.E. Gastineau, F.A. Sharpton, L.W. Anderson, and C.C. Lin, Electronn-Impact Excitation of the Xenon Atom, 35th Gaseous Electronics Conference, 1982.

5. W. Williams, S. Tra mar, and A. Kuppermann, Jour. of Chem. Phys. 62, 3031 (1975).

6. C.E. Kuyatt, J. Arol Simpson, and S.R. Mielczarek, Phys. Rev. 138, 2A, A385 (1965).

7. L. Sanche and G.J. Schulz, Phys. Rev. A 5, 1672 (1972).

8. R.M. St.John, "Calibration of a Spectroscopic System", Methods of Experimental Physics, Vol. 8, 27 (1969), Academic Press, Inc., N.Y.

9. R.M. St.John and R.G. Fowler, Phys. Rev. 122, 1813 (1961).

10. U. Fano and J.W. Cooper, Phys. Rev. 138, 2A, A400 (1965).

11. L. Allen, D.G.C. Jones, and D.G. Schofield, Jour. Opt. Soc. Am. 59, 842 (1969).

12. R.E. Huffman, Y. Tanaka, and J.C. Larrabee, Jour. of Chem. Phys. 39, 902 (1963).

13. H.E. White, Phys. Rev. 38, 2016 (1931).

14. Yueh-Jaw Shiu, Manfred A. Biondi, and Dwight P. Sipler, Phys. Rev. A 15, 494 (1977).

15. J.A. Hornbeck and J.P. Molnar, Phys. Rev. 84, 621 (1951).

16. N.E. Small-Warren and L.Y. Chow Chiu, Phys. Rev. A 11, 1777 (1975).

17. R. Shuker, A. Szoke, E. Zamir, and Y. Binur, Phys. Rev. A

11, 1187 (1975).

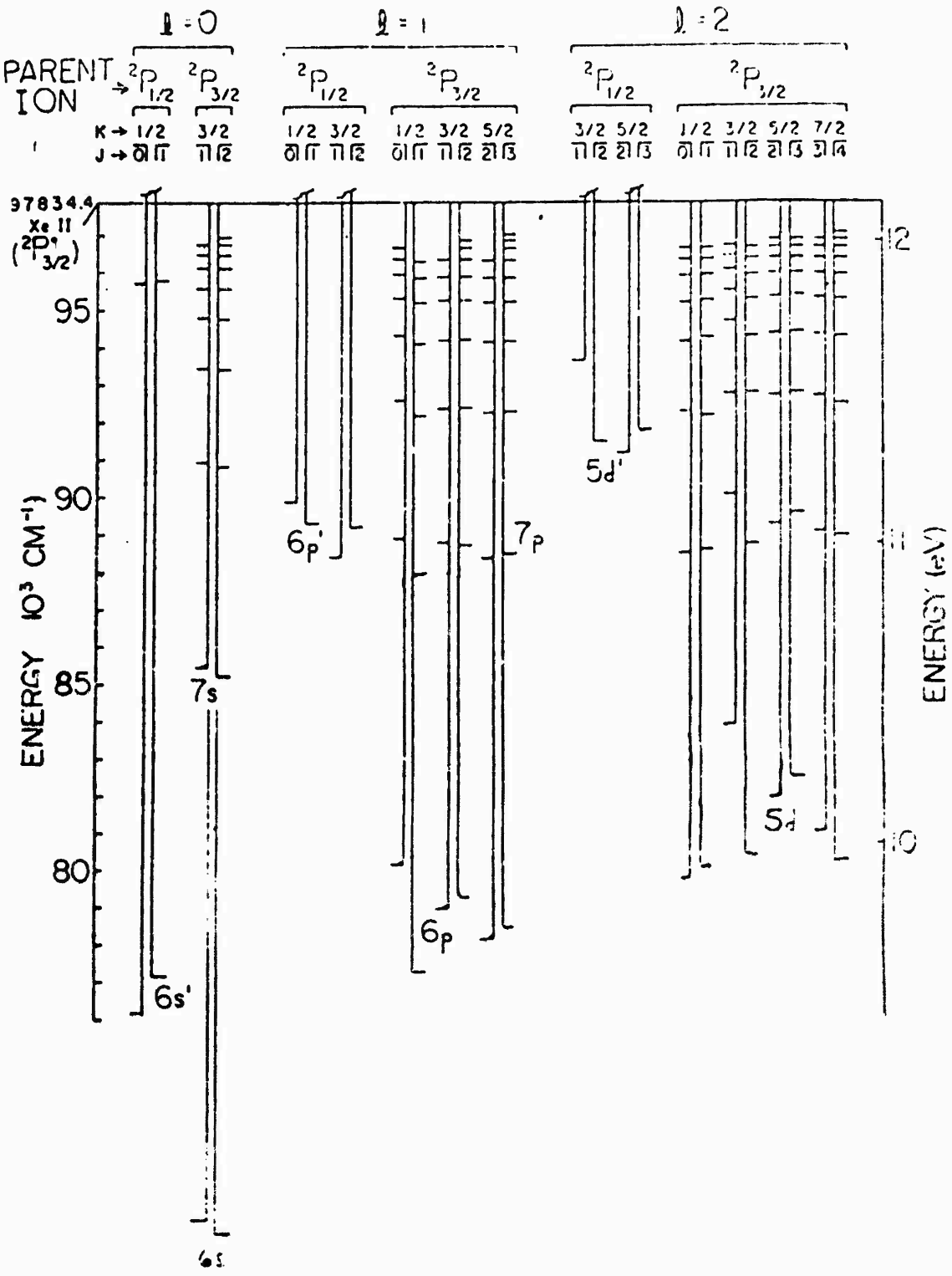
18. D.W.O. Heddle, Proc. Phys. Soc. 90, 81 (1967).

19. R. Shuker, Y. Binur, and A. Szoke, Phys. Rev. A 12, 515 (1975).

20. J. Sabbagh and N. Sadeghi, J. Quant. Spectrosc. Radiat. Transfer, Vol. 17, 297, Pergamon Press (1977).

21. M. Aymar and M. Coulombe, Atomic Data and Nuclear Data Tables, Vol. 21, 537, Academic Press (1978).

FIGURES


ℓ = 0
ℓ = 1
ℓ = 2
PARENT ION
$^2P_{1/2}$
$^2P_{3/2}$
K → 1/2 3/2
J → 0Π1 1Π2
1/2 3/2 1/2 3/2 5/2
0Π1 1Π2 0Π1 1Π2 2Π3
3/2 5/2 1/2 3/2 5/2 7/2
1Π2 2Π3 0Π1 1Π2 2Π3 3Π4
97834.4
Xe II
($^2P'_{3/2}$)
95
90
85
80
ENERGY 10^3 CM^{-1}
12
11
10
ENERGY (eV)
6s'
7s
6s
6p'
6p
7p
5d'
5d

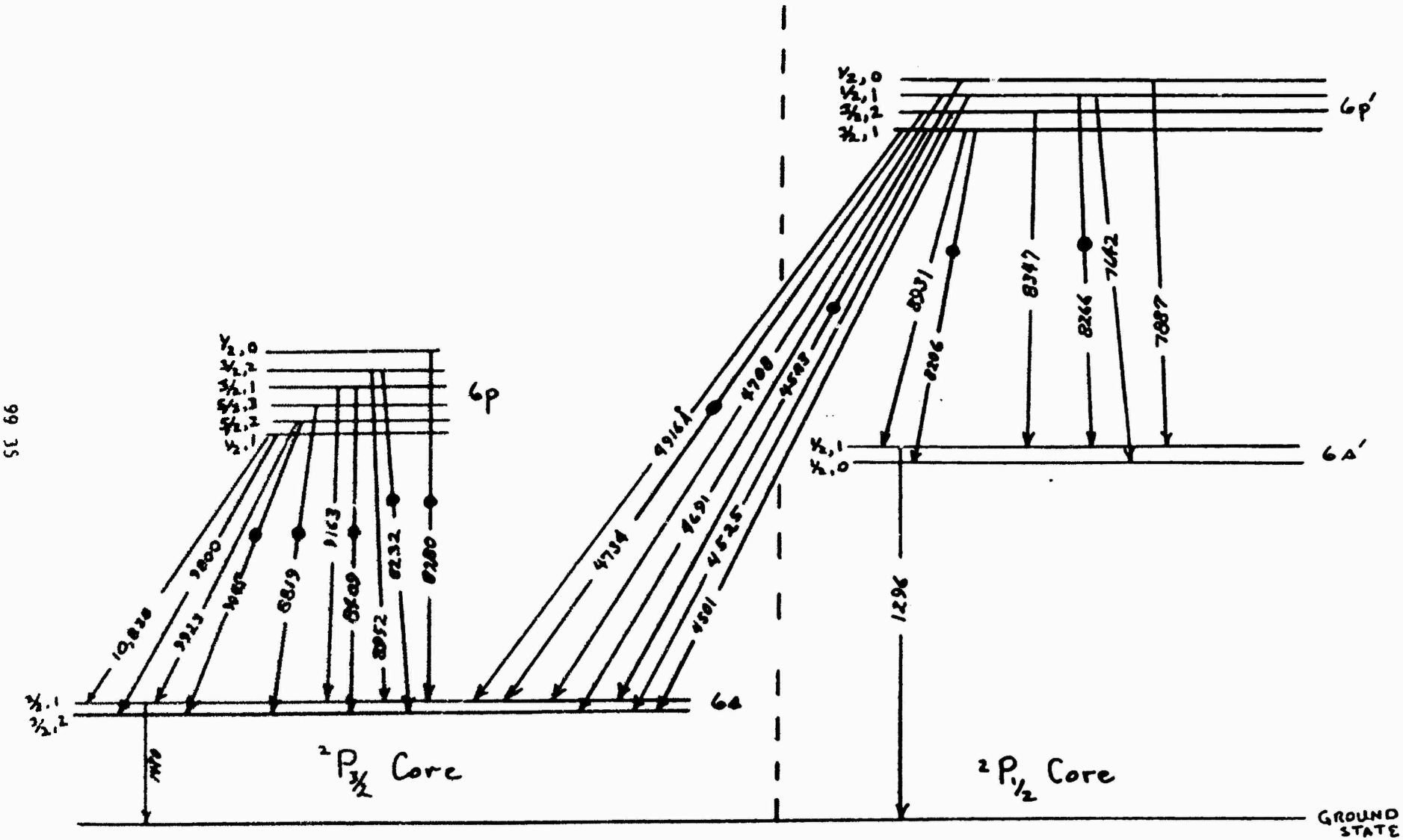

6p
6p'
6s
6s'
10,830
9123
9465
6869
6163
6560
6352
6232
6150
4916
4734
4708
4691
4583
4525
4501
8447
8266
7662
7087
1296
$^2P_{3/2}$ Core
$^2P_{1/2}$ Core
GROUND STATE

FIG 2

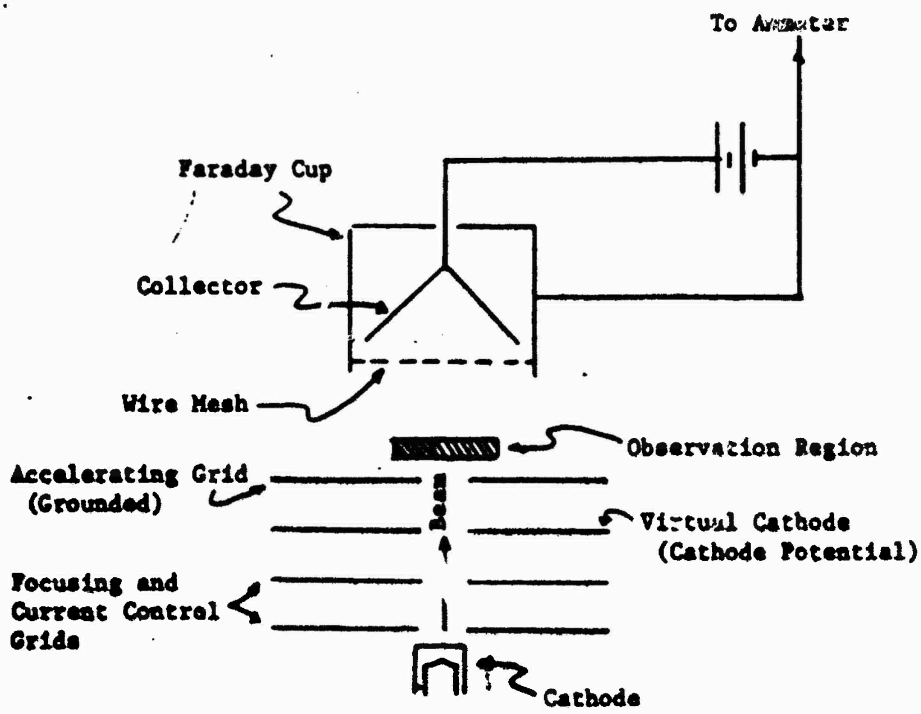

To Ammeter
Faraday Cup
Collector
Wire Mesh
Observation Region
Accelerating Grid
(Grounded)
Beam
Virtual Cathode
(Cathode Potential)
Focusing and
Current Control
Grids
Cathode

FIG 3

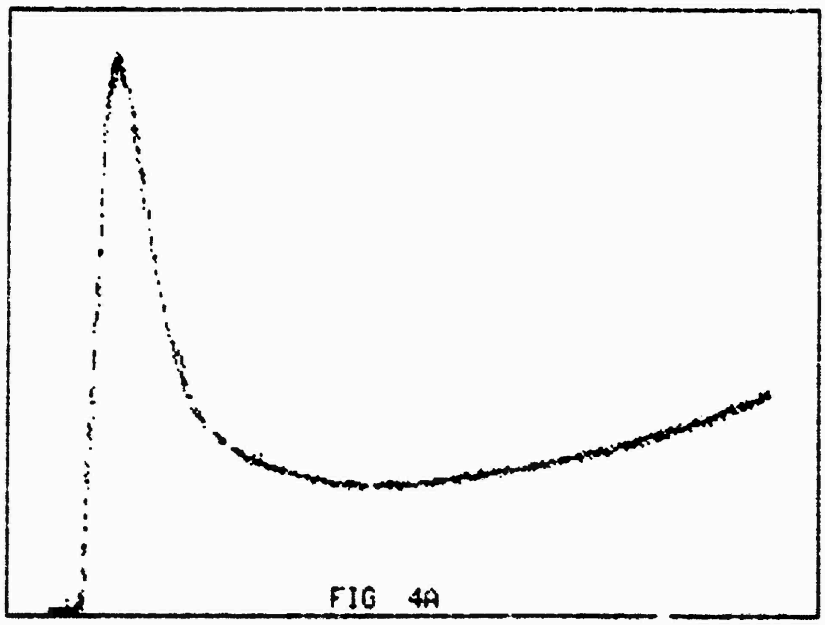


FIG 4A

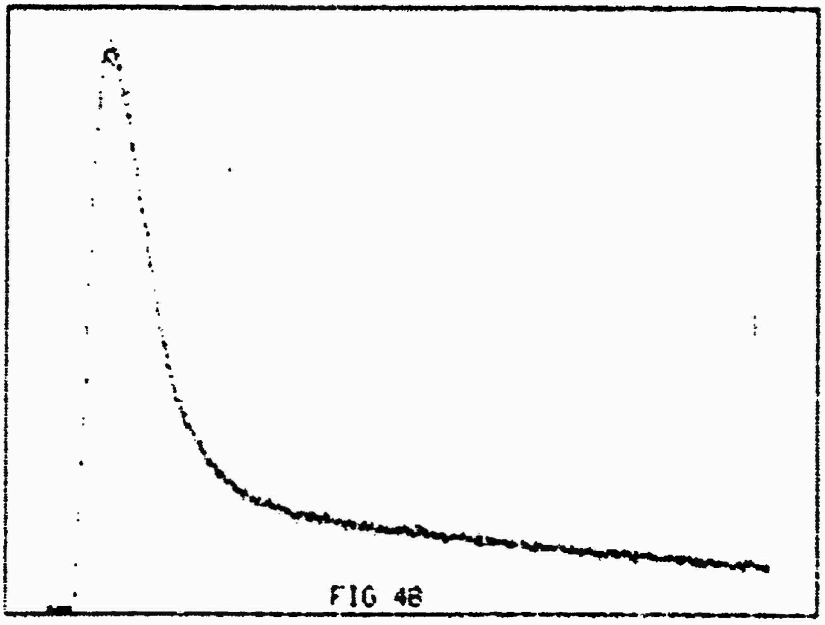


FIG 4B

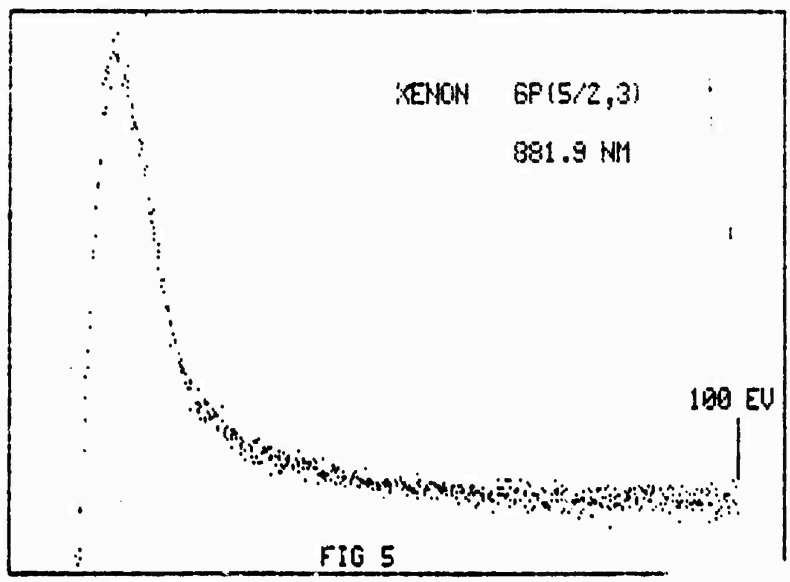


FIG 5

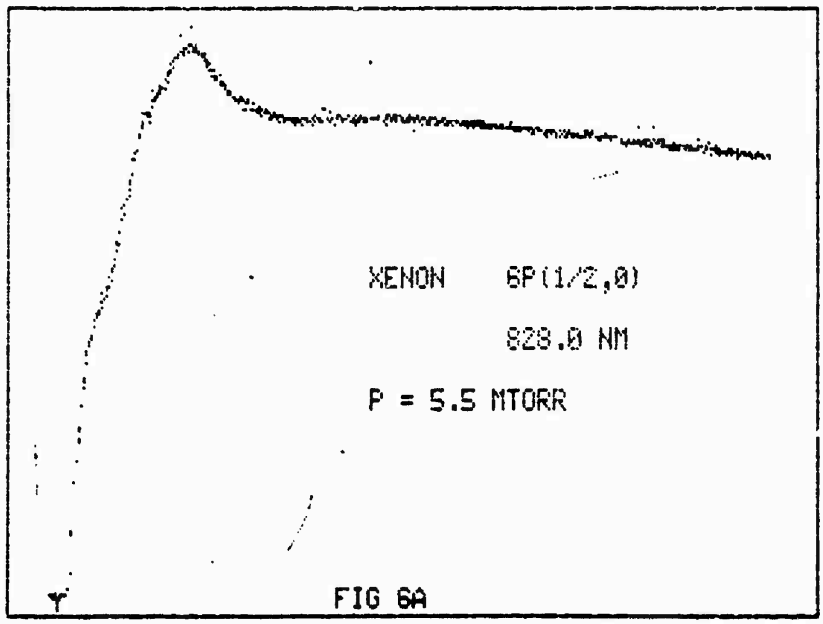


FIG 6A

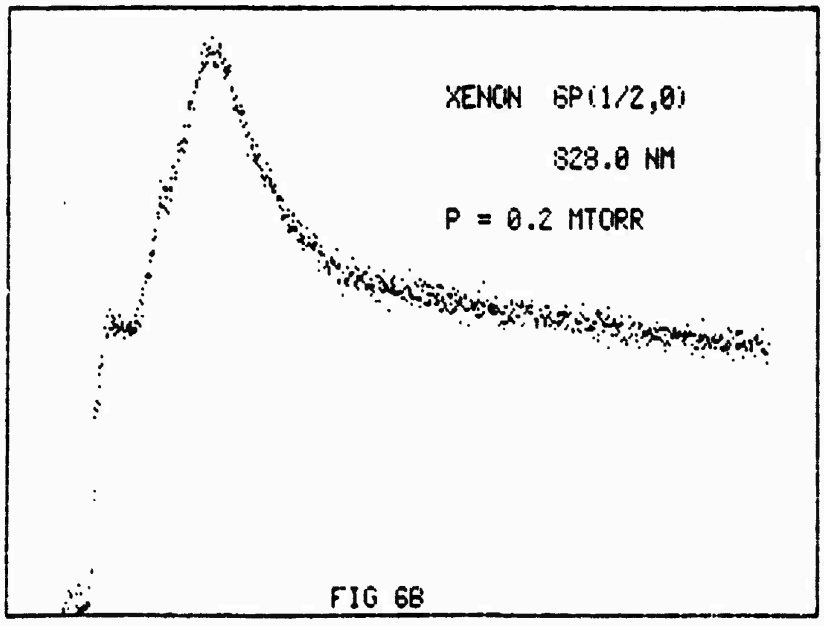


FIG 6B

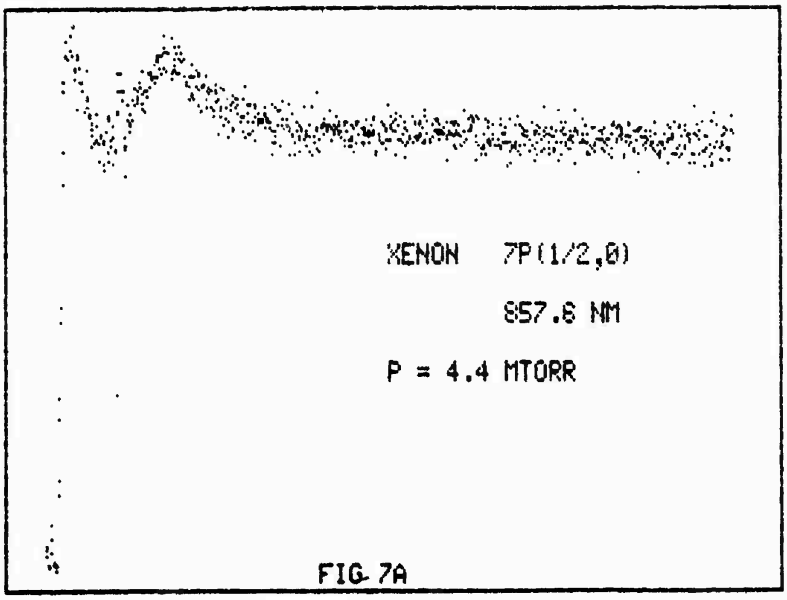


FIG 7A

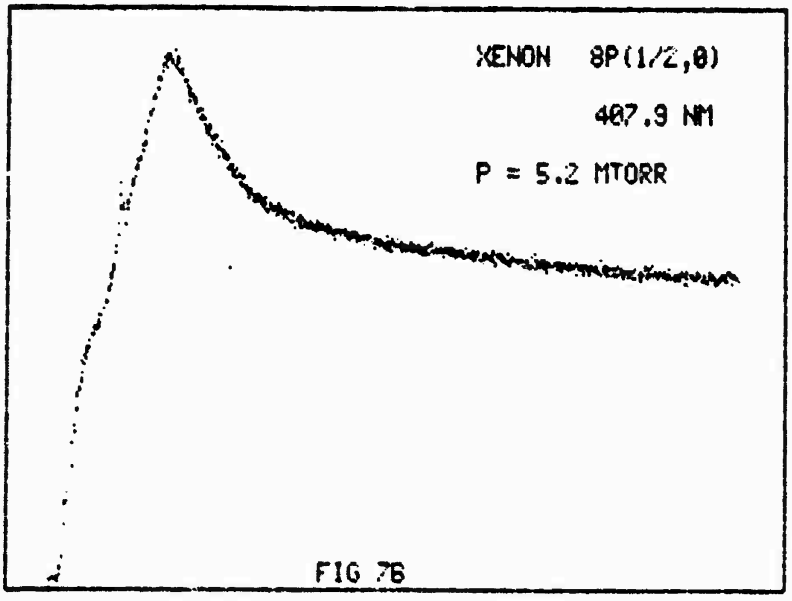


FIG 7B

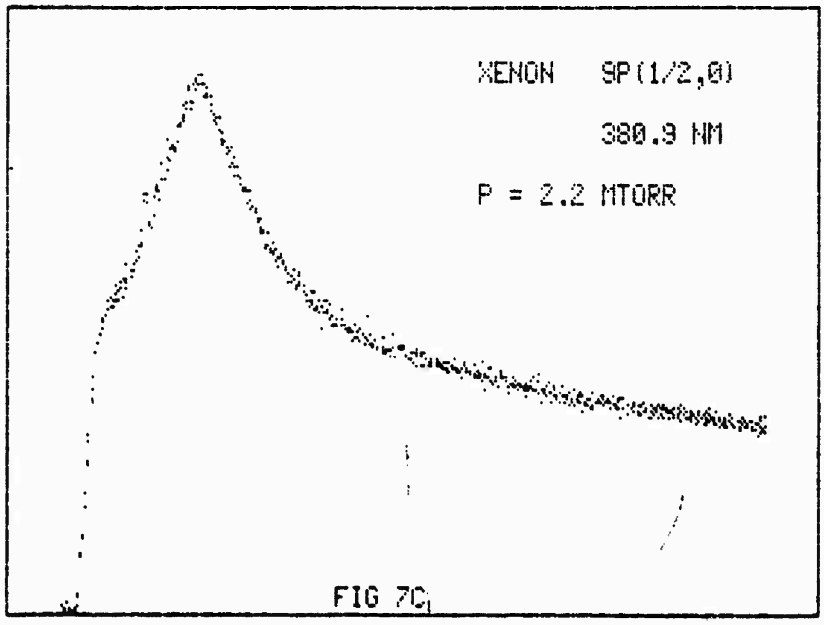


FIG 7C

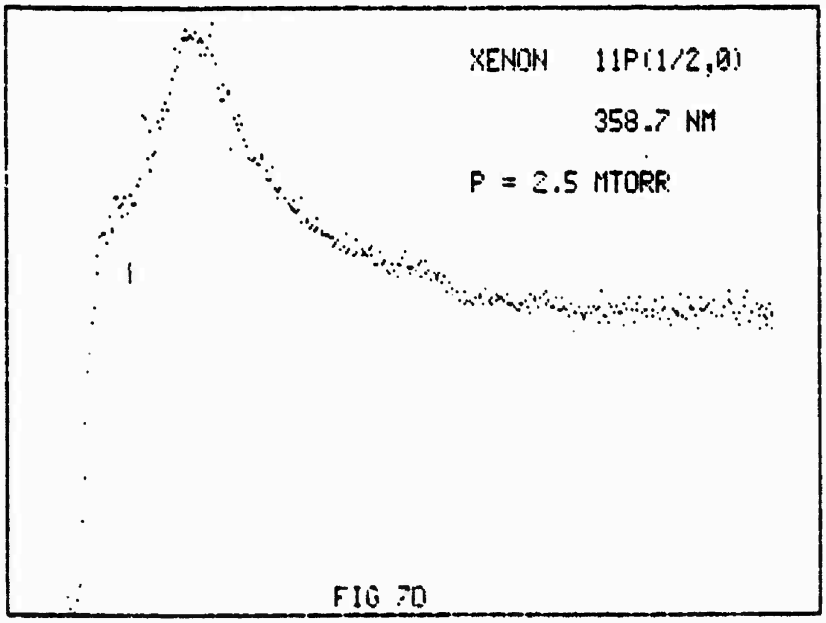


FIG 7D

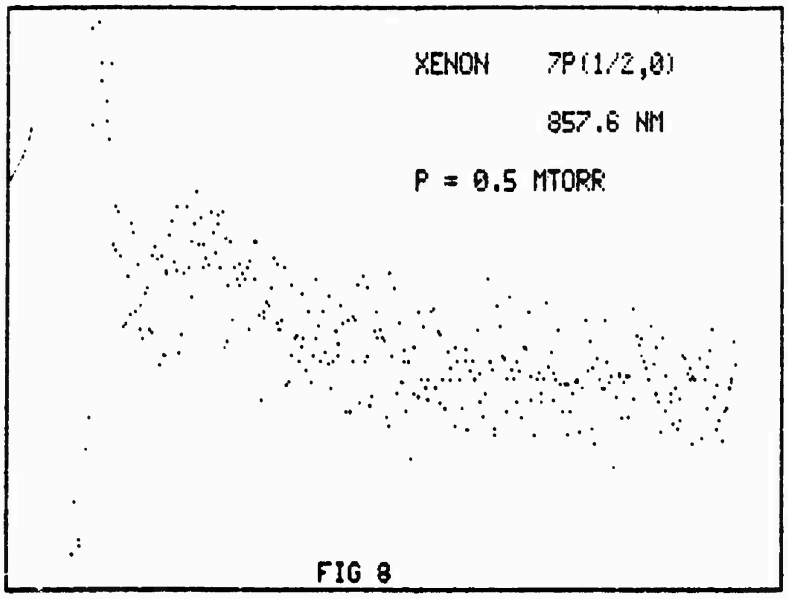


FIG 8

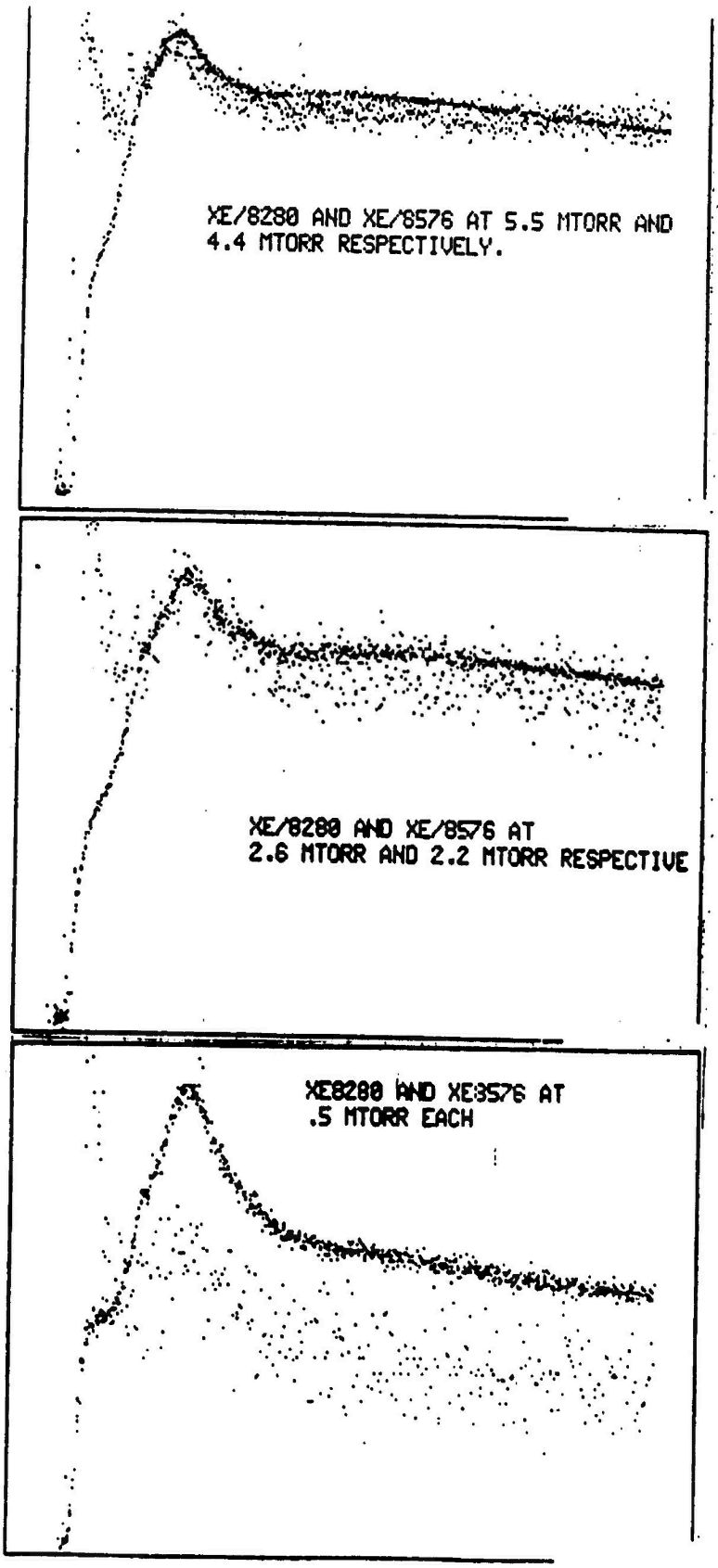

XE/8280 AND XE/8576 AT 5.5 MTORR AND
4.4 MTORR RESPECTIVELY.
(a)
XE/8280 AND XE/8576 AT
2.6 MTORR AND 2.2 MTORR RESPECTIVE
(b)
XE8280 AND XE3576 AT
.5 MTORR EACH
(c)

FIG 9

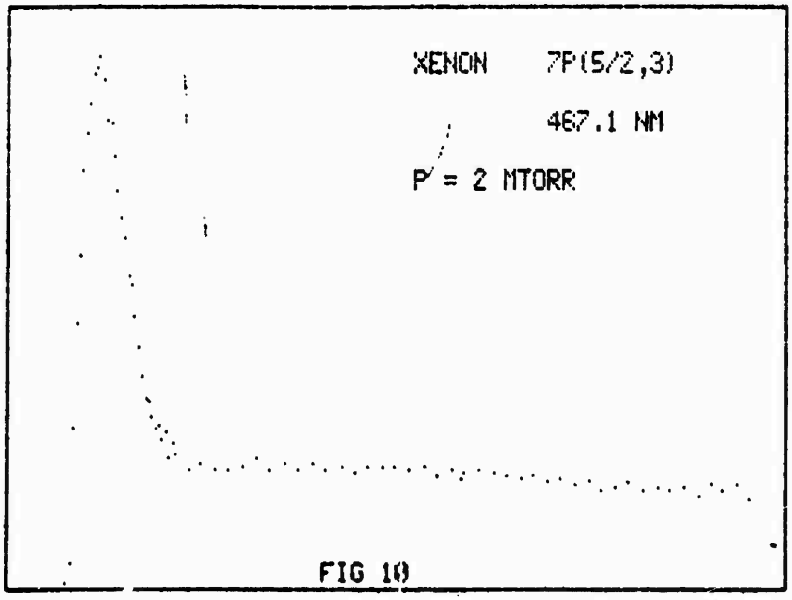


FIG 10

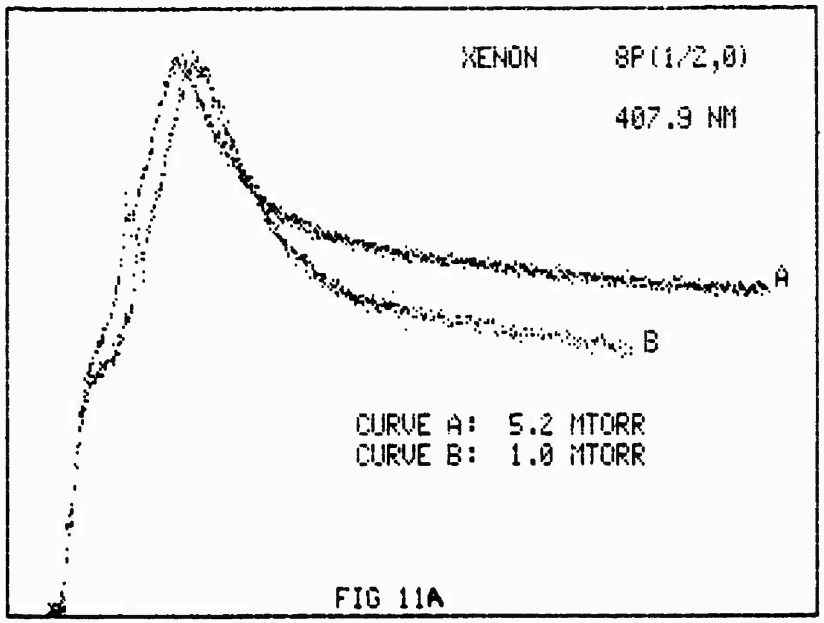


FIG 11A

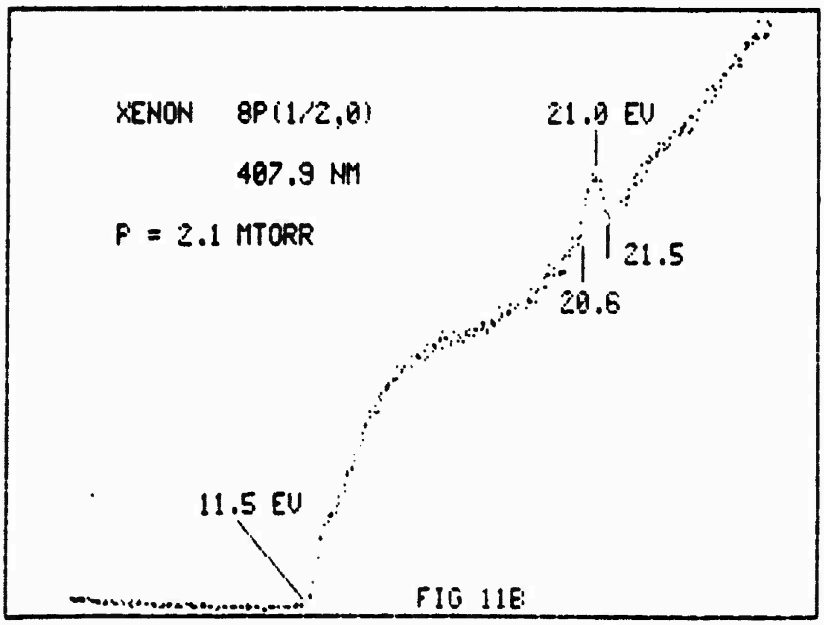


FIG 11B

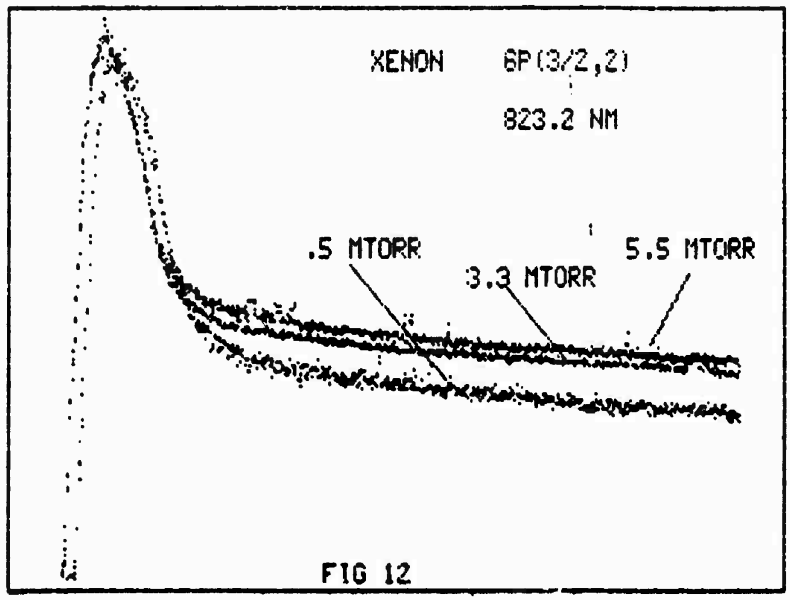


FIG 12

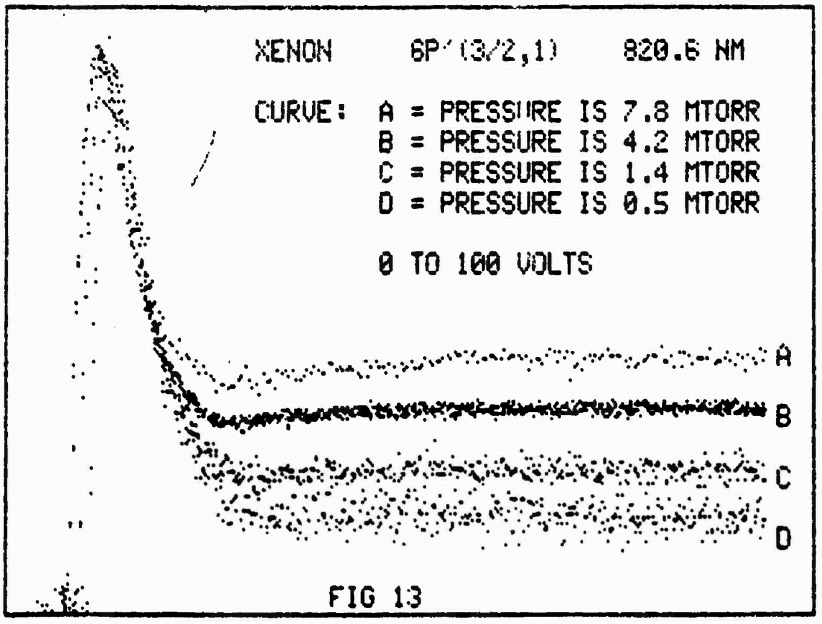

XENON 6P'(3/2,1) 820.6 NM
CURVE: A = PRESSURE IS 7.8 MTORR
B = PRESSURE IS 4.2 MTORR
C = PRESSURE IS 1.4 MTORR
D = PRESSURE IS 0.5 MTORR
0 TO 100 VOLTS
A
B
C
D

FIG 13

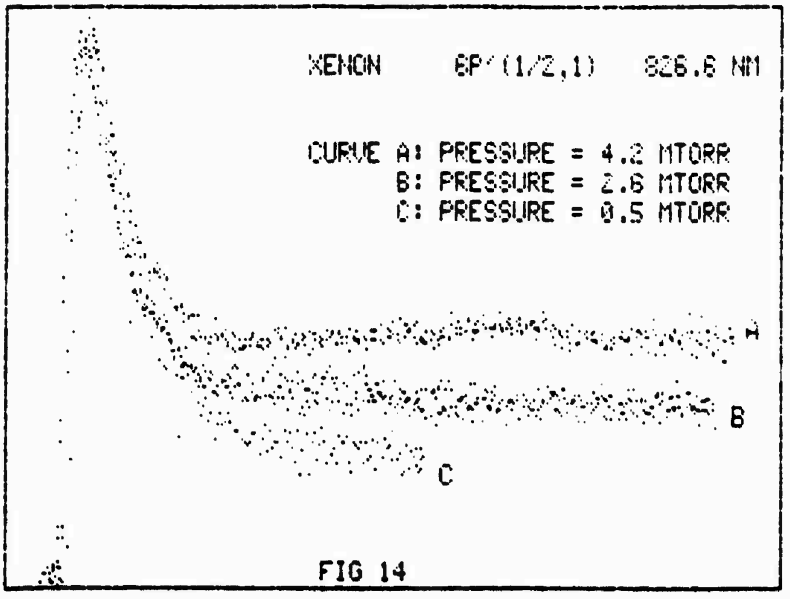

XENON 6P'(1/2,1) 826.6 NM
CURVE A: PRESSURE = 4.2 MTORR
B: PRESSURE = 2.6 MTORR
C: PRESSURE = 0.5 MTORR
A
B
C

FIG 14

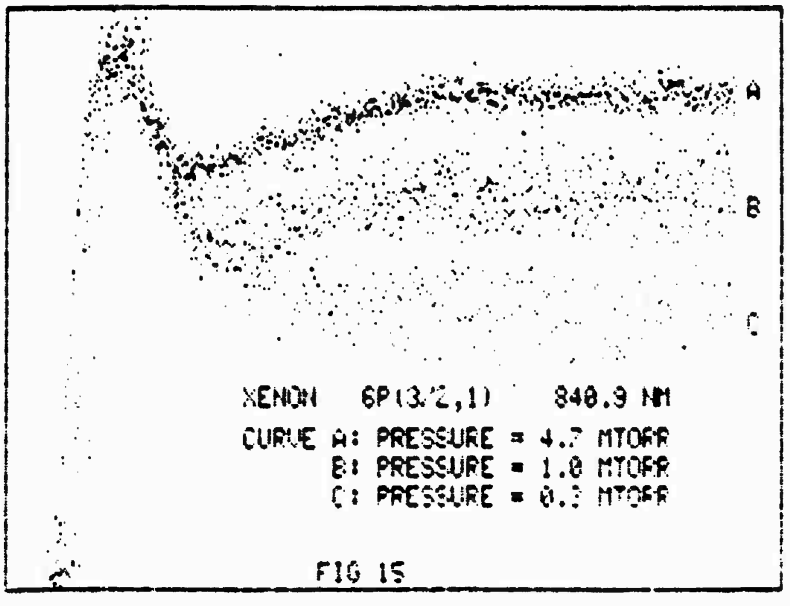

A
B
C
XENON 6P(3/2,1) 840.9 NM
CURVE A: PRESSURE = 4.7 MTORR
B: PRESSURE = 1.0 MTORR
C: PRESSURE = 0.3 MTORR

FIG 15

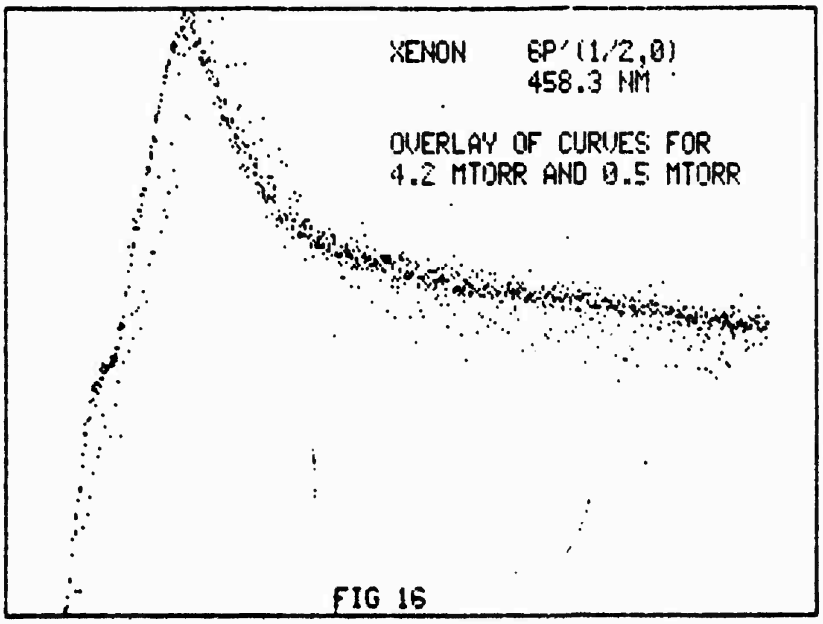

XENON 6P'(1/2,0)
458.3 NM
OVERLAY OF CURVES FOR
4.2 MTORR AND 0.5 MTORR

FIG 16

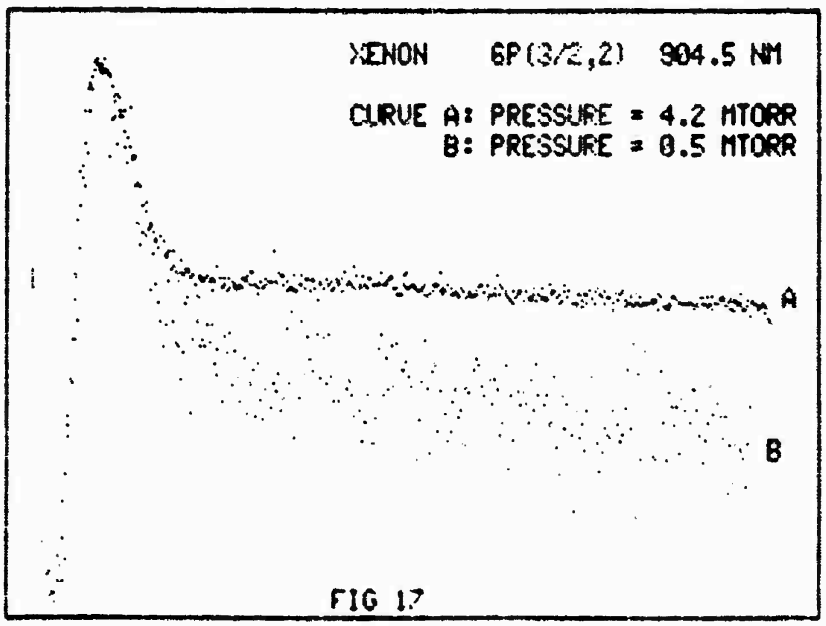

XENON 6P(3/2,2) 904.5 NM
CURVE A: PRESSURE = 4.2 MTORR
B: PRESSURE = 0.5 MTORR
A
B

FIG 17

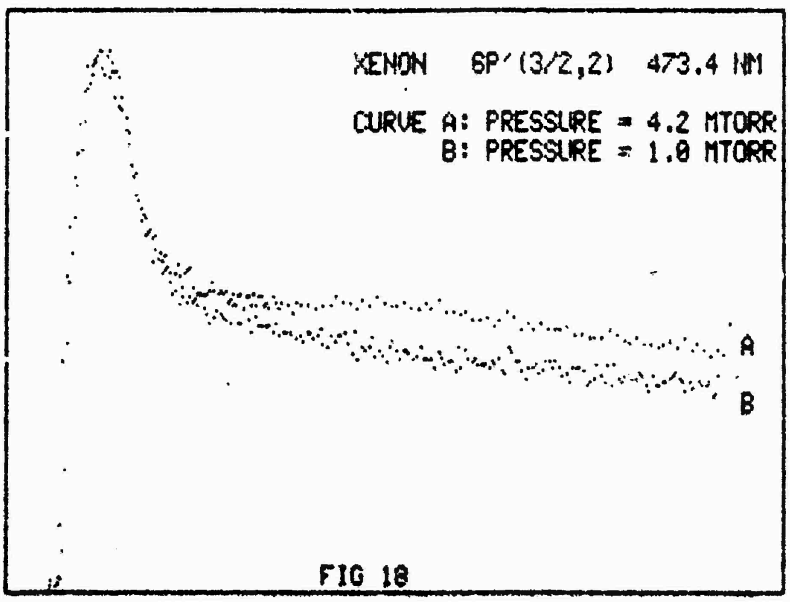

XENON 6P'(3/2,2) 473.4 NM
CURVE A: PRESSURE = 4.2 MTORR
B: PRESSURE = 1.0 MTORR
A
B

FIG 18

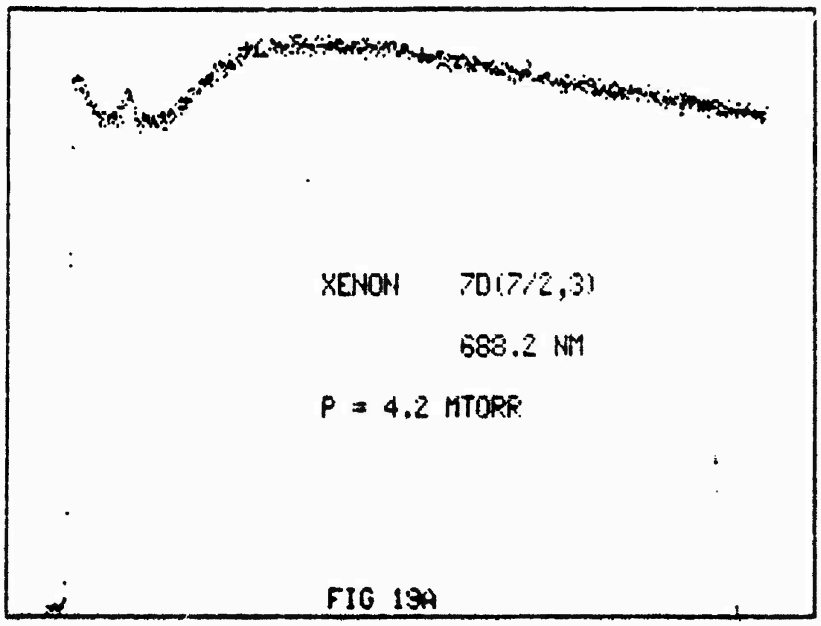


FIG 19A

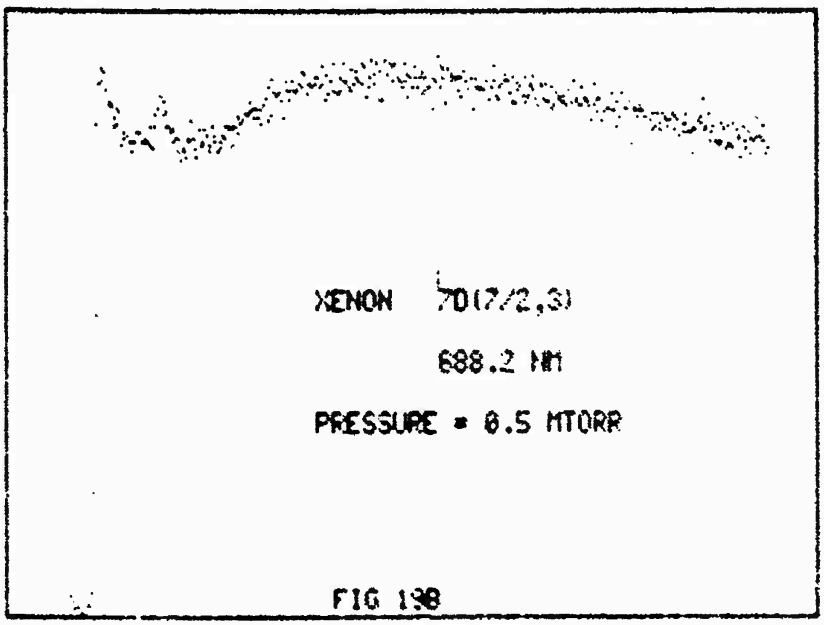


FIG 19B

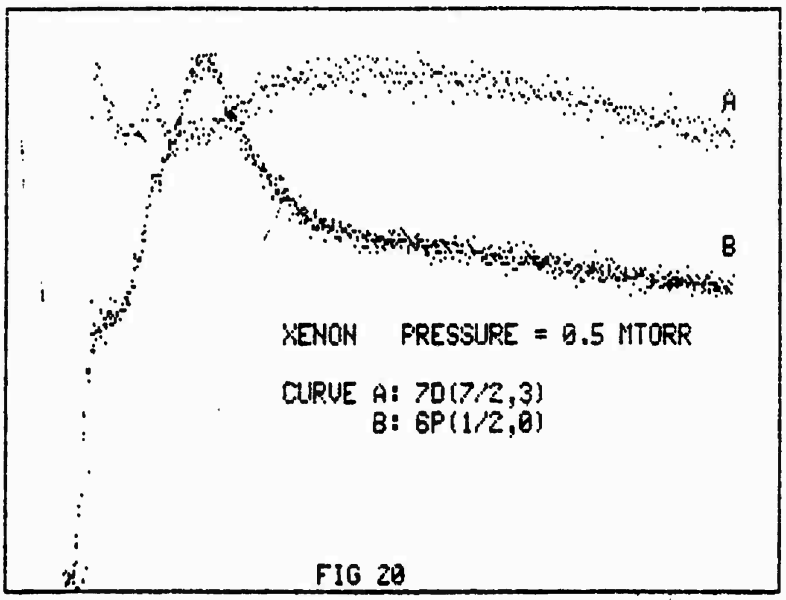


FIG 20

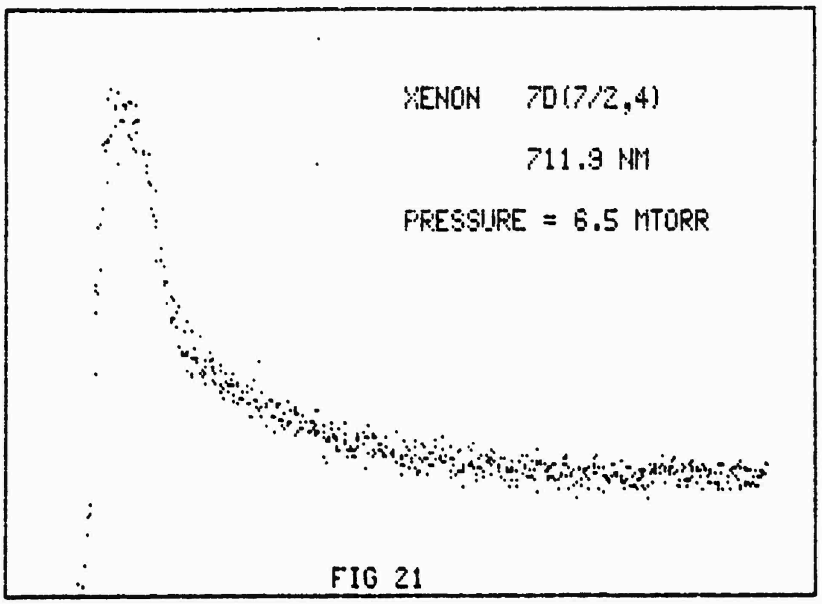

XENON 7D(7/2,4)
711.9 NM
PRESSURE = 6.5 MTORR

FIG 21

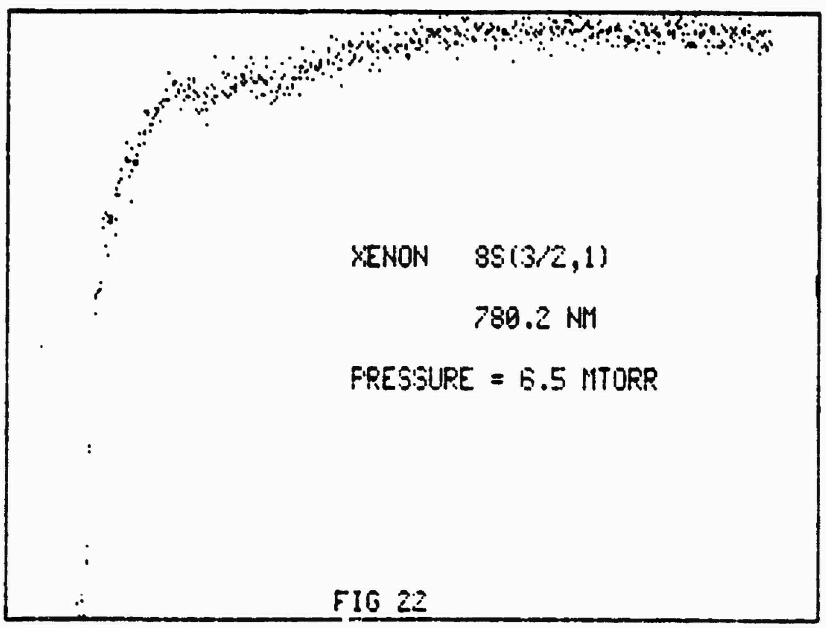

XENON 8S(3/2,1)
780.2 NM
PRESSURE = 6.5 MTORR

FIG 22

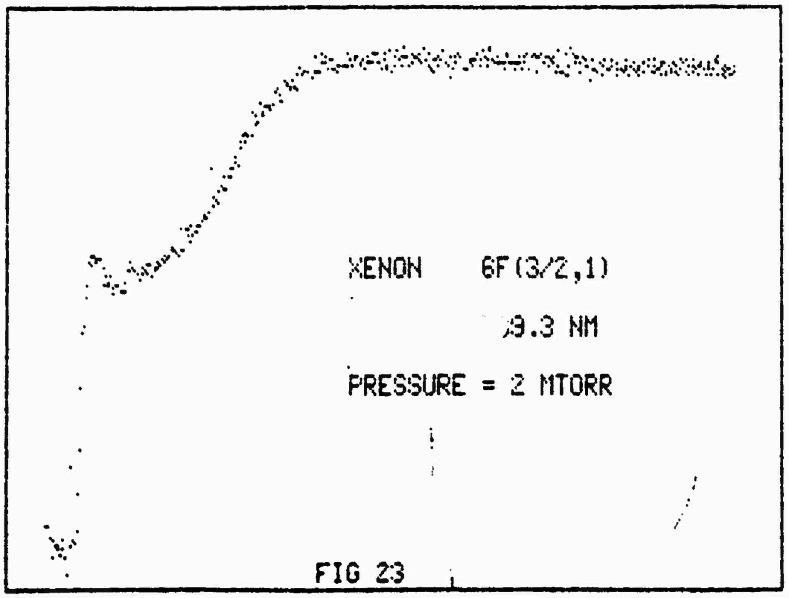

XENON 6F(3/2,1)
λ3.3 NM
PRESSURE = 2 MTORR

FIG 23

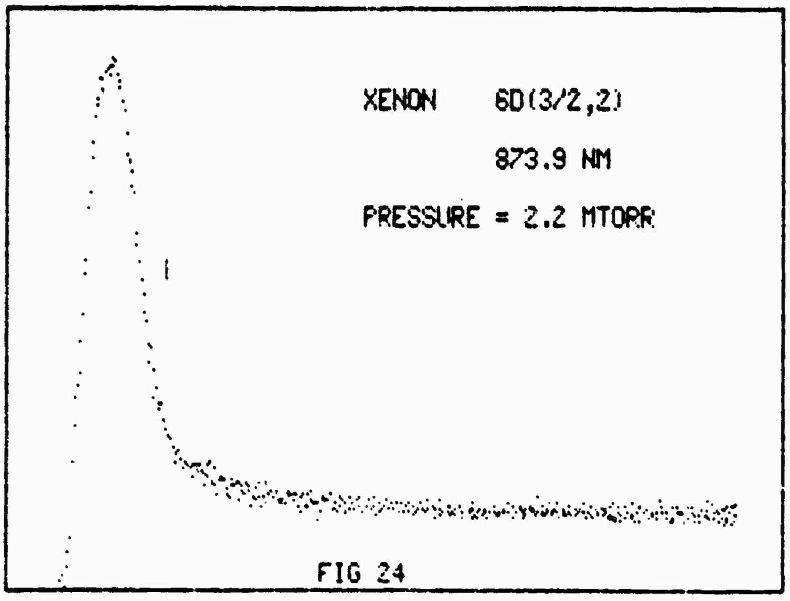

XENON 6D(3/2,2)
873.9 NM
PRESSURE = 2.2 MTORR

FIG 24

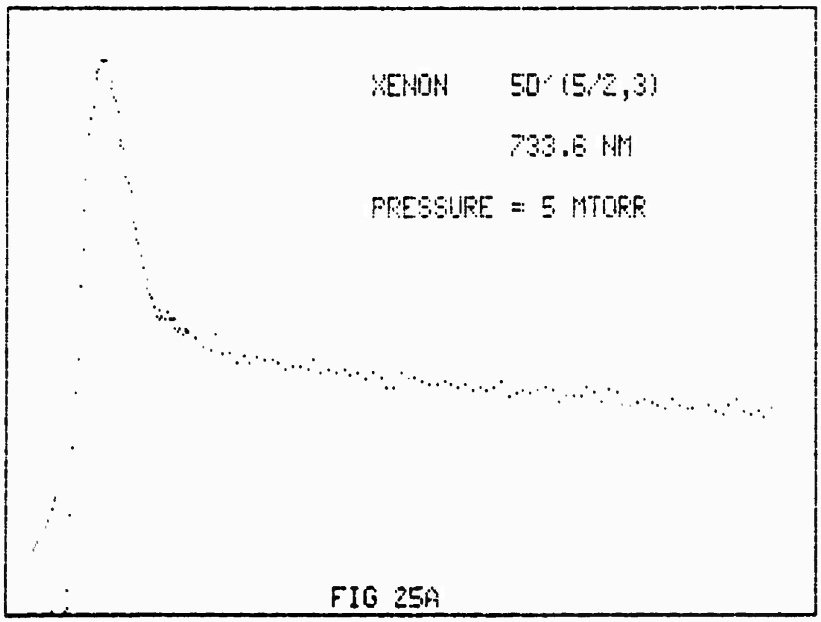


FIG 25A

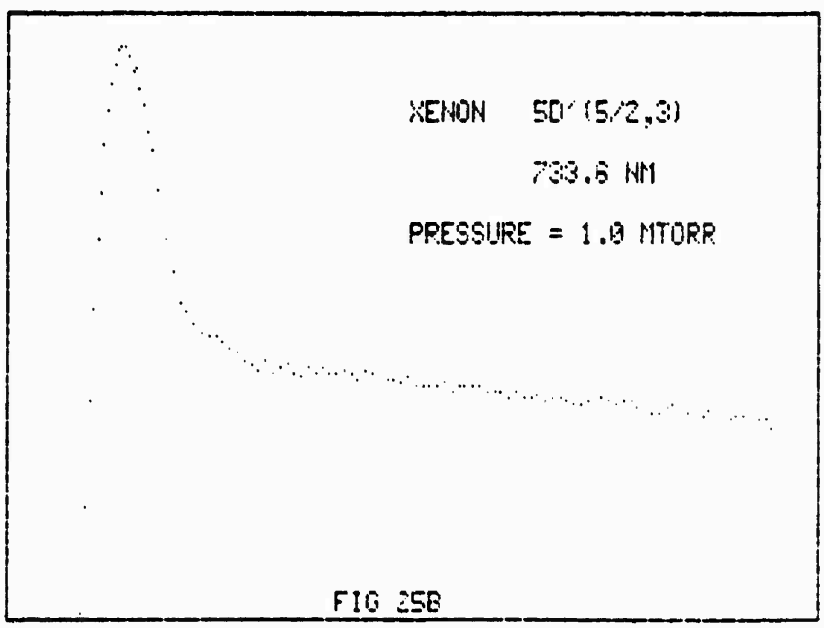


FIG 25B

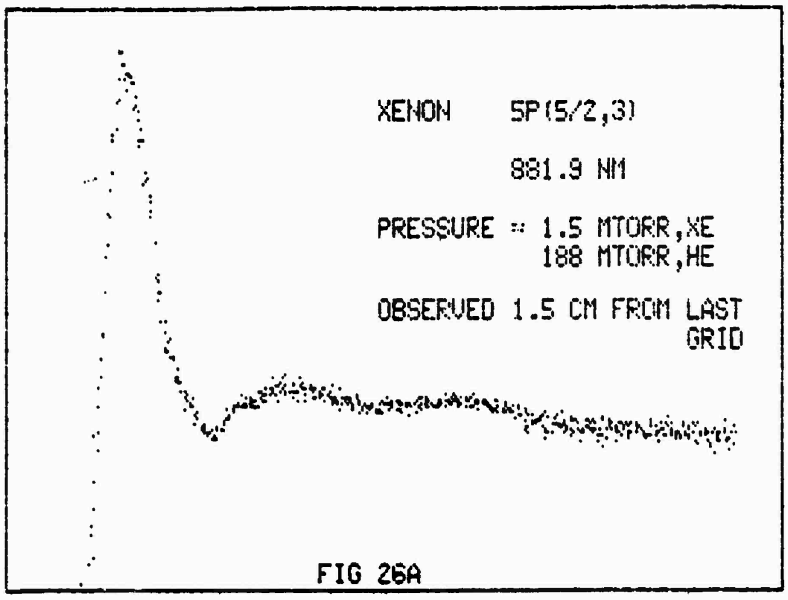


FIG 26A

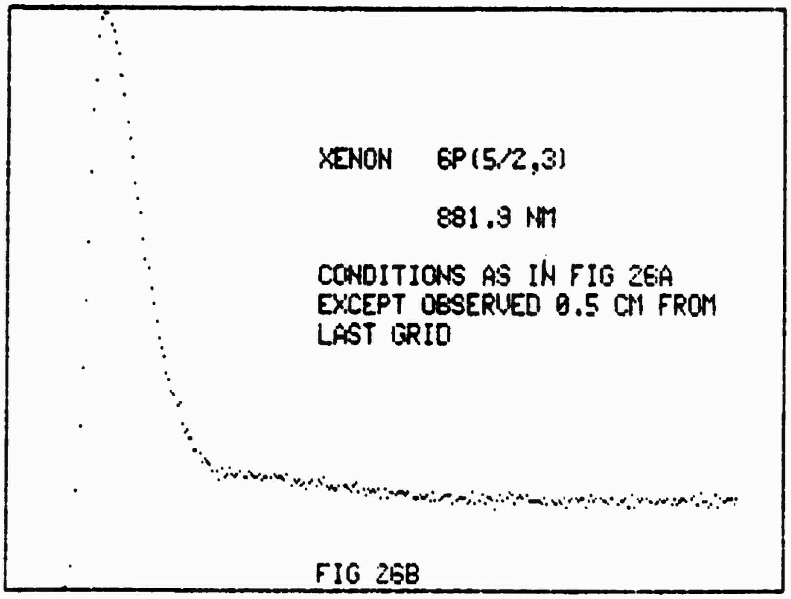


FIG 26B

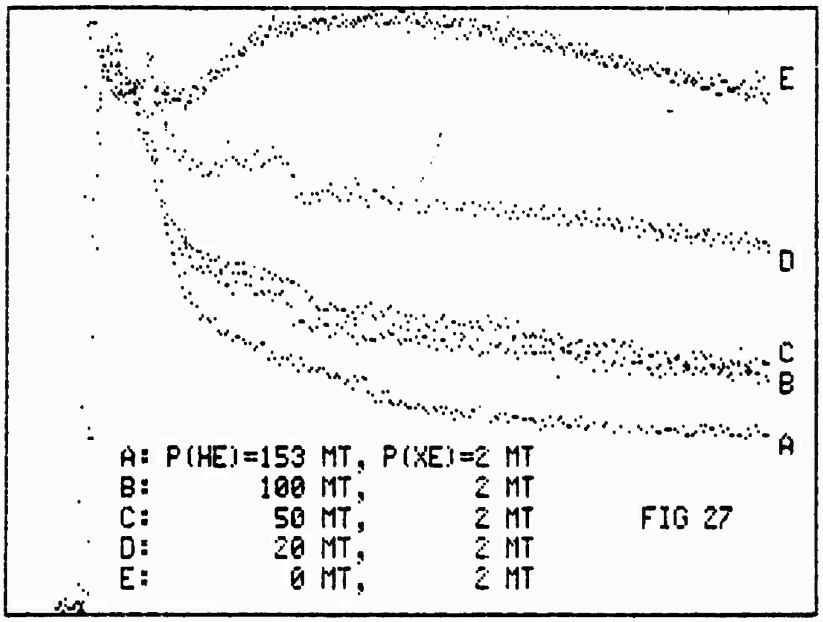

E
D
C
B
A
A: P(HE)=153 MT, P(XE)=2 MT
B: 100 MT, 2 MT
C: 50 MT, 2 MT
D: 20 MT, 2 MT
E: 0 MT, 2 MT

FIG 27

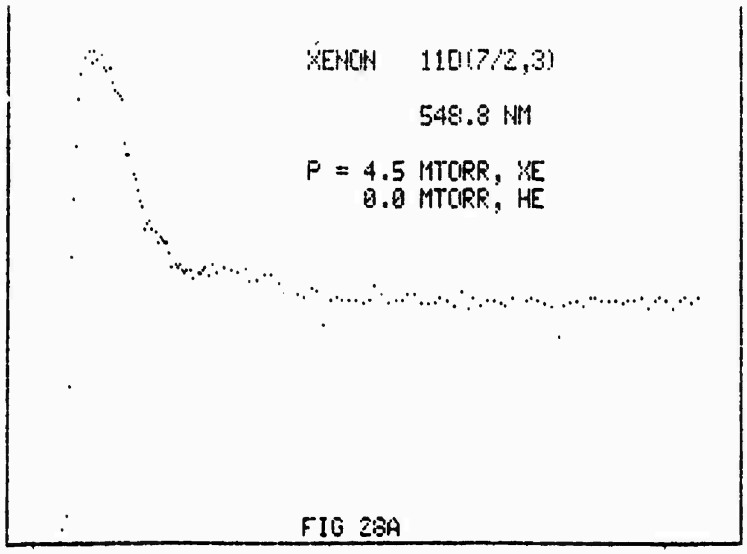


FIG 28A

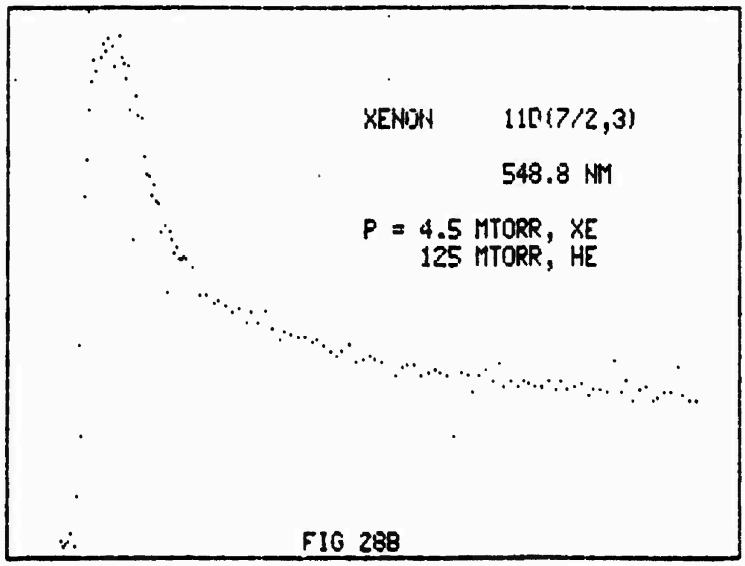


FIG 28B

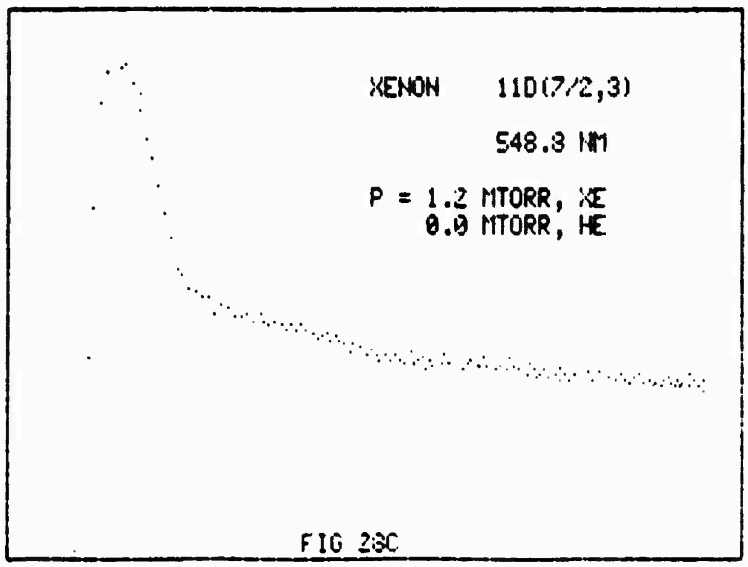


FIG 28C

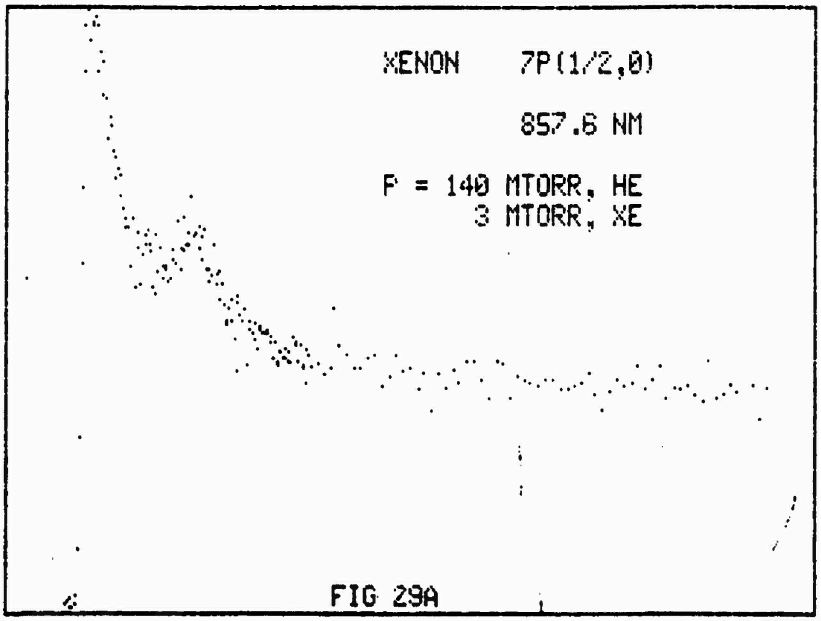


FIG 29A

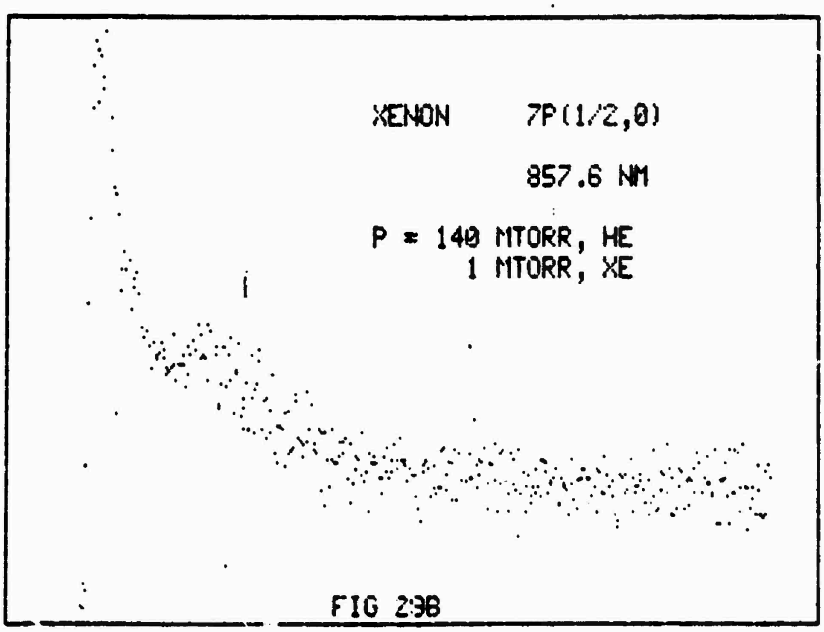


FIG 29B

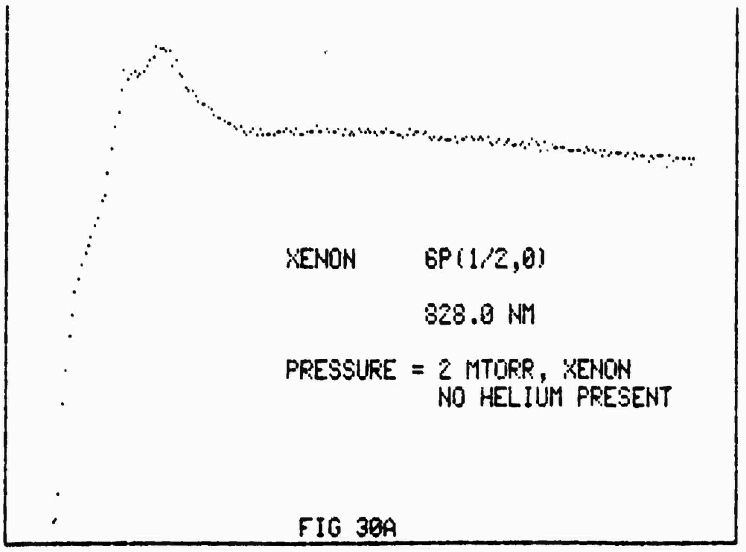


FIG 30A

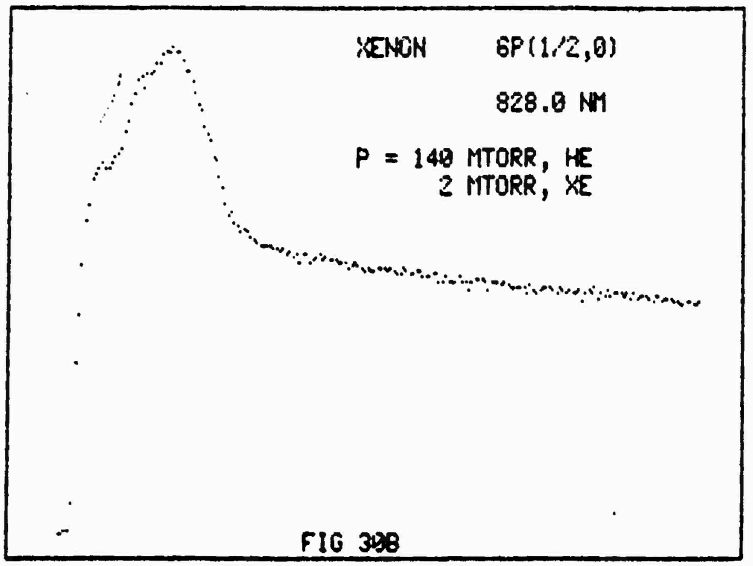


FIG 30B

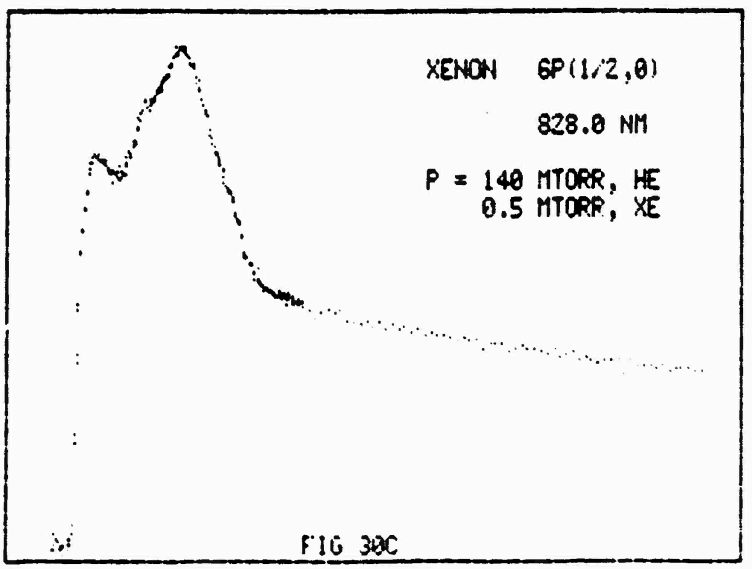


FIG 30C

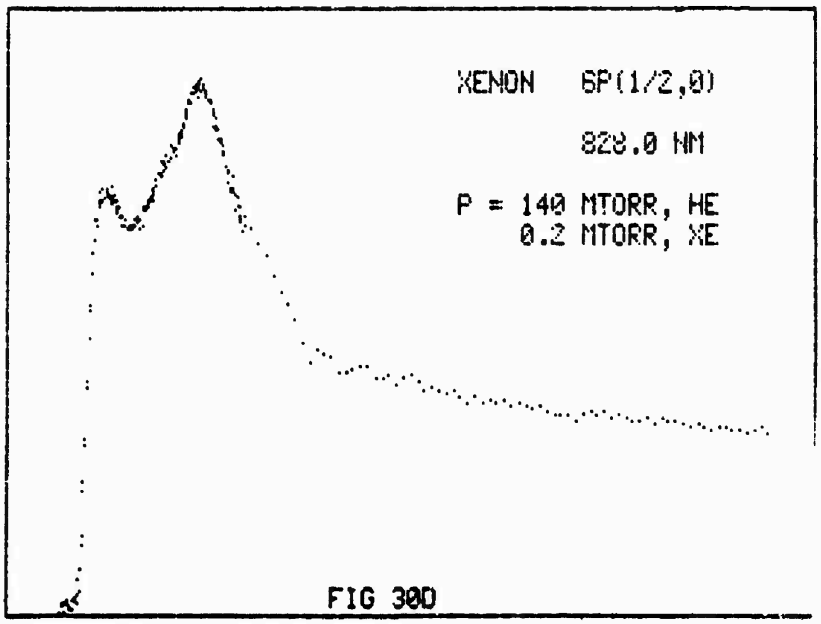


FIG 30D

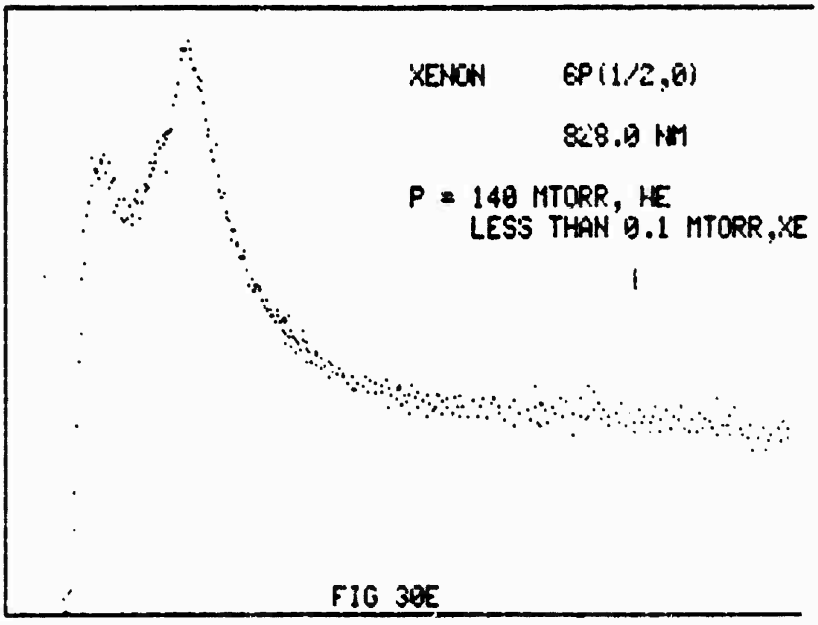


FIG 30E

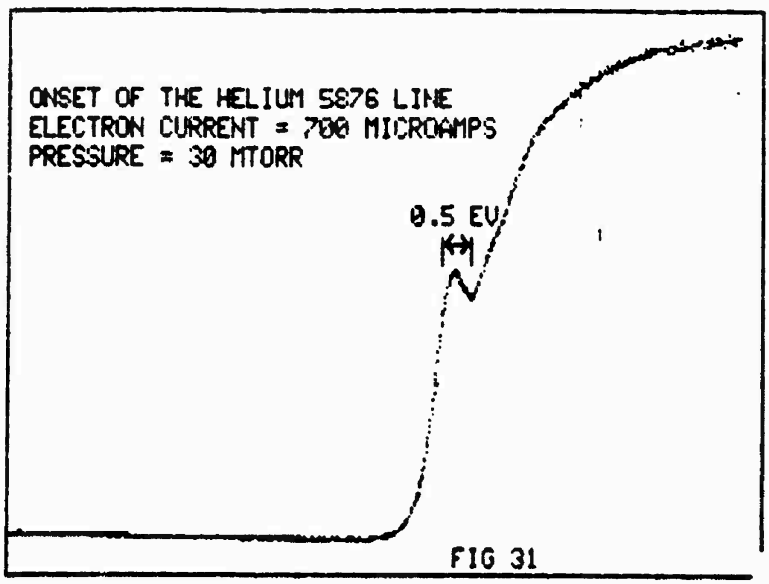

ONSET OF THE HELIUM 5876 LINE
ELECTRON CURRENT = 700 MICROAMPS
PRESSURE = 30 MTORR
0.5 EV

FIG 31

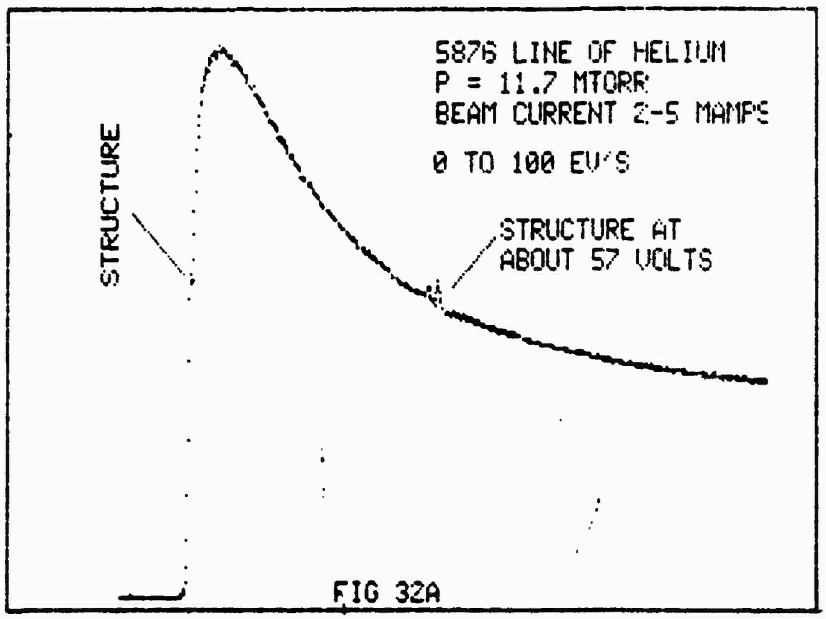


FIG 32A

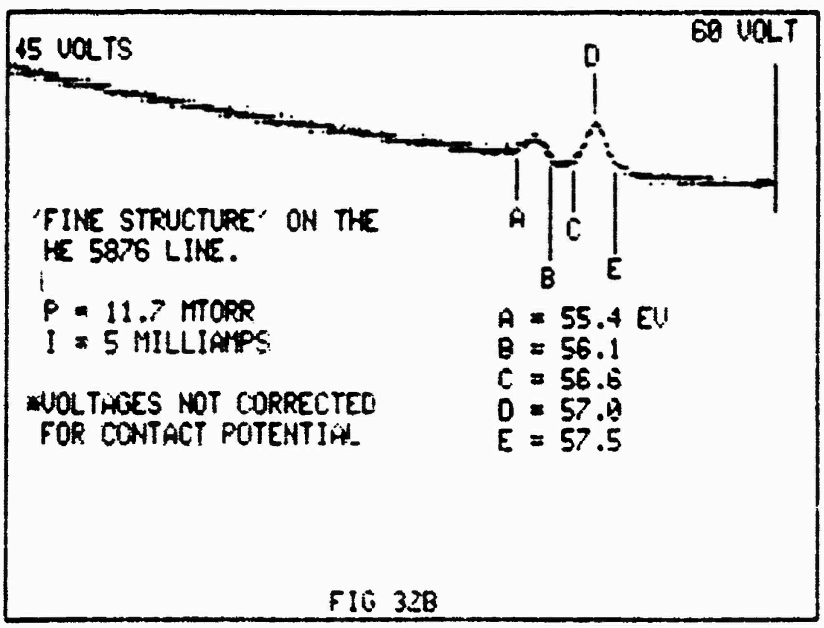


FIG 32B

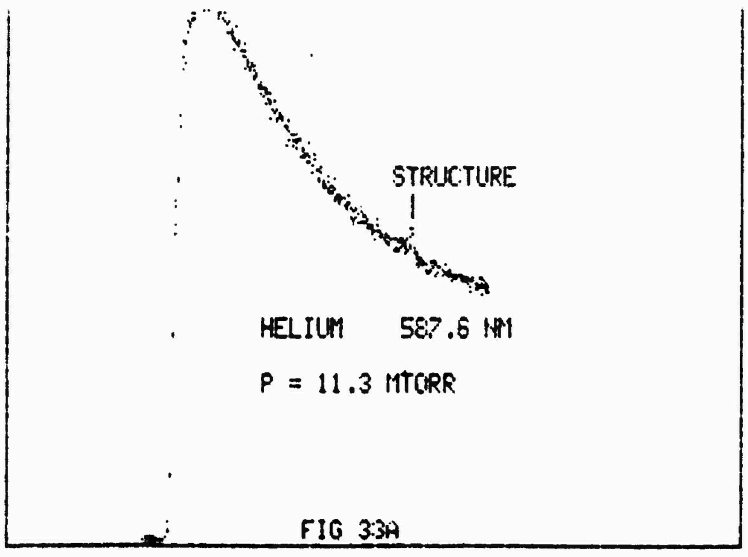


FIG 33A

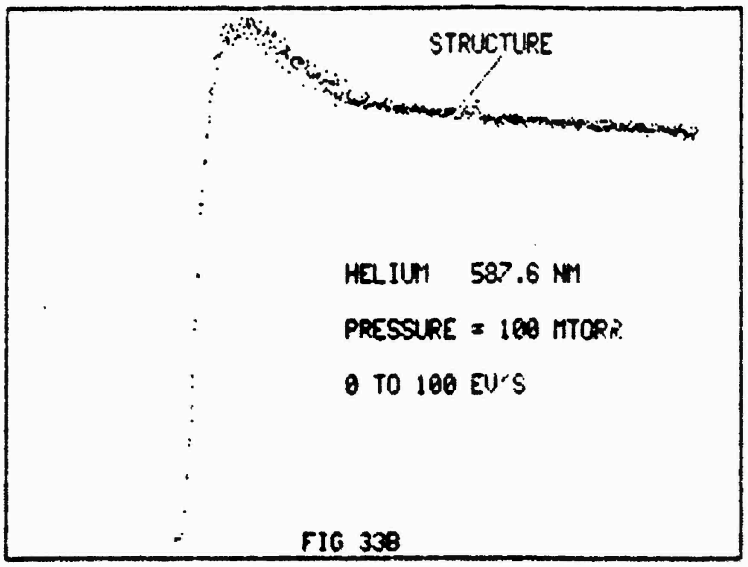


FIG 33B

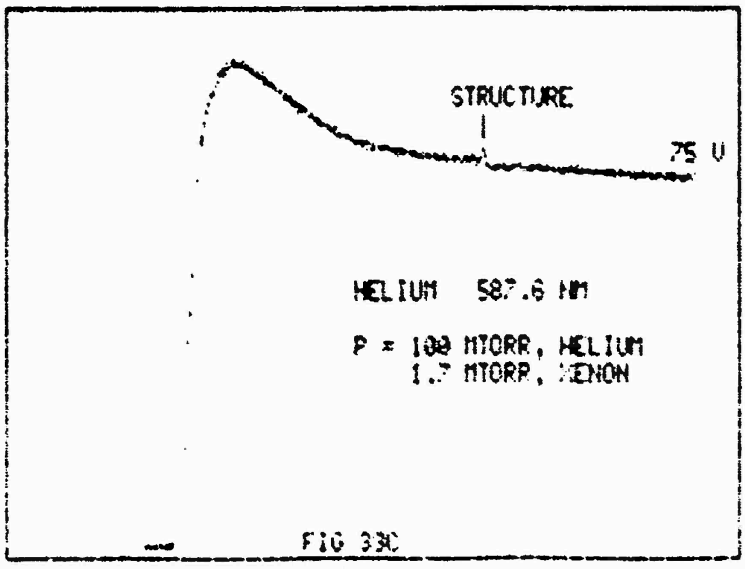


FIG 33C

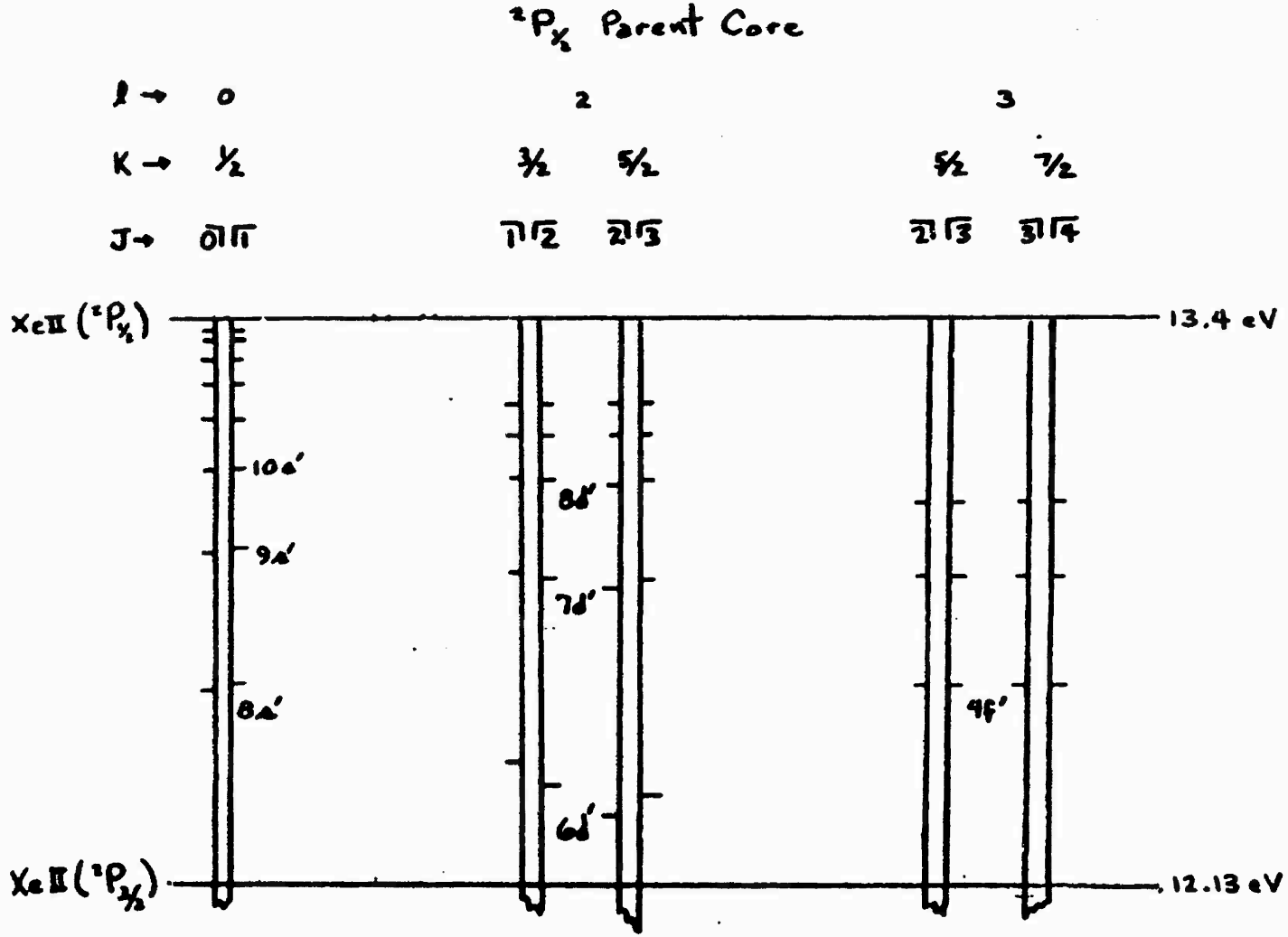

$^2P_{1/2}$ Parent Core
l → 0 2 3
K → 1/2 3/2 5/2 5/2 7/2
J → 0|1 1|2 2|3 2|3 3|4
Xe II ($^2P_{1/2}$)
13.4 eV
10d'
9d'
8d'
8d'
7d'
6d'
4f'
Xe II ($^2P_{3/2}$)
12.13 eV

FIG 34

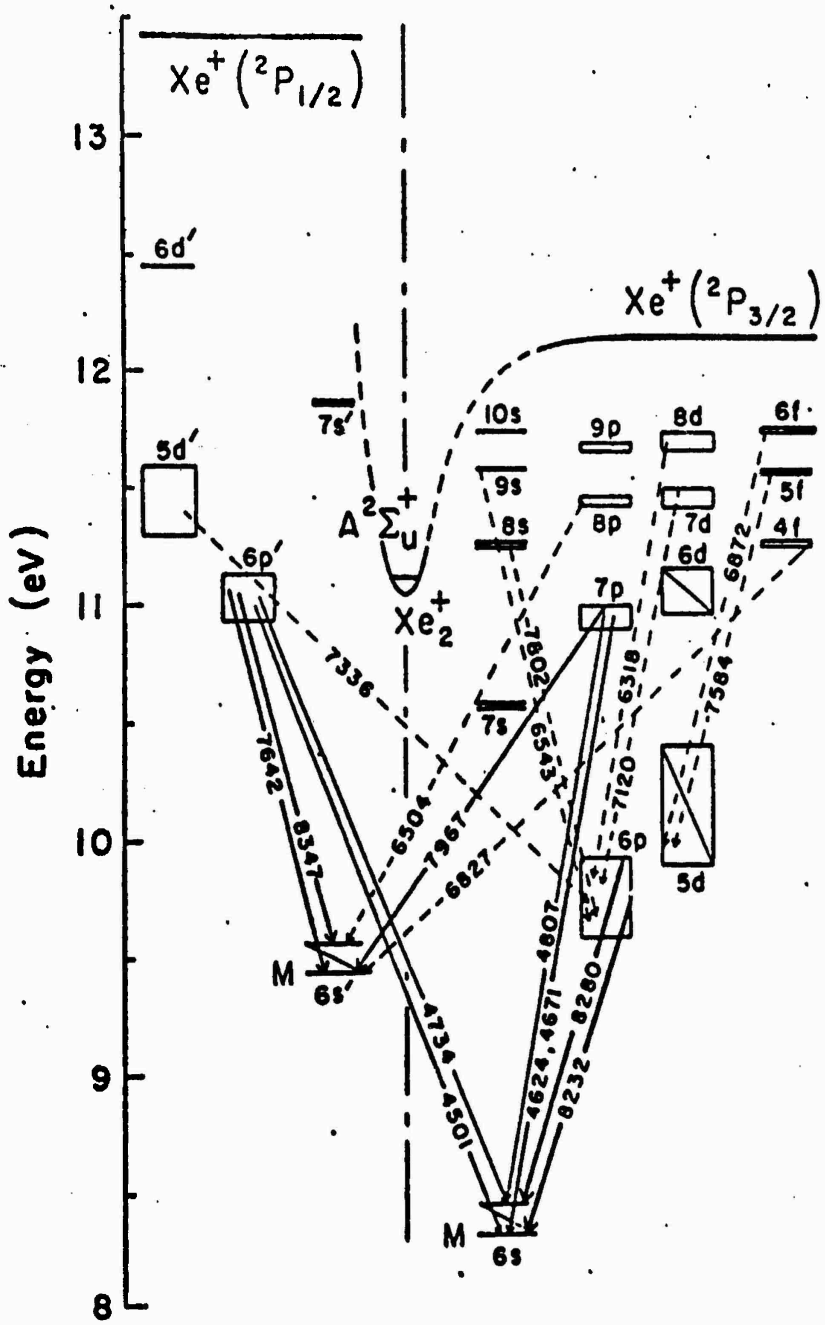


FIG 35 (From Ref. 15)

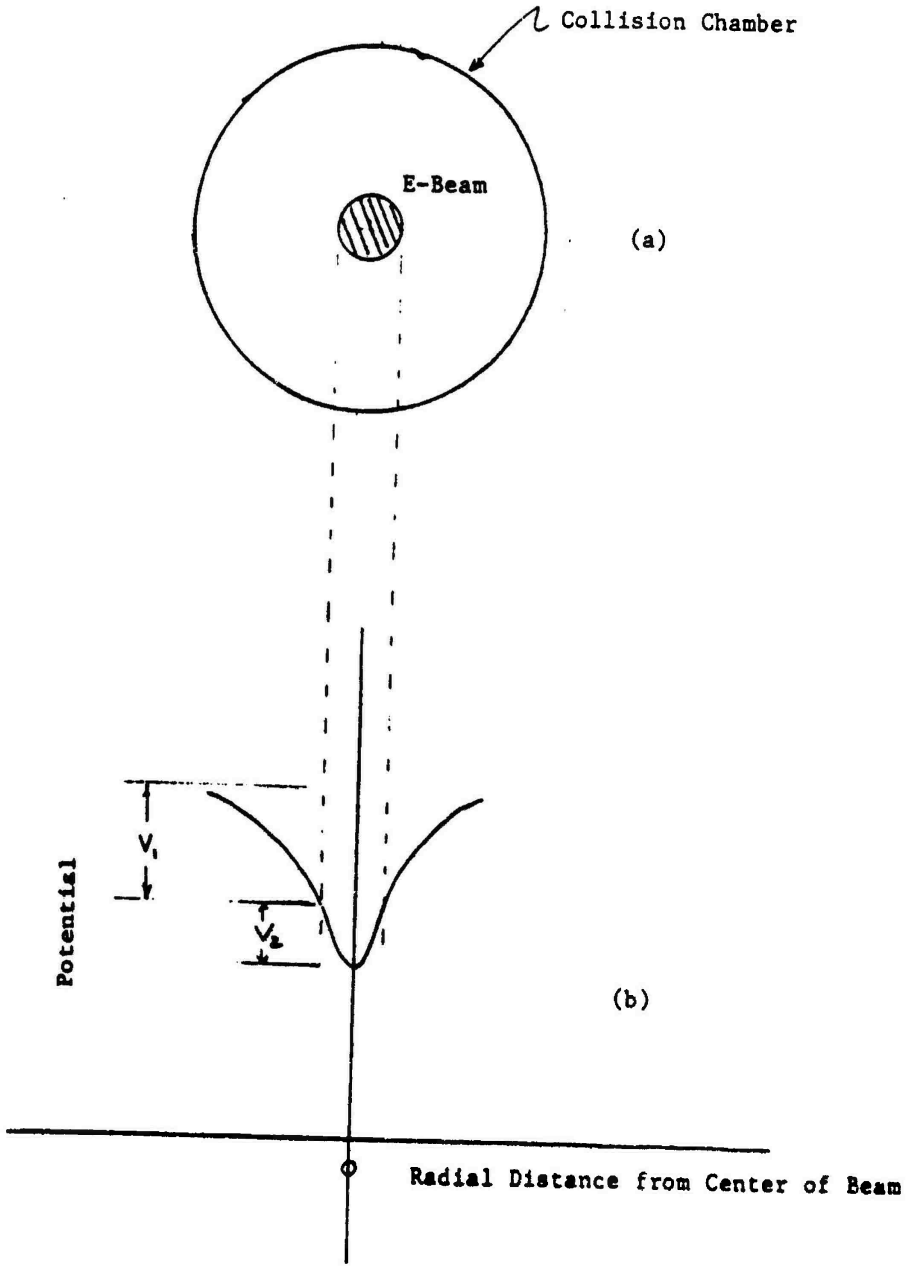

Collision Chamber
E-Beam
(a)
Potential
V_1
V_2
(b)
Radial Distance from Center of Beam

FIG 36

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the
SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

A FINAL REPORT

EFFECT OF GRAPHIC INFORMATION OF READING COMPREHENSION: EYE MOVEMENTS IN READING VERBAL TEXT WITH GRAPH

Prepared by	:	Dr. Shih-sung Wen
Academic Rank	:	Professor of Psychology
Department and University:		Department of Psychology Jackson State University
Research Location	:	Human Resources Laboratory Williams AFB, AZ 85224
Effort Focal Point	:	Dr Thomas Longridge
Date Submitted	:	August 11, 1983
Contract No	:	F49620-82-C-0035

EFFECT OF GRAPHIC INFORMATION ON READING COMPREHENSION:

EYE MOVEMENTS IN READING TEXT WITH GRAPH

by

Shih-sung Wen

ABSTRACT

This study compared reading comprehension between reading verbal text with a graph, and reading verbal text without a graph. Eye movements in reading were monitored, recorded, and analyzed to examine patterns of interaction between text reading and graph viewing. Results showed that (a) text reading with graph slightly but not significantly improved reading comprehension of the text, (b) reading text with graph decreased both reading time and the number of eye fixations on the text and (c) the graph was attended primarily in the second half a reading period, especially in the final quarter. An implication of using eye tracking techniques to study reading was discussed. Some limitations of the study were observed and a follow-on study with changes in sampling and design was recommended.

ACKNOWLEDGEMENT

The author wishes to express gratitude to the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for providing him with this meaningful summer research experience at the Air Force Human Resources Laboratory, Williams AFB, AZ.

The author would like to thank Dr. Thomas Longridge for his valuable suggestions on areas of research and generous assistance in providing equipment and human resources. The author is also indebted to Dr. George Geri for laboratory and technical assistance and helpful comments on manuscripts. Thanks are also due to Paul Wetzel for his suggestions and encouragement. Finally, the author would like to thank Dr. Milton E. Wood for his personal interest in the present study.

It is a common practice for authors of textbooks and scientific papers to include graphs, diagrams, charts, tables and pictures in the text. The underlying rationale of the practice appears to be that pictures are vivid, spatial and are better-recalled than words (Dallett and Wilcox, 1968; Haber, 1970; Nelson, Metzler and Reed, 1974; Pavio, 1974; Pavio, Rogers and Smythe, 1968). It is, therefore, expected by text authors that properly designed and displayed pictorial materials will assist readers in reading and recalling verbally presented texts. Little attention, however, has been given to ascertaining the extent to which reading comprehension is improved due to inclusion of pictorial materials.

In view of recent progress in using sophisticated eye tracking equipment to study perceptual processes in reading, Well (1983) suggested that contrasting the findings from work in text reading and in complex scene perception should lead to deeper insights in both domains. In line with Well's suggestion, Loftus (1983) has compared the similarities and differences between the two domains in terms of the "useful field-of-view", "sequential predictability" and "temporal requirement for completing perceptual processing." In fact, almost all activities involve some interaction between the two systems (Spoehr and Lehmkuhle, 1982). However, the limited number of studies on coordinating text reading and picture viewing (Carpenter and Just, 1975; Clark and Chase, 1972; Glushko and Cooper, 1978; Pezdek, 1977; Potter and Faulconer, 1975) have not identified a mechanism by which the two systems interact in comprehension.

Eye movements in reading reflect visual attention and cognitive processes (Just and Carpenter, 1980; Rayner, 1978; Remington, 1978). Eye fixations, saccades and regressions, though indirect measures of cognitive processes, are so far the best indices of attentional and hence information processing in reading. The technology of monitoring eye movements can, therefore, be used to record a reader's attentional distribution when pictorial illustrations are presented along with printed verbal reading materials.

The purpose of the present study was to examine the effect of graphic data on reading comprehension. Specifically, a comparison was made between reading verbal texts with and without an accompanying graph. To identify patterns of visual interaction between text reading and graph viewing, eye movements including fixations, saccades and regressions were recorded and analyzed.

METHOD

Subject

Six adults, two females and four males, were requested to participate in various reading tasks. Their educational attainment ranged from BS to PhD degrees. All subjects had normal vision and no indication of any eye diseases.

Material and Design

The reading materials consisted of three sets of verbal text (V1, V2, V3) with their corresponding graphs (G1, G2, G3). The length of the three reading texts was 83 words for V1, 94 words for V2, and 108 words for V3. The V1 and G1 were selected and revised from Bell, Loomis and Cervone (1982), the V2 and G2 from Foster and Gavelek (1983) and the V3 and G3 from Bayley (1970). An

additional reading task (PG) was selected and modified from Tilly, et al, (1982) for practice before formal reading took place. All reading materials were taken from scientific research reports whose level of reading difficulty was appropriate for the subjects selected.

All texts were typed with prestige elite letters, 12 characters per inch. The texts were left-margin justified. Since the study required a text-only reading task and text-graph combination reading task, the displacement of the two reading tasks varied. For the text-only reading task, there was no graph and no reference to any figure in the text. For the text-graph combination reading task, a graph was placed beneath the text and a "see Figure X" was indicated early in the text. In either case, the center of reading materials always overlapped the center of the calibration chart.

Reading comprehension was measured with a multiple-choice quiz for each test. The number of items was five for Quiz V1, seven for Quiz V2 and seven for Quiz V3. The 10-point scale was used to convert test scores.

To test the effect of presence or absence of graphic information on text reading comprehension, each subject went through three reading conditions: Condition 1 consisted of repeated readings of the same text-only reading material, Condition 2 consisted of a text-only reading followed by a text-graph combination of the same text; Condition 3 consisted of a text-graph combination reading followed by the same text without a graph. The order of the three reading conditions for each subject was randomized to minimize any specific order effect.

Instrument

A Gulf and Western Eye-Movement Monitor Model 200 was used to monitor eye fixations and saccades in reading. The device uses low level, invisible infrared eye illumination and synchronous detection of the reflected signals. The instrument measures horizontal position of the eye over a range of approximately $\pm$20 degrees with an accuracy of one degree. For measuring vertical eye movements, the range is +10 degrees (up) and -20 degrees (down) with an accuracy of approximately 1 degree.

Other instruments, such as a Hitachi CCTV camera, an ITC Ikegami 9-inch TV monitor and a strip-chart recorder were also used to monitor and record the data.

Procedure

The subject was told that the purpose of the study was to identify eye movement patterns in normal reading. The subject was also informed about the reading tasks, calibration procedure and required activities during the study. During the initial warm-up conversation (about the subject's interests, career, etc.), the bitebar for stabilizing head position was prepared. When the subject was seated in front of the eye-movement monitor, alignments and adjustments were made for calibration and reading. A 12-point (30 deg X 30 deg) calibration chart was used for each of the reading materials. The horizontal eye movement of the left eye and the vertical eye movement of the right eye were simultaneously monitored.

The reading materials were placed 14 inches from the subject's eyes. The size and visual angle of the three reading materials were:

	Size (in)	Visual angle (deg)		Size (in)	Visual angle (deg)
V1	6.75 X 2.33	27.1 X 9.5	V1G1	6.75 X 6.33	27.1 X 25.5
V2	6.75 X 3.00	27.1 X 12.2	V2G2	6.75 X 7.00	27.1 X 28.1
V3	6.75 X 3.33	27.1 X 13.6	V3G3	6.75 X 7.00	27.1 X 28.1

After calibration and performance of the practice task, the subject proceeded with three reading conditions in random order. At the end of the each reading condition the subject was removed from the apparatus and given the appropriate quiz. Then the subject returned to the monitor and resumed his/her reading position for recalibration and reading. Concurrently, the recording of eye movements was carried out throughout the calibration and reading.

RESULTS

Reading Comprehension

The means and standard deviations of the reading comprehension score across the six subjects were respectively 7.5 and 3.20 for reading Condition 1 (text-text), 9.0 and 1.15 for reading Condition 2 (text-text/graph), and 9.5 and 0.50 for reading Condition 3 (text/graph-text). The data show that under Conditions 2 and 3, in which a graph was once presented along with the text, the mean comprehension scores were higher and scores were less variables than the comprehension score under Condition 1.

The trend in favor of Conditions 2 and 3 on reading comprehension, however, was not statistically significant by a one-way ANOVA, within-subjects repeated measures design, the results of which are presented in Table 1. That is, for the data collected in the present study, the mean differences in reading comprehension were not significant between presence and absence of a graph in the reading text.

TABLE 1

ANOVA Of Reading Comprehension Scores In Three Reading Conditions

Source of Variance	SS	df	MS	F
Between Subjects	20	5	4.0	
Within Subjects	64	12	5.3	
Reading Conditions	13	2	6.5	1.27 (n.s.)
Residual	51	10	5.1	
Total	84	17		

Eye Movements

An analysis of analog eye movements data from the strip chart recordings was conducted in terms of the mean of (a) total time spend in reading, (b) number of fixations, (c) number of regressions and (d) fixation duration.

As shown in Table 3, reading time for all three conditions decreased substantially as the result of the first reading. Although this was normally expected, the initial reading time for the text itself was remarkably short for Condition 3, wherein a text-graph combined reading material was presented first. Similarly, reading time for the second reading in Condition 3 was also much shorter than in the second reading of the other two conditions.

TABLE 2

Reading Time, Number of Fixations and Regressions and Fixation Duration

	Reading Condition		
	1	2	3
Total Reading Time (sec)			
First Reading	43.8	46.7	33.9*
Second Reading	30.2	31.1*	22.8
Number of Fixations			
First Reading	148	163	121
Second Reading	103	109	83
Number of Regressions			
First Reading	23	27	22
Second Reading	14	17	11
Mean Fixation Duration (ms)			
First Reading	234.7	218.2	219.5
Second Reading	191.7	220.0	213.3

*For the text only

The advantage of Condition 3 over Conditions 1 and 2 was further demonstrated by virtue of the smaller number of fixations on the text evident during reading. Data in Table 3 showed a clear numerical advantage in fixations for both first and second readings in Condition 3 over those of other reading conditions. The fixation frequencies of the second reading under both Conditions 1 and 2 were similar, as were their total reading times.

The number of regressions decreased from first to second reading for all reading conditions although differences among conditions were negligible.

The mean fixation time or fixation duration was computed by dividing the total reading time less saccade time by the number of fixations. As shown in Table 3, the fixation duration was relatively constant for both first and second readings across the three conditions.

A question of interest concerns the relative degree of attention accorded the graph during the course of reading. To answer this question, the vertical eye movement data, which were used to detect changes in visual attention between the text and the graph, were further analyzed. Results of the analysis are displayed in Figure 1.

Insert Figure 1 Here

As can be seen in the Figure, about 36% (for Condition 2) and 43% (for Condition 3) of total reading time was spent in viewing graphs and approximately 80% of graph viewing took place in the second half of the reading period, particularly in the final quarter.

Within this general pattern of the text-graph eye-view interaction, an individual idiosyncracy appeared to be functioning across different contents of the text. In Conditions 2 and 3, for example, subjects C and E both looked at the graph about twice: one in an earlier stage and the other at the end of the reading period.

DISCUSSION

The results of an analysis of reading comprehension, coupled with examinations of eye movement patterns, indicated that given the stimulus materials and measurement design employed in the present study, presence of a graph did not significantly increase comprehension of text, but did result in less reading time and fewer eye fixations for the text.

A finding of particular interest is that an initial reading with text and graph in Condition 3 was characterized by having shorter reading time and fewer eye fixations while maintaining relatively high reading comprehension.

Since the time spent in viewing the graph was shorter for the first reading of Condition 3 (18.8 seconds) than for the second reading of Condition 2 (23.2 seconds), it is reasonable to speculate that in Condition 2 the graph appears to serve the reader primarily for verification purposes rather than for comprehension of the text.

Another interesting finding is that most subjects delayed their attention to the graph until the text reading was completed, although their attention to the graph was called for (by "see Figure X") early in the reading. Apparently, the subject did not want to be disrupted during reading. Or, the role of a graph in text reading has not been properly defined.

One plausible reason for nonsignificant differences in comprehension among three reading conditions could be the lack of a specific instruction on time allowed to finish a reading. Subjects were told to "read normally" and "not to study." They were, however, reminded that a comprehension quiz would be

administered upon completion of every second reading. Apparently, having a quiz after reading influenced subjects' desire to read carefully and thus increased their reading time enormously. Some subjects, for example, went over a paragraph more than three times and exceeded 60 seconds for a paragraph of less than ten lines. Thus a possible graphic effect on comprehension was "swamped" in high comprehension test scores in the two conditions.

Another plausible reason was that the subject appeared to have difficulties in getting used to reading under laboratory conditions. The subject was surrounded by instruments, being on a bitebar, being bodily still for minutes and burdened with tedious calibrations before reading.

Although the above methodological problems and the artificiality of the laboratory situation remain to be improved, one implication from the present study can be drawn. Research on how to integrate graphic information and the verbal text to enhance readers' comprehension and short-term and long-term memory has been astonishingly lacking. The present study, an effort to investigate the text-graph interaction in reading through current eye tracking and recording techniques, is a step toward better understanding of reading behavior. For an optimal level of comprehension and memory in reading, strategies should be developed to induce proper interactions between the verbal text and the picture. Studies of this kind should have a great impact on existing editorial guidelines which commonly neglect how pictures can be much more intelligently used to improve comprehension and memory of the text reading.

RECOMMENDATIONS

In the present study, the following limitations were observed: (a) The population from which the sample was drawn was limited and, therefore, biased; (b) The duration of the study was too short to allow adequate data collection, analysis, interpretation; (c) Other strategies (different experimental designs and data analyses) for conducting this research were unable to be tested due to time and resource limitations.

It is, therefore, recommended that a follow-on study be conducted to (a) adequately cope with the above limitations and (b) further enhance our understanding of the interactive effects of text and pictorial representation of information.

REFERENCES

Bayley, N. Behavioral Correlates of Mental Growth: Birth to Thirty-six Years. In B. L. Kintz and J. L. Bruning (Eds), Research in Psychology. Glenview, IL: Scott, Foresman and Co., 1970.

Bell, P. A., Loomis, R. J. and Cervone, J. C. Effects of Heat, Social Facilitation, Sex Differences and Task Difficulty on Reaction Time. Human Factors, 1982, 24, 19-24.

Carpenter, P. A. and Just, M. A., Sentence Comprehension: A Psycholinguistic Processing Model of Verification. Psychological Review, 1975, 82, 45-73.

Clark, H. H. and Chase, W. G. On the Process of Comparing Sentence Against Pictures. Cognitive Psychology, 1972, 3, 472-517.

Dallett, K. and Wilcox, S. G. Remembering Pictures Vs Remembering Descriptions. Psychonomic Science, 1968, 11, 139-140.

Foster, R. N. and Gavelek, J. R. Development of Intentional Forgetting in Normal and Reading-delayed Children. Journal of Education Psychology, 1983, 75, 431-440.

Glushko, R. J. and Cooper, L. A. Spatial Comprehension and Comparison Processes in Varification Tasks. Cognitive Psychology, 1978, 10, 391-421.

Haber, R. N. How We Remember What We See. Scientific American, 1970, 222, 104-112.

Just, M. A. and Carpenter, P. A. A Theory of Reading: From Eye Fixations to Comprehension. Psychological Review, 1980, 87, 329-354.

Loftus, G. R. Eye Fixations on Text and Scenes. In K. Rayner (Eds), Eye Movements in Reading: Perceptual and Language Processes. New York: Academic Press, 1983.

Nelson, T. O., Metzler, J. and Reed, D. A. Role of Details in the Long-term Recognition of Pictures and Verbal Descriptions. Journal of Experimental Psychology, 1974, 102, 184-186.

Pavio, A. Spacing of Repetitions in the Incidental and Intentional Free Recall of Pictures and Words. Journal of Verbal Learning and Verbal Behavior, 1974, 13, 497-511.

Pavio, A., Rogers, T. B. and Smythe, P. C. Why Are Pictures Easier to Recall Than Words? Psychonomic Science, 1968, 11, 137-138.

Pezdek, K. Cross-modality Semantic Integration of Sentence and Picture Memory. Journal of Experimental Psychology: Human Learning and Memory, 1977, 3, 515-524.

Potter, M. C. and Faulconer, B. A. Time to Understand Pictures and Words. Nature, 1975, 253, 437-438.

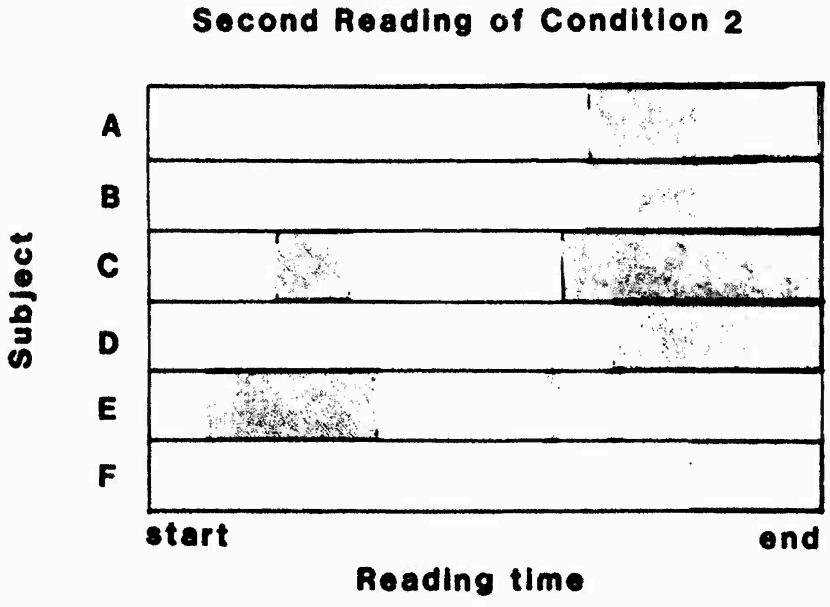
Rayner, K. Eye Movements in Reading and Information Processing. Psychological Bulletin, 1978, 85, 618-660.

Remington, R. W. Attention and Saccadic Eye Movements. Journal of Experimental Psychology: Human Perception and Performance, 1980, 6, 726-744.

Spoehr, K. T. and Lehmkuhle, S. W. Visual Information Processing. San Francisco: Freeman, 1982.

Tilley, A. J., Wilkinson, P. T., Warren, P. S. G., Watson, B. and Drud M. The Sleep and Performance of Shift Workers. Human Factors, 1982, 24, 629-641.

Well, A. D. Perceptual Factors in Reading. In K. Rayner (Eds), Eye Movements in Reading: Perceptual and Language Processes. New York: Academic Press, 1983.



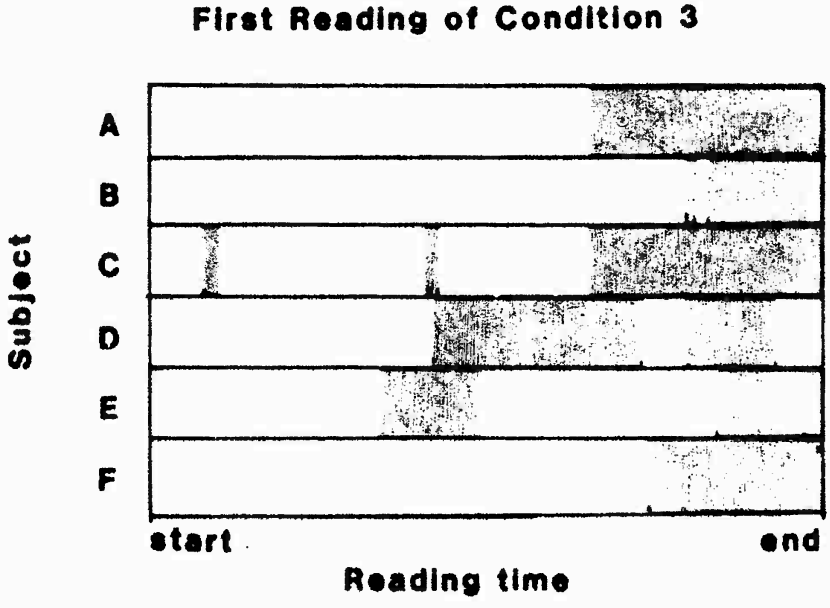


Fig. 1. Time Spent on Viewing Graph (shaded area)

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

INTRAGASTRIC ADMINISTRATION OF DIBROMOMETHANE TO RATS

Prepared by:	Kenneth J. Williamson and Lynn E. Cashion
Academic Rank:	Associate Professor and Graduate Student, respectively
Department and University:	Department of Civil Engineering Oregon State University
Research Location:	Toxic Hazards Division AF Aerospace Medical Research Laboratory Wright-Patterson AFB, Ohio 45433
USAF Research Contact:	Lt Col M. G. MacNaughton
Date:	August 13, 1983
Contract No:	F49620-82-C-0035

INTRAGASTRIC ADMINISTRATION OF DIBROMOMETHANE TO RATS

by

KENNETH J. WILLIAMSON AND LYNN E. CASHION

ABSTRACT

Partition coefficients for dibromomethane for saline, rat bile, corn oil, and mineral oil were determined by gas absorption. Dibromomethane was orally dosed to rats in saline at 10, 20, 50, and 100 mg/kg and the resulting time course for blood concentrations of the parent compound and carboxyhemoglobin were measured. The time traces were simulated using a physiological toxicokinetic model including partitioning to various body compartments, first-order absorption from the gut lumen, liver metabolism, kidney excretion, and exhalation. The model could not accurately simulate the experimental data with a single first-order absorption coefficient applied to all the dosages.

ACKNOWLEDGMENTS

We wish to acknowledge the encouragement and support offered by Lt Col M. G. MacNaughton, Deputy Director, Toxic Hazards Division; Dr. Melvin Andersen, Chief, Biochemical Toxicology Branch, for his continued technical advice; and Maj H. J. Clewell for his assistance with the computer simulations. Special thanks are offered to Lt G. Jepson, SSgt D. Voisard, SSgt B. Burgess, and Ms. Betty Anderson for their technical support with the laboratory experiments. Financial support for Dr. Williamson and Ms. Cashion was obtained through the Summer Faculty and Graduate Student Research Program administered by the Southeastern Center for Electrical Engineering Education (SCEEE) under the direction of Dr. Warren D. Peele. Research support was provided by the Toxic Hazards Division.

I. INTRODUCTION

A physiologically-based (PB) pharmacokinetic model was developed by Ramsey and Andersen (1983) to describe the behavior of styrene in rats and humans. The model proved capable of relating the applied dosages of styrene to measured internal concentrations in blood and tissues. Good agreement was achieved between predicted and measured styrene levels in blood and fat for rats and humans with a single set of physiological and biochemical constants.

Ramsey and Andersen (1983) also attempted to model the uptake of styrene after both iv and oral dosing. The iv dosage experiment of Withey and Collins (1977) was successfully described with the physiological and biochemical constants used to model Ramsey and Andersen's data from inhalation experiments. In addition, the results of the po dose experiment of Withey (1976) were described by the model assuming first-order absorption from the gut lumen (first-order absorption constant, ka, equal to 5.5/hr). However, styrene administered in vegetable oil (Withey, 1976) could not be successfully modelled assuming either first-order absorption, zero-order absorption, or a combination of the two.

The PB pharmacokinetic modeling approach has been further developed to describe the kinetics of halogenated methanes. In particular, dibromomethane has been chosen for further research because of its lipophilic nature, known metabolic pathways, and ease of measurement of both the parent compound and the metabolic products of bromide and carbon monoxide (as carboxyhemoglobin). This reseach problem was similiar to the work conducted by the senior author for the Environmental Protection Agency concerning absorption of priority pollutants by invertebrates, and as such, he was assigned to the Air Force Aerospace Medical Research Laboratory (AFAMRL), Wright-Patterson AFB, Ohio, for his summer faculty appointment.

II. OBJECTIVES

The objective of this research was to further examine the mechanisms of absorption of dibromomethane (DBM) after intragastric administration. This research was designed to complement on-going studies of DBM administration by inhalation and intravenous injection at the AFAMRL. Due to time constraints, the DBM was administered only in saline and distilled water.

III. METHODS

Dibromomethane

The halogenated methanes are widely used in medicine and industry, and have recently been shown to be produced in the chlorination of fresh and saline waters. Dibromomethane has low mammalian toxicity; it is rapidly metabolized by the P-450 mixed function oxidase system to carbon monoxide, and by a glutathione transferase pathway to carbon dioxide (Anders et al., 1977).

Animals

Male Fischer 344 (CDF) rats weighing 190 to 250 g were maintained in a constant temperature and humidity chamber. Water and feed were provided ad libitum during holding. Food was removed for at least 16 hours prior to dosing.

To allow serial blood sampling, a polyethylene cannula was inserted into the right jugular vein after anesthesia with an ip dose of 70% ketamine-30% xylazine. The cannula was filled with a 10% heparin-saline solution to prevent clotting.

For experiments involving dosage into the small intestine, a midline incision was made and a flexible tubing cannula inserted into the duodenum just below the entrance of the bile duct. The cannula was sutured to the intestine wall and exteriorized through the peritoneal cavity wall. For experiments involving ligation of the small intestine, a loop of suture was placed around the duodenum and tied with a slip knot; just before dosing the slip knot was held in place with a piece of small diameter Tygon tubing and tightened securely.

Partition Coefficients

Partition coefficients for DBM in saline, rat bile, corn oil, and mineral oil were determined by gas absorption as described by Sato and Nakajima (1979). Bile was obtained by cannulation of the bile duct and collection overnight. The partition coefficients of various dilutions of corn oil and mineral oil with bile were also measured to determine if the emulsification of the oils by bile altered the DBM partitioning.

Experimental

DBM was dosed at levels from 10 to 100 mg/kg. Stock dosing solutions were made in saline and distilled water. The DBM was administered in a total volume of 2 ml with a glass syringe and stainless steel gavage needle. DBM concentrations of the stock solutions were measured before each dose by gas chromatography after dilution in hexane. DBM was dosed at levels from 10 to 100 mg/kg.

After dosing, blood was withdrawn at specified time intervals for up to 6 hr from the jugular cannula and analyzed for DBM and carboxyhemoglobin. The DBM was extracted from a 20 µl blood sample with 500 µl of hexane in a glass vial. One µl samples were then

quantified by GC analysis (HP 5880, 10% SE 30 stainless steel column, EC detector). Carboxyhemoglobin levels were determined with 4 µl samples by the method of Rodkey et al. (1979). At the end of each run, 2 ml of blood were drawn from each animal and centrifuged; 1 ml of plasma was decanted and frozen. These plasma samples were later analyzed for bromide with a specific ion electrode; samples were prepared by precipitating soluble protein with 3 ml of 10% TCA and ionic strength adjustment with sodium nitrate. Total body burden of bromide was calculated assuming a volume of distribution of 0.34 liters/kg.

IV. PHYSIOLOGICAL MODEL

The physiological model used to model the experimental data is described in detail by Ramsey and Andersen (1983). The DBM is absorbed from the gut lumen assuming first-order kinetics and distributed to body compartments of fat, muscle, rapidly-perfused tissue, and liver. Metabolism occurs in the liver or other metabolizing tissues by Michaelis-Menten kinetics. DBM entering each tissue by arterial blood is assumed to be rapidly dispersed throughout the tissue volume; venous blood leaving each tissue is assumed to be in equilibrium with the tissue concentration based upon the blood to tissue partition coefficient. DBM in alveolar air is assumed to equilibrate with the arterial blood and the alveolar air leaving the lung carries an amount of DBM based upon the blood to air partition coefficient. The coefficients and their values used in the models are listed in Table 1.

V. RESULTS

Partition Coefficients

The partition coefficients for saline, rat bile, mineral oil, and corn oil were determined to be (mean ± 1 standard error) 14.1 ± 2.0,

Table 1

Parameter Used for the Physiological Model

Parameter	Value
Maximum utilization rate	20 mg/kg/hr
Half velocity coefficient	0.36 mg/l
Liver partition coefficient	0.919
Fat partition coefficient	10.6
Slowly perfused tissue partition coefficient	0.546
Rapidly perfused tissue partition coefficient	0.9
Blood partition coefficient	74.1
Liver weight	4% of body weight
Organ weight	5% of body weight
Fat weight	7% of body weight
Muscle weight	75% of body weight
Haldane coefficient for rat hemoglobin	150
Pulmonary ventilation	11 1/hr
Cardiac output	13.1 1/hr
Blood volume	59 ml
Lung diffusivity	0.05 1/hr mm Hg
Liver blood flow	25% of cardiac output
Fat blood flow	9% of cardiac output
Slowly perfused tissue blood flow	15% of cardiac output
Rapidly perfused tissue blood flow	9% of cardiac output
Organ blood flow	51% of cardiac output

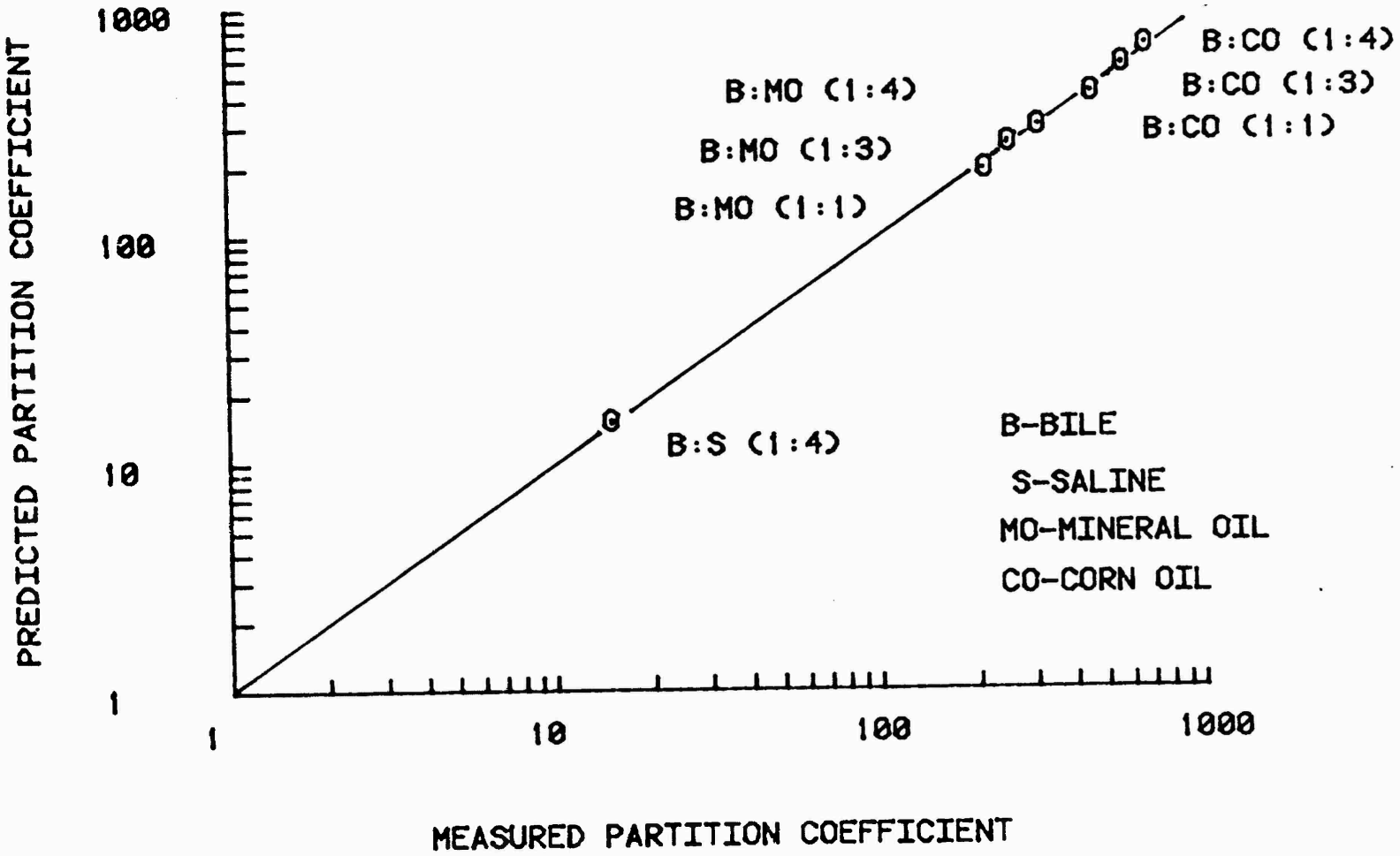
18.1 ± 4.5, 374 ± 4, and 837 ± 5, respectively. The saline value compared closely with previously determined values at AFAMRL of 14.4 ± 0.4.

The measured partition coefficients for the mixtures of saline, corn oil, and mineral oil with bile are shown in Figure 1. The predicted values are based upon linear addition of the partition coefficients. The close agreement between measured and predicted values indicates that emulsification of fats by bile does not alter the partition coefficients for DBM.

Dosing in Saline

DBM and HbCO Levels. The measured blood concentrations for DBM and carboxyhemoglobin (HbCO) for dosages of 10, 25, 50 and 100 mg/kg are shown in Figures 2 through 9. All curves show a rapid increase in DBM concentration with the maximum blood level achieved within 15 minutes. Peak DBM blood levels were a function of the dosage in mg/kg (6, 14, 40, and 60 mg/l for dosages of 10, 25, 50 and 100 mg/kg, respectively). The carboxyhemoglobin levels exhibited a characteristic bell-shaped curve, following the blood level with a significant lag period. Liver metabolism of DBM saturates at blood levels of about 2 mg/l (6 times Km of 0.34 mg/l), thus significant decreases in HbCO are not achieved until the DBM blood levels are below this value.

Bromide Body Burdens. The total body bromide measured at the end of each experiment varied from 43 to 160 per cent of the equivalent bromide in the administered DBM assuming a volume of distribution of 0.34 liters/kg (Table 2). Values less than 100 could be due to incomplete absorption or exhalation; values greater than 100 may be due to inaccuracies in dosage or decreased volumes of distribution due to bleeding and dehydration. Using the same techniques, Archer (1983) obtained a range of bromide levels from 58 to 141 per cent of bromide



01-101

Figure 1. Measured versus Predicted Partition Coefficients for DBM for Various Mixtures of Bile with Saline and Mineral and Corn oil

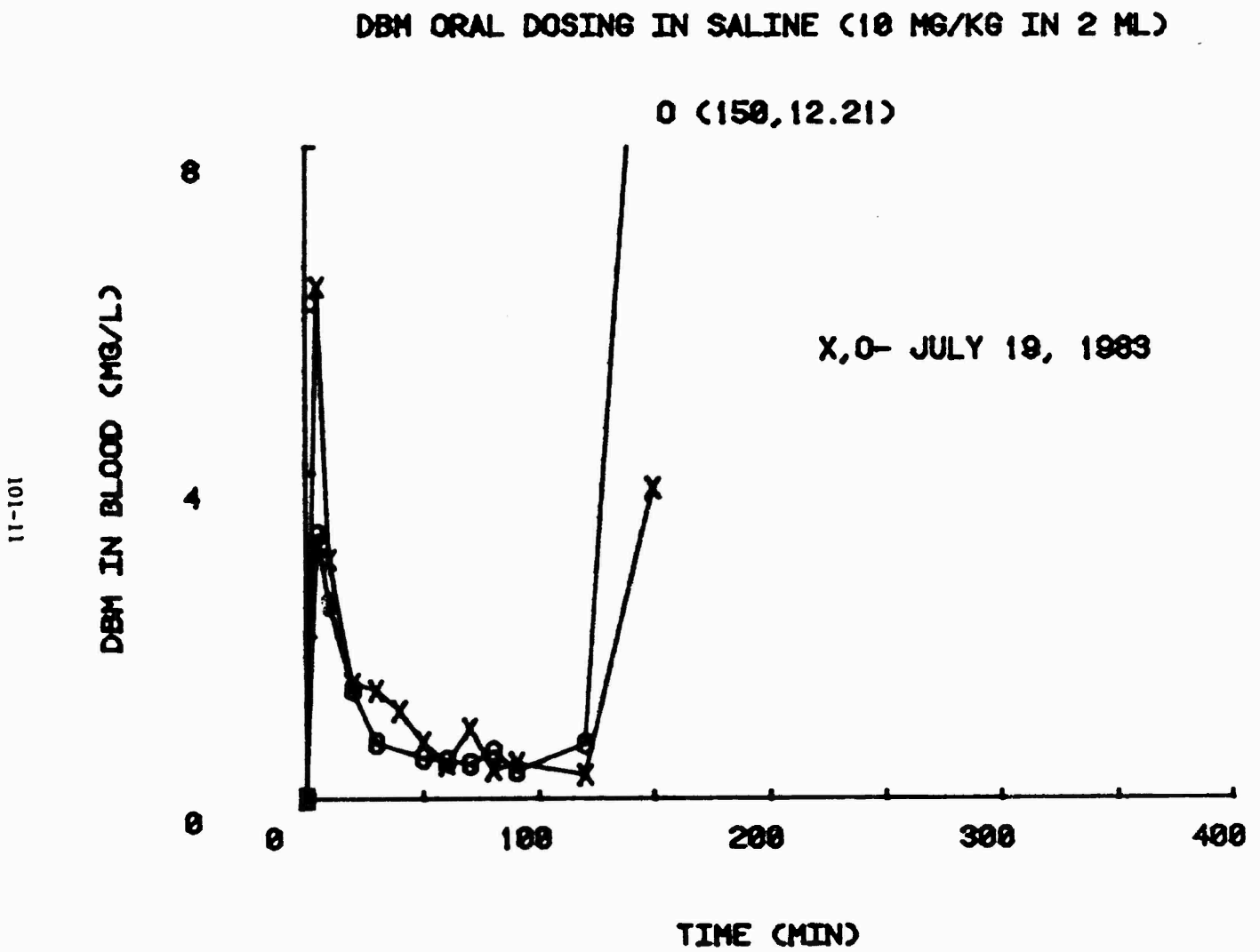


Figure 2. Dibromomethane Blood Level versus Time for 10 MG/KG Dose

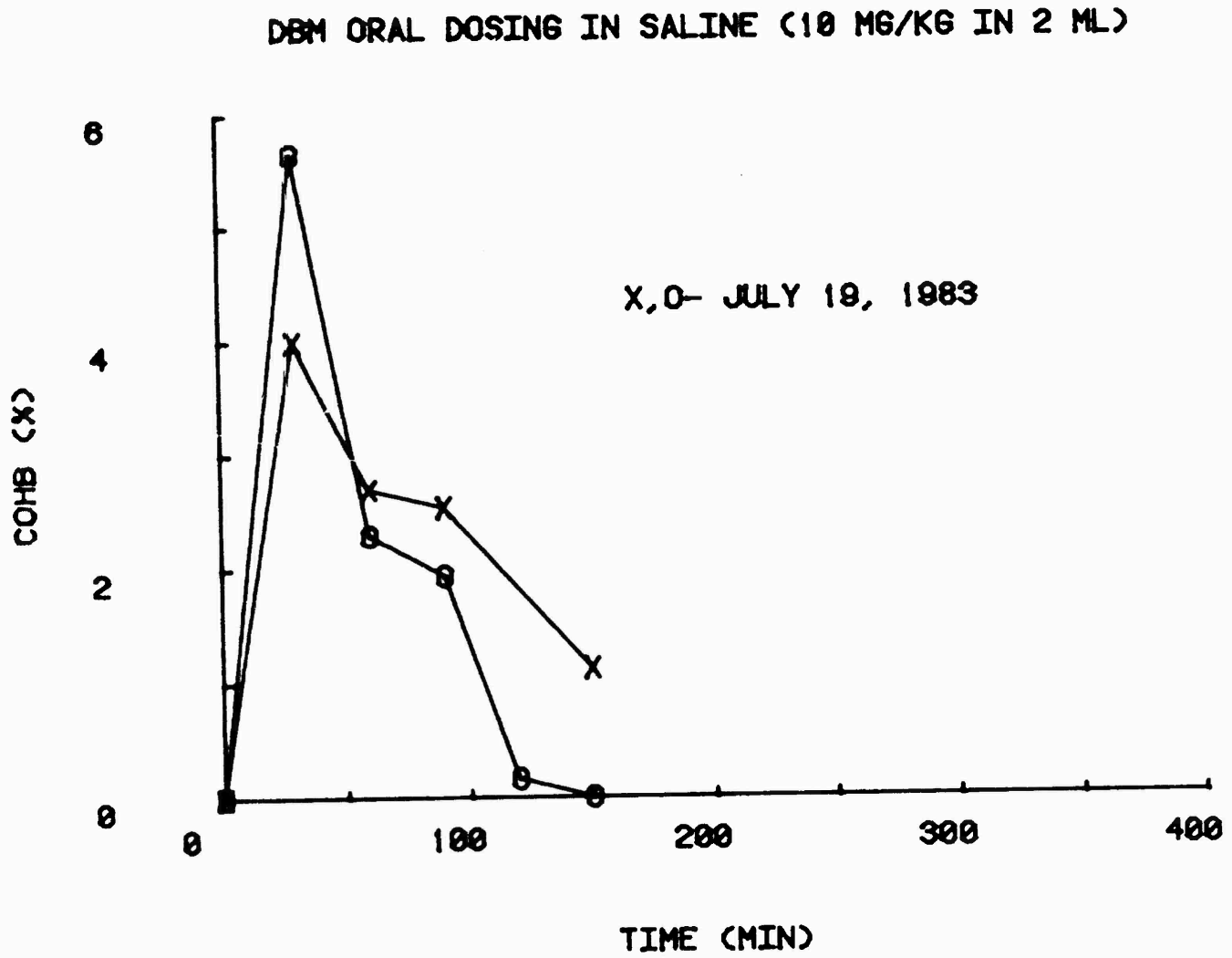


Figure 3. Carboxyhemoglobin versus Time for 10 MG/KG Dose

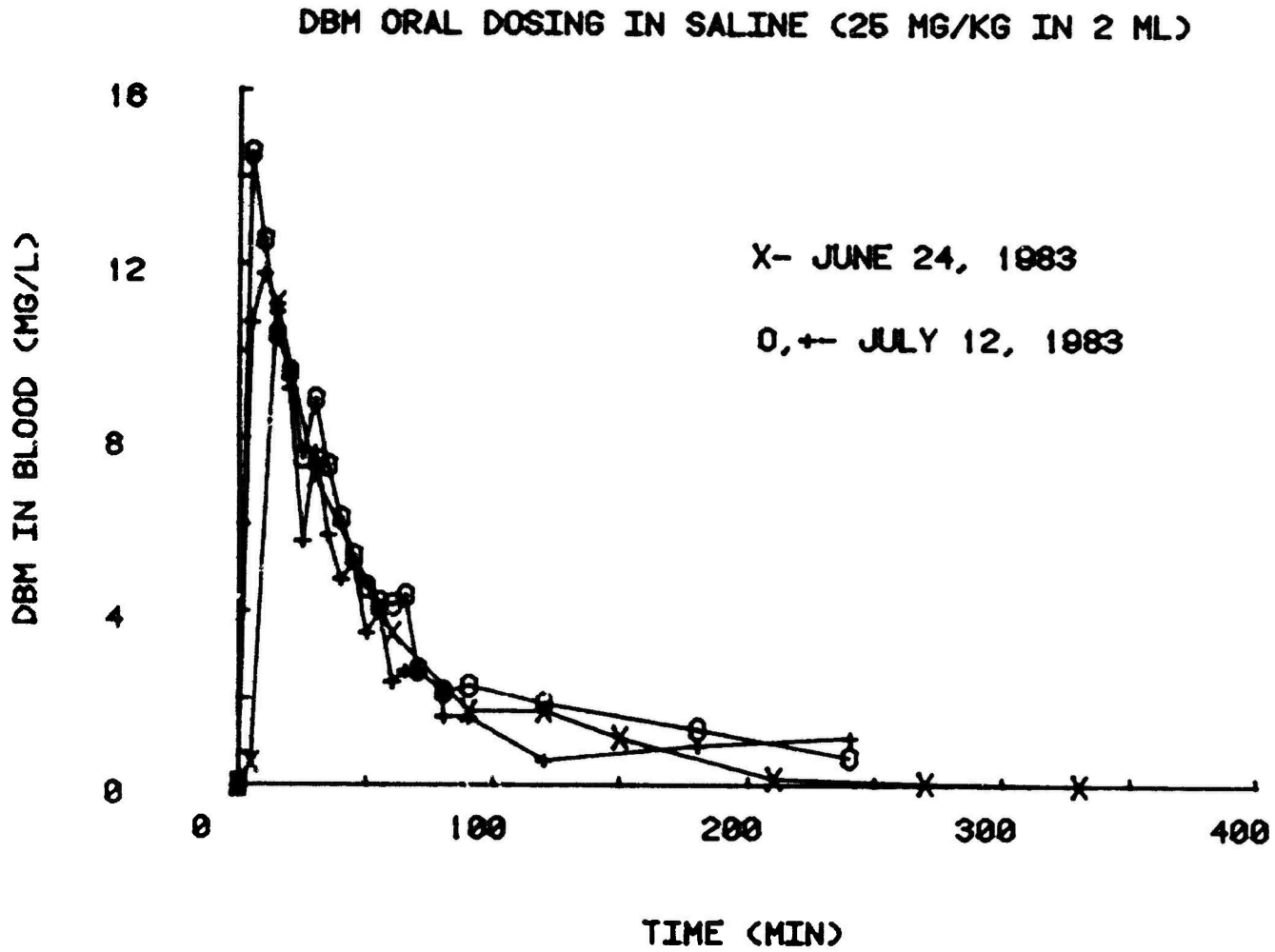


Figure 4. Dibromomethane Blood Level versus Time for 25 MG/KG Dose

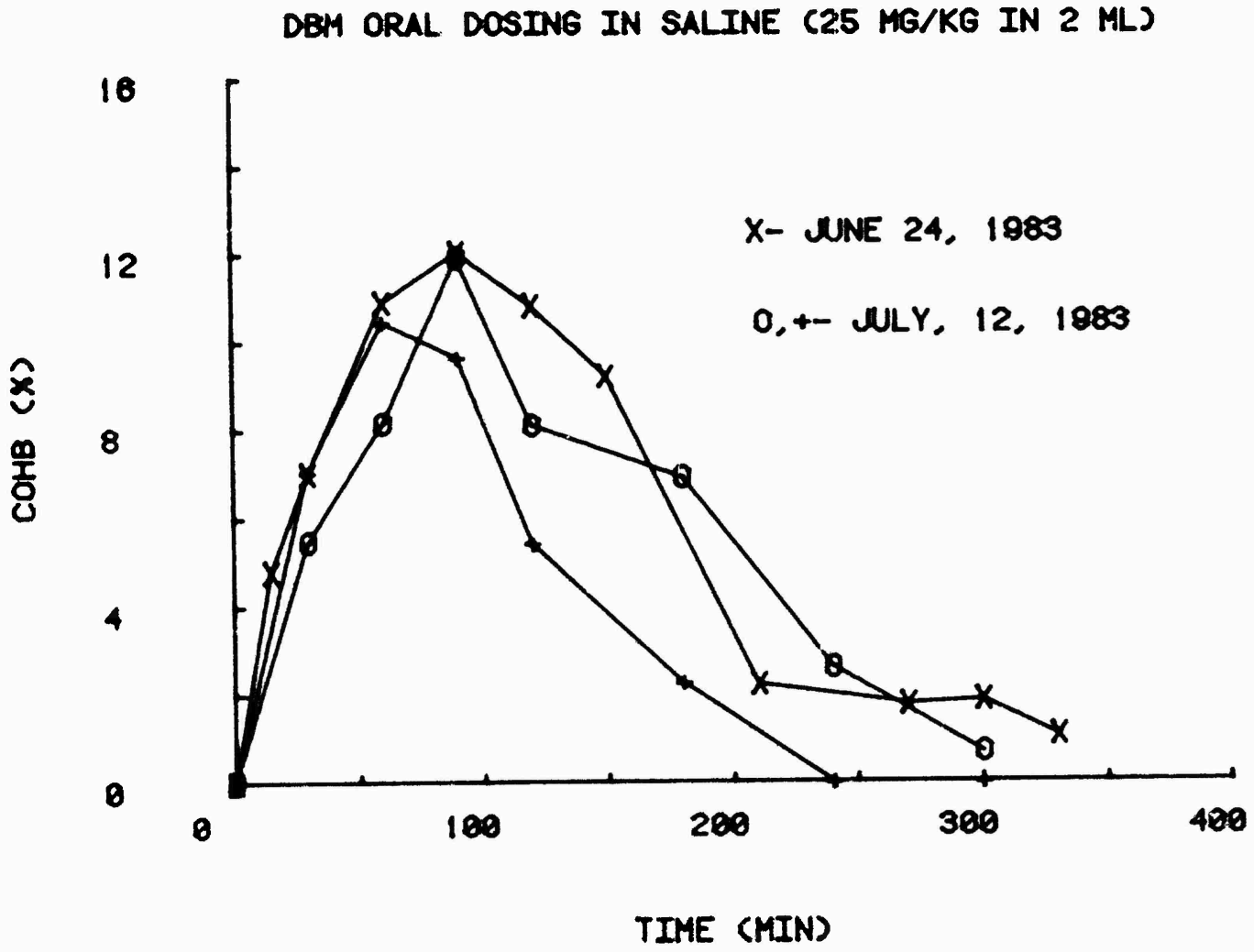


Figure 5. Carboxyhemoglobin versus Time for 25 MG/KG Dose

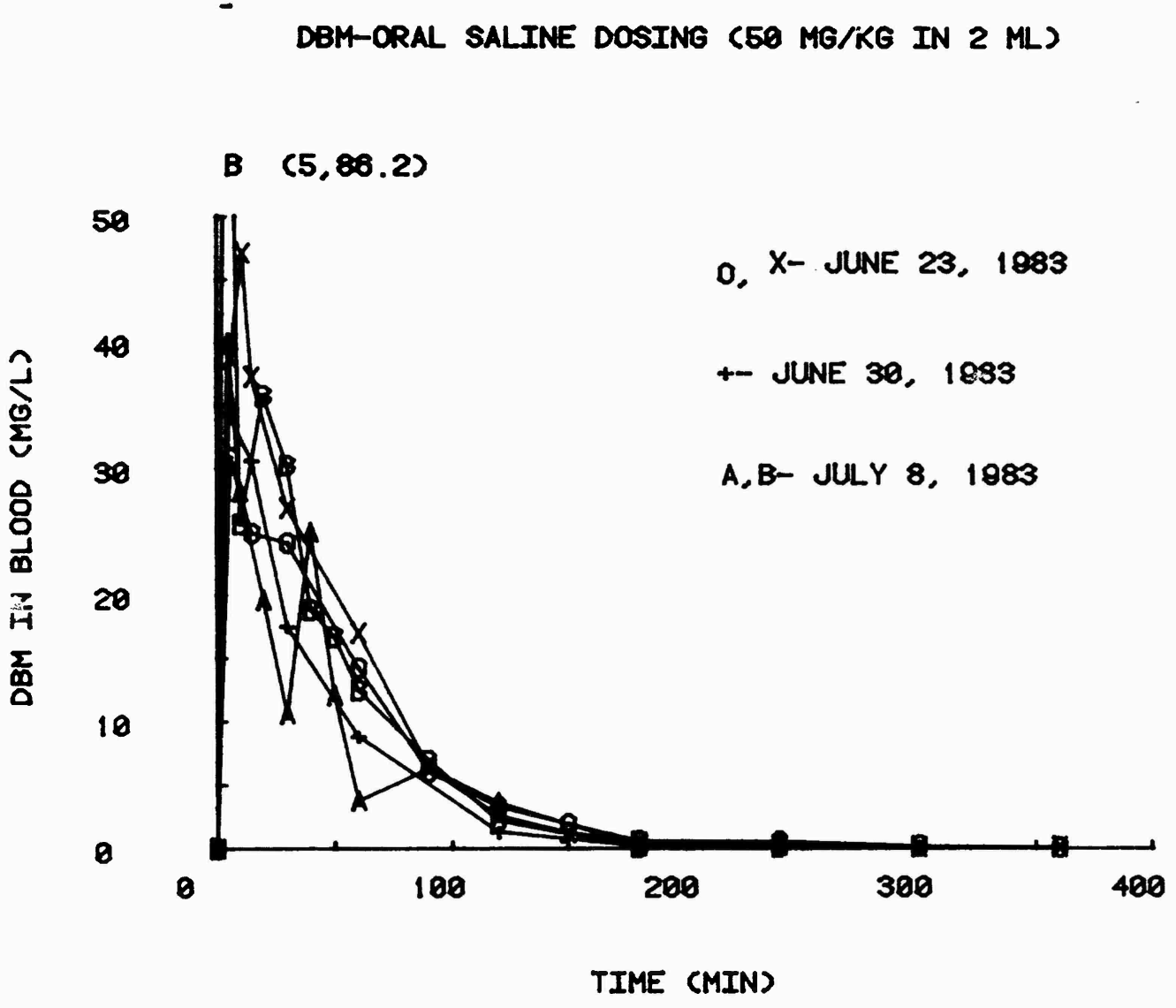


Figure 6. Dibromomethane Blood Level versus Time for 50 MG/KG Dose

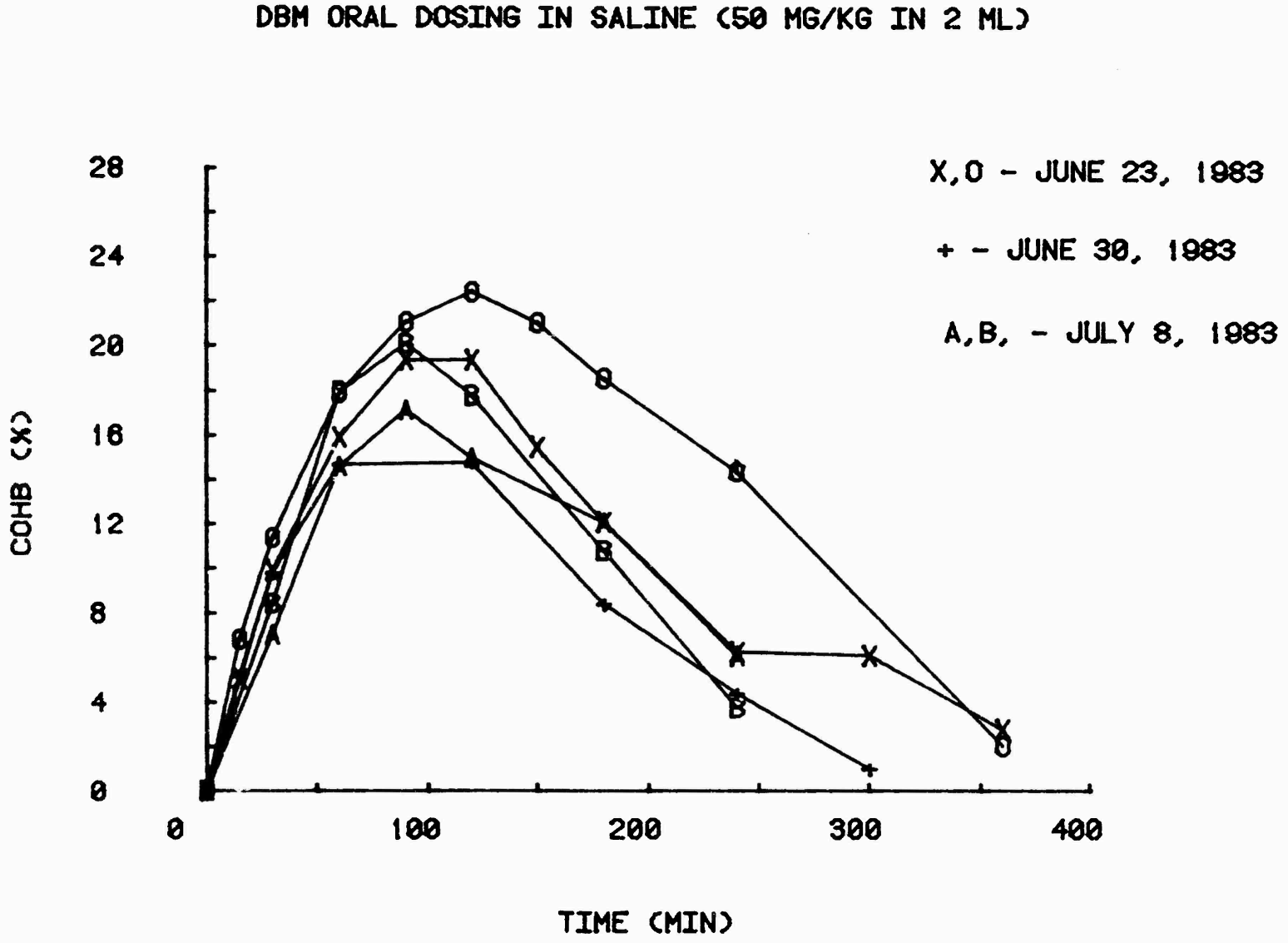


Figure 7. Carboxyhemoglobin versus Time for 50 MG/KG Dose

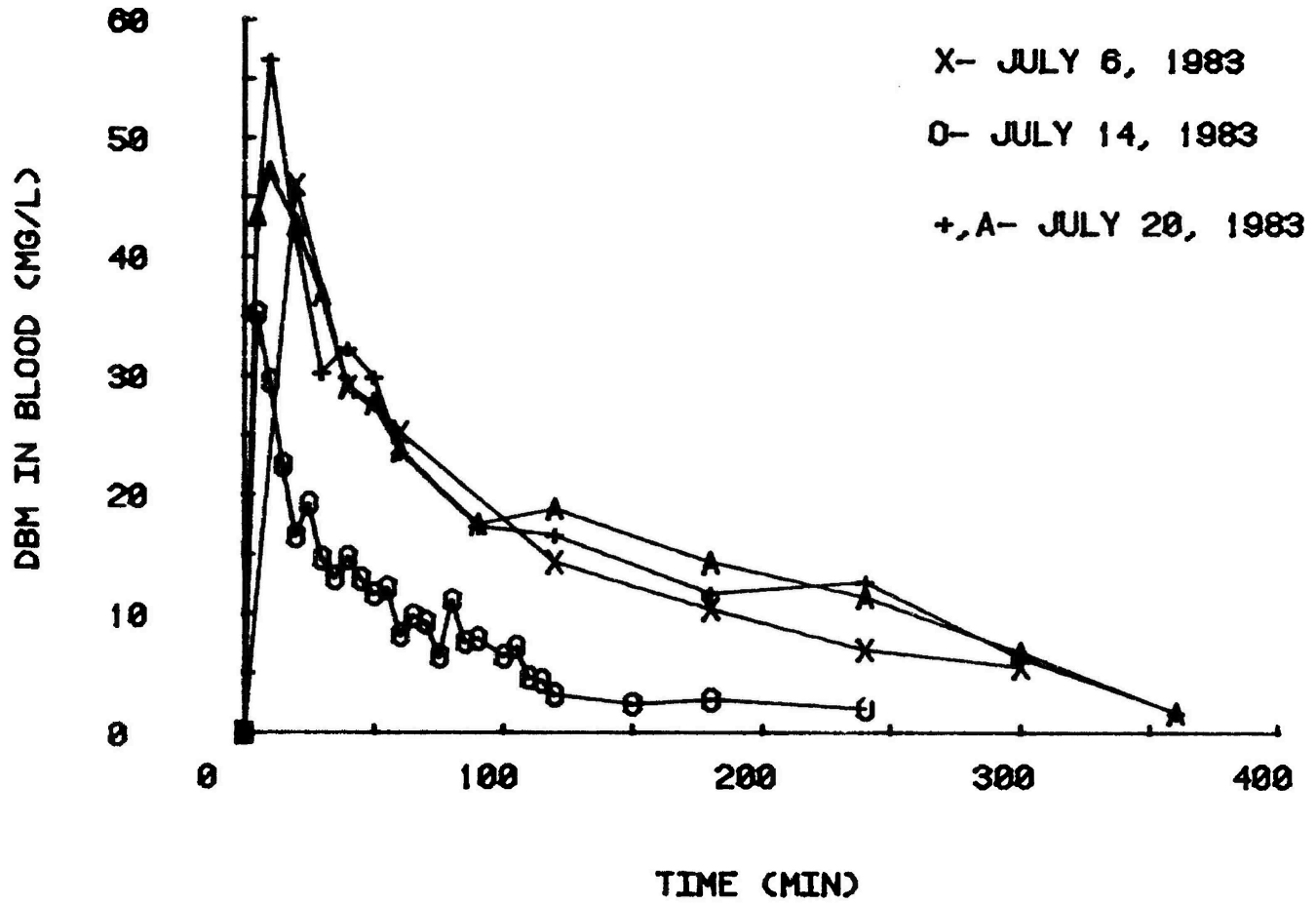


Figure 8. Dibromomethane Blood Level versus Time for 100 MG/KG Dose

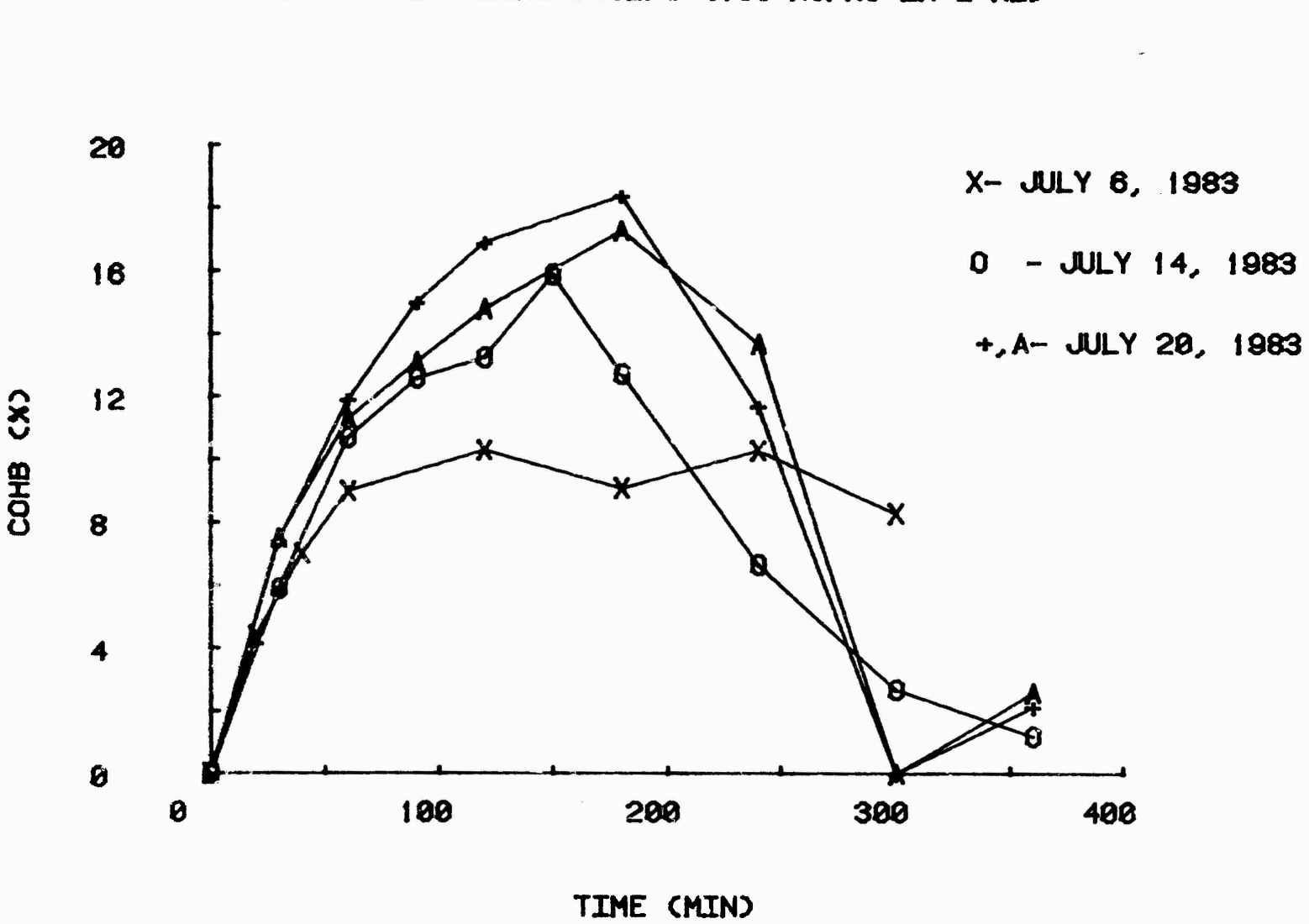


Figure 9. Carboxyhemoglobin versus Time for 100 MG/KG Dose

Table 2
Measured and Predicted Bromide Body Burdens

Date (see Figs. 2-9)	Dosage (mg/kg)	Measured Bromide (% of dose)	Predicted Bromide (% of dose)
July 19-x	10	103	99
-o	10	125	99
June 24-x	25	160	96
July 12-o	25	138	96
-+	25	119	96
June 23-+	50	-	90
-o	50	-	90
June 30-+	50	150	90
July 08-A	50	116	90
-B	50	116	90
July 06-x	100	43	80
July 14-o	100	66	80
July 20-+	100	80	80
-A	100	65	80

equivalent dosage for iv administration (dosages of 15 to 65 mg/kg). Exhalation can result in a loss of up to 20 per cent of the DBM at large blood levels as shown by the predicted bromide levels in Table 2. These values were obtained from simulations assuming first-order uptake from the gut lumen using the physiological model; these simulations predicted blood levels of DBM approximately equal to the measured values, and as such, should closely model the exhalation of DBM. In general, lower bromide recoveries were observed at higher dosages where greater exhalation and less complete metabolism would be expected. It does appear, however, that with negligible exhalation bromide levels in the blood are greater than 100 per cent of the dosed bromide, probably due to incomplete diffusion of the bromide throughout the volume of distribution.

Gastric Emptying. The release of fluids from the stomach is thought to significantly affect the absorption of chemicals from the gut lumen. This occurs due to the specialized physiology of the small intestine for absorption including high surface areas and extensive capillary networks that results in much more rapid uptake than occurs through the stomach wall. Clements et al. (1978) showed that gastric emptying in humans was an exponential function, but could be mono- or bi-phasic with significant intervals of no emptying. The emptying function significantly influenced the kinetics of acetaminophen absorption by delaying transport to the small intestine. Withey (1976) noted significant fluctuations in blood concentration versus time traces for po dosages to rats of vinyl chloride; he attributed such transients to gastric emptying.

The induction of transients in the DBM blood levels after dosage in saline was observed in some runs. Two of the most significant examples from the July 8 and July 12 experiments are plotted again in Figure 10; in both cases, large fluctuations are observed which could be attributed to periodic release from the stomach. The lack of such fluctuations in other data may be due to smaller and more frequent

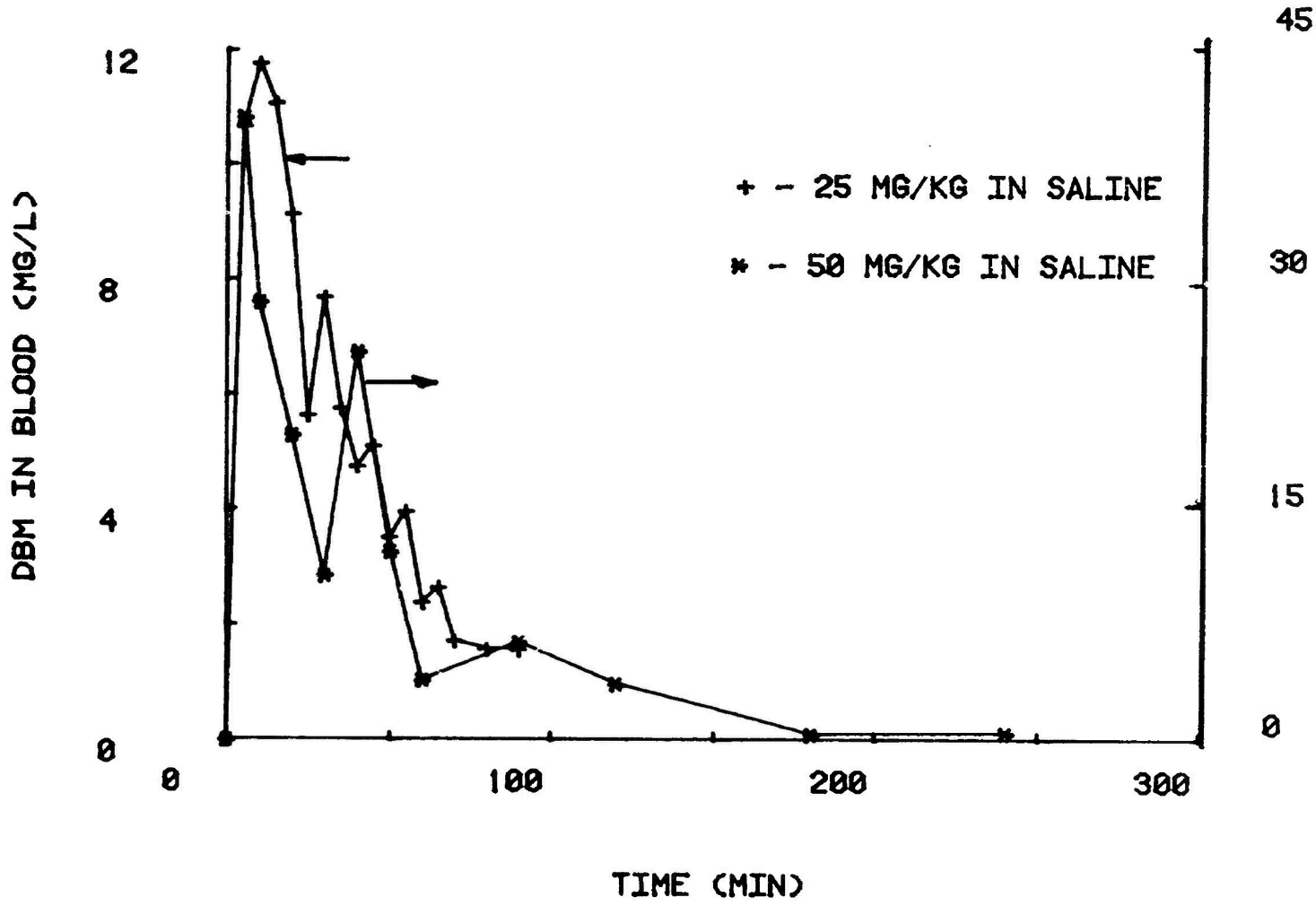


Figure 10. Examples of Transient DBM Concentrations in the Blood

emptying or the fluctuations may have been missed due to the large sampling intervals. Any physiologically based model of absorption from the gut should include gastric emptying; the appropriate mathematical formulations are described in detail by Clements et al., 1983.

In several instances sharp increases in DBM blood concentrations were measured 2 to 4 hr after dosage; specific examples are shown in Figure 2 where the DBM levels rose to 3.8 and 12.2 mg/l. Such large increases would not be expected from gastric emptying 2.5 hr after dosing. These results suggest that some mechanism exists to store DBM in a pool that then discharges into the blood; most likely these inputs are DBM in the lymph system that discharges from the intestines into the vena cava just below the point of sampling from the jugular cannula.

Further research probably should include experiments to quantify the effect of lymph transport on the distribution of the DBM. Specifically, it should be possible to cannulate the intestinal lymph duct and sample parent compound concentrations directly (Warshaw, 1972).

Physiological Modelling. The physiological model described previously was used to simulate the four dosages of DBM in saline (Figures 11 through 14). The absorption from the gut was assumed to be first-order which would most closely approximate a blood flow limited case. The first-order absorption coefficients that provided a reasonable fit to the experimental data were 10/hr for the 10 mg/kg dosage, 5/hr for the 25 and 50 mg/kg dosages, and 2/hr for the 100 mg/kg dosage. Examination of Figures 11 through 14 clearly shows that the model gives the best predictions of both the DBM and HbCO levels at the lower dosages. If the first-order absorption coefficients for the higher dosages are chosen to allow the initial peak of DBM in the blood to match the experimental data, then the DBM levels in the blood

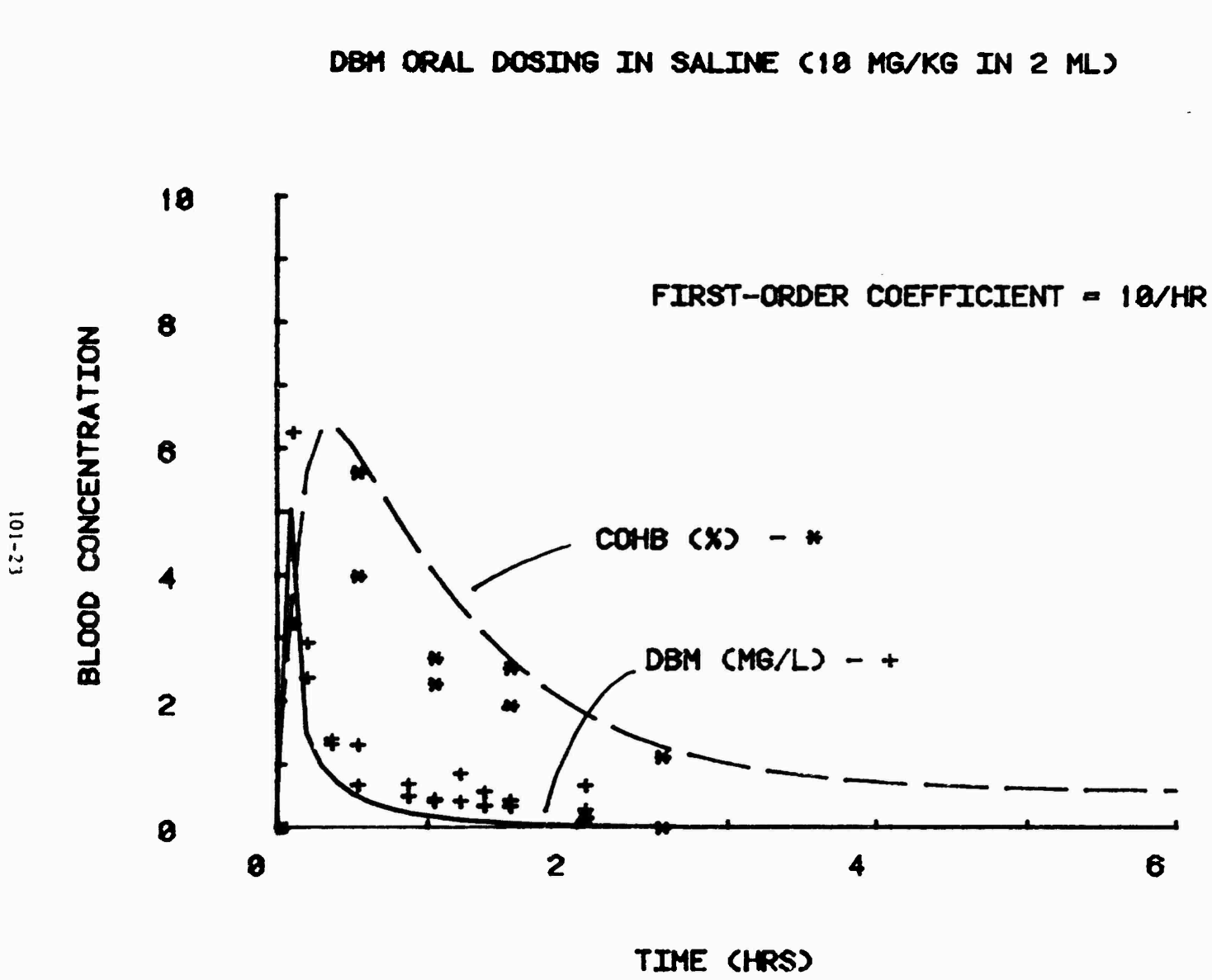


Figure 11. Model Prediction of Dibromomethane and Carboxyhemoglobin for 10 MG/KG Dose

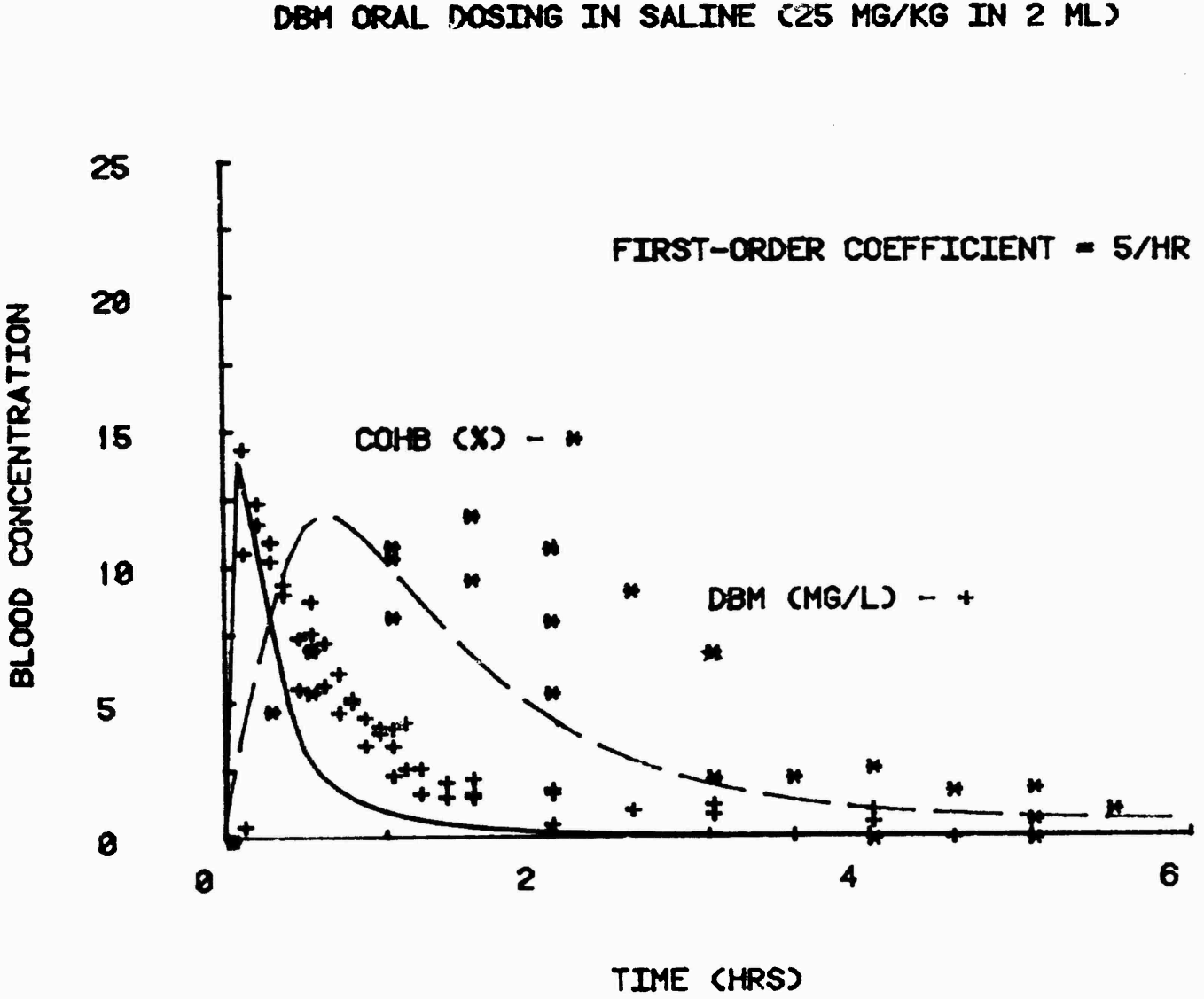


Figure 12. Model Prediction of Dibromomethane and Carboxyhemoglobin for 25 MG/KG Dose

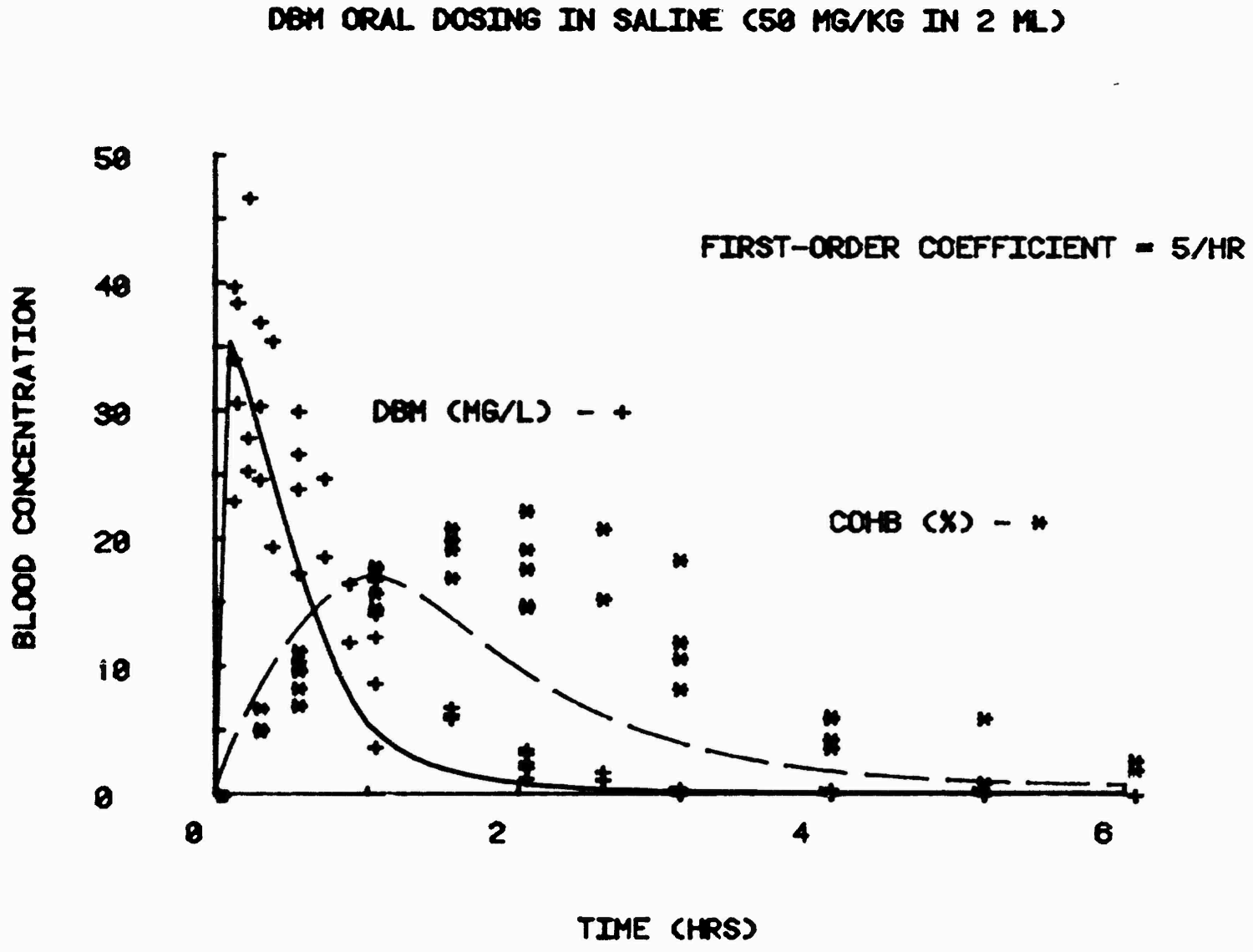


Figure 13. Model Prediction of Dibromomethane and Carboxyhemoglobin for 50 MG/KG Dose

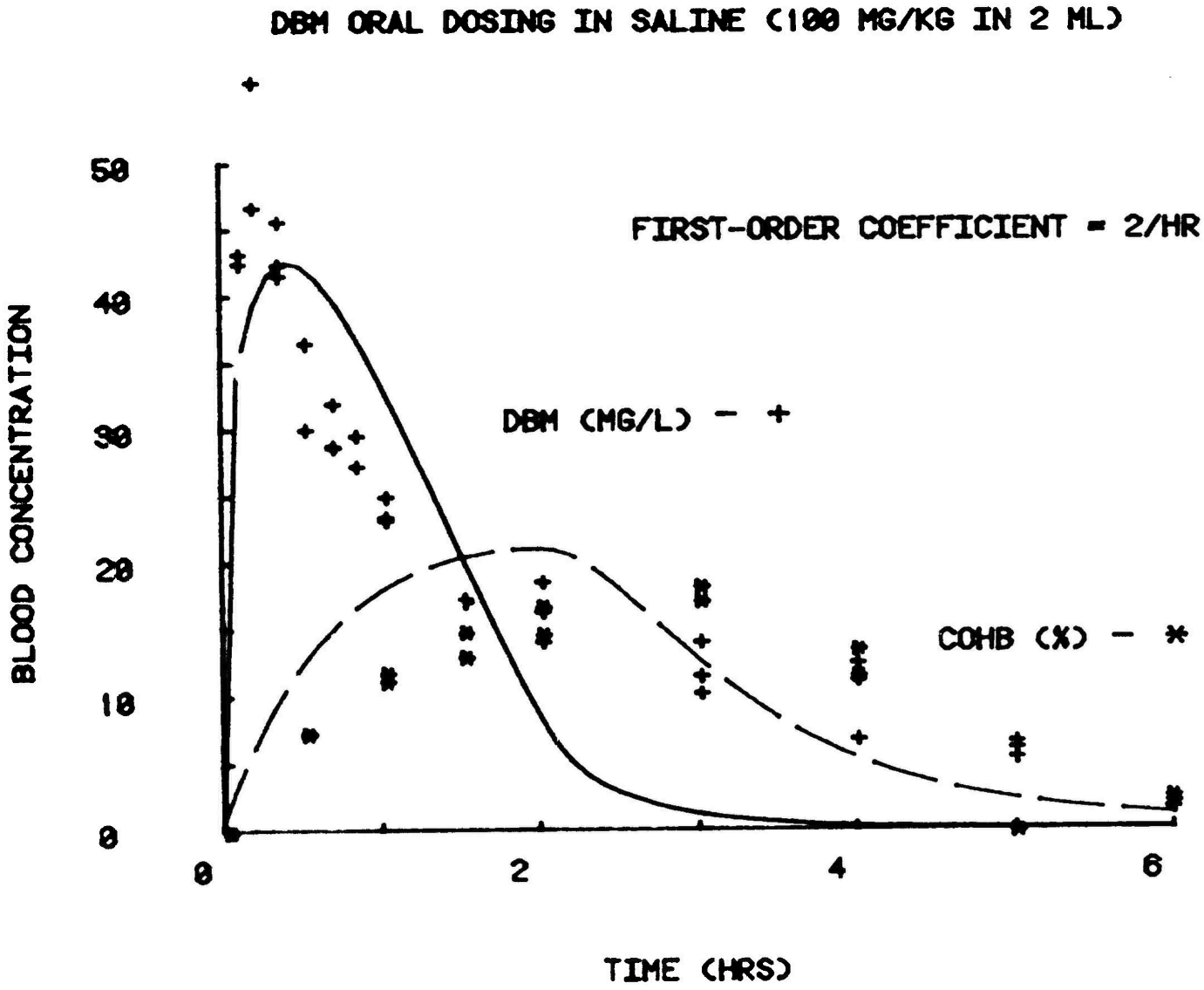


Figure 14. Model Prediction of Dibromomethane and Carboxyhemoglobin for 100 MG/KG Dose

drop too quickly, which in turn causes the HbCO levels to decrease too rapidly as compared to the data. If the physiological model was altered to include diffusional effects so that the transport of the DBM became perfusion limited and not exactly first-order (see Winne, 1978 for a review of this approach), the opposite effect would be expected in that transport would be retarded at lower dosage levels. This retardation effect may be due to transport in the intestinal lymph system; the incorporation of the lymph system in the physiological model has been initiated at AFAMRL and the data in Figures 2-9 will be used to verify the model.

Dosage into the Small Intestine

In an attempt to overcome the complexities of gastric emptying upon the resulting blood level concentrations, DBM in 2 ml of saline was dosed directly into the duodenum through an intestinal cannula. The DBM in blood versus time curve (Figure 15) was somewhat less than the corresponding curve for oral dosage (Figure 6); less total metabolism also occurred as indicated by the measured bromide levels of 88 and 76 per cent of the equivilant dosed bromide levels. The reasons for the smaller blood levels of DBM and the bromide body burdens are unknown. It was noted, however, that all the animals where the peritoneal cavity was opened showed high levels of post-operative stress; this may have resulted in less blood flow to the entire digestive system, and lower absorption.

Dosage after Intestinal Ligation

To determine the rate of DBM absorption directly from the stomach, experiments were conducted with oral dosing of DBM in 2 ml of saline after ligation of the intestine with suture. This procedure resulted in severe distress in the animals and corresponding lack of metabolism of the DBM; the DBM levels in the blood rose to about

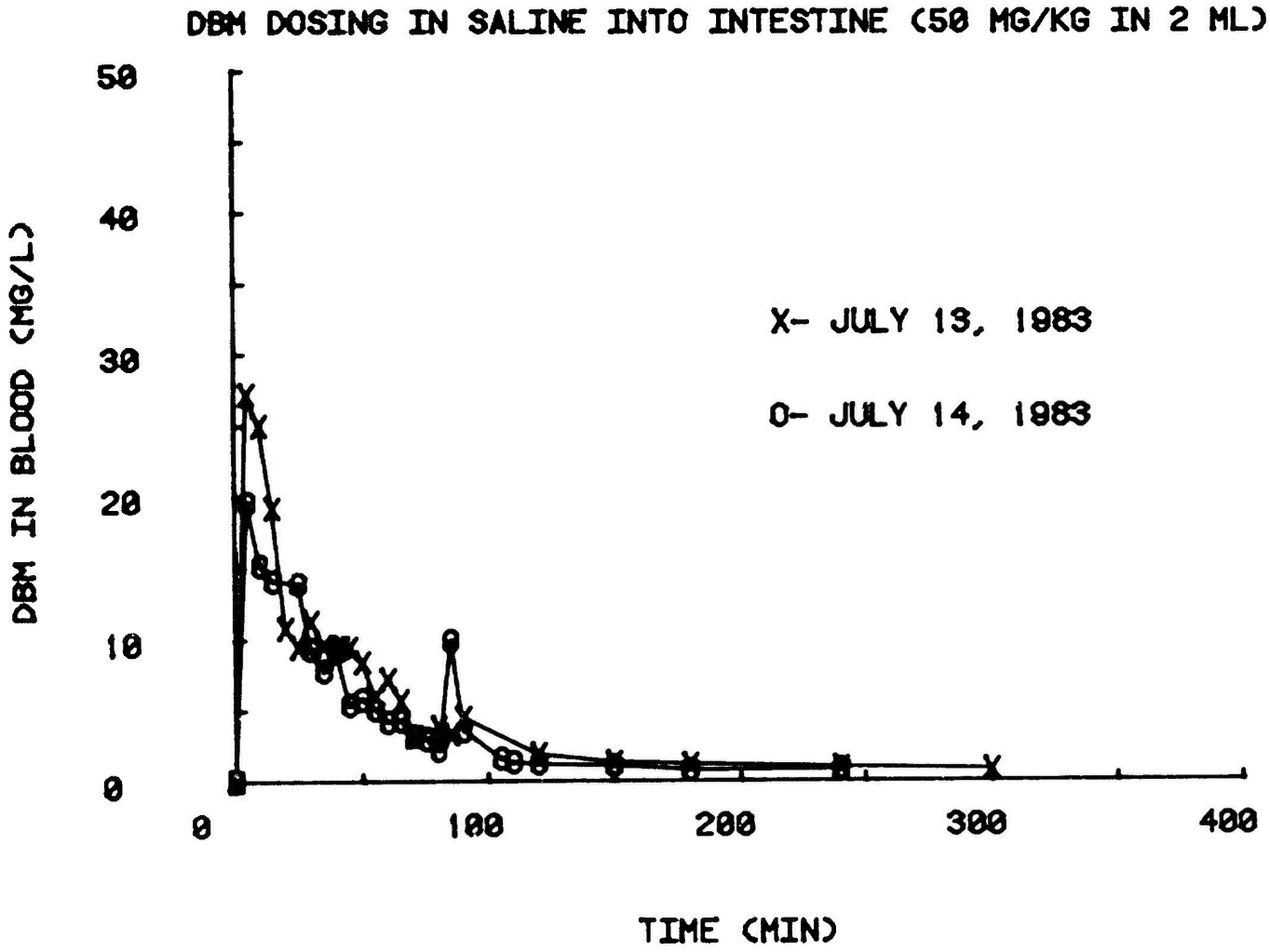


Figure 15. Dibromomethane Blood Level versus Time for 50 MG/KG Dose into Intestine

10 mg/l and remained constant as the HbCO levels decreased to 0 per cent (Figures 16 and 17). These experiments were conducted with animals from Lot CO-5 which was found to produce erratic results for the oil dosages. It is unknown whether the inconclusive results occurred from the surgical stress or from the lot of defective rats. These results did show that DBM can be rapidly absorbed from the stomach with levels reaching about 15 mg/l within a few minutes. However, the stress on the animals precludes any interpretation of the importance of the stomach in overall absorption from the gut.

Dosage in Distilled Water

The absorption of a water soluble chemical from the gut involves the two processes of passive diffusion and convection in the water absorption. Water is absorbed directly through pores in the intestine wall, and as such, could act as a carrier of small soluble chemicals. Distilled water is known to remain in the gut longer than saline as salts are absorbed until isotonic conditions are achieved. The resulting DBM blood concentration and HbCO levels versus time for dosage of 50 mg/kg are shown in Figures 18 and 19; comparison to Figures 6 and 7 shows lower DBM and HbCO levels than resulted from equivalent saline dosages. Lower metabolism was also shown in the bromide levels of 77 and 82 per cent (no bromide sample obtained from the July 15 run) of the equivalent bromide dose (Table 2). The reason for this lower bioavailability is unknown; however, distilled water is known to stimulate intestinal lymph flows which may contribute to storage of a portion of the DBM in the lymph system.

VI. CONCLUSION

Based upon this research, it is concluded that:

1. The vehicle:air partition coefficients of saline, rat bile,

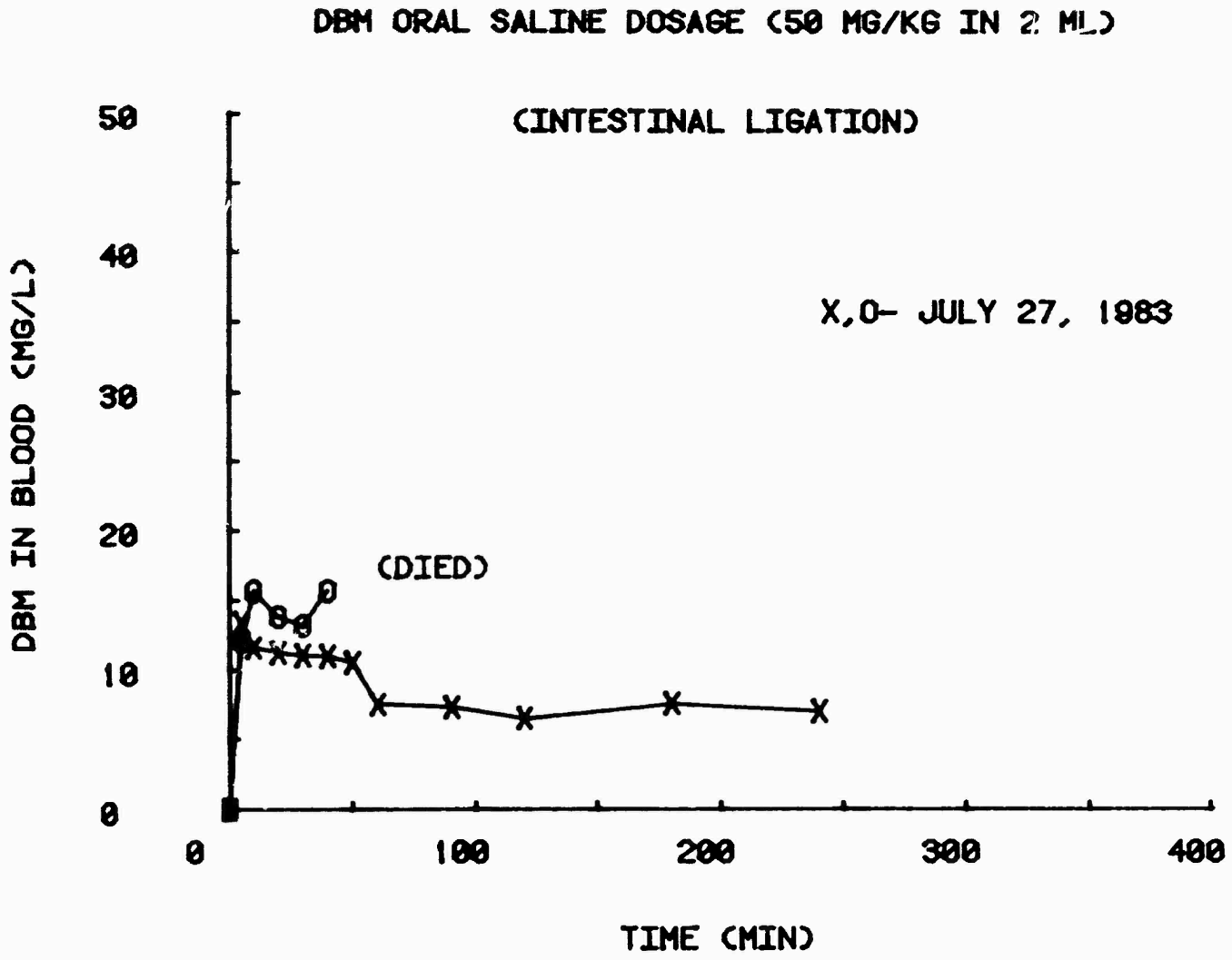


Figure 16. Dibromomethane Blood Level versus Time after Intestinal Ligation

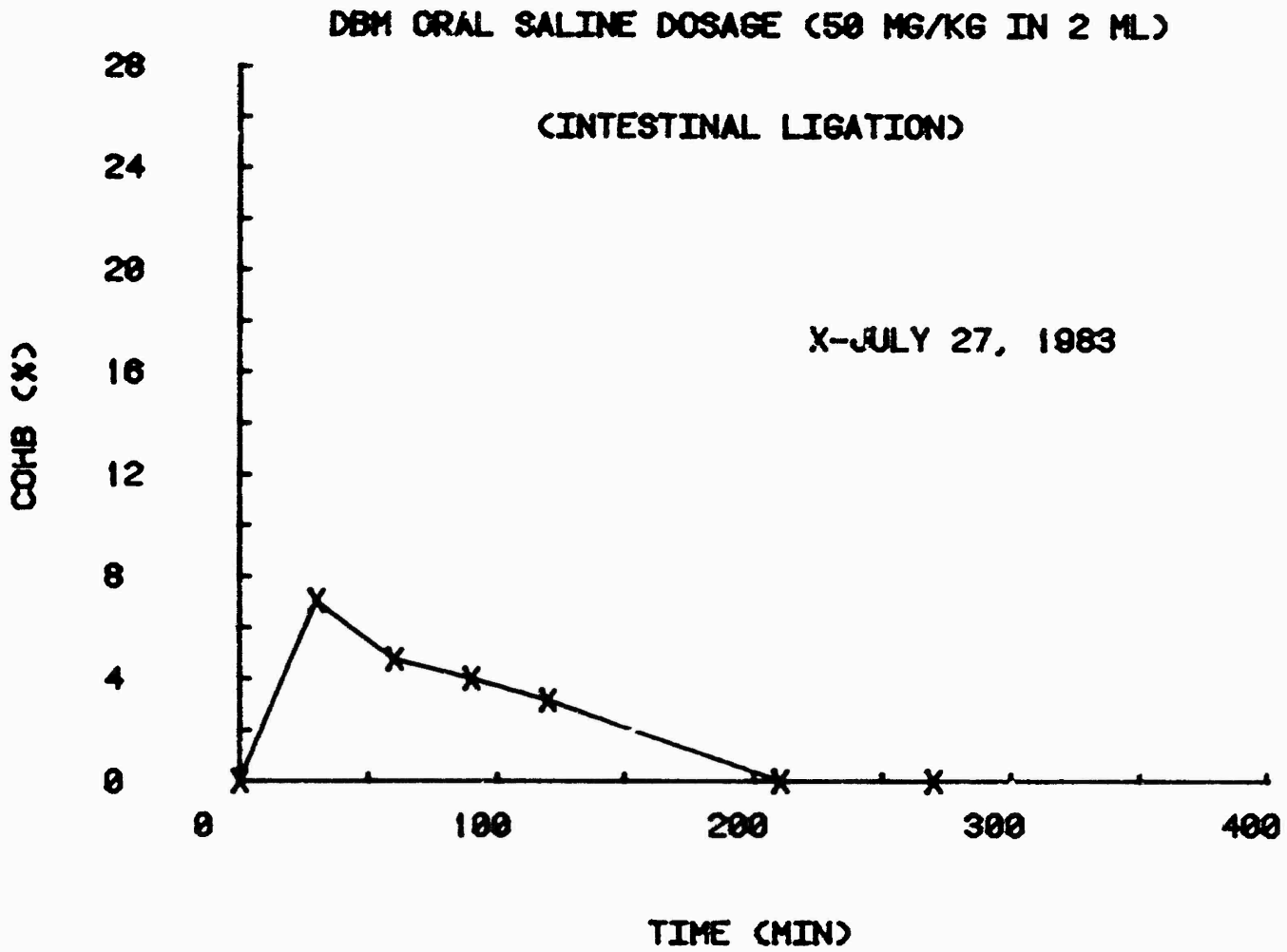


Figure 17. Carboxyhemoglobin Level versus Time after Intestinal Ligation

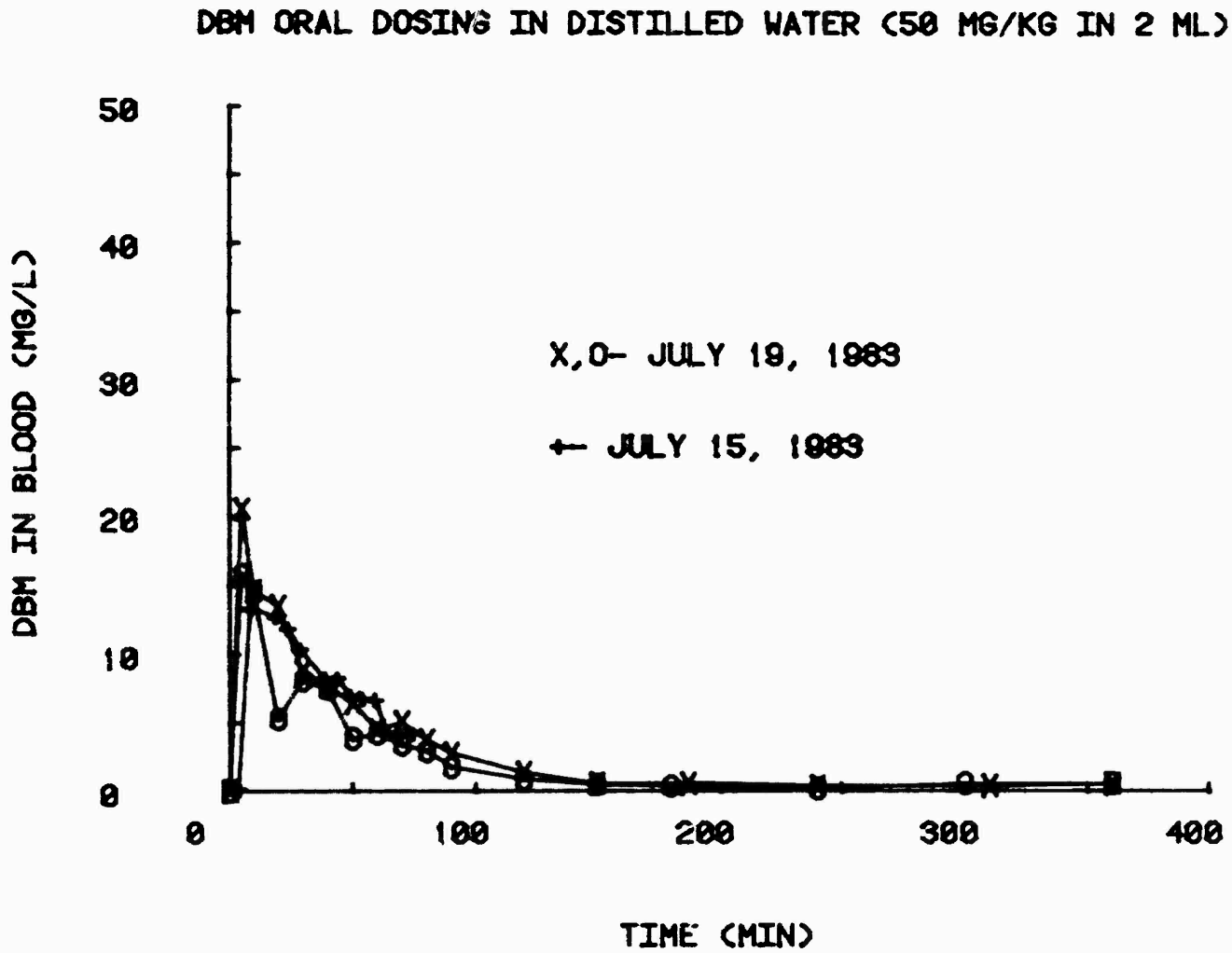


Figure 18. Dibromomethane Blood Level versus Time after Distilled Water Dosage

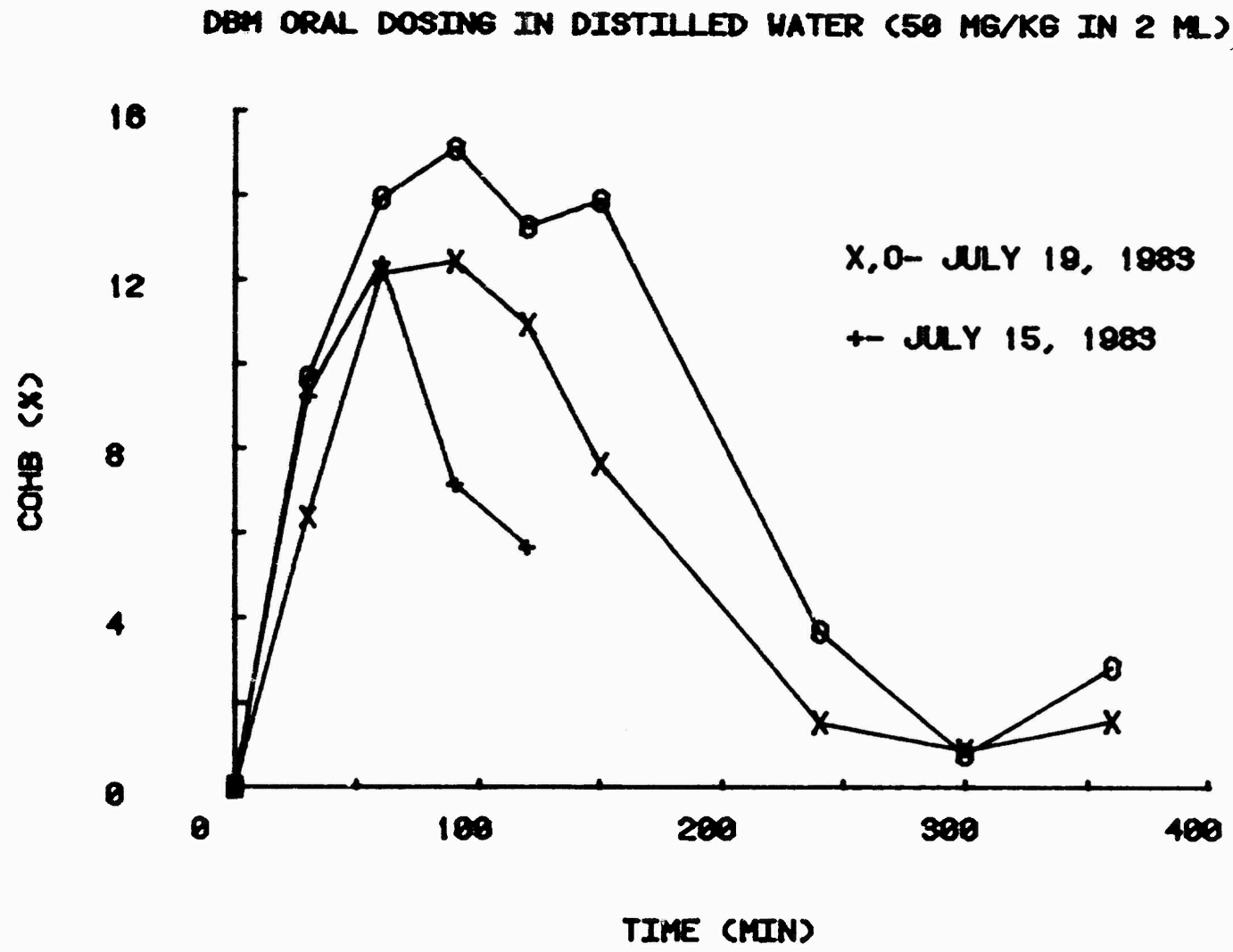


Figure 19. Carboxyhemoglobin Levels versus Time after Distilled Water Dosage

corn oil, and mineral oil for DBM are 14, 18, 374, and 837 g/g, respectively.

2. Emulsification by bile does not alter the partition coefficients of DBM to corn or mineral oil.

3. DMB is rapidly absorbed from the gut lumen after dosing in saline with no significant lag periods.

4. Measured bromide levels exceed by about 25 per cent the predicted levels based upon the bromide equivalents in DBM.

5. The physiological model assuming first-order absorption from the gut is inadequate to describe the DBM blood and carboxyhemoglobin concentration versus time for oral dosage in saline at all four doses.

VII. RECOMMENDATIONS

1. Further research is needed to determine the time course of gastric emptying in rats and the influence of this phenomenon upon the absorption from the gut lumen. Present physiological models have an indequate data base to include this phenomenon to describe chemical transport.

2. Physiological models need to be expanded to include chemical transport through the lymph system. Such alteration needs to include intestinal lymph flow rates, partition coefficients to lymph fluids, and delays between dosing and transport of the lymph fluid to the blood.

3. Other dosing vehicles need to be examined to detemine the effect on the absorption of DBM from the gut lumen of partition coefficients, degradability, and lymphatic transport.

REFERENCES

Anders, M.W., Kubic, V.L., and Ahmed, A.E. (1977). Metabolism of halogenated methanes and macromolecular binding. J. Environ. Pathol. Toxicol. 1, 117-124.

Archer, R. (1983). The pharmacokinetics of dibromomethane and metabolites. Final report submitted to Southeastern Center for Electrical Engineering Education under 1983 USAF-SCEEE Graduate Student Summer Support Program.

Clements, J.A., Heading, R.C., Nimmo, W.S., and Prescott, L.F. (1978). Kinetics of acetaminophen absorption and gastric emptying in man. Clin. Pharmacol. Ther. 24(4), 420-431.

Ramsey, J.C. and Andersen, M.E. (1983). A physiologically based description of the inhalation pharmacokinetics of styrene in rats and humans. Toxicol. Appl. Pharmacol., accepted for publication.

Sato, A. and Nakajima, T. (1979). Partition coefficients of some aromatic hydrocarbons and ketones in water, blood and oil. Brit. J. Indust. Med. 36, 231-234.

Warshaw, A.L. (1972). Technique. A simplified method of cannulating the intestinal lymphatic of the rat. Gut 13, 66-67.

Winne, D. (1978). Blood flow in intestinal absorption models. Pharmacokinetics and Biopharmaceutics 6(1), 55-78.

Withey, J.R. (1976). Pharmacodynamics and uptake of vinyl chloride monomer administered by various routes to rats. J. Toxicol. Environ. Health 1, 381-394.

Withey, J.R. and Collins, P.G. (1977). Pharmacokinetics and distribution of styrene monomer in rats after intravenous administration. J. Toxicol. Environ. Health 3, 1011-1020.