

AD A139403

UNITED STATES AIR FORCE

SUMMER FACULTY RESEARCH PROGRAM

1983

PROGRAM TECHNICAL REPORT

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

Volume I of II

Program Directors, 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

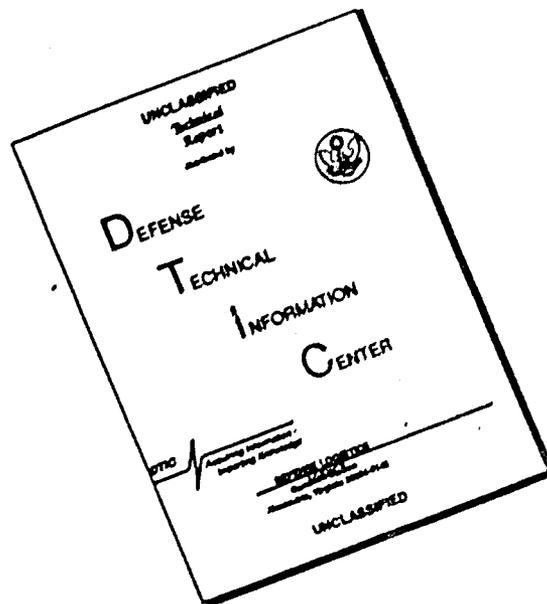
December 1983



Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By _____	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

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

DISCLAIMER NOTICE



THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 84-0154	2. GOVT ACCESSION NO. <i>AD A139403</i>	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) UNITED STATES AIR FORCE SUMMER FACULTY RESEARCH PROGRAM(1983), TECHNICAL REPORT, VOL 1 of 2		5. TYPE OF REPORT & PERIOD COVERED FINAL
7. AUTHOR(s) Warren D. Peele Earl L. Steele Maj. Amos L. Otis		6. PERFORMING ORG. REPORT NUMBER
9. PERFORMING ORGANIZATION NAME AND ADDRESS Southeastern Center for Electrical Engineering Education 11th & Massachusetts Ave; St. Cloud, FL 32769		8. CONTRACT OR GRANT NUMBER(s) F49620-82-C-0035
11. CONTROLLING OFFICE NAME AND ADDRESS AFOSR/XOT Building 410 Bolling AFB, DC 20332		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2301/D5
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		12. REPORT DATE December 1983
		13. NUMBER OF PAGES
		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

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

↓
#20.) ABSTRACT CONT'D:

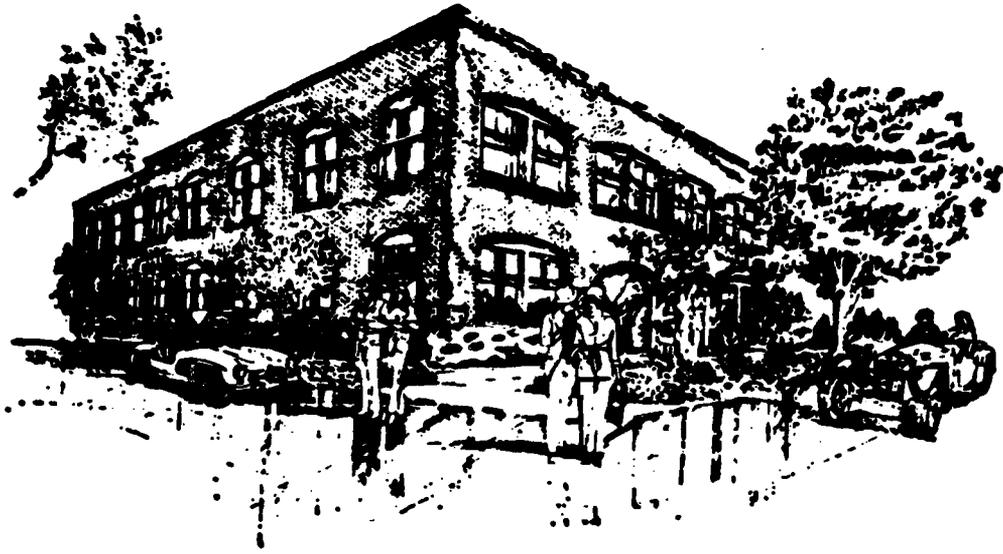
at an Air Force facility mutually agreed upon by the faculty members and the Air Force. In addition to compensation, travel and cost of living allowances are also paid. The USAF-SFRP is sponsored by the Air Force Office of Scientific Research, Air Force Systems Command, United States Air Force, and is conducted by the Southeastern Center for Electrical Engineering.

↑

UNCLASSIFIED

FOSR-TR- 84 - 0154

Approved for public release;
distribution unlimited.



SCEEE
©
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

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

List of Participants (continued: page 2)

Dr. Stanley Bashkin Professor Univeristy of Arizona Physics Department Tucson, AZ 85711 (602) 626-2322	<u>Degree:</u> Ph.D., Physics, 1950 <u>Specialty:</u> Accelerator-based Atomic Physics, Atomic Spectro- scopy <u>Assigned:</u> GL
Dr. Joseph Baumgarten Professor Iowa State University Mechanical Engineering Department Ames, Iowa 50010 (515) 294-1380	<u>Degree:</u> Ph.D., Mechanical Engineering, 1958 <u>Specialty:</u> Vibrations, Machinery Dynamics <u>Assigned:</u> AEDC
Dr. Michael Becker Associate Professor University of Texas Electrical Engineering Department Austin, TX 78712 (512) 471-3628	<u>Degree:</u> Ph.D., Electrical Engineering, 1973 <u>Specialty:</u> Lasers, Laser Material Interactions, Nonlinear Optics <u>Assigned:</u> WL
Dr. Henry Bertoni Professor Polytechnic Institute of NY Electrophysics Department Brooklyn, NY 11201 (212) 643-4832	<u>Degree:</u> Ph.D., Electrophysics, 1967 <u>Specialty:</u> Ultrasonic and Electro- magnetic Propagation and Diffraction <u>Assigned:</u> RADC
Dr. Willie Bragg Program Manager University of Cincinnati College of Education Cincinnati, OH 45221 (513) 475-4542	<u>Degree:</u> Ph.D., Special Education, 1979 <u>Specialty:</u> Mental Retardation, Early Childhood Education, Speech and Hearing <u>Assigned:</u> SAM
Dr. M. Quinn Brewster Assistant Professor University of Utah Mechanical Engineering Department Salt Lake City, Utah 84106 (801) 581-7105	<u>Degree:</u> Ph.D., Mechanical Engineering, 1981 <u>Specialty:</u> Thermal Radiation in Particulate Media <u>Assigned:</u> RPL
Mr Louis W. Buckalew Associate Professor Alabama A&M University Psychology Department Normal, AL 35762 (205) 322-2489	<u>Degree:</u> MS, Psychology, 1969 <u>Specialty:</u> Drug Response Factors, Human Engineering, Alcohol <u>Assigned:</u> SAM

List of Participants (continued: page 3)

Dr. Chester Canada Associate Professor Oklahoma State University Computer Science Department Stillwater, OK 74078 (409) 624-5724	<u>Degree:</u> Ph.D., Physics, 1968 <u>Specialty:</u> Materials Response to Dynamic Loading <u>Assigned:</u> ESC
Dr. Jack Chatelain Professor Utah State University Physics Department Logan, UT 84321 (801) 750-2936	<u>Degree:</u> Ph.D., Physics, 1957 <u>Specialty:</u> Theoretical Physics <u>Assigned:</u> WL
Mr. John Cicero Instructor Illinois Institute of Technology Electrical Engineering Department Chicago, IL 60532 (312) 567-3400	<u>Degree:</u> MSEE, Electrical Engineering, 1978 <u>Specialty:</u> Microprocessor Hardware and Software Design. Multiple Access Satellite Network Protocols <u>Assigned:</u> HRL/W
Dr. Frank Collins Professor University of Tennessee Aerospace Engineering Department Tullahoma, TN 37388 (615) 455-0631	<u>Degree:</u> Ph.D., Mechanical Engineering 1968 <u>Specialty:</u> Physical Fluid Dynamics <u>Assigned:</u> AEDC
Dr. Richard Conte Assistant Professor Manhattan College Mechanical Engineering Department Bronx, NY 10471 (212) 920-0148	<u>Degree:</u> Ph.D., Mechanical Engineering, <u>Specialty:</u> Heat Transfer, Numerical Computer Analysis, Solar Energy <u>Assigned:</u> APL
Dr. Billy Covington Assistant Professor Sam Houston State University Physics Department Huntsville, TX 77340 (713) 294-1606	<u>Degree:</u> Ph.D., Physics, 1978 <u>Specialty:</u> Solid State Physics <u>Assigned:</u> ML
Dr. Peter Crane Assistant Professor University of Pittsburgh Psychology Department Johnstown, PA 15905 (814) 266-9661	<u>Degree:</u> Ph.D., Experimental Psychology, 1979 <u>Specialty:</u> Cognitive Psychology/ Human Factors <u>Assigned:</u> 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, DMS,
Information Systems
Assigned: AD

Dr. Carol Deaknye
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 Psycho-
metrics
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: Radiation 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	<u>Degree:</u> Ph.D., Electrical Engineering, 1970 <u>Specialty:</u> Computer Interface Design for Data Acquisition and Digital Control <u>Assigned:</u> AL
Dr. Richard Gill Assistant Professor Wright State University Engineering Department Dayton, OH 45435 (513) 873-2701	<u>Degree:</u> Ph.D., Mechanical Engineering, 1981 <u>Specialty:</u> Human Factors Engineering <u>Assigned:</u> AMRL
Dr. John Giolma Assistant Professor Trinity University Science Department San Antonio, TX 78209 (512) 736-7563	<u>Degree:</u> Ph.D., Electrical Engineering, 1975 <u>Specialty:</u> Signal Processing, Modeling and Simulation <u>Assigned:</u> SAM
Dr. Samuel Green Associate Professor Auburn University Psychology Department Auburn, Alabama 36830 (205) 826-4412	<u>Degree:</u> Ph.D., Psychology, 1975 <u>Specialty:</u> Statistics, Measurement Theory, Industrial Psychology <u>Assigned:</u> LMDC
Dr. Arthur Gutman Associate Professor Florida Inst. of Technology Psychology Department Melbourne, Fla 32901 (305) 723-3701	<u>Degree:</u> Ph.D., Psychology, 1975 <u>Specialty:</u> Experimental and Bio- psychology; Computer Applications <u>Assigned:</u> HRL/L
Dr. Terry Herdman Associate Professor Virginia Tech Mathematics Department Blacksburg, VA 24060 (703) 961-5279	<u>Degree:</u> Ph.D., Mathematics, 1974 <u>Specialty:</u> Volterra Integral Equations Ordinary and Functional Differential Equations <u>Assigned:</u> FDL
Dr. Stuart Hirschfield Assistant Professor Hamilton College Mathematics Department Clinton, NY 13323 (315) 859-4136	<u>Degree:</u> Ph.D., Computer and Information Sciences, 1978 <u>Specialty:</u> Artificial Intelligence, Software Engineering <u>Assigned:</u> RADC

List of Participants (continued: page 7)

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 Astro-
nautics, 1983
Specialty: Orbit Mechanics, Space-
craft 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	<u>Degree:</u> Ph.D., Mathematical Sciences, 1980 <u>Specialty:</u> Mathematical Statistics <u>Assigned:</u> SAM
Dr George Kirby Assistant Professor West Virginia University Mech. & Aerospace Eng. Department Morgantown, West Virginia 26505 (304) 293-4111	<u>Degree:</u> Ph.D., Engineering Mechanics, 1982 <u>Specialty:</u> Experimental Mechanics, Fracture Mechanics, Vibrations <u>Assigned:</u> FDL
Mr. James Kirkpatrick Associate Professor Alabama A&M University Mathematics Department Normal, Alabama 35762 (205) 859-7239	<u>Degree:</u> MS, Physics, 1981 <u>Specialty:</u> Applied Mathematics <u>Assigned:</u> AEDC
Dr. Stephen Krause Assistant Professor Arizona State University Mech. & Aerospace Eng. Department Tempe, AZ 85287 (602) 965-2050	<u>Degree:</u> Ph.D., Engineering Materials, 1981 <u>Specialty:</u> Electron Microscopy, X-Ray Diffraction, & Properties of Polymers <u>Assigned:</u> ML
Dr. Madakasira Krishna Associate Professor South Carolina State University Computer Science Department Orangeburg, SC 29117 (803) 536-7120	<u>Degree:</u> Ph.D., Fluid Mechanics, Numerical Analysis, Computer Science <u>Specialty:</u> Computational Fluid Mechanics <u>Assigned:</u> FDL
Dr. William Kyros Associate Professor University of Lowell Mechanical Engineering Department Lowell, Mass 01854 (617) 459-9357	<u>Degree:</u> Ph.D., Education, 1980 <u>Specialty:</u> Mechanical Behavior of Materials <u>Assigned:</u> ML
Dr. David C. Lai Professor University of Vermont Comp. Sci. & E.E. Department Burlington, VT 05405 (802) 656-3330	<u>Degree:</u> Ph.D., Electrical Engineering, 1960 <u>Specialty:</u> Signal Processing and Pattern Recognition <u>Assigned:</u> 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 Chemistry,
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
Dayton,

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

List of Participants (continued: page 12)

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) 873-2855

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 Szydluk
Professor
State University of New York
Physics Department
Plattsburgh, NY 12901
(518) 564-2048

Degree: Ph.D., Physics, 1964
Specialty: Theoretical and Compu-
tational 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 Manu-
facturing
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 University/Fresno
6. Dr. Paul Szydlak - 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

PARTICIPANT LABORATORY ASSIGNMENT (Continued: page 2)

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 Institute of New York
2. Dr. Charles Ih - University of Delaware

PARTICIPANT LABORATORY ASSIGNMENT (Continued: page 4)

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 Lewittes - 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

<u>Volume I Report Number</u>	<u>Title</u>	<u>Research Associate</u>
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

<u>Report Number</u>	<u>Title</u>	<u>Research Associate</u>
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 Transfer 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 Evaluation 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 , $\text{H}^+(\text{CH}_3\text{CN})_k(\text{H}_2\text{O})_m$, And $\text{H}^+(\text{HCN})_n$ Cluster Ions	Dr. Carol A. Deakne
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

<u>Report Number</u>	<u>Title</u>	<u>Research Associate</u>
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

<u>Report Number</u>	<u>Title</u>	<u>Research Associate</u>
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 Microscopical 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 Ki-kpatrick
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

<u>Report Number</u>	<u>Title</u>	<u>Research Associate</u>
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

<u>Report Number</u>	<u>Title</u>	<u>Research Associate</u>
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, <u>Gammarus Tigrinus</u>	Dr. Gerald V. Poje
80	Simplification of Nonlinear Systems	Dr. L. Rai Pujara

<u>Report Number</u>	<u>Title</u>	<u>Research Associate</u>
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. Szydluk

<u>Report Number</u>	<u>Title</u>	<u>Research Associate</u>
93	Building a Multiple Regression Equation When Many Variables Are Available	Dr. Enoch C. Temple
94	A Preliminary Investigation 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

FEASIBILITY OF RECORDING ATMOSPHERIC INTERFEROGRAMS

Prepared by: Dr. Richard Anderson

Academic Rank: Professor

Department and University: Department of Physics
University of Missouri-Rolla

Research Location: Rome Air Development Center,
Strategic Surveillance Section, Griffis, AFB

USAF Research: Dr. Norman Chonacky

Date: August 8, 1983

Contract No: F49620-82-C-0035

THE FEASIBILITY OF RECORDING

ATMOSPHERIC INTERFEROGRAMS

by

Richard Anderson

ABSTRACT

A survey of the literature indicates that atmospheric interferograms were recorded and have been used to verify certain aspects of the theory of turbulence. Self-referencing interferograms have not been recorded. They could be used in reproduction to produce distorted wavefronts. The intent of the research was to record self-referencing interferograms under conditions of various degrees of turbulence. The amount of turbulence would be measured at certain points along the laser beam propagation path, so C_n^2 will be known for each interferogram.

After some initial experimentation it is evident that a Mach-Zehnder will be difficult to align and presents fringe visibility problems. A Smartt interferometer is an alternate instrument for the recording of self-referencing interferograms. Problems related with its use are discussed in this paper. Also, an alternate interferometer is discussed.

Acknowledgement

The author would like to thank the Air Force Systems Command, Air Force Office of Scientific Research, Southeastern Center for Electrical Engineering Education, and Rome Air Development Center for this opportunity to perform interesting and worthwhile research.

He would like to express his appreciation to Dr. Norman Chonacky, Donald Stebbins, Todd Updike, Dr. Jerry Knopp, and Dr. James Cusack for suggesting an area of research and providing the collaboration, guidance, and facilities to perform this research. The entire support staff of the Strategic Surveillance Section must be thanked for the excellent administrative and technical support.

I. INTRODUCTION:

When a light beam traverses a turbulent atmosphere, it suffers distortions produced by random refractive-index variations. These variations produce fluctuations in the phase of the optical signal sensed by a detector. This has detrimental effects on the resolution of the imaging system. This has importance in optical image quality in photography, in optical communications, and in optical tracking systems.

Atmospheric turbulence has been characterized over a horizontal path by the magnitude of the refractive index structure parameter C_n . Some investigators define the degree of turbulence in terms of C_n^2 instead of C_n . An integrated C_n can be measured over a path using a scintillation measuring system.^{1,2} Alternately, C_n may be measured at discrete points along the propagation path of a laser beam by using temperature measuring microprobes.³ The temperature measurements yield point values of the temperature structure parameter C_T which can be converted to C_n with a knowledge of the atmosphere pressure, temperature, and the laser wavelength.⁴

Above a flat field with nearly uniform vegetation, turbulence can be characterized as a break up of the air into patches with a lower size limit ℓ_0 and an upper limit L_0 . The cross section of the laser beam will be filled with patches of varying size. Depending upon the degree of turbulence, the patches may vary from meters to millimeters in diameter. The path size will be important in determining our spatial frequency of the interference fringe and in defining a self-referencing aperture diameter.

Turbulence has been characterized by the measured value of C_n .^{5,6} The exact boundaries between weak, moderate, and strong turbulence are not well defined. The values in reference (5) will be used to define turbulence.

A review of the literature⁵⁻¹¹ on interferometric measurements of turbulence is of value to this experiment, even though, these interferograms were not self-referencing. These experiments demonstrate photographic exposure times and the degree turbulence to produce

blurring. These experiments give an idea of the spatial frequency of the fringes required for different turbulence and indicate the size of lasers needed in these experiments.

The self-referencing interferograms will be recorded and the phase interferogram will be replayed to produce a distorted laser beam which can be used in laboratory experiments. If these interferograms are correlated with field C_n measurements, in replay distorted laser beams corresponding to different degrees of turbulence can be produced.

II. OBJECTIVES:

The principal objective of this research will be to produce interferograms which may be used in reproduction to produce an aberrated or distorted laser beam. This distorted beam may be used in other laboratory experiments. Each interferogram will be correlated with field measurements for the amount of turbulence. In this way, each interferogram may be used to produce a distorted beam corresponding to various degrees of turbulence.

The second objective will be to determine the most satisfactory method of recording the interferograms. They shall be self-referencing interferograms and they have never previously been produced.

Interferograms of this type may be recorded using different experimental arrangements and these techniques must be evaluated. Virtually none of the experimental parameters are known to produce these interferograms.

III. IMPORTANT THEORY:

The refractive index structure parameter C_n will be evaluated from temperature differences measured by microprobe resistance thermometers placed near the laser beam path. The temperature measurements with the microprobes allows the evaluation of the temperature structure parameter which is defined as

$$C_T^2 = \frac{\langle (T_2 - T_1)^2 \rangle_{\text{ave}}}{r^\nu}, \quad \ell_0 \leq r \leq L_0, \quad r \geq (\lambda L)^{1/2} \quad (1)$$

where r is the distance between the probes, ν is general equal to $2/3$, and L is the laser beam path length. Also r is greater than the Fresnel zone size of the laser light, $r \geq (\lambda L)^{1/2}$. $\langle (T_2 - T_1)^2 \rangle_{\text{ave}}$ is a reasonably long time average of the square of the measured temperature differences. The refractive index structure parameter C_n is defined as

$$C_n^2 = \frac{\langle (n_2 - n_1)^2 \rangle_{\text{ave}}}{r^{2/3}}, \quad \ell_0 \leq r \leq L_0, \quad r \geq (\lambda L)^{1/2}, \quad (2)$$

where n_1 and n_2 are the indices at two points a distance r apart perpendicular to the propagation direction of the laser beam and $\langle (n_2 - n_1)^2 \rangle_{\text{ave}}$ is a time average of the square of the difference of indices. The index of refraction of air can be related to the atmospheric pressure, mean air temperature, and laser wavelength⁴

$$n = 1 + (77.6 P/T) (1 + .00753/\lambda^2) \times 10^{-6} \quad (3)$$

As a result, the temperature and refractive index structure parameters are related to each other

$$C_n^2 = [77.6 P/T^2 (1 + 0.00753/\lambda^2) \times 10^{-6}] C_T^2 \quad (4)$$

In order to observe interference fringes between two partially coherent beams of light with an interferometer it is necessary to consider fringe visibility.¹² In the experiment the laser source has a definite bandwidth and is called a quasimonochromatic source. P_1 and P_2 are two sources at distances s_1 and s_2 from a point Q where they interfere. The field at the point Q is

$$E(Q, t) = E(P_1, t - s_1/c) + E(P_2, t - s_2/c) \quad (5)$$

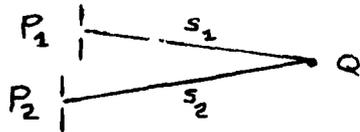


Figure 1. Two Source Interference.

The time averaged intensity is

$$\langle I(Q) \rangle = \langle I_1(Q) \rangle + \langle I_2(Q) \rangle + 2 \operatorname{Re} \Gamma_{12}(\tau), \quad (6)$$

where the first and second terms are the intensity at Q produced by source P_1 and P_2 and the last terms are the complex conjugates of each other and are the real part of the mutual intensity function so

$$\begin{aligned} \langle E(P_1, t-s_1/c) E^*(P_1, t-s_1/c) \rangle &= I_1(Q) \\ \langle E(P_2, t-s_2/c) E^*(P_2, t-s_2/c) \rangle &= I_2(Q) \end{aligned} \quad (7)$$

$$\langle E(P_1, t-s_1/c) E^*(P_2, t-s_2/c) \rangle = \langle E(P_1, t+\tau) E^*(P_2, t) \rangle = \Gamma_{12}(\tau)$$

In the equation above $\tau = (s_2 - s_1)/c$ and $\Gamma_{12}(\tau)$ is defined by the equation

$$\Gamma_{12}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T E(P_1, t+\tau) E^*(P_2, t) dt = \langle E(P_1, t+\tau) E^*(P_2, t) \rangle \quad (8)$$

The normalized mutual intensity function is

$$\gamma_{12}(\tau) = \frac{\Gamma_{12}(\tau)}{\sqrt{\Gamma_{11}(0)\Gamma_{22}(0)}} = \lim_{T \rightarrow \infty} \frac{\frac{1}{2T} \int_{-T}^T E(P_1, t+\tau) E^*(P_2, t) dt}{\sqrt{\Gamma_{11}(0)\Gamma_{22}(0)}} \quad (9)$$

where $\gamma_{12}(\tau)$ has a value from 0 to 1. Also $\Gamma_{ii}(0)$ is

$$\Gamma_{ii}(0) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T E(P_i, t) E^*(P_i, t) dt = \langle E(P_i, t) E^*(P_i, t) \rangle = \langle I_i(Q) \rangle \quad (10)$$

$$\text{so} \quad \Gamma_{12}(\tau) = \sqrt{\langle I_1(Q) \rangle \langle I_2(Q) \rangle} \gamma_{12}(\tau) \quad (11)$$

Equation (6) becomes

$$\langle I(Q) \rangle = \langle I_1(Q) \rangle + \langle I_2(Q) \rangle + 2[\langle I_1(Q) \rangle \langle I_2(Q) \rangle]^{1/2} |\gamma_{12}(\tau)| \cos \phi(\tau), \quad (12)$$

where $\phi(\tau)$ is the phase difference.

The visibility of the interference fringes is

$$V = \frac{\langle I(Q) \rangle_{\max} - \langle I(Q) \rangle_{\min}}{\langle I(Q) \rangle_{\max} + \langle I(Q) \rangle_{\min}} \quad (13)$$

The maximum and minimum intensity depends on the phase angle difference, where $\cos \phi$ equals +1 and -1, so

the visibility is

$$V = \frac{2[\langle I_1(Q) \cdot I_2(Q) \rangle]^{1/2} |\gamma_{12}(r)|}{\langle I_1(Q) \rangle + \langle I_2(Q) \rangle} \quad (14)$$

In the proposed experiment, using interferometers of various design, the incident atmospherically distorted wavefront is divided so a small portion of this wavefront can be isolated. This portion of the beam may directly interfere or may be expanded and collimated before it interferes with the total original atmospherically distorted wavefront. Since a small portion of the incident beam interferes with the total distorted beam, this portion is the reference for the interferogram, so the interferogram is self-referencing. The visibility will depend directly upon the intensity of the interfering beams and on the spatial coherence of the partial wavefront.

In this discussion two points (1) and (2) are separated a distance r_{12} in the plane of the aperture which selects the portion of the wavefront used for self-referencing. There will be a difference in the average index of refraction and phase for light has traveled through the turbulent atmosphere from the laser transmitter to the points (1) and (2) in the aperture. In turbulence theory a quantity called the phase structure function is

$$D_\phi = \langle (\phi_2 - \phi_1)^2 \rangle_{\text{av}}, \quad \ell_0 \leq r_{12} \leq L_0, \quad (15)$$

where r_{12} lies within the subrange of turbulence. Tatarski¹³ and Lutomirski and Yura¹⁴ have related the phase structure function to the refractive index structure parameter. Their equations within the subrange of turbulence have slightly different forms and are

$$D_\phi = 2.91 k^2 C_n^2 r_{12}^{5/3} L \quad (16)$$

$$\text{or } D_{\phi} = 2.91k^2 C_n^2 r_{12}^{5/3} L [1 - 0.8 (2\pi r_{12}/L_0)^{1/3}], \quad (17)$$

where $k = 2\pi/\lambda$ and L is the laser beam path length. These equations hold for plane waves which is the form of our wavefront at the telescope. At very small scales $r_{12} \ll \ell_0$ the relation is

$$D_{\phi} = 3.44k^2 C_n^2 r_{12}^2 L \ell_0^{-1/3} \quad (18)$$

For a Gaussian beam shape the normalized mutual coherence function is

$$\gamma_{12}(r_{12}, L) = \exp[-\frac{1}{2}D_w(r_{12}, L)] = \exp[-\frac{1}{2}(D_{\phi} + D_{\ell})] = \exp[-\frac{1}{2}D_{\phi}] \exp[-\frac{1}{2}D_{\ell}], \quad (19)$$

where $D_w(r_{12}, L) = D_{\phi}(r_{12}, L) + D_{\ell}(r_{12}, L)$, D_{ϕ} is the phase structure function, and D_{ℓ} is the log-amplitude structure function. At large L and/or strong turbulence D_{ℓ} saturates and this is the scintillation saturation region so $D_w \approx D_{\phi}(r_{12}, L) = D_{\phi}(r_c, L)$

$$\text{and } \gamma_{12}(r_c, L) = \exp[-1/2 D_{\phi}(r_c, L)] \quad (20)$$

The length $r_{12}\ell = (\lambda z)^{1/2}$ is the correlation length and r_c is the coherence length, which can be evaluated from equation (20).

The equations in the last paragraph assume a nearly parallel laser beam has been transmitted through a turbulent atmosphere. Depending upon the amount of turbulence, it may have or not have undergone beam spreading and this beam is collected directly by the interferometer and r_c is a distance between two points in the collected beam. In our experiment the laser beam slightly diverges and the telescopic system reduces the image size by 1/10th.

The size of the aperture through which light passes depends upon our definition of coherence. Some investigators define a coherent patch when $[\langle(\phi_2 - \phi_1)^2\rangle]^{1/2}$ changes by π , others when γ_{12} equals $1/e$, or 0.88.¹⁵ This means D_{ϕ} may be taken as π^2 , 2, or 0.256. Assuming a value for L , C_n , and λ , it is possible to calculate values of r_c by accepting a coherence definition for D_{ϕ} .

A review of the literature⁹ on interferometric turbulence measurements indicates that the spatial frequency of the fringes must

exceed $1/\ell_0$. This will be the criterion accepted in this experiment. The spatial frequency of the fringes is determined by the interferometer and is

$$f = 2\theta/\lambda, \quad (21)$$

where θ is the angle of intersection of the interfering beams.

IV. DISCUSSION OF THE EXPERIMENT:

The major facility to be used is the 1m Cassegrain telescope equipped with a coelostat at the RADC/Verona site. The telescope may receive light along a nearly horizontal path from a laser transmitter in small buildings at approximately 0.3, 0.6, and 2km from the receiver. The laser transmitter may be a triple etalon tuned dye laser of approximately 5 to 10 watts average power and bandwidth and coherence length at 500nm of 3.6 GHz and 8.3cm or an etalon prism tuned argon or krypton ion laser which can radiate near a watt in each discrete lasing line. High power ion lasers have a bandwidth of ± 22 MHz and a coherence length of 7m. The bandwidth of the ion laser is measured using a Fabry-Perot interferometer. In the case of dye laser, hardened Fabry-Perot mirrors are required and the nominal operating wavelength, λ_0 , must be measured using a separate apparatus.¹⁶

The transmitter is mounted on an optical bench on a concrete platform. A beam expander produces a slightly divergent wavefront which symmetrically fills the 1m plane mirror of the coelostat. The transmitter laser is approximately 1.5m above a flat horizontal field covered with grass from 0.6 to 1.2m tall. The coelostat mirror is approximately 12m above the field.

The coelostat directs the collected beam vertically into the 1m Cassegrain telescope. The back focal length and f number of the telescope remains constant by changing the position of the secondary mirror, even though finite distant objects are observed. The f number of the telescope is f/18. The light from the telescope is collimated using a reflective optics collimator. This collimator produces a 10cm beam which is deflected by a 45° mirror into the interferometer.

One type of interferometer which will be used to produce interferograms is the Smartt interferometer.¹⁷ A Smartt interferometer has a f-number of $f/2$ while the telescope is $f/18$. The collimated beam is focused using a corrected $f/2$ lens upon the point-diffraction film. The Smartt interferometer consists of a thin partially transparent non-scattering film placed on an optical flat. A precise, very small circular aperture is present in this coating. The quality of the image formed by the telescope system viewing through a turbulent atmosphere is controlled by the turbulence and without turbulence depends on the alignment of the telescope and associated optics. In this experiment the $f/2$ lens will probably dominate the aberrations of the optical system which gives a central spot of radius $R = 2.44\lambda$.

The aberrated light is focused to a blur spot and the smallest blur spot is the circle of least confusion. The precise aperture is moved by micrometer screws and is placed in the center of this circle. Light is diffracted by the aperture producing a large central Airy pattern while the remainder of the beam passes through the semitransparent film where it is attenuated and it has all the phase distortions of the atmosphere and optics. In the Smartt interferometer the plate may be moved along the optic axis or in a direction perpendicular to it by micrometer drives. When the precise aperture is on the optic axis and slightly displaced from the circle of least confusion, circular fringe patterns are observed in the central Airy ring pattern, since diffracted light from the aperture and the beam diverging through the semi-transparent film interfere. When the precise aperture is moved laterally to the optic axis with the fine adjustment drive, straight line fringes are observed on a plane surface placed perpendicular to the optic axis. The system is equivalent to the interference pattern produced by two point sources of light and straight line fringes are observed except for the atmospheric distortion of the fringes. The size of the precise aperture is not known so it is impossible to correlate this aperture size to the coherence patch in the atmosphere. As a result, the degree of coherence of the self-referencing wavefront produced by the aperture cannot be calculated.

The lateral displacement of the aperture should not be greater than the radius of the Airy central spot produced by the $f/2$ lens. This assures that adequate energy is still focused on the aperture and fringes can still be observed. In our experiment at a wavelength $\lambda = 500\text{nm}$ this corresponds to a lateral displacement of $1.22 \times 10^{-3}\text{mm}$. The literature¹⁸ on the interferometer indicates that the instrument has a minimum back focal length of 2mm. This indicates that a lens projects the interference pattern on a ground glass or the fringes may be photographed with a camera focused at infinity. Using this minimum focal length, the spatial frequency of the fringes is 1.22 fringes/mm.

The subspace of turbulence extends from ℓ_0 measured in millimeters to L_0 in meters. If one wishes to use the interferograms in playback to produce a distorted laser beam, the fringe frequency, $f > 1/\ell_0$, must be greater than 1 fringe per mm. The telescope system collects light in a 1m aperture and it emerges in a 10cm aperture. This means linear dimensions are reduced by 0.1 and areas by 0.01. As a result, the interferometer should have a spatial frequency $f > 10$ fringes per mm. The Smartt interferometer will only be able to record on the interferogram details of turbulence where $\ell_0 > 1\text{cm}$. This is adequate for the case of weak turbulence.

In reproduction of the distorted laser beam it will be necessary to separate the reference beam from the phase information beam. For the Smartt interferometer they have an angular separation of 0.61×10^{-4} radians. Assuming the full 10cm incident beam is used, this will place severe conditions in separating the beams in the reproduction. Separation occurs at $\sim 170\text{m}$.

In optics a spatial coherence occurs where the term $\exp[-\frac{1}{2} D_\phi]$ = 0.88 or $D_\phi = 0.256$. Using the simplified theory of Tatarski¹³ given in equation (16) for path lengths of 0.3, 0.6, and 2km for weak, moderate, and strong turbulence, the coherence lengths are evaluated. They are shown in Table 1 where r_c' is the unreduced coherence radius, r_c is the reduced radius, A_c is the reduced coherence area, and A_c/A is the ratio of the reference beam area to the unrestricted beam area.

Table 1. Coherence Lengths, Areas and Area Ratios

Path Length L(km)	Turbulence $C_n^2(m^{-2})$	r'_c (cm)	r_c (cm)	A_c (cm) ²	A_c/A_{total}
0.3	6.4×10^{-17}	12.00	1.20	4.52	5.76×10^{-2}
0.6	6.4×10^{-17}	7.88	$.38 \times 10^{-1}$	1.90	1.49×10^{-2}
2.0	6.4×10^{-17}	3.83	3.83×10^{-1}	4.61×10^{-1}	5.87×10^{-3}
0.3	1.6×10^{-15}	1.73	1.73×10^{-1}	9.40×10^{-2}	1.20×10^{-3}
0.6	1.6×10^{-15}	1.14	1.14×10^{-1}	4.08×10^{-2}	5.19×10^{-4}
2.0	1.6×10^{-15}	5.55×10^{-1}	5.55×10^{-2}	9.68×10^{-3}	1.23×10^{-4}
0.3	2.5×10^{-13}	8.37×10^{-2}	8.37×10^{-3}	2.20×10^{-4}	2.80×10^{-6}
0.6	2.5×10^{-13}	5.51×10^{-2}	5.51×10^{-3}	9.54×10^{-5}	1.21×10^{-6}
2.0	2.5×10^{-13}	2.68×10^{-2}	2.68×10^{-3}	2.26×10^{-5}	2.88×10^{-7}

As the light travels through the telescope, there are ten reflections of light with losses at each. Assuming 90% reflectivity the amount of light at the output is 1/3 the incident input. This seems to agree with observations of Dr. Jerry Knopp. As a result, all of the light should be collected and used.

If a Mark-Zehnder interferometer is used to record the interferograms, it will be of conventional design and to utilize the full beam, so 15cm mirrors and beam splitters are required. They should have a surface flatness of $\lambda/20$. The self-referencing aperture should be placed in the cross arm and aperture sizes, r_c , are given in Table 1. Turbulence measurements will be made at points along the optical path using resistance thermometer microprobes.³ The beam from the aperture must be expanded and collimated to the incident beam size using a corrected optical telescope. The aperture and telescope must have x and y translational motions transverse to the beam and must be rotatable through small angles about a central axis. The aperture must be aligned with the beam expander to avoid pronounced coma and astigmatism and the collimator must give a plane wavefront to avoid a circular fringe pattern.

The intensity of the self-referencing and full size distorted beams are controlled by the coatings on the beam splitters. The first beam splitter should have a high reflection coating on the first surface and an antireflecting coating on the second surface designed for the wavelength used in the experiment. The ratio of the aperture area to total beam area is indicative of the division of energy in both arms of the interferometer. For good visibility these intensities should be nearly equal. If s wave polarized radiation is used and this polarization is maintained in the atmosphere and optics, beam splitters at 45° angle of incidence can be designed with 99.99% reflective coatings and 99% antireflective coatings for a wavelength λ_0 .²⁰ If weak and moderate turbulence are studied, different beam splitters must be used. The first surface of the first beam splitter must have reflectivities from 90% to 99.9% and its second surface must have an AR coating with a transmission of 99%. The second beam splitter should have AR coatings on both surfaces of about 90% transmission. Since different beam splitters must be used depending upon the amount of turbulence, this will present frequent alignment problems.

In the earlier discussion it was assumed that the lower limit of the size of the subspace of turbulence was 1mm and this required resolving 40 fringes/mm. The resolution of high and low contrast Plus X and Tri X film are 125 lines/mm, 50 lines/m and 100 lines/mm and 50 lines/mm, respectively. Their speed (ASA) are 125 and 400, respectively. All of these films are at least 100 times faster than the best holographic film (Agfa Geavert 10E56). If one assumes the flashlamp pumped dye laser was used as the transmitter and an ideal pulse energy of 0.1 joules was obtained at some known wavelength λ_0 , then on a 10cm diameter circle in the film plane approximately 2.5×10^4 ergs/cm² could be delivered using an ideal unrestricted Mach-Zehnder. In measuring weak and moderate turbulence the aperture and beam splitters attenuate the beam from 10^{-2} to 10^{-4} so 2.5×10^2 to 25 erg/cm² are delivered at the film plane. These films can be exposed with ~ 0.1 erg/cm² so some attenuation may be needed.

If a spatial frequency of fringes of 40/mm is specified at 500 nm, the beams intersect at an angle of 1×10^{-2} radians. In reproduction the distorted beam may be resolved from the reference beam in a distance of 10m. As a result, both reproduction problems and availability of film of proper resolution seem ideal.

The limiting conditions on the Mach-Zehnder will be the production of the beam splitters with various specified coatings. Also there will be severe alignment problems. From my limited experience, the self-referencing Mach-Zehnder will be difficult to align. Figure 2 is a sketch of the Mach-Zehnder.

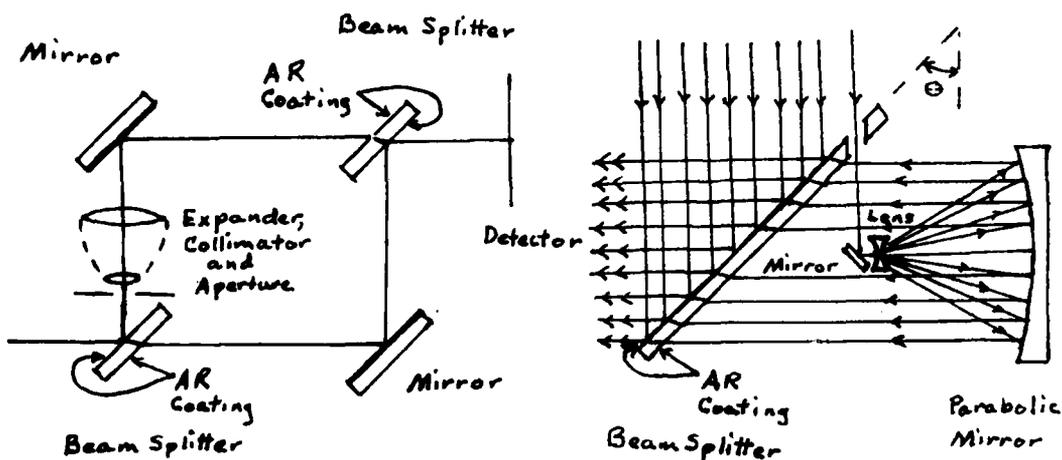


Figure 2. Mach-Zehnder Interferometer. Figure 3. Special Interferometer.

An alternate interferometer is sketched in Figure 3. The lens will be a diverging lens with an aspheric surface to correct for spherical aberration. The spatial frequency of the fringes is changed by making small changes in the angle $\theta = 45 \pm \Delta\theta$ and $f = 2\Delta\theta/\lambda$. The intensity of the beams is controlled by the transmittance (99%) of the antireflecting dielectric coatings placed on the beam splitter. The reference aperture must be a compromise for it is controlled by the hole placed in the beam splitter. So, only weak turbulence can be studied with an aperture of about 1cm diameter. The beam splitter has a 15cm diameter.

There will be stringent design conditions placed on the lens-mirror beam expander and collimator for the self-referencing beam. Again the nature of the fringes is effected by this alignment, but it should be possible to permanently align it and the beam splitter can be rotated through small angles and tilted to produce the desired fringe spacing and pattern. It should be an improvement over the Mach-Zehnder where the aperture-expander-collimator and two totally reflecting mirrors must be aligned and tilted to produce the desired pattern.

This interferometer can only be used with the ion lasers with coherence length of several meters, since the total optical path through the telescope system must be less than the coherence length of the laser. The optical path length difference between the reference beam passing through the telescope and that reflected from the beam splitter is greater than the coherence length of the dye laser ($\sim 10\text{cm}$).

The fringes are equally spaced and can be adjusted for a density of 40 to 50/mm. Any phase difference between the self-referencing beam and the beam reflected from a given position on the beam splitter appears as a distortion in the fringe. This is the same as observed with the Mach-Zehnder. Since only a beam splitter with antireflection coatings can be used to control beam intensities, only weak turbulence conditions may be studied with the interferometer.

It must be mentioned that a camera shuttered argon or krypton ion laser must be used in the place of the flashlamp pumped dye laser. This laser emits discrete lines with improved bandwidth. The output power of an etalon-prism tuned laser's output is 60%²¹ of the simple prism tuned laser. Certain lines for the 18W laser may have outputs of 2.4 watts and after losses in the telescope 0.8 watt reaches the interferometer. For exposure times of 10^{-3} sec and with an attenuation for weak turbulence of 10^{-2} about 80 ergs reach the 10cm diameter exposed area on the film. An exposure of 1 erg/cm^2 is incident on the film plane which is 10 times more than required to expose Plus X or Tri X film.

V. RECOMMENDATIONS:

This study indicates that the experiment is entirely feasible. The problem is much more complicated than originally proposed and will require careful study and design. Several optical components must be special ordered and the design carefully specified.

Either an ion laser or dye laser may be used. The ion laser has improved bandwidth and this allows variations in the design of the interferometer. With exposures of 1/1000 sec weak turbulence can be studied using the ion laser. With exposures of 1/100 sec, the ion laser may be used to study moderate turbulence.

Many additional changes and recommendations are contained in the report. This report is a feasibility study which evaluates and points out problems related to recording self-referencing interferograms by three techniques.

It appears that the Mach-Zehnder or a special interferometer can be used to achieve self-referenced interferograms. Inadequate information is available on the Smartt interferometer to completely evaluate it but it may be used in weak turbulence where $\ell_0 > \text{cm}$.

REFERENCES

1. G. R. Ocks, W. D. Cartwright, and D. D. Russell, "Optical C_n^2 Instrument Model II," NOAA Tech. Mem., ERL-WPL-51, 1979.
2. G. R. Ocks and R. J. Hill, "A Study of Factors Influencing the Calibration of Optical C_n^2 Meters," NOAA Tech. Mem., ERL-WPL-106, 1982.
3. D. P. Greenwood and D. B. Youmans, "A Fine-Wire Microtemperature Probe for Atmospheric Measurements," Rome Air Development Center Report, RADC-TR-75-240, 1975.
4. S. L. Valley, Handbook of Geophysics and Space Environment, McGraw-Hill, New York, 1965.
5. M. Bertolotti, M. Carnevale, L. Muzii, and D. Sette, "Interferometric Study of Phase Fluctuations of a Laser Beam Through the Turbulent Atmosphere," Appl. Optics, Vol. 7, pp. 2246-2251, 1968.
6. H. J. Pfeifer, M. König, and B. Koch, "Effects of Atmospheric Turbulence on the Visibility and Spacing of Interference Fringes," J. Opt. Soc. Am., Vol. 70, pp. 163-167, 1980.
7. M. Bertolotti, M. Carnevale, L. Muzii, and D. Sette, "Atmospheric Turbulence Effects on the Phase of Laser Beams," Appl. Optics, Vol. 13, pp. 1582-1585, 1974.
8. G. M. B. Bouriciuss and S. F. Clifford, "Experimental Study of Atmospherically Induced Phase Fluctuations in an Optical Signal," J. Opt. Soc. Am., Vol. 60, pp. 1484-1489, 1970.
9. R. G. Buser and G. K. Born, "Determination of Atmospherically Induced Phase Fluctuations by Long-Distance Interferometry," J. Opt. Soc. Am., Vol. 60, pp. 1079-1083, 1970.
10. R. G. Buser, "Interferometric Determination of the Distance Dependence of the Phase Structure Function for Near-Ground Horizontal Propagation at 6328 A," J. Opt. Soc. Am., Vol. 61, pp. 488-491, 1971.
11. G. K. Born, R. Rosenberger, K. D. Erben, F. Frank, F. Mohr, and G. Sepp, "Phase-Front Distortion of Laser Radiation in a Turbulent Atmosphere," Appl. Optics, Vol. 14, pp. 2851-2863, 1975.
12. M. Lurie, Effects of Partial Coherence on Holography, University Microfilms, Ann Arbor, Michigan, pp. 21-35, 1967.

13. V. I. Tatarski, Wave Propagation in a Turbulent Medium, McGraw-Hill, New York, 1961.
14. R. F. Lutomirski and H. T. Yura, "Wave Structure Function and Mutual Coherence Function of an Optical Wave in a Turbulent Atmosphere," J. Opt. Soc. Am., Vol. 61, pp. 482-487, 1971.
15. M. Born and E. Wolf, Principles of Optics, p. 511, Pergamon Press, Oxford, 1965.
16. R. A. Anderson, "The Theory of the Fabry-Perot Interferometer and Its Use with Wave Trains of Short Temporal Coherence," Internal Report RADC/OCSP, 1983.
17. R. N. Smartt and J. Strong, "Point-Diffraction Interferometer," J. Opt. Soc. Am., Vol. 62, p. 737, 1972.
18. Ealing Optics Catalog, pp. 253-255, The Ealing Corp., South Natick, MA.
19. J. I. Davis, "Considerations of Atmospheric Turbulence in Laser System Design," Appl. Optics, Vol. 5, pp. 139-147, 1966.
20. R. Carnes, McDonnell Douglas Corp., Astronautics Division, Laser Communications Section, private communication.
21. Spectra Physics, Mountain View, CA, private communication.

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

TESTING THE REPRESENTATIVENESS OF THE SUPPLY DATA BANK

Prepared by: Richard W. Andrews and Frederick J. Gentner
Academic Rank: Associate Professor and Ph.D. Student
Department and University: Statistics and Management Science
Graduate School of Business, The University of Michigan
Research Location: Air Force Logistics Management Center, Supply Directorate
USAF Research: Captain Ronald E. Travis, USAF
Date: September 24, 1983
Contract No: F49620-82-C-0035

TESTING THE REPRESENTATIVENESS OF THE SUPPLY DATA BANK

by

Richard W. Andrews
and
Frederick J. Gentner

ABSTRACT

The objective of this project was to investigate the representativeness of the Air Force Supply Data Bank (SDB) with respect to the Air Force Standard Base Supply System (SBSS). As a working definition of representativeness, we require that a sample include at least one unit from each stratum. In addition, a sample should be balanced. That is, the sample mean and standard deviation should be close to the population mean and standard deviation on all test variables. By interviewing Air Force management, familiar with the SBSS, 78 supply test variables were subjectively selected. A statistic B, which measures closeness was used to compare alternative data banks.

The statistical methodology was accomplished by writing and executing two FORTRAN programs, BASEREP and BCOMP. BASEREP tests the representativeness of a SDB by random generating 100 alternative SDB's with the same coverage and comparing their values of B. BCOMP enables sensitivity analysis to be performed on the values of B.

Two SDB's are considered, one with 12 bases and one with 6 bases. Both of these have the same coverage. The SDB with 6 bases is better balanced for 11 of the 12 months tested. By using BCOMP the bases and the test variables that cause balancing problems are determined.

It is recommended that the 6 base SDB be augmented with one additional base so as to cover the entire Air Force, and that any data bank be tested monthly for representativeness.

Acknowledgment

The authors thank the Air Force Systems Command, the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for the opportunity to spend a challenging and rewarding summer with the Supply Directorate of the Air Force Logistics Management Center. We especially appreciate the cooperation of Lt. Col. John C. Lauer, Jr., who provided a working environment that enabled this research to be carried out.

In addition we would like to acknowledge the close working relationship that we were able to establish with Captain Ronald E. Travis. Our knowledge of a new substantive area, the Air Force Standard Base Supply System, is predominantly due to the patience of Major Steven M. Orenstein, Captain Travis, and CMSergeant William G. Hargrave, Jr. They answered many detailed questions. Our data handling chores could not have been accomplished without the help of Mr. Charles Miller, MSergeant William L. Hause, and MSergeant Joseph Fuller. Frequent discussions with Mr. Wayne Faulkner kept the direction of the research focused. In general, the Air Force Logistics Management Center provided an environment for productive work.

I. Introduction

The Air Force Logistics Management Center (AFLMC) maintains the Air Force Transaction History and Stock Control Data Bank (Supply Data Bank).¹ The purpose of maintaining the Supply Data Bank (SDB) is to provide data for stockage policy analysis. The SDB consists of extensive stock number supply records from 12 bases. These 12 bases were chosen so that six of the major commands are represented. Table 1 provides a list of the twelve participants and their corresponding commands.

Table 1
Stock Control Data Bank Participants

<u>Base Name</u>	<u>Corresponding Command</u>
* Minot	Strategic Air Command (SAC)
Langley	Tactical Air Command (TAC)
* England	Tactical Air Command (TAC)
* Little Rock	Military Airlift Command (MAC)
Dover	Military Airlift Command (MAC)
* Randolph	Air Training Command (ATC)
* Upper Heyford	U.S. Air Force Europe (USAFE)
Bitburg	U.S. Air Force Europe (USAFE)
Bentwaters	U.S. Air Force Europe (USAFE)
Clark	Pacific Air Force (PACAF)
* Kunsan	Pacific Air Force (PACAF)
CIRF (Kadena)	Pacific Air Force (PACAF)

The choice of bases to include in the data bank did not take into account base level supply characteristics. The purpose of this project is to address the representativeness of this data bank with respect to the Air Force Standard Base Supply System (SBSS). Aggregate supply data at the base level is available for all bases through the monthly M32 reports. It is appropriate to require that the bases in the SDB should be representative with respect to this monthly data. The quantities chosen from the M32 report for the purpose of testing the representativeness of the data bank are referred to as the supply test variables.

When choosing a sample from a population of items, there is a multitude of designs that can be used to obtain the selection. Sometimes a random sampling design is appropriate and sometimes it is more appropriate to use a purposive selection. In the case of choosing the bases to be in the SDB, a purposive selection was used within the six major commands that are represented.²

Regardless of the manner in which the sample is selected, it is appropriate to test the resulting sample for representativeness.^{3,4} Since the supply data bank is a selection of 12 bases from all Air Force bases, the representativeness of that data bank will be investigated. Because of the large volume of data in the SDB, the AFLMC is considering using 6 of the 12 SDB bases for some of their analyses. The 6 bases under consideration are designated with an asterisk in Table 1. Notice that the use of these 6 bases will still provide representation for the same 6 major commands. The representativeness of these 6 bases will also be investigated.

In Section II the objective of this research is explicitly stated. The supply data bank and associated supply variables are discussed in Section III. Section IV gives a working definition of representativeness and cites the appropriate literature. In Section V the statistical methodology used in this investigation is described. The results of the research are given in Section VI.

Recommendations are given in the last section in two areas. First, specific recommendations concerning the existing data bank are given. Second, recommendations concerning the use of the data bank in future studies are stated.

II. Objectives of Research Effort

The three objectives of this project are:

- (1) to develop and present the methodology to test the representativeness of a subset of Air Force bases with respect to SBSS characteristics,
- (2) to make a statement concerning the representativeness of SDB(12), the Supply Data Bank with all 12 bases, and
- (3) to compare the representativeness of SDB(6), the Supply Data Bank with just 6 bases with SDB(12).

In order to accomplish these objectives five steps had to be completed.

1) The principle investigators had to become familiar with the Standard Base Supply System. This was accomplished by reading and studying parts of Air Force Manual¹ 67-1, Volume II, Part 2 and by having extensive discussions with supply personnel at the Air Force Logistics Management Center. In addition, the principle investigators interviewed supply personnel at Maxwell Air Force Base and Eglin Air Force Base. The purpose of those interviews was to identify important base level supply characteristics. 2) The supply test variables had to be chosen and classified. In Section III the choice of these test variables is discussed. 3) The supply test variable data had to be acquired and converted for use with the testing system. This time consuming and onerous task is briefly discussed in Section III. 4) The statistical methodology had to be researched, developed, and programmed. A description of the methodology is given in Section V. The calculations were accomplished by using two FORTRAN programs: BASEREP and BCOMP. Documentation for these two programs are contained in a supplemental report.⁵ 5) The results of the computer runs had to be evaluated with respect to their impact on representativeness. This is discussed in the results section.

III. Standard Base Supply System Data

The Air Force SBSS keeps detailed records on every supply transaction. Every base keeps an item record file, a detail record file, a repair cycle record file and a transaction file. For a complete description of the copious details contained in these files see U.S.A.F. Manual¹ 67-1, Vol. II, Part 2. Twelve bases have been designated to participate in the Supply Data Bank maintained by AFLMC. Each of these 12 bases send a monthly record dump of the complete transaction file to AFLMC. In addition each of these 12 bases send to AFLMC a semi-annual record dump of the item record, detail record, and repair cycle record files.

Table 2 gives the reader an appreciation of the amount of supply data available in the SDB. This data is stock number specific, that is each record is identified by a stock number. The record dumps are loaded from tape to disc packs on the AFLMC computer system in a manner that makes them available for stockage policy analysis.⁶

Table 2

Supply Data Bank (12 participating bases)

<u>Data Type</u>	<u>Frequency of Dump</u>	<u>Approximate Average Number of Records per Base</u>	<u>Number of Fields in Each Record</u>
Transaction	Monthly	126,000	34
Item Record	Semi-Annual	65,000	52
Detail Record	Semi-Annual	122,000	25
Repair Cycle Record	Semi-Annual	8,000	56

In order to test for representativeness with respect to the SBSS, base level supply test variables must be specified. Furthermore, the values of these base level supply variables must be available for the period of interest.

Through discussions and interviews with personnel knowledgeable in the SBSS, 78 supply test variables were identified. The values of these variables are available monthly from the M32 reports.¹

Table 3 gives a list of the 78 supply test variables that were chosen. These specific variables were subjectively chosen so that the important base supply characteristics are captured. By using these variables the differences and similarities between bases with respect to the supply system can be analyzed. In Table 3 these 78 variables are classified into 16 categories.

Magnetic tapes consisting of 23 of the 44 pages of the M32 report were obtained from the Air Force Data Systems Design Center. Each of the 12 months consisted of several seven-track tapes. These seven-track tapes were obtained from the Univac 1050 computing system. For each month the seven-track tapes were converted to a nine-track tape compatible with IBM type systems.

This conversion program was written in Assembler language. For each base and for each month, 115 numbers were extracted from the nine-track tapes through the use of a COBOL program. This is the data set which provides the input into the programs BASEREP and BCOMP. A more detailed explanation is contained in the documentation.⁵

IV. Representativeness

In a series of four articles, Kruskal and Mosteller^{4,7,8,9} discuss thoroughly the use of the terms "representative sample" and "representative sampling." They state that the use of these terms has spawned nine categories of meanings. They provide evidence that these terms are often used without being explicitly defined. For this report, two of their categories are combined to provide a working definition of representativeness.

First, for a sample to be representative it must provide coverage of the population. In addition, the sample should in some way be a miniature version

Table 3

Supply Test Variables
(By Category)

Category 1: Measures of Size

Number of Items Records - Overall
Total (Repair Cycle)

Number of Item Records - Overall Total
(EOQ)

Number of Item Records - Overall Total
(Equipment in Stock)

Dollar Value On-Hand Balance - Overall
Total (Repair Cycle)

Dollar Value On-Hand Balance - Overall
Total (EOQ)

Dollar Value On-Hand Balance - Overall
Total (Equipment in Stock)

Category 2: Activity Measures

Total Transactions (Supplies)

Total Transactions (Equipment)

Total Issues (Supplies)

Total Issues (Equipment)

Total Due-Outs (Supplies)

Total Due-Outs (Equipment)

Total Receipts (Supplies)

Total Receipts (Equipment)

Total DOR - Recoverable

Total DOR - EOQ

Total DOR - Equipment

Total Overall Requisitions - Total
Number

Category 2: Activity Measures (Con't)

Total Overall Requisitions - Dollar
Value Total

Category 3: Effectiveness Measures

Recoverable Issue Effectiveness

EOQ Issue Effectiveness

Recoverable Stockage Effectiveness

EOQ Stockage Effectiveness

Recoverable Release Effectiveness

EOQ Release Effectiveness

Total Item Records with Requisition
Objective, Zero Accessible Assets -
Overall Total

Category 4: MICAP

Total Number MICAP Cause Code A

Total Number MICAP Cause Code B

Total Number MICAP
Total All Other Causes

Total Number Deletes
Delete Code 0

Total Number Deletes
Delete Code 9

Total Number Deletes
Total All Other Codes

Category 5: Repair Cycle Information

Average RCT for Total RTS
Total All Organizations

Table 3 (Con't)

Category 5: Repair Cycle Information
(Con't)

Average RCT for Total NRTS
Total All Organizations

Average RCT for Total Condemned
Total All Organizations

Average AWP for RTS
Total All Organizations

Average AWP for NRTS
Total All Organizations

Average AWP for Condemned
Total All Organizations

Total Units: RTS
Total All Organizations

Total Units: NRTS
Total All Organizations

Total Units: Condemned
Total All Organizations

Category 6: EOQ/Recoverable Mix

EOQ Line Items Requested As a Percent
of Total EOQ and Recoverable

Category 7: Urgency Mix

Line Items Requested, Urgency Need B
As a Percent of Overall Total

Line Items Requested, Urgency Need C
As a Percent of Overall Total

Category 8: Priority Mix

Total Overall Requisitions
Priority Group II As a Percent of Total

Total Overall Requisitions
Priority Group III As a Percent of Total

Category 9: Funding Mix

Percent of Total Item Records with
Requisition Objective - System
Support Division

Percent of Total Item Records with
Requisition Objective - General
Support Division

Category 10: Bench Stock

Bench Stock Line Items Authorized
Total of Overall Summary

Bench Stock Line Items Due Out
Total of Overall Summary

Bench Stock Issues and DOR Totals

Category 11: Excess

Dollar Value Excess Supplies
General Support Division

Dollar Value Excess Supplies
System Support Division

Dollar Value Excess Supplies
Non-Stock Fund

Dollar Value Excess Equipment
General Support Division

Dollar Value Excess Equipment
Non-Stock Fund

Total Excess Detail Records

Category 12: Special Level

Total Number of Item Records with
Special Level - Repair Cycle

Total Number of Item Records with
Special Level - EOQ

Table 3 (Con't)

Category 12: Special Levels (Con't)

Total Number of Item Records with
Special Level - Equipment in Stock

Total Number of Item Records with
Special Level, Date of Last Demand > 365

Category 13: Minimum Level

Overall Total Number of Details
Justification O-Life Cycle Ret

Overall Total Number of Details
Justification A ISSL/FOSSL

Overall Total Number of Details
Justification - All Others

Total GSD Minimum Level Dollar Value

Category 14: WRSK/MSK

Total WRSK Transactions

Total Type of Details - MSK

Total Type of Details - WRSK

Category 15: Receipts on Time

Total Priority Group 1 - Percent on Time

Total Priority Group 2 - Percent on Time

Total Priority Group 3 - Percent on Time

Category 16: Inventory Accuracy

Total Line Items Counted - Complete

Total Line Items Counted - Special

Total Line Items Counted - Identity Change

Total Line Items In Lot

Category 16: Inventory Accuracy
Con't

Accuracy Percent of Units -
Complete

Accuracy Percent of Units - Special

Accuracy Percent of Item Records
Sampled

of the population. It is necessary that both of these components be explicitly defined so that the SDB is testable with respect to the definition of representativeness.

The Kruskal and Mosteller definition of coverage will be used: "A sample containing at least one item from each stratum of a partition of the population is said to cover it for that partition." This definition of coverage will be used as one of the two components of representativeness. Therefore, if a representative sample is required, it can only be representative with respect to the collection of stratum that have at least one item in the sample. This has immediate impact on the SDB. Obviously, the 12 bases were chosen so that each of 6 major commands had at least one base included. Therefore, at best, the SDB can only be representative of the subset of the entire Air Force which is contained in these 6 major commands. Because these 6 major commands comprise 97 of the 116 Air Force bases, this is not a major restriction.

The concept of a balanced sample³ will be used to provide the miniature component of representativeness. In a strict definition of a balanced sample, the sample moments must be equal to the corresponding population moments for all test variables. This is an overly restrictive requirement when one is working with 78 test variables. In this report we provide the miniature concept by stating that a sample is balanced if the sample mean and sample standard deviation are close to the corresponding population mean and population standard deviation of the test variables. In the next section the statistic that measures closeness is defined. Its value is compared with the same statistic for other samples that provide the same coverage.

V. Statistical Methodology

The two purposes of this section are (1) to define a statistic which measures the balance of the data bank with respect to the population of interest and (2) to provide a methodology which evaluates the computed statistic.

Let X_{1j} = the value of the i^{th} supply test variable for the j^{th} base ($i = 1, 2, \dots, 78; j = 1, 2, \dots, N$). The population mean and variance of the i^{th} supply test variable are given by:

$$\bar{X}_1(P) = N^{-1} \sum_{j=1}^N X_{1j},$$

$$\text{and } S_1^2(P) = (N-1)^{-1} \sum_{j=1}^N (X_{1j} - \bar{X}_1(P))^2.$$

Let n be the number of bases in the data bank. For this study n will be either 12 or 6. Let D represent the subset of the integers from 1 to N that correspond to the bases in the supply data bank. The data bank mean and variance of the i^{th} test variable are given by:

$$\bar{X}_1(D) = n^{-1} \sum_{j \in D} X_{1j}$$

$$\text{and } S_1^2(D) = (n-1)^{-1} \sum_{j \in D} (X_{1j} - \bar{X}_1(D))^2.$$

The statistic that is proposed to measure balance is designated B and given by:

$$B = \sum_{i=1}^{78} \left\{ \left(\frac{\bar{X}_1(P) - \bar{X}_1(D)}{\bar{X}_1(P)} \right)^2 + \left(\frac{S_1(P) - S_1(D)}{S_1(P)} \right)^2 \right\}$$

The following points should be noted concerning the statistic B . Dividing by $\bar{X}_1(P)$ and $S_1(P)$ indicates that the statistic B measures the difference between the moments relative to the population. This feature will allow comparison between computed B values. Note that the difference in means and difference in standard deviations are equally weighted. Also note that the relative differences are squared. This allows both a positive and a negative difference to contribute to an increase in B . Finally, note that if a data bank is perfectly balanced with respect to its mean and standard deviation on all test variables, then the computed value of B would be zero. Smaller

values of B indicate a better balanced data bank.

The statistic B can be calculated for each of the two proposed data banks and for each of the 12 months. What methodology should be used to evaluate these computed B values? For each month and for each of the two proposed data banks, 100 other data banks with the same coverage properties were randomly generated. That is, each of the major commands that are represented in the proposed data bank have identically the same number of representatives in each of the randomly chosen data banks.

For each of the 100 generated data banks, a B value is calculated. A count is made of the number of times that the value of B for the proposed data bank is smaller than the random data bank value of B. This count measures how well the SDB is balanced with respect to others with the same coverage.

VI. Results and Conclusions

By executing the program BASEREP the statistical methodology yields the results given in Table 4.

Table 4
Counts For Which Supply Data Bank Is Better Balanced
(Out of 100)

<u>Month</u>	<u>SDB (12 bases)</u>	<u>SDB (6 bases)</u>
April, 1982	19	94
May, 1982	46	88
June, 1982	33	96
July, 1982	37	85
August, 1982	46	93
September, 1982	32	97
October, 1982	10	99
November, 1982	23	80
December, 1982	58	93
January, 1983	16	5
February, 1983	56	96
March, 1983	86	95

Out of the 100 replications the count is given as to how many times the SDB being tested had a smaller B value than the randomly selected SDB. An

explanation on how to interpret this count is as follows: By calculating 100 B-values from stratified random samples, the analysis has estimated the sampling distribution of B. In hypothesis testing the user specifies an acceptable size of the type I error. This is done based on the user's awareness of the context of the test. In a similar manner the acceptable count will depend on the context of this test of representativeness.

In many statistical tests of this type, a size of the type I error is set at 5 or 10 percent. This would indicate that a SDB was acceptable for any count larger than 5 or 10. Because the sampling distribution is being estimated, a more stringent 20 percent size of the type I error is suggested. Therefore, one should require at least a count of 20 before accepting a SDB as balanced. However, if the user wishes this criteria to be even more stringent, a larger count can be required. The counts in the 80's and 90's in Table 4 indicate exceptionally well balanced SDB's.

Three conclusions are obvious from Table 4. 1) For several of the months the SDB(12) is not well balanced. 2) In all but one month the SDB(6) is better balanced than the SDB(12). This will be further explained below. 3) The variability in the monthly counts indicate that any SDB should be checked for balance each month.

There are two questions that arise from Table 4 that need to be investigated. First, why is SDB(6) better balanced than SDB(12)? To investigate this question a special analysis was executed for the month of October, 1982. That month produced a low count of 10 for SDB(12) and a high count of 99 for SDB(6).

To investigate this situation each of the 12 bases in SDB(12) were eliminated from the data bank one at a time. A value of B was calculated for each resulting SDB(11). These B values are given in Table 5. In a similar manner,

the SDB(6) was augmented with one base from the 6 eliminated bases. These B values were calculated by using BCOMP. The results are given in Table 5.

Table 5

Investigation of the Difference in SDB(12) and SDB(6), October 1982

<u>SDB</u>	<u>B-Value</u>	<u>SDB</u>	<u>B-Value</u>
All 12	23.01	6 only	14.78
Excluding Randolph	28.04	Augmented by Dover	22.17*
Excluding Little Rock	26.88	Augmented by Langley	14.52
Excluding Dover	20.56*	Augmented by CIRF (Kadena)	17.69
Excluding Minot	26.04	Augmented by Clark	27.58*
Excluding Langley	25.07	Augmented by Bitburg	18.31
Excluding England	28.22	Augmented by Bentwaters	15.22
Excluding CIRF (Kadena)	24.31		
Excluding Clark	17.86*		
Excluding Kongsan	25.68		
Excluding Upper Heyford	25.49		
Excluding Bitburg	25.24		
Excluding Bentwaters	25.50		

As can be seen the B value for SDB(12) is reduced the most when Dover and Clark are excluded. Likewise, the B value for SDB(6) is increased the most when Dover and Clark are included. This indicates that these two bases strongly degraded the balance of SDB(12) for October 1982.

The second question that arises from Table 4 is: Why is the count for the SDB(6) in January, 1983 so low? To investigate this question each one of the 6 bases were eliminated in turn and then each one of the variable categories were eliminated in turn. The resulting B values as computed by BCOMP are given in Table 6.

Table 6

Investigation of Low Count for SDB (6), January, 1983

<u>Base Investigation</u>		<u>Variable Investigation (All 6 bases)</u>		
<u>SDB</u>	<u>B-Value</u>	<u>Variable Categories</u>		<u>B-Value</u>
		All 16		
All 6	149.24	Excluding Cat. 1		148.41
		Excluding Cat. 2		146.42
Excluding Randolph	211.47	Excluding Cat. 3		148.18
		Excluding Cat. 4		148.02
Excluding Little Rock	205.92	Excluding Cat. 5		13.27
		Excluding Cat. 6		149.19
Excluding Minot	208.46	Excluding Cat. 7		148.78
		Excluding Cat. 8		149.18
Excluding England	206.03	Excluding Cat. 9		149.00
		Excluding Cat. 10		148.19
Excluding Kunsan	208.74	Excluding Cat. 11		147.74
		Excluding Cat. 12		148.74
Excluding Upper Heyford	20.67	Excluding Cat. 13		148.74
		Excluding Cat. 14		148.81
		Excluding Cat. 15		148.89
		Excluding Cat. 16		147.08

It is obvious from Table 6 that the Repair Cycle data (Category 5) from Upper Heyford is strongly degrading the balance of SDB(6) for the month of January, 1983. This type of investigative analysis with the B values gives insight into which bases should be used in the SDB and in which months they should be used.

The resulting B values for all randomly generated data banks of size 6 for January, 1983 were observed. The lowest value of B was 14.69 compared to 149.24 for SDB(6). This lowest value came from a databank consisting of the following bases: Randolph, Hurlburt, Vandenberg, Langley, Kunsan, and Alconbury. The methodology of this report is not an optimizing procedure; however, the B values may suggest a data bank which is better balanced.

VIII. Recommendations

The first three recommendations pertain to the use of the Supply Data Bank as presently configured. The last two recommendations are suggestions for follow-on research.

1) Based on the analysis in this report we would recommend using the SDB(6) in lieu of the SDB(12). One should be aware that a balanced SDB with 12 bases is better than a balanced SDB with 6 bases, but the monthly counts on SDB(12) as indicated by Table 4 suggest it is not as consistently balanced as SDB(6). Furthermore the use of any SDB should be done with caution. Even though SDB(6) is clearly balanced for 11 of the 12 tested months, it severely lacks balance in January, 1983.

2) Based on the principle of coverage, SDB(6) covers only the 97 bases in the 6 major commands listed in Table 1. There are 7 other major commands which have a total of 19 bases. If those 19 bases can be considered to be homogeneous with respect to the SBSS then I recommend forming a single stratum out of those 19 bases. Then one base from that stratum could be added to SDB(6). The choice of this base should be made by calculating B values. This scheme would provide coverage for the entire Air Force.

3) Any supply data bank that is used should be tested for balance each month. Any irregularity should be investigated before that data is used for stockage policy analysis.

4) Because of the immense volume of data in any SDB, subsampling within bases will be necessary for most analysis questions. The subsampling should be designed in a manner so that the resulting estimates can be evaluated. The statistical concepts of bias and minimum variance should be considered in the design. The estimated variance of the estimators will have to include both the first and second stage sampling variation. Because of the purposive selection of the bases, the first stage variation will have to be evaluated via a superpopulation model.

5) Important supply descriptive statistics only available from the SDB should be designated. These quantities should be calculated monthly as

estimates of population parameters. The associated precision of these estimates should be calculated by considering the two stages of variation mentioned above.

REFERENCES

1. United States Air Force Manual 67-1, Volume II, Part 2.
2. Minutes for FY83 USAF Retail Stockage Advisory Group (RSAG) Meeting, May 2-5, 1983, Gunter Air Force Station, AL.
3. R. M. Royall and J. Herson, "Robust Estimation in Finite Populations," Journal of the American Statistical Association, Vol. 68, pp. 880-889, 1973.
4. W. Kruskal and F. Mosteller, "Representative Sampling, II: the Current Statistical Literature," International Statistical Review, Vol. 47, pp. 245-265, 1979.
5. R. W. Andrews and F. J. Gentner, Documentation for Software Developed as a Supplement to the Summer Faculty Research Project, September 1983.
6. T. James, "RAMAS II Supply Data Bank (SDB) Users/Maintenance Guide," AFLMC Report #990022, AFLMC, Gunter Air Force Station, AL.
7. W. Kruskal and F. Mosteller, "Representative Sampling, I: Non-scientific Literature," International Statistical Review, Vol 47, pp. 13-24, 1979.
8. W. Kruskal and F. Mosteller, "Representative Sampling, II: Scientific Literature, Excluding Statistics," International Statistical Review, Vol. 47, pp. 111-127, 1979.
9. W. Kruskal and F. Mosteller, "Representative Sampling, IV: the History of the Concept in Statistics, 1895-1939," International Statistical Review, Vol. 48, pp. 169-195, 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

INTELLIGENT CONTROLLER FOR SPACE EXPERIMENTS

Prepared by:	Dr. Francesco L. Bacchialoni
Academic Rank:	Associate Professor
Department and University:	Department of Electrical Engineering University of Lowell
Research Location:	Air Force Geophysics Laboratory, Aerospace Instrumentation Division, Sounding Rocket Branch
USAF Research Colleague:	Jack R. Griffin
Date:	13 September 1983
Contract No:	F49620-82-C-0035

INTELLIGENT CONTROLLER FOR SPACE EXPERIMENTS

by

Francesco L. Bacchialoni

ABSTRACT

This document reports the investigation on operational requirements and architecture of an intelligent controller designed to support AFGL scientific experiments in space. This controller is designed to minimize the workload of the operator, by testing automatically various subsystems of the experiment hardware, and conducting the experiment.

ACKNOWLEDGEMENT

The author would like to thank 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, Hanscom AFB, Massachusetts. He would like to acknowledge the laboratory, in particular the Sounding Rocket Branch for its hospitality and excellent working conditions. Finally, he would like to thank Mr. Edward F. McKenna for suggesting this area of investigation and for his guidance and collaboration. The author would also like to acknowledge many helpful discussions with Jack Griffin, Russell Steeves, Ray Wilton, William Miller, Roger Jacobs, Kenneth Walker, and several other members of the staff and the help of Cheryl Carpenter in typing this report.

I. INTRODUCTION

In recent years scientific experiments in space have become a frequent event, and it is reasonable to expect an ever increasing occurrence of them, considering the unique advantages of space for many types of scientific experiments.

The space vehicles mostly used for flights into the space near the earth are the Sounding Rockets and the Space Transportation System (STS), and it is anticipated that STS-based experiments will become common in the near future.

The scientific experiments flown in space have become rather elaborate and it is expected that more and more complicated ones will be attempted in the near future. These experiments obviously require a very careful organization and planning, and extensive pre-flight testing. This is usually done by human operators, who carefully follow detailed pre-flight assembly and test procedures. The experiment itself is eventually run in flight by an appropriate automatic sequencer.

An intelligent controller can greatly help the operator(s) in testing the scientific experiment hardware and in running the experiment. Important functions that the controller should perform are:

1. Test the hardware for proper operation, and check the results against pre-set values. Issue comprehensive results and detailed diagnostics, as deemed necessary.
2. Run the experiment, possibly in an adaptive fashion, keeping in account the results of the test.
3. Perform appropriate data processing to reduce the amount of data to be stored or transmitted.

The present study involves considerations on an intelligent controller specifically designed for AFGL experiments flown on sounding rockets and on the Space Transportation System.

II. OBJECTIVES

The main objective of this project was to investigate functional characteristics and realization of an experimental intelligent controller, to be utilized in connection with AFGL space experiments.

Detailed objectives of this investigation were:

1. Selection of the actions that the controller is expected to take.
2. Preliminary hardware architecture of a first generation controller.
3. Preliminary software considerations for the same.

III. GENERAL CHARACTERISTICS

The structure of both hardware and software of an intelligent controller capable of executing tests and running experiments may range from simple to extremely complicated. The establishment of its functional characteristics and specs is rather similar to the establishment of management procedures, therefore, it is based mostly on engineering judgment and not particularly on scientific determination.

As a starting point, it seems reasonable to develop the intelligent controller as an item supplementary to the existing manually operated control and test hardware used in connection with the sounding rockets, leaving the development of a controller specifically designed for STS-based experiments to a later time, with the anticipation that it will not be substantially different from that designed for sounding rocket experiments. This controller should work independently, but should not prevent or interfere with

the hardware tests as they are presently conducted. For this first generation controller, the operational scenario is assumed as follows:

1. The different subsystems of the scientific experiment hardware are assembled on the sounding rocket and tested as usual by human operators, using remote control from the blockhouse, through hardwired connections. These tests are performed when useful, and if necessary repeated several times during the assembly of the experiment hardware to be flown.

2. After all the instruments are assembled and ready to be launched, the operator may activate the intelligent controller for an automatic test. The test controller performs sequentially several specific tests and compares the results with stored (ROM) values, eventually sending a comprehensive diagnostic message to the blockhouse. This automatic test may be repeated several times.

3. The system operator activates the experiment controller, which operates as a sequencer designed to direct the particular experiment on board at that time. This experiment run may be repeated several times, like the automatic test, but independently from it.

Figure 1 indicates in a comprehensive block diagram the relations among the human operator, the intelligent controller and the scientific experiment.

IV. CONTROLLER SPECIFICATIONS

The usual method of design for microprocessor controlled systems is an iterative procedure starting from the functional specifications and going through a process of hardware/software selection and considerations. The performance of the first tentative design is then estimated and checked against the original specs, and the entire design process is possibly repeated one or more times until the final result is considered satisfactory. Notice that this loop is closed through the original specs which may be modified in this process, to accommodate hardware/software constraints.

A few (fuzzy) specifications of our controller can be established rather quickly, since the background reasoning is quite simple. Here they are:

1. Power consumption should be as small as possible. This dictates the adoption of low-power integrated circuits, which almost invariably means CMOS technology.

2. The same basic controller should be adaptable to different scientific experiments. This fact strongly recommends a modular structure, where hardware "modules", specifically designed for individual experiments and subsystems are inserted into the basic controller configuration. Controller software should also be organized in modular form.

3. Processor size. Selection of an 8-bit processor appears to be adequate, at least for the design of a first generation controller. Several CMOS 8-bit processors are presently available, while 16-bit CMOS processors will become available in the near future. These will be considered at a later time, when there will be accumulated experience on the operation of a controller based on an 8-bit processor.

4. Communications with subsystems. Due to the complexity of the overall configuration of the scientific experiment hardware and the controller itself, both serial and parallel communications links should be available.

5. Speed of operation. Speed of all the existing processors is considered adequate, since it seems that high speed is not strictly necessary for the controller. This estimate may have to be revised at a later date when the actual detail design is implemented.

6. Reliability. Since the controller is going to direct the experiment, its reliability is extremely important. This point is not examined in this document, but clearly must be considered later. Different techniques can be adopted to enhance the final reliability of the controller, but the inherent complications are beyond the scope of this investigation.

V. HARDWARE ARCHITECTURE

The hardware architecture has been established keeping in mind the factors of flexibility to accommodate different experiments, and modularity as means to obtain flexibility.

The envisioned design separates three functions that the intelligent controller is expected to accomplish. They are:

1. Health test of the various subsystems of the experiment hardware.
2. Experiment execution.
3. Data processing.

In reference to Figure 1, the test controller and the experiment controller are shown as separate blocks, while the "data processor" is assumed to be implemented as a "subsystem", therefore is not otherwise indicated.

Some consideration was given to the idea of adopting separate hardware for the test controller and the experiment controller. This leads to duplication of most hardware, which for this design is undesirable; therefore this configuration was discarded.

The adopted design utilizes common hardware for the test controller and the experiment controller, and the different functions are going to be provided by different software.

The controller design centers around an 8-bit CMOS microprocessor, with memory up to 64 Kbytes, parallel I/O, serial I/O, and interface for the MIL-STD-1553B serial bus. Memory and I/O address maps may be tentatively established, and they are shown in Figure 2. It is clear that at each

subsystem location there must be an appropriate interface capable of communicating with the central controller. At this point it is also clear that these interfaces have to be rather complex to be able to perform all the necessary operations.

Figure 3 shows a simplified block diagram of the hardware architecture of the controller. In order to have maximum convenience for the software design, many controller operations are to be executed through interrupts, requested by the human operator or by the subsystems.

The RS-232C serial I/O, the parallel I/O and the MIL-STD-1553B I/O are to be implemented as three separate modules which may or may not all be present in any specific experiment. Since the I/O speed is relatively slow, no Direct Memory Access (DMA) has been provided.

Several microprocessors have been considered, all 8-bit and CMOS. They are listed here: G-65SCXXX (family), 1802, 80C85, NSC 800, S6809. Selection of the NSC 800 has been based on its architecture, interrupt handling and instruction set, which are all very similar to those of the Z-80. This offers flexibility in the design and possible expansion of the related hardware and software modules.

Selection of memory and peripheral chips has been done somewhat arbitrarily, keeping in mind our background rules of low power and modularity of realization. As Interrupt Controller, the type 82C59 has been adopted. It can be programmed to handle masking and priority on eight interrupt request lines, and it also supplies "pointers" to the CPU. It can be programmed accessing it as a parallel I/O port. If more than eight interrupts are later desired, three additional ones can be handled directly by the processor; if even more are desired (up to 64), additional 82C59's may be added to the system.

Figure 4 shows a section of the controller including the microprocessor and up to 64 Kbytes of memory. It is clear that the memory size can be easily adjusted by varying the number of chips, without basically changing

the architecture. Also, RAM may be easily exchanged with ROM and vice versa, if desired. A possible variation of the memory design consists in replacing the entire memory hardware with one of the newly developed nonvolatile RAM modules (ex.: NVRD64, which is a 64K x 8 module). Notice that DMA has been excluded from the design, therefore, there is no provision for floating the system busses. Notice also that only some significant connections have been drawn, and that the I/O address decoder and interfaces are not shown.

VI. INTERRUPT SYSTEM

The following interrupts are planned to request these system operations:

- A. Requests CPU to check initialization of all subsystems.
- B. Requests CPU to read from any subsystem.
- C. Requests CPU to write to any subsystem.
- D. Requests CPU to execute Test.
- E. Requests CPU to execute Experiment.
- F. Requests CPU to stop and be available to repeat any action

With the adoption of the 82C59 Programmable Interrupt Controller, these six interrupts can be easily handled, and additional ones accommodated, if later desired. Masking and priority setting by software allows to have a reduced number of hard wire interrupt request lines between the subsystems and the controller, obviously simplifying wiring.

The interrupts related to the action taken by the human operator, that is "Start Test", "Start Experiment" and "Stop" must be requested through separate lines, in order to allow a high degree of freedom in deciding the

action. It seems reasonable, for these three interrupts, to assign highest priority to "Stop", second highest to "Start Test" and lowest to "Start Experiment", so that if three or two actions are asked at the same time by mistake, the controller will stop or execute a test but not execute the experiment.

VII. FLOW CHARTS OF OPERATIONS

Before proceeding to the design of various interfaces, it is useful to collect and show on flow charts all the actions that the test controller is expected to exercise. Figures 5 and 6 show these flow charts. Notice that handshaking through interrupts is used throughout.

First action of the human operator is to turn on power, by means of an on/off switch dedicated to the controller and subsystem interfaces. This action must also activate automatic initialization of the controller and subsystems. Initialization is applied to all the subsystems by means of a hard wired connection carrying an appropriate control signal (RESET IN, active low). Each interface goes through its own initialization procedure and then issues an interrupt request on a hardware line, common to all the interfaces (OR gates). It is sufficient that one interface issue an interrupt request to activate the handshaking; if the processor does not receive an interrupt request before completing execution of "Timeout 1", then a message to this extent is stored and a diagnostic message is issued to the blockhouse. When power is on, initialization may be requested again by the operator, but during the initialization procedure, the operator cannot request any other action: the switches provided to start Test or Experiment or to Stop are deactivated by software.

After initialization is completed, the operator may select "Start Test", "Start Experiment" or "Stop". If by mistake he requests more than one action, the previously indicated interrupt priorities will take care of it and allow only one action.

The interrupt request for Test activates the D vector address and there an appropriate routine takes care of testing sequentially each of the subsystems. Subsystem identification and data are read by the processor using interrupt handshaking. Notice that depending upon the particular subsystem, the "Test Program for Subsystem N" (See Figure 6) may involve activating other subsystems, also. Finally, the processor evaluates the collected data and issues a diagnostic message.

It is clear that the detailed hardware and software design will have to be custom made for each experiment. It seems at this time that the specialized design for the hardware part of the system will be considerably easier than that of the software.

In the case of experiments flown up to now on sounding rockets, the successive steps of the experiment have been controlled in each case by an automatic sequencer, which operates by closing or opening switches wired into the different subsystems. The subsystems are turned on and off at certain times, without any consideration given to their data or health conditions. For really complicated, long-lasting and expensive experiments, it may be useful to modify the experiment step sequence according to signals issued by the experimental hardware, obtained from measured values and subsystem health. This envisioned adaptive operation clearly has an enormous potential for improving the execution of automatically run experiments. The price is greatly increased operational complexity, including the risk of system instability.

At this point, our engineering judgment is that to keep things relatively simple, the adaptive feature should not be included into the design of our first generation controller. It may be gradually introduced into subsequent designs, after having gained experience step by step. Therefore, our experiment controller software will consist of an "open loop" sequencer, similar in action to the ones used up to now. The program to run the experiment is started only by the operator requesting the interrupt E,

which directs the CPU to run the experiment. If necessary, during critical parts of the Test or Experiment program execution, the interrupt F (Stop) may be disabled by software masking, in order to prevent the operator from stopping the procedure at a critical point.

As indicated previously, reliability of operation is extremely important. No special consideration has been given to it for this first generation controller, therefore it remains one of the items to be reconsidered for a second generation controller.

VIII. TYPICAL SUBSYSTEM INTERFACE

Each subsystem interface must perform the following duties:

1. Initialization of the local hardware, requested by the controller.
2. Receive bytes from the controller. Format may be parallel or serial (RS 232C or MIL-STD-1553B).
3. Transmit bytes to the controller. Format may again be parallel or serial (RS 232C or MIL-STD-1553B).
4. Recognize the bytes received from the controller as parts of a command program.
5. Issue command bytes to a local decoder designed to issue command signals to the local hardware.
6. Read data bytes from the local hardware.
7. Handle signals for communications with the controller.

Hardware design of a typical subsystem interface is illustrated in

simplified form in Figure 7, where serial communications (RS 232C) have been assumed.

First action of each subsystem interface is to confirm completed initialization: local processor initializes all hardware then sends out "initialization completed" byte to UART and Interrupt (A) request to CPU (by way of RTS pin of UART). Controller CPU selects interfaces, one by one, sequentially, using subsystem command decoder (which does not appear in the central controller block diagram). Signal applied to each UART is CTS (Clear to Send).

When the controller CPU wants to send a byte to a specific subsystem, an interrupt request is applied to that subsystem processor (coming from the central controller command decoder). The subsystem processor responds through the local command decoder, issuing an interrupt request to the central controller (only Interrupt C should be enabled).

IX. SOFTWARE CONSIDERATIONS

Only some general remarks can be made here, since the development of the controller software is definitely a long and intricate process, beyond the time limits of this project.

Software for our intelligent controller should be organized in modular form, as previously indicated, in order to simplify debugging and to allow relatively easy insertion of additional new modules and/or replacement of existing modules. The starting point should be a detailed flow chart of the desired operation and a set of detailed flow charts of the related modules. Usual programming techniques are to be used, including the use of a microprocessor "development system".

Software for the subsystem interfaces is going to be definitely simpler than that of the controller. This software may also be subdivided into modules,

for easy debugging. Same design techniques apply here, and again the use of a development system is considered extremely helpful.

X. RECOMMENDATIONS

Continued effort is necessary to reach the point where a prototype intelligent controller can be built.

In particular, the following subjects should be carefully considered and completed:

- Interface between main controller and subsystem interfaces
- Subsystem interfaces
- Reliability
- Software

This paper reports the preliminary study concepts and begins to develop the hardware architecture. The remaining steps should be executed and eventually the whole design reconsidered from the beginning and modified if necessary. This design loop should be repeated several times until the result is considered satisfactory.

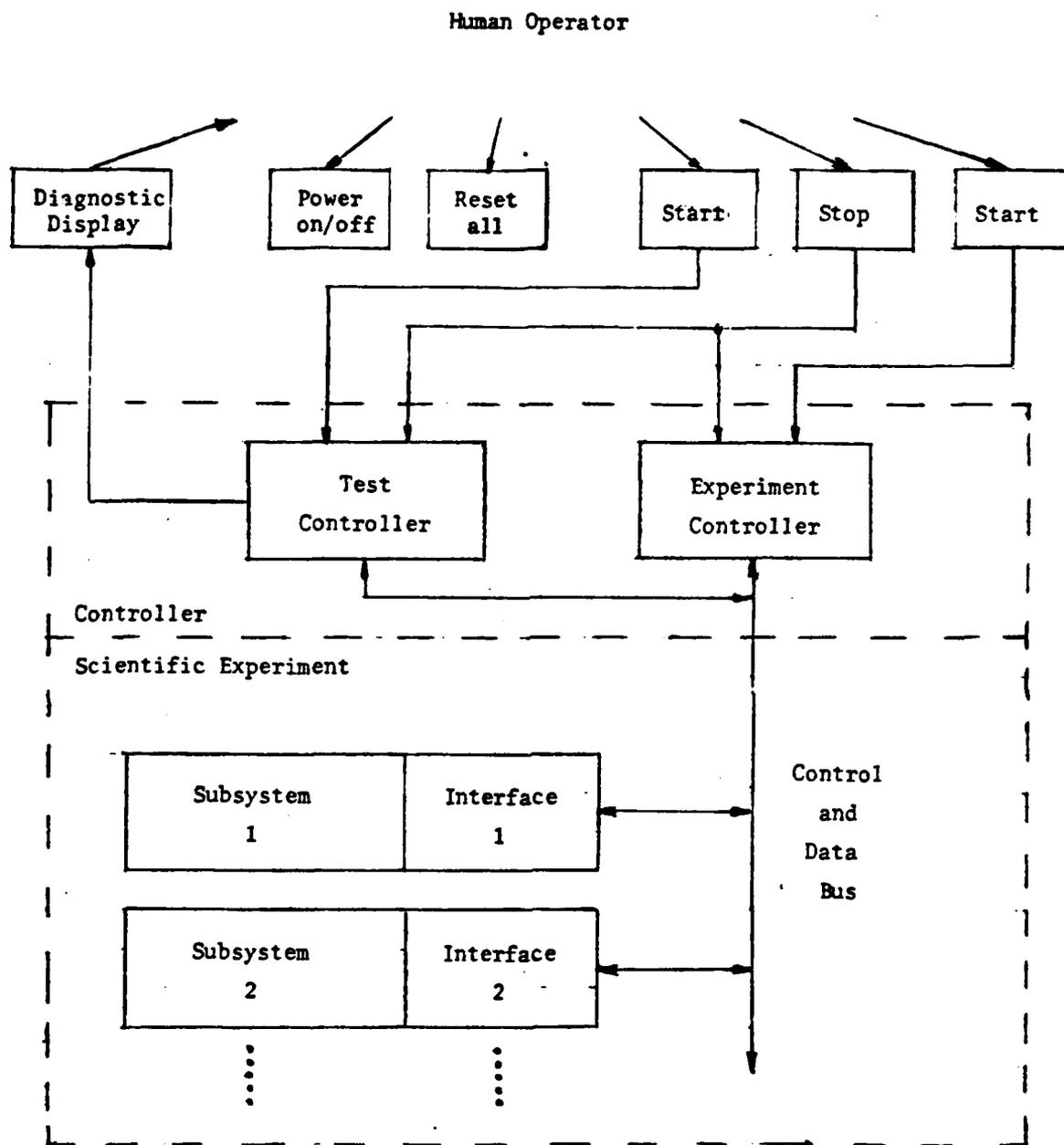


FIGURE 1. COMPREHENSIVE BLOCK DIAGRAM OF INTELLIGENT CONTROLLER

Memory Map

Decimal Address	Address Bus																
	A ₁₅	A ₁₄	A ₁₃	A ₁₂	A ₁₁	A ₁₀	A ₉	A ₈	A ₇	A ₆	A ₅	A ₄	A ₃	A ₂	A ₁	A ₀	
0000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	ROM
4095	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1		
4096	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	RAM
8191	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1		
8192	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	RAM
65535	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		

I/O Address Map

Decimal Address	Address Bus																
	A ₁₅	A ₁₄	A ₁₃	A ₁₂	A ₁₁	A ₁₀	A ₉	A ₈	A ₇	A ₆	A ₅	A ₄	A ₃	A ₂	A ₁	A ₀	
00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Parallel I/O
63	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1		
64	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	Serial I/O
127	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1		
128	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	MIL-STD-1553B
255	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1		

FIGURE 2. POSSIBLE MEMORY AND I/O ADDRESS MAPS

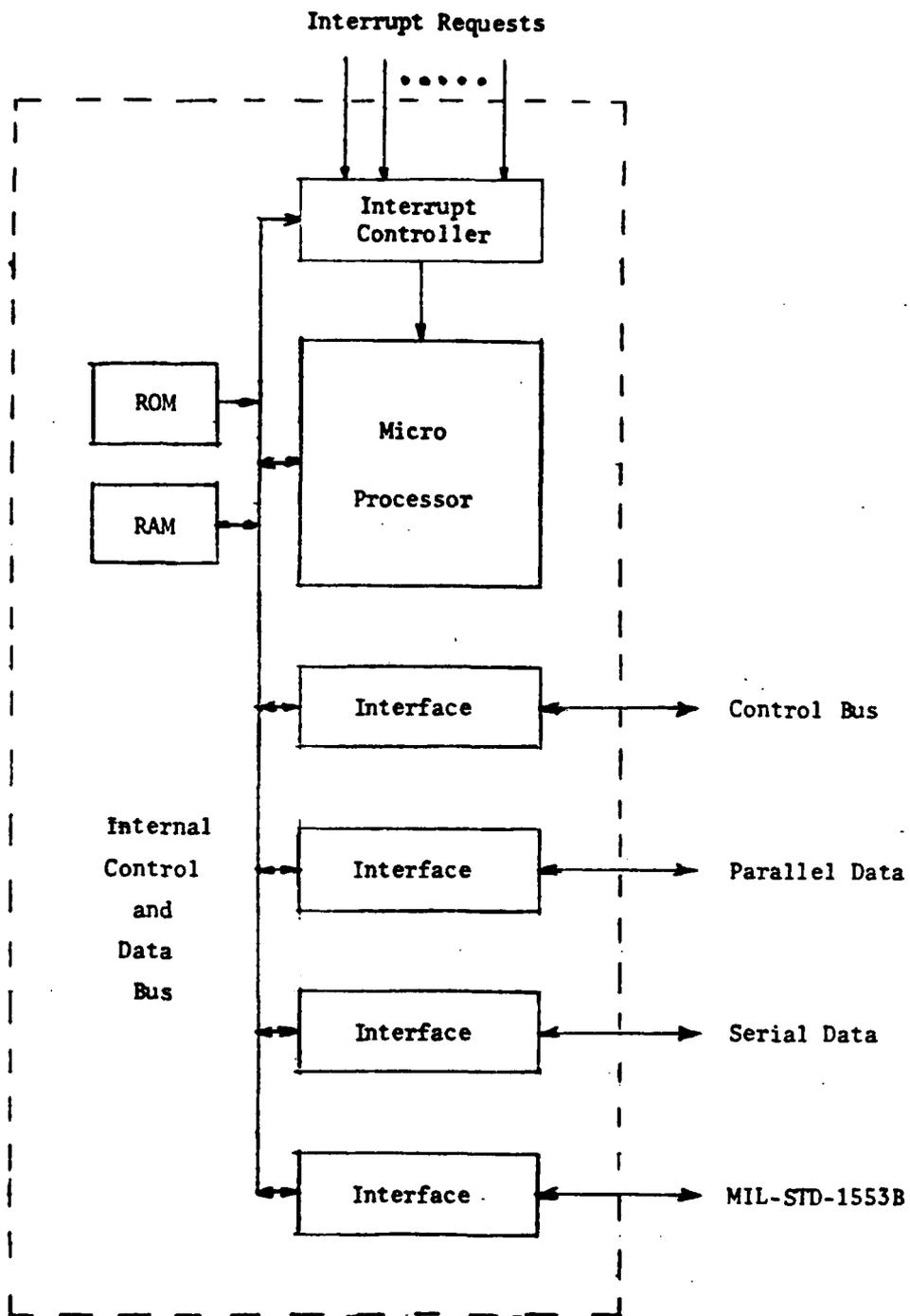


FIGURE 3. SIMPLIFIED BLOCK DIAGRAM OF INTELLIGENT CONTROLLER

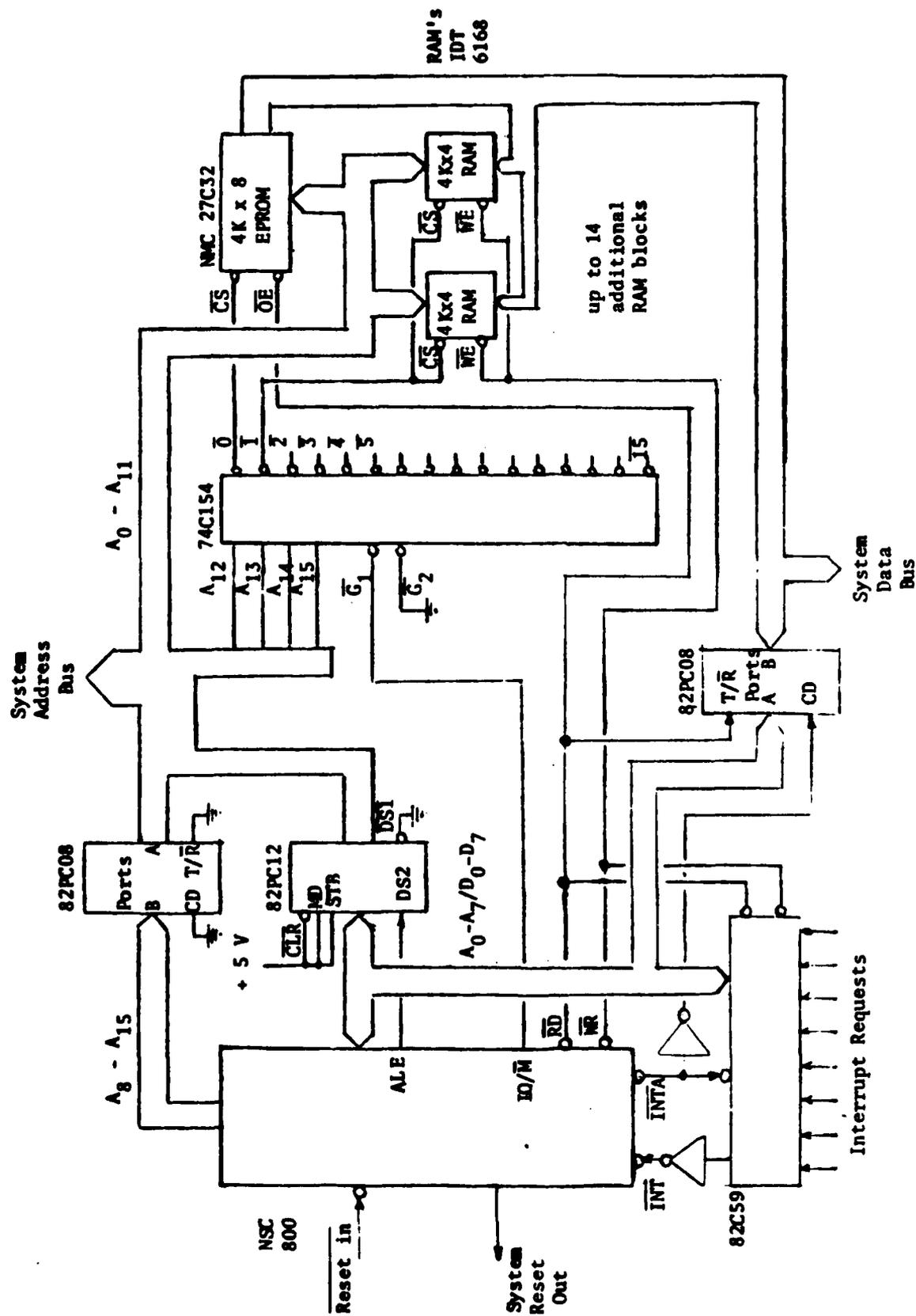


FIGURE 4. PARTIAL CONTROLLER HARDWARE

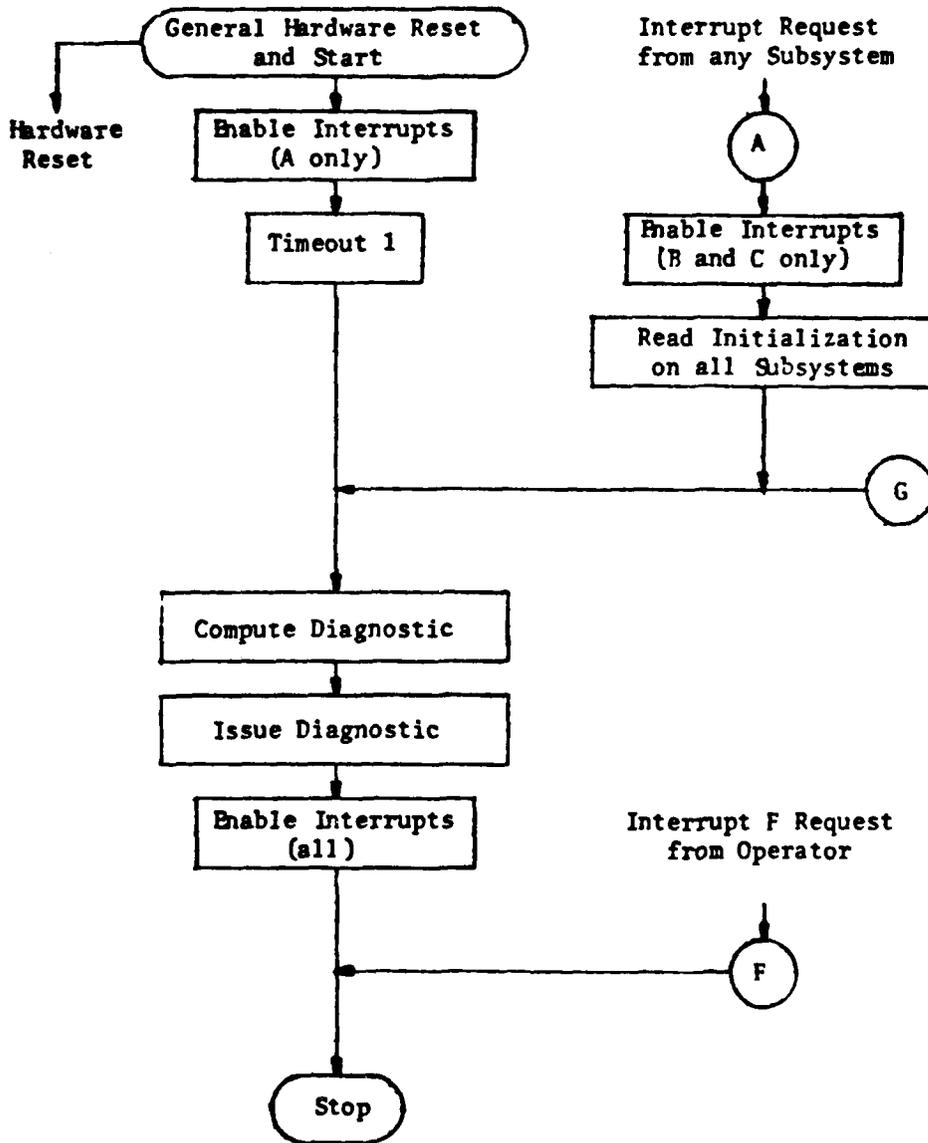


FIGURE 5. SKELETON FLOW CHART OF SOFTWARE/HARDWARE ACTIVITIES

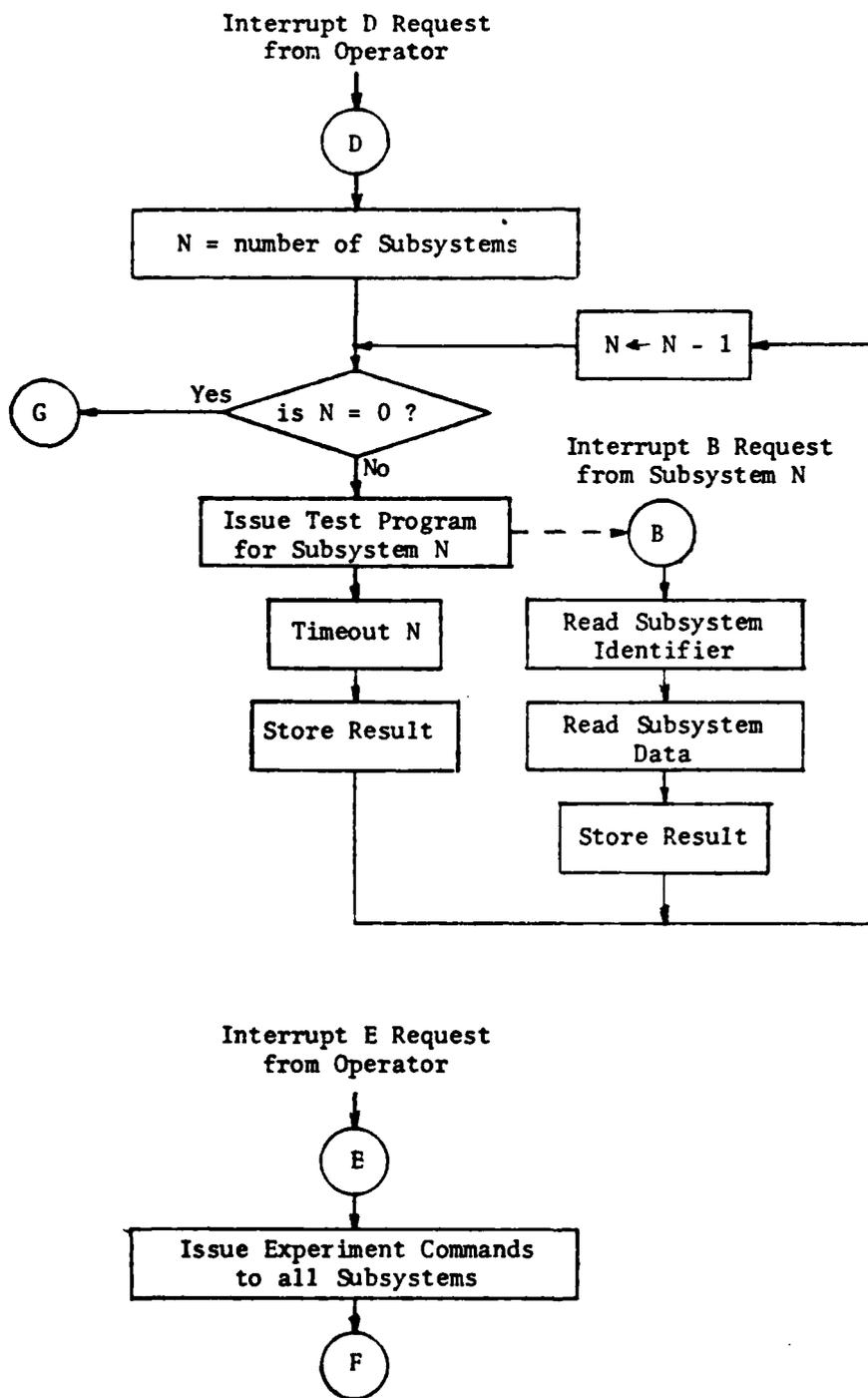


FIGURE 6. SKELETON FLOW CHART OF SOFTWARE/HARDWARE ACTIVITIES

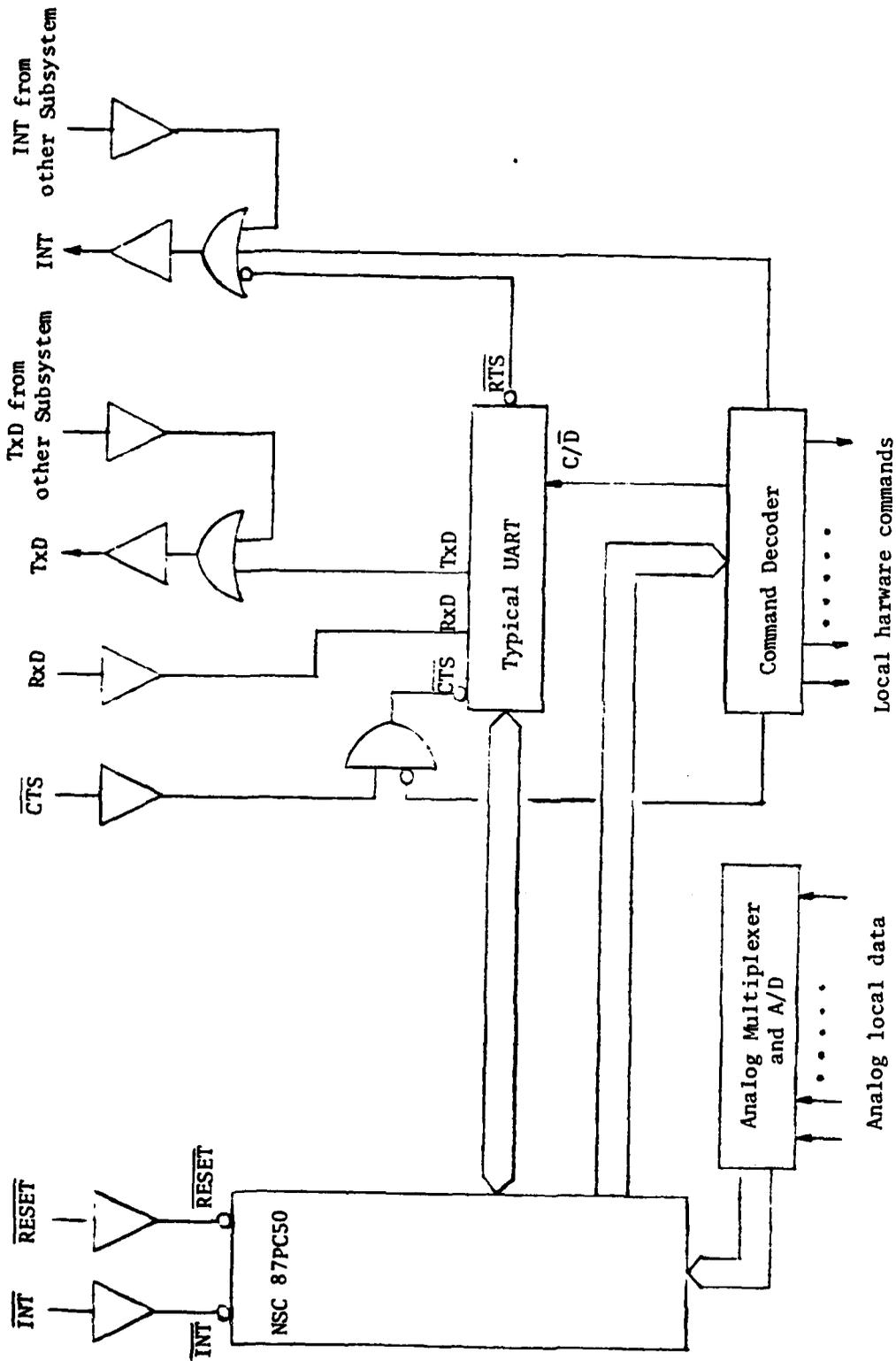


FIGURE 7. PRELIMINARY HARDWARE FOR SUBSYSTEM INTERFACE WITH SERIAL LINK

REFERENCES

1. Harold S. Stone, Microcomputer Interfacing, Addison Wesley Publishing Company, Reading, Massachusetts, 1982.
2. James W. Coffron, Practical Hardware Details for 8080, 8085, Z80 and 6800 Microprocessor Systems, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1981.
3. Adam Osborne, An Introduction to Microcomputers, Volumes 1 and 2, Adam Osborne and Associates, Inc., Berkeley, California, 1976.
4. E. Yourdon and L.L. Constantine, Structured Design, Yourdon Press, New York, New York, 1978.
5. NSC800 Microprocessor Family Handbook, National Semiconductor.
6. MCS-80/85 Family User's Manual, Intel Corporation.
7. Bipolar and CMOS Digital Data Book, Harris Corporation.

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

INFRARED EARTHLIMB EMISSION LINESHAPES

AS SIGNATURES OF ATMOSPHERIC PARAMETERS

Prepared by:	Dr. Pradip M. Bakshi
Academic Rank:	Research Professor
Department and University:	Department of Physics Boston College
Research Location:	Air Force Geophysics Laboratory, Optical Physics Division, Radiation Effects Branch
USAF Research Colleague:	Dr. R. Armstrong
Date:	October 31, 1983
Contract No.:	F4 9620-82-C-0035

INFRARED EARTHLIMB EMISSION LINESHAPES

AS SIGNATURES OF ATMOSPHERIC PARAMETERS

by

Pradip M. Bakshi

ABSTRACT

The relationship of infrared earthlimb emission lineshapes to atmospheric parameters is investigated. An analytical treatment is developed for the calculation of emission lineshapes for isolated lines, including Doppler, collisional and self-absorption effects, in a planetary atmosphere viewed in the limb as a function of tangent height. The line spectral radiance is shown to be represented by a product of two factors, the first a weighted average of the ratio of the upper to lower level populations of the emitter and the second a simple function of the total absorption along the line of sight. The analytical forms, with suitable approximations, (i) make more transparent the relationship between various features of the spectral lineshapes and the underlying atmospheric parameters, and (ii) enable one to infer many important features of the lineshape with considerable saving in computing time as compared to direct numerical integration of the radiation transport equation.

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 the opportunity to participate in a very worthwhile summer program at the Air Force Geophysics Laboratory. He would like to thank Dr. R. Armstrong for the kind hospitality at the laboratory.

The author also wishes to thank Dr. M. Bullit, Dr. R. Picard and Dr. R. Sharma for valuable discussions and collaboration on the main topic of this report. Thanks are also due to Dr. R. Armstrong, Dr. W. Blumberg, Dr. E. Lee and Dr. J. Lurie for very helpful, extensive discussions on the other two topics mentioned in this report.

I. INTRODUCTION

Infrared airglow has been the subject of a number of recent studies¹⁻³. A comprehensive review of this rapidly changing field, with special emphasis on current advances at Air Force Geophysics Laboratory, has been given by Sharma⁴.

The observations of the infrared airglow are made from rockets⁵⁻⁸, or satellites⁹. Of particular interest are the earthlimb observations such as those from the SPIRE rocket^{1,6}. The limb-looking observations offer specific advantages over zenith-looking experiments. The line of sight can be varied by changing the telescope angle, and one can select a day, a twilight or a night line of sight. By increasing the tangent height (the distance of closest approach of the line of sight to the earth's surface), large path lengths can be sampled and the limb experiments can thus detect emissions which might be missed in the zenith experiments. Also, the limb-viewing allows one to select the altitude dependence of the volume probed, and the observed line profiles taken at different tangent heights contain detailed, though indirect, information about the density and temperature variations of the observed species.

The theoretical analysis addresses the problem of predicting the emission lineshape for isolated lines, including the Doppler, collisional and self-absorption effects, and taking into consideration the effects of variations in the atmospheric conditions along the line of sight which spans a wide range of altitudes for earthlimb observations. A computer program had been developed at AFGL¹⁰ to obtain the emission lineshapes as a function of the tangent height for given

atmospheric parameter profiles. The program uses as input atmospheric parameter profiles, given as functions of altitude, and achieves a numerical integration of the radiation transport equation to determine the emission lineshape for various lines of sight characterized by different tangent heights.

Specific calculations were carried out for the CO_2 (01101+00001) vibrational transition for tangent heights ranging through the mesosphere and the lower thermosphere, and using either the Doppler or the Voigt line profile. The resultant radiance profiles displayed the following characteristics: (1) The intensity at the line center was almost constant, independent of the tangent height. (2) They all exhibit a maximum at about three Doppler widths away from the line center. (3) Beyond the maximum, they all fall off rapidly for the Doppler profile, but develop a wing for the Voigt profile. (4) The integrated radiances for individual lines show qualitatively different behaviors for the Doppler and the Voigt profiles. The former leads to almost constant values of the integrated radiance for all except very weak lines. The Voigt profile calculation matches the Doppler result for weaker lines, but shows correlation with the line strength for the stronger lines.

My contribution during the Summer Program was the development of an alternative analytical approach for the calculation of the earthlimb emission lineshapes. These analytical forms already enable us to achieve a qualitative understanding of the salient features of the radiance profiles, and additional refinements and extensions may provide a complete quantitative description, with considerable saving in

computing time as compared to the earlier numerical approach.

The main objectives of the program are described in Section II. The mathematical development is given in Section III. The main analytical approximations are presented in Section IV, and their implications discussed in Section V. Two other areas were also the subject of a preliminary investigation during the Summer Program; these are mentioned in Section VI. Finally, the recommendations for future work are presented in Section VII.

II. OBJECTIVES

The main objective was to develop an analytical formula for the earthlimb spectral radiance which exhibits in a compact form the essential role played by the various atmospheric parameter distributions in the formation of the spectral line profile. Additional goals were to examine the possibility of simplifying the analytical expression for specific spectral domains such as the central region or the wing region and to gain insights about the relationship between the lineshape and the atmospheric parameters. A longer range objective would be to develop, using these analytical relations, an inversion technique for the determination of atmospheric parameter profiles from the observations of spectral radiances of various lines at various tangent heights. One can also expect the analytical approach to offer significant computational advantages.

III. PATH SPECTRAL RADIANCE

The limb radiance I_ν at a frequency interval ν from the line center is given by

$$I_\nu = \frac{A}{4\pi} \int_{-\infty}^{\infty} ds g(\nu, z(s)) n_u(z(s)) e^{-\int_s^{\infty} k_0 g(\nu, z(s')) n_l(z(s')) ds'} \quad (1)$$

where A is the Einstein coefficient for spontaneous emission, k_0 is the integrated absorption coefficient, $n_l(z)$ and $n_u(z)$ are the densities of the lower and upper state at altitude z , and the integrals are performed over s , the distance along the line of sight as measured from the tangent point. The function $g(\nu, z)$ is the normalized lineshape function,

$$\int_{-\infty}^{\infty} g(\nu, z) d\nu = 1. \quad (2)$$

One can choose g to be a Lorentzian, Doppler or Voigt profile.

Introducing the ratio of the upper and lower level densities

$$R(z) = \frac{n_u(z)}{n_l(z)} \quad (3)$$

we can rewrite eq. (1) in the form

$$I_\nu = \frac{A}{4\pi k_0} \int_{-\infty}^{\infty} ds R(z(s)) F(\nu, s), \quad (4)$$

where

$$F(\nu, s) = f(\nu, s) e^{-\int_s^{\infty} ds' f(\nu, s')} \quad (5)$$

and

$$f(\nu, s) = k_0 g(\nu, z(s)) n_l(z(s)). \quad (6)$$

We note that F is always a positive quantity, and is easily integrable,

$$\int_{-\infty}^{\infty} F(\nu, s) ds = 1 - e^{-k(\nu)}, \quad (7)$$

in terms of the total absorption along the line of sight,

$$k(\nu) = k_0 \int_{-\infty}^{\infty} ds g(\nu, z(s)) n_{\ell}(z(s)). \quad (8)$$

We can then express the limb radiance as

$$I_{\nu} = \frac{A}{4\pi k_0} \langle R \rangle (1 - e^{-k(\nu)}), \quad (9)$$

where,

$$\langle R \rangle = \frac{\int_{-\infty}^{\infty} ds R(z(s)) F(\nu, s)}{\int_{-\infty}^{\infty} ds F(\nu, s)} \quad (10)$$

is an averaged R with the positive measure F. Eq. (9) is the central analytical result of this work, which reduces the task of computing I_{ν} to simply evaluating the total absorption $k(\nu)$ and finding the weighted average of the density ratio along the line of sight. Various approximations or simplifications are now feasible, depending on the domain of interest.

IV. ANALYTICAL APPROXIMATIONS

For thick lines, the line center is characterized by $k \gg 1$, and eq. (9) reduces to

$$I_{\nu} = \frac{A}{4\pi k_0} \langle R \rangle. \quad (11)$$

The average $\langle R \rangle$ can be estimated by applying the method of steepest descent to evaluate the numerator in eq. (10),

$$\langle R \rangle \approx R(z(s_0)) = R(z_0), \quad (12)$$

where s_0 , the location along the line of sight where F achieves its peak value, is given by

$$\frac{df}{ds} = -f^2. \quad (13)$$

As we move away from the line center, s_0 moves closer to the tangent point, and z_0 moves towards the tangent height. Thus variation of I_ν near the center reflects the variation of k with altitude.

The wing region is characterized by little or no absorption, $k \ll 1$ and $F \rightarrow f$, so that

$$\langle R \rangle_{\text{wing}} = \frac{\int ds g n_u}{\int ds g n_l}. \quad (14)$$

For the Doppler profile, there are no pronounced wings due to the fast Gaussian decline of $g(\nu)$. For the Voigt profile, when the Lorentzian width $\alpha_L \ll \alpha_D$ the Doppler width, one can show that g is separable,

$$g(\nu) \approx \frac{\alpha_L(z(s)) B(\nu)}{\pi \nu^2}, \quad (15)$$

where

$$B(\nu) = 1 + \frac{3}{2} \frac{\alpha_D^2}{\nu^2} + \frac{15}{4} \frac{\alpha_D^4}{\nu^4} + \dots \quad (16)$$

This separability renders $\langle R \rangle_{\text{wing}}$ in (14) almost independent of ν ,

$$\langle R \rangle_{\text{wing, Voigt}} = \frac{\int ds \alpha_L B n_u}{\int ds \alpha_L B n_l} \approx \frac{\int ds \alpha_L(z(s)) n_u(z(s))}{\int ds \alpha_L(z(s)) n_l(z(s))} \quad (17)$$

where in the last step we ignore the slight dependence of B on $z(s)$ through $\alpha_D(z)$. We also note that k can be expressed as

$$k(\nu) = \frac{k_0}{\pi \nu^2} \int ds B \alpha_L n_l \approx \frac{k_0}{\pi \nu^2} \int ds \alpha_L n_l. \quad (18)$$

These results allow us to provide an explicit result for the spectral radiance in the wing region,

$$I_{\nu}^{\text{Voigt}} = \frac{A}{4\pi^2\nu^2} \int ds \alpha_L n_u + O(\nu^{-4}), \quad (19)$$

where the higher order terms can be developed systematically in an inverse power series in ν .

This brief discussion illustrates only some of the techniques, and some of the results of the analytical approach. More detailed results are to be presented¹¹ at the AGU meeting in December 1983 (the abstract is reproduced in the References). A journal article is also currently under preparation¹².

V. IMPLICATIONS

We now discuss the implications of the analytical results of the preceding two sections. Eq. (9) provides a convenient, exact procedure for calculating the spectral radiance I_{ν} in terms of the atmospheric parameters. k depends on the lower level density profile of the emitter and on the lineshape function g . The latter depends on the pressure profile through the Lorentzian width α_L . The ratio R reflects the vibrational temperature. k and $\langle R \rangle$ can be computed from eqs. (8) and (10), and together they provide the spectral radiance I_{ν} as a function of frequency. Evaluating I_{ν} in this fashion, for each frequency, would entail almost the same amount of computation as a direct numerical integration of the transport equation.

The advantage of the analytical approach arises by noting that the

structure of eq. (9) makes it unnecessary to evaluate k when it is significantly larger than one; the factor containing k then simply drops out in eq. (9). For thick lines, this implies that the region near the line center requires the evaluation of $\langle R \rangle$ alone. At the other end, in the wings, the asymptotic result eq. (17) shows the independence of $\langle R \rangle$ from ν ; $\langle R \rangle_{\text{wing}}$ is easily evaluated by simple quadratures involving only the Lorentzian width, and the upper and lower level density profiles. Even $k(\nu)$ has a very simple form, eq. (18) and the lead term in eq. (19) for I_ν in the wing involves only a simple quadrature involving the Lorentzian width and the density profile of the upper level population of the emitter. These features provide a very significant computational advantage, without any noticeable loss of accuracy. The lineshape at 70km tangent height was evaluated in this fashion from eq. (9), using eq. (12) for approximating $\langle R \rangle$ in the central region, using eq. (17) for $\langle R \rangle$ in the wing region and evaluating k by eq. (18). The resulting line profile was found to be in remarkable agreement with the computer generated exact profile.

At the line center, eq. (9), along with eq. (12) explains the relative independence of the radiance from tangent height. The effective altitude z_0 which determines $\langle R \rangle$, and thus I_0 , is where the function F attains its peak. It can be shown from eq. (13) that z_0 is a rather slowly varying function of the tangent height of the line of sight. Thus $\langle R \rangle$ and I_0 do not vary much with variation in tangent height. This analysis provides an explanation for item (1) noted in

the Introduction. The other features of the line profiles noted in the Introduction can also be explained by extending the analysis given in the previous two sections. The details are to be reported in a journal article currently under preparation¹². We indicate here in section VII some of the analytical extensions and generalizations that should be undertaken to achieve a complete quantitative description of the spectral line profiles with considerable saving in computing time as compared to the earlier numerical approach.

VI. OTHER TOPICS

Investigation of two other topics was also initiated during the Summer Program. (1) An assessment of the current theory of beam plasma discharge and its applicability to the Excede observations and the Labcede experimental results. (2) Examination of the role of the radial expansion of the post-laser-burst plasma in the Linus experiment. An "Initial Charge Separation" model was proposed. However, detailed calculations remain to be carried out.

Both these topics will be pursued further if appropriate resources become available.

VII. RECOMMENDATIONS

The analytical results developed in Sections III and IV provide a convenient and useful representation for the spectral radiance. It would be worthwhile to further pursue this approach along the following lines:

1. The analytical result, eq. (9), with further approximations and refinements, offers the possibility of achieving a quantitative agreement with the computer calculations for the emission lineshape. We would like to complete this program and develop simple approximations in various parameter domains that make this approach sufficiently accurate (to within 5 or 10% of the exact results).

2. The wing region can now be calculated fairly accurately by the methods described in section II with considerable saving in computing time. We propose to refine the calculation of $\langle R \rangle$ in the vicinity of the line center by expanding R inside the integral sign in a Taylor series. This leads to a moments description,

$$\langle R \rangle = R_0 + aR_0' + bR_0'' + \dots$$

in terms $R_0 = R(z_0)$ and derivatives of R with respect to the altitude. Preliminary results indicate that adding the first two moments provides good agreement with the exact results.

3. After developing adequate techniques for an accurate determination of the lineshape for a given tangent height, we propose to develop interpolation formulas based on the smooth variation of functions like z_0 , a and b (in item 2 above) as functions of tangent height. Since the same atmospheric parameter profiles play a role in determining the analytical forms for different tangent heights, it is reasonable to expect a slow and smooth variation in z_0 , a , b , etc. Then it would be possible to study the lineshape at two tangent heights and be able to predict the profiles for all intervening tangent heights by interpolation.

4. We would then like to predict the integrated radiances in terms of

the given atmospheric parameter profiles by using the analytical forms.

5. Finally, one should consider the inversion problem, using the analytical tools developed above. The objective would be to predict the atmospheric parameter profiles from the observed lineshapes at different tangent heights.

REFERENCES

1. A.T. Stair, Jr., R.M. Nadile, J.C. Ulwick, K.D. Baker and D.J. Baker, "Infrared Measurements of Aurora, Airglow and the Upper Atmosphere", AIAA Reprint 81-0421, Aerospace Sciences Meeting, January, 1981.
2. B.F. Gordiets, M.N. Markov and L.A. Shelepin, "IR Radiation of the Upper Atmosphere", Planet. Space Sci., 26, pp. 933-942, (1978).
3. T. Ogawa, "Excitation Processes of Infrared Atmospheric Emissions", Planet. Space Sci. 24, pp. 749-756, (1976).
4. R.D. Sharma, "Infrared Airglow from the Undisturbed Atmosphere", AFGL Handbook of Physics, Chapter II, (1983).
5. A.T. Stair, Jr., J.C. Ulwick, D.J. Baker, C.L. Wyatt and K.D. Baker, Geophys. Res. Letters, 1, 117, (1974).
6. R.M. Nadile, N.B. Wheeler, A.T. Stair, Jr., D.C. Frodsham and L.C. Wyatt, "SPIRE - Spectral Infrared Experiment, SPIE vol. 24, pp. 18-24, (1977).
7. M.N. Markov, Appl. Opt. 8, 887, (1969).
8. D. Offermann and K.V. Grossmann, Geophys. Res. Letters, 5, 387, (1978).
9. M.N. Markov, G.M. Grechko, A.A. Gubanev, Y.S. Ivanov and V.S. Petrov, Cosmic Research 15, 102, (1977).
10. R.D. Sharma, R.D. Saini, M.K. Bullitt and P. Wintersteiner, "A Computer Code for Calculation of IR Radiation Through Non-Equilibrium Atmosphere", AFGL-TR-83(in press).

11. M. Bullit, P. Bakshi, R.M. Picard and R.D. Sharma, "Relationship of Infrared Earthlimb Emission Lineshape to Atmospheric Parameters", Abstract submitted to the AGU Meeting, San Francisco, December 1983.

12. M. Bullit, P. Bakshi, R.H. Picard and R.D. Sharma, Paper under preparation.

Relationship of Infrared Earthlimb Emission Lineshape
To Atmospheric Parameters

M. BULLITT (AFGL, Hanscom AFB, Mass. 01731)

P. BAKSHI* (Boston College, Chestnut Hill, Mass. 02167)

R.H. PICARD (AFGL, Hanscom AFB, Mass. 01731)

R.D. SHARMA (AFGL, Hanscom AFB, Mass. 01731)

We calculate the emission lineshape for isolated lines, including self absorption, in a planetary atmosphere viewed in the limb as a function of tangent height. A general numerical and analytical treatment is developed and is illustrated through the case of the CO_2 ν_2 ($01^1_0 \rightarrow 00^0_0$) vibrational transition in the earthlimb for tangent heights ranging through the mesosphere and the lower thermosphere. The lineshape can be altered dramatically from that of an optically thin Voigt line, and for some tangent heights the relative importance of the Doppler center portion of the line and of the Lorentzian line wings can be reversed, so that nearly the entire measured line intensity is due to the Lorentzian wings. We examine the dependence of the line and band profiles on the characteristics of the radiative transfer process and on atmospheric parameters, such as the translational and vibrational temperatures and the emitter density. The analytical expression for the line spectral radiance has the form $I_\nu \sim \langle R \rangle (1 - \exp[-K(\nu)])$, where $\langle R \rangle$ is a weighted average of the ratio of the upper to lower level populations of the emitter and $K(\nu)$ is the total absorption along the line-of-sight. In the wings, the variation is primarily due to $K(\nu)$ while in the central region it is primarily governed by $\langle R \rangle$. The analytical forms enable one to calculate all the important features of the self-absorbed line and agree remarkably well with the more time-intensive numerical integration of the radiation transport equation.

*Supported by AFOSR/SCEEE Summer Faculty Research Program.

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 EVALUATION OF A CANTILEVER BEAM SOLID STATE ACCELEROMETER

Prepared by: Daniel W. Barr, Ph.D
Academic Rank: Assistant Professor
Department and University: Department of Electrical Engineering
Virginia Military Institute
Research Location: Air Force Armament Laboratory
Eglin Air Force Base, Florida
USAF Research: Mr Richard Mabry
Date: 20 July 1983
Contract No: F49620-82-C-0035

AN EVALUATION OF A CANTILEVER BEAM SOLID STATE ACCELEROMETER

by

Daniel W. Barr

ABSTRACT

Solid state accelerometers produced in a batch process are characterized for potential application in Safe and Arming and Fuzing systems. The devices sense positive and negative accelerations by monitoring stress in an etched silicon beam which bends with G loading. The stress is detected by resistors diffused in the beam whose values change with stress due to the piezoresistive effect. The device uses a bridge configuration of resistors placed to enhance sensitivity. Eight devices representative of the process are tested and reveal the acceleration detection sensitivity is linear and averages 8.4×10^{-5} volts/G for a 10 volt bridge bias between -100 and +100 G. Acceleration sensitivity shows degradation with time. Transverse sensitivity is directional and does not exceed 8 percent of major axis sensitivity. The devices show offset voltages due to slight bridge imbalance which are temperature dependent. Mechanically the beams behave as second order systems subject to a resonant mode which is excited by impulsive loading. Beam resonances range from 1180 to 5319 hertz and show damping ratios from 3.4×10^{-3} to 1.57×10^{-2} . The accelerometer in its present form is not suited for application in safe and arming and fuzing systems due to low sensitivity, temperature dependent offset, susceptibility to resonance induced by mechanical shock and time dependent sensitivity drift. Proposed design changes which enhance the applicability of the device include the use of Ion Implantation and a fully active bridge structure. Time dependent sensitivity drift and variance in the temperature coefficients of the resistors in the bridge remain the primary limitations of the cantilever beam solid state accelerometer.

ACKNOWLEDGMENT

I wish to thank Mr Richard Mabry who suggested this topic and Mr Roy Stables who served as the effort focal point on the project. This research was conducted in the Fuzes and Sensors Branch (DLJF) of the Air Force Armament Laboratory (AFATL), Eglin Air Force Base, Florida. Assistance on the part of the members of DLJF commanded by Lt Col Hubert L. Redmon made this effort possible. Special thanks go to Mrs Helen Noiseux for typing this document on a tight time schedule. This research was funded through the Air Force Systems Command (AFSC), Air Force Office of Scientific Research (AFOSR) on Contract F49620-82-C-0035.

I. INTRODUCTION: On 18 January 1983 the Air Force Armament Laboratory at Eglin Air Force Base awarded a contract to Insouth Microelectronics, Inc. of Auburn, Alabama, to produce 100 solid state accelerometers. The accelerometers were ultimately intended for use in safe and arming and fuzing systems.

Several accelerometers representative of the technological approach were available for testing. These devices were produced as an offshoot of an Army solid state accelerometer contract. Due to my background in solid state device fabrication and testing work at Harry Diamond Laboratories on electronic fuzing packages I was given the assignment of characterizing the accelerometers for potential use in safe and arming and fuzing systems.

During the 10-week period a trip was made to Insouth where final approaches to Insouth's accelerometer design were reviewed. On the basis of this information comments on each approach were made in this report.

The ultimate objective of my assignment was to provide third party feedback on two topics. Were beam type accelerometers acceptable telemetry devices in the safe and arming and fuzing environment, and was Insouth's design approach sound?

Section III summarizes the visit to Insouth which provides a history of the accelerometer design and a review of the four new design approaches. After this section the results of the characterization are reported for device design SAC 5.

II. OBJECTIVES: The objective of this research effort is the evaluation of solid state accelerometers for potential application in safe and arming and fuzing systems. The areas examined are:

1. Acceleration Detection Sensitivity.
2. Temperature Effects.
3. Mechanical Impulse Response.
4. Effects of Mechanical Stress
5. Design Considerations.

III. INSOUTH DESIGN AND FABRICATION: In June 1983 a trip was taken to Insouth Microelectronics, Inc., in Auburn, Alabama, for the purpose of discussing the solid state accelerometer design and fabrication procedure.

APPROACH

The approach to acceleration detection in the silicon based devices produced by Insouth remains using diffused piezoresistive sensors in a bridge configuration to sense stress in a preferentially etched beam. A top view of the accelerometer is shown in Fig. 1. The cantilever beam is outlined by a dark U shaped border formed by an etch from the top to bottom side of the wafer that defines the length and width of the beam. The beam is attached to the silicon chip on the left where four resistors are diffused. The two resistors on the beam sense stress as the beam bends and the remaining two complete a bridge circuit. The resistors on the beam alternate position around the bridge with those off the beam to double the output caused by $\Delta R/R$ changes in the stressed resistors on the beam.

To date Insouth has produced 6 designs (SAC1-SAC6) with sensitivities ranging from 1×10^{-5} to 60,000 G's. Early devices revealed three potential problem areas. First, uniformity of the beam thickness across a given wafer was a function of thickness variations in the original silicon wafer. Next, slight variation in resistors composing the bridge caused a device offset voltage which may hamper direct coupled operations. Finally, the resistors which compose the bridge had different temperature coefficients. That caused the output voltage of the accelerometer to change with temperature.

BEAM THICKNESS

Beam thickness is an important processing parameter for its effects on device sensitivity and the natural resonant frequency of the accelerometer. Thin beams maximize sensitivity but lower resonant frequency. As the beam becomes thicker, sensitivity decreases as resonant frequency increases. Application dictates optimum beam thickness which is controlled by etch duration during processing. Uniformity of beam thickness across a given wafer is a function of front to rear side parallelism. Wafer surface variation of 25 microns can affect accelerometer sensitivity and resonant frequency by a factor of 5 across the wafer.

OFFSET VOLTAGE AND TEMPERATURE EFFECTS

Offset voltage and temperature effects in the accelerometer output are dependent on the conduction mechanisms of the resistors composing the bridge circuit. The ideal bridge is composed of resistors which are equal in value over the usable temperature range. Designs SAC 1 - SAC 6 all used resistors emplaced by a gaseous diffusion process. This process produces devices showing offsets which exceed output at 100 G's by as much as 100 percent as seen in Table 1. This is primarily due to lack of precise control of doping profiles across the wafer. Resistive variation with temperature is not well understood and involves an interplay of such factors as carrier mobility and concentration, lattice uniformity, doping profile, and crystal defect density. All of these factors are interdependent in at best a nonlinear fashion.¹

Carrier concentration and doping profiles can be more closely controlled using Ion Implantation rather than gaseous diffusion. This control is particularly pronounced where sheet resistivities in excess of 500 ohms/square are desired with assurance of repeatability. Design data supplied by Insouth indicates that the ratio of temperature coefficient to piezoresistive sensitivity is optimum at a surface concentration of approximately 10^{19} atoms/cc. This high level of implanted ions and the associated defect damage may influence carrier mobility and piezoresistive strain coefficient to the point that optimum surface concentration will require empirical evaluation.² If concentrations of 10^{19} atoms/cc do optimize the temperature-sensitivity ratio this level of doping will decrease sheet resistivity by a factor of approximately 5 over the current design. As sheet resistivity decreases, the bridge resistance will fall requiring increased bias potential to raise the output signal level of the accelerometer. With increased bias potential and lower device impedance, power requirements will increase to the point that V^2/R self-heating may become a new problem to contend with.

SAC 7 - SAC 10 DESIGN

In order to improve performance Insouth has designed four new accelerometers:

1. Half Active, Full Bridge, with Trimming Resistors - Single Beam Support.
2. Full Active, Full Bridge - Double Beam Support.

3. Full Active, Full Bridge, Double Beam Support with Stress Concentrators.

4. Half Active, Full Bridge with Trimming Resistors - Single Beam Support with Stress Concentrators.

These designs incorporated four new concepts; the full active bridge, double beam support, trimming resistors and stress concentrators. The projected performance benefits and potential processing problems associated with each change follow.

FULL ACTIVE BRIDGE

The present design places two of the bridge resistors on the beam to sense longitudinal stress and two off the beam which serve as a reference. The full active bridge will utilize the transverse piezoresistive coefficient of the beam by placing the resistors which were off the beam onto the beam positioned to sense transverse stress. The piezoresistive coefficient for the longitudinal orientation is opposite in magnitude to the transverse direction, thus this approach should increase the accelerometer output for a given bias. This technique should not introduce any processing problems for diffused and implanted resistors have successfully been placed on the beam's surface. The sensitivity of the accelerometer to torsional loads will increase with the addition of the transverse strain sensors.

DOUBLE BEAM SUPPORT

A double supported beam is a more stable configuration than a beam with a single support. Doubling the support to the beam will increase the natural resonant frequency of the device and lessen its sensitivity. Sensitivity may be increased by loading the beam center with a small mass but would lower the resonant frequency. Widening the beam at the center of its length to accommodate loading will increase the devices's sensitivity to torsional forces. This device is a more rugged device than the single support beam. A limited range of excursion obviates the mechanical stops which prevent beam destruction in the single support arrangement.

TRIMMING RESISTORS

Offset voltage in the designs SAC 1 - SAC 6 arises due to slight differences in the values of the resistors composing the bridge. Trimming resistors added in series to each leg of the bridge will allow compensation of slight bridge imbalance. This technique is accomplished by placing trimming resistors which are shorted with a fuze link in the bridge. The trimming resistance is added by blowing the fuze links with the energy supplied by laser light. Conceptually, this idea is sound. In order to obtain precise nulls a wider range of values of trimming resistors may be necessary and the nulling process will be time consuming.

STRESS CONCENTRATORS

In an effort to increase the sensitivity of the accelerometers, small V grooves are etched in the resistors on the beam which localize stress in the implanted resistor. This localized stress should cause an increased change in resistance when compared to resistors without the V grooves.

V grooves etched in silicon have a depth equal to 70 percent of the surface width. Ion implantation produces annealed junction depths of about a micron. Thus, the photolithographic process must produce window widths less than 1.4 microns or the implanted resistors will be physically severed by the V grooves. Photolithographic masking and etching processes are not well controlled below 1.5 microns. Small variations in depth of the V groove will produce great changes in the bridge leg resistance as the semiconductor junction is approached. This process is likely to introduce offset voltage problems that trimming resistors could not compensate. The V groove process would be better applied to resistors formed by diffusion techniques where junction depths can approach two microns.

IV. DEVICE CHARACTERIZATION (SAC 5):

ACCELERATION SENSING

The device output response to major axis acceleration was tested by placing the devices in a centrifuge. The accelerometers were mounted with the major axis in line with centrifugal acceleration and the G loads

varied by changing the angular velocity of the rotating arm. The bridge was biased at 10 volts and the output signal was amplified 100 times by circuitry mounted on the rotating arm next to the accelerometers. A digital multimeter indicated the amplified output outside of the centrifuge.

The accelerometers detect both positive and negative G loads as seen in Fig. 2 where output is plotted against load. The major axis sensitivities of several devices to static G loads is shown in Table 1. These sensitivities are between 60 and 90 percent lower than those reported by Insouth using a shake table. This may indicate an aging effect in the accelerometers in that the tests were made over two years apart. Device 050782-64 was manufactured 6 months after the other accelerometers. Its sensitivity is nearly twice that of accelerometers with similar beam resonances.

An accelerometer was also tested for major axis sensitivity on a shake table. The G load was changed by setting the displacement and varying the frequency of vibration. The G load was monitored using a commercial accelerometer.

Figure 3 shows the accelerometer output spectrum for frequencies from 200 to 2000 hertz for 20, 50, and 100 G loads. This range of frequencies is seen in the fuzing environment and the data indicates that the output for the devices as packaged is nonlinear with respect to frequency well below resonance; 3304 hertz in this case. The sharp peaks in the data may be caused by resonances occurring in the package and not in the accelerometer. This may also account for differences in the shake table and centrifuge sensitivity results. No unpackaged devices were available for testing.

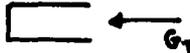
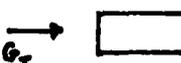
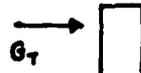
At 450 hertz the device sensitivity was calculated to be 5.5×10^{-5} Volts/G. This is 73 percent less than Insouth's shake table result.

ACCELEROMETER TRANSVERSE SENSITIVITY

The accelerometers are most sensitive to acceleration sensed along the major axis which is perpendicular to the broad faces of the beam. The beam sensitivity perpendicular to the major axis is defined as the minor axis or transverse sensitivity. Minor axis sensitivity can hamper fuzing applications where the environment contains off-axis accelerations caused by spin, vibration, and concussion. Due to the beams asymmetry four directions of minor axis sensitivity exist; two are perpendicular to length and two are perpendicular to the width of the beam. In order to compare directional sensitivity, a transverse sensitivity coefficient is defined as the ratio of the minor axis sensitivity divided by the major axis sensitivity.

To evaluate the transverse sensitivity of the accelerometer it was mounted in the centrifuge so the normal component of the angular acceleration acted along the minor axis. Tangential acceleration effects were minimized by stabilizing the centrifuge's angular velocity as data points were taken and mounting the devices with the major axis perpendicular to the direction of rotation. In this position, the major

axis was sensing 1 G. The devices were placed in each of the 4 positions of the minor axis and data points were taken in each of the positions from 0 to 100 G's. Table 2 summarizes the results of these measurements.

Device Position	Transverse Sensitivity (Volts/G)	Transverse Sensitivity Coefficient (Volts/G)
(Major Axis Sensitivity - 1.14×10^{-4})		
	9.2×10^{-6}	8.1×10^{-2}
	9.2×10^{-6}	8.1×10^{-2}
	3.92×10^{-7}	3.43×10^{-3}
	3.92×10^{-7}	3.43×10^{-3}

ACCELEROMETER TRANSVERSE SENSITIVITY (DEVICE # 12098100)

TABLE 2

This data indicates the greatest transverse sensitivity occurs with loading perpendicular to the width of the beam and represents 8 percent of the major axis sensitivity. Off axis acceleration should be oriented perpendicular to the length of the beam to minimize transverse effects.

EFFECTS OF TEMPERATURE ON OFFSET VOLTAGE

The semiconductor resistors composing the bridge are not linear with respect to temperature as seen in Fig. 4. Here the resistance of a single element of the bridge is measured using a digital multimeter as temperature is varied. The parabolic shape is primarily due to changes in the mobility of the carriers with temperature.³ The offset voltage of the bridge biased at 10 volts appears nearly linear with temperature as seen in Fig. 5. It is felt that this near linear behavior is due to a compensating effect of individual temperature coefficients of the bridge resistors which masks the expected parabolic shape. This change in offset over the range of temperatures plotted represents 2.5 times the acceleration signal at 100 G's. This bridge (050782-64) contained no temperature compensation circuitry.

In an effort to improve temperature effects, Insouth placed a thin film resistor in parallel with a resistor in the bridge and another in series with the bridge. Due to the difference in the resistive temperature coefficients of the diffused and thin film resistors a temperature compensation of sorts is effected. Total offset over the -40 to +40°C temperature range is still comparable to the magnitude of the acceleration signal at 100 G's in these devices. The offset with respect to temperature in the compensated device is shown in Fig. 6.

ACCELEROMETER RESPONSE TO MECHANICAL IMPULSES

Accelerometers used in fuzing applications are prone to mechanical shock. Solid state accelerometers using piezoresistive sensors on a beam are particularly sensitive to mechanical impulses which excite the beam to oscillate at its resonant frequency. Mechanical shock transmitted to the beam can cause the detected signal to be obscured by the device output during the resonant transient.

In order to test the effects of mechanical impulses on the devices they were tapped with a plastic rod and the amplified output signal was recorded on a digital storage scope. Figure 7 shows the output of an accelerometer amplified 100 times immediately following an impulse to the top of the package. This signal is of the form:

$$V_{out}(t) = V_{max} e^{-\alpha t} \sin(\omega t + \beta)$$

where $V_{out}(t)$ = Device output voltage wrt time

Bias voltage = + 10 volts

V_{max} = Maximum induced voltage

α = Damping coefficient = damping ration X natural radian frequency

ω = Radian frequency of damped beam resonance

β = phase constant

This equation represents the response of an underdamped second order dynamic system. The damping ratios (actual damping/critical damping) of the accelerometers ranged from 1.93×10^{-3} to 1.57×10^{-2} . Damping ratios in this range alter the damped frequency of oscillation negligibly compared to the natural resonant frequency of the beam. Thus, the frequency of the signal caused by tapping is an accurate indication of the natural resonant frequency of the device. Figure 8 is an expanded segment of Fig. 7 from 6.2 to 26 milliseconds. The period of the wave is 0.617 milliseconds indicating a resonant beam frequency of 1620 hertz.

In the equation V_{max} is proportional to the maximum beam displacement caused by the impulse. Tapping the package with the end of a 6 gram plastic rod moving at approximately 20cm/sec produced peak voltages ranging from 11 to 17 millivolts using 10 volts of bias. Given the detection sensitivities this represents a peak load of 100 - 500 G's.

The damping ratios and resonant frequencies of the accelerometers are shown in Table 1. The transient signal is attenuated by 99 percent in the device with the smallest damping coefficient in 164 milliseconds. Resonant frequencies range from 1180 to in excess of 5000 hertz. The phase constant β is used in the equation to synchronize the mechanical impulse occurrence to the initial maximum induced voltage in the accelerometer.

EFFECTS OF MECHANICAL STRESS ON ACCELEROMETER OUTPUT

The accelerometer output signal is effected by mechanical stresses which cause voltage levels that are superimposed on the acceleration induced output of the device. These stresses are a function of device mounting configuration, package rigidity and abrupt temperature cycling. All of these factors influence potential device application and dictate operating environment.

In the routine handling of the accelerometers, it was noted that finger pressure applied perpendicular to the broad faces of the accelerometer package caused the output voltage of the device to change as much as 100 percent. To investigate this phenomenon the devices were placed on a triple-beam balance with a stationary pointed rod positioned at the geometric center and perpendicular to the broad face of the accelerometer package. The devices were loaded by balancing the beam with the devices in place then moving the weights on the beam to the desired load. The amplified output of the device was monitored using a digital multimeter.

In this load test the side of the devices labeled with the MEC logo was referenced as the top and the opposite side as the bottom. This labeling was consistent with the position of the output wires on all devices.

The test revealed the device output is relatively constant for loads applied to the top side of the devices. The output variation was less than 3 percent for loads to 500 grams. When the load was applied to the bottom side the output was a linear function of the load from 0 to 500

grams as seen in Fig. 9. Sensitivities were between .01mV/gram and .028mV/gram. This data is presented in Table 1. The data is consistent with the device mounting on the bottom side of the package where stresses developed in the package are coupled to the accelerometer which acts like a load cell.

In the process of studying the sensitivity of the offset voltage to temperature the devices were abruptly cycled from room temperature (23.5°C) to an ice bath (0°C). During this abrupt temperature change it was noted that immediately following the immersion of the devices in the ice bath that the offset voltage first increased for a few seconds and then decreased as expected. This initial increase is attributed to stresses in the accelerometer caused by it coupling to the ceramic substrate and metal package. Figure 10 shows this phenomenon as the device output voltage is plotted as a function of time following an abrupt immersion in an ice bath from a room temperature ambient.

V. RECOMMENDATIONS: The solid state accelerometer in its present form (SAC 5) is not well suited for low G sensing applications in safe and arming or fuzing systems requiring a direct coupled output. The device operation is hampered by low sensitivity, temperature dependent offset, susceptibility to resonance induced by mechanical shock and time dependent sensitivity drift. Temperature dependence of the offset may be improved by electronic support circuitry; although due to the complex variation of resistance with temperature in the bridge elements this circuitry would require calibration for each accelerometer. Of the new designs the full active bridge without stress concentrators holds promise of increased sensitivity although temperature dependent offset will remain a problem.

The new designs which use ion implanted resistors will be a valuable tool in evaluating the piezoresistivity of implanted semiconductor layers. Devices with increased sensitivity and resonances above 2000 hertz could be used in an AC coupled mode for applications where sensing durations are less than several seconds.

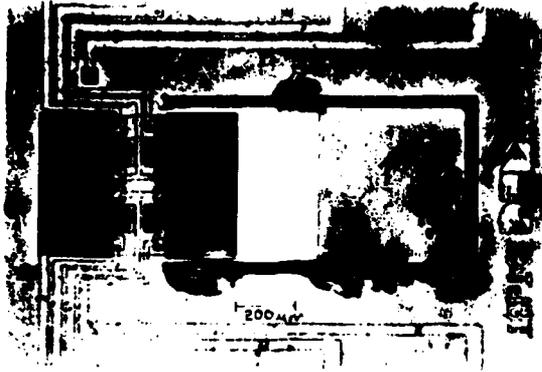
Time dependent sensitivity drift and variance in the temperature of the resistors in the bridge remain the primary limitations of the cantilever beam solid state accelerometer. Further research in these areas is advised.

Device Number	Resonant Frequency (Hertz)	Acceleration Detection Sensitivity (Centrifuge) (Volts/G)	Insouth's Acceleration Detection Sensitivity (Shake Table) (Volt/G)	Package Stress Sensitivity (Volts/Gram)	Beam Damping Ratio (Dimensionless)	Offset Voltage (Volts)	Comments
12098100	1180.	1.13×10^{-4}	3.84×10^{-4}		6.07×10^{-3}	4.79×10^{-3}	
050782-64	1235.	2.15×10^{-4}			4.21×10^{-3}	6.51×10^{-2}	No temperature compensation
12098101	1620.	1.26×10^{-4}	3.0×10^{-4}		3.24×10^{-3}	6.12×10^{-3}	
12188116	1919.	5.55×10^{-5}	2.0×10^{-4}	2.22×10^{-5}		0.65×10^{-3}	Lead breaks flush to package during shake table test
11098128	3304.	3.02×10^{-5}	2.0×10^{-4}	2.78×10^{-5}	1.93×10^{-3}	4.55×10^{-3}	Package destroyed on removal from shake table
11098109	3740.	3.25×10^{-5}	1.5×10^{-4}	1.05×10^{-5}	1.57×10^{-2}	1.20×10^{-2}	
11098117	5319.	1.98×10^{-5}	1.94×10^{-4}		3.78×10^{-3}	1.00×10^{-2}	

*Shake Table Test at Eglin AFB showed 5.5×10^{-5} on 5/6/82
 All measurements taken at 22 - 24 °C

ACCELEROMETER CHARACTERISTICS

TABLE 1



Photograph of Accelerometer
Figure 1

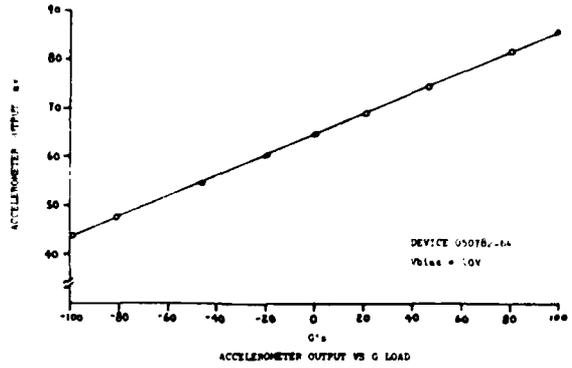


Figure 2

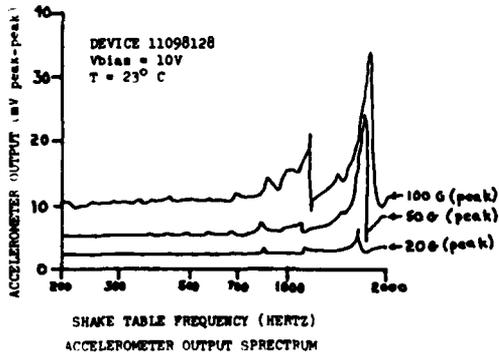


Figure 3

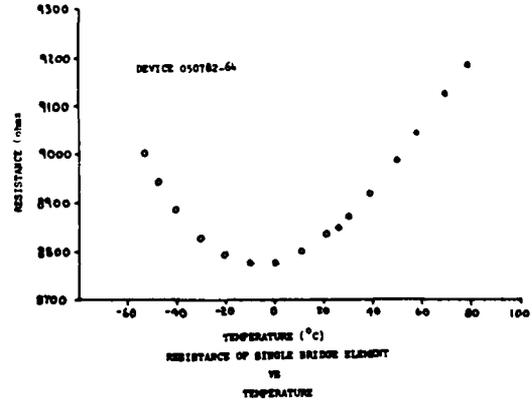


Figure 4

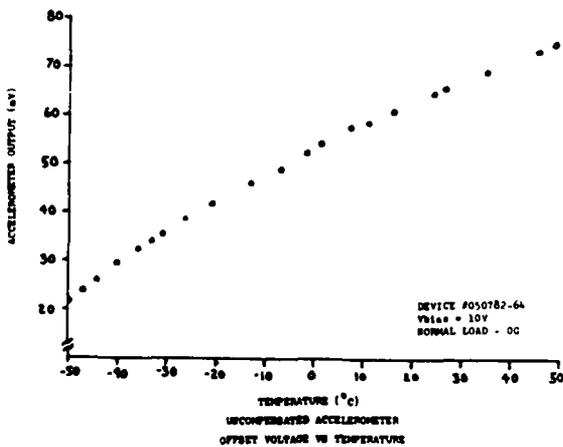


Figure 5

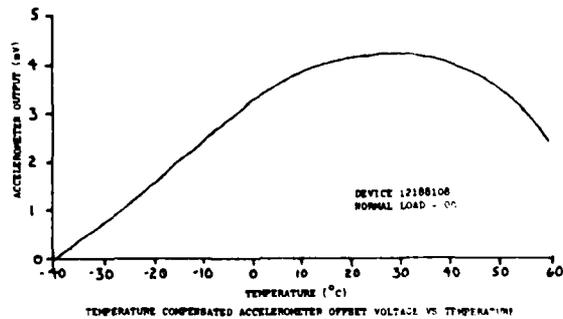
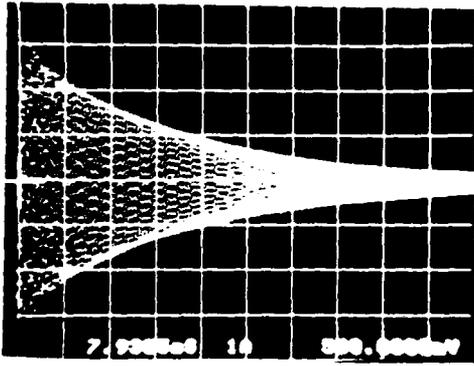
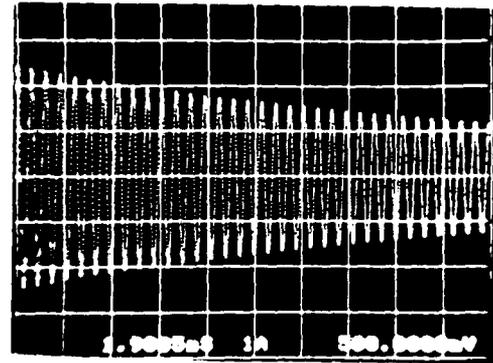


Figure 6



VERT - ACC. OUTPUT (X100), 500mV/div
 HORIZ - TIME (sec), 7.94 msec/div



VERT - ACC. OUTPUT (X100), 500 mv/div
 HORIZ - TIME (sec), 1.98 msec/div

DEVICE 12098101

ACCELEROMETER MECHANICAL TRANSIENT RESPONSE

Figure 7

Figure 8

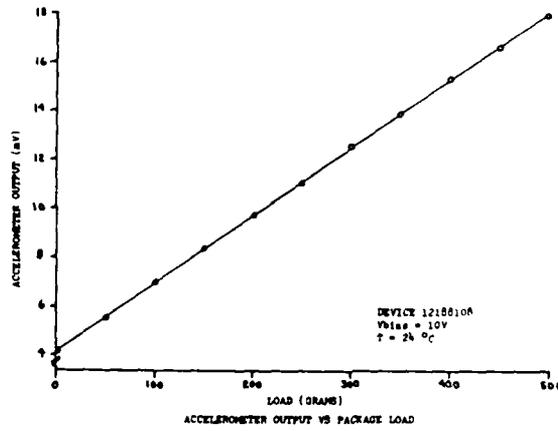


Figure 9

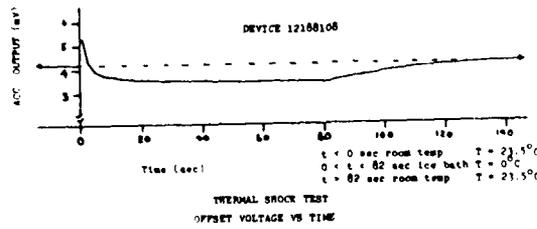


Figure 10

REFERENCES

1. Camenzind, Hans, Circuit Design for Integrated Electronics, Reading, Massachusetts, Addison - Wesley Publishing Co., 1968.
2. Beanland, D. G., "Damping Effects Associated with High Dose Implantation of Silicon," IREECON International Convention Digest, Institute of Radio and Electronic Engineers, Sydney, Australia, 1979, pp 480 - 491.
3. Camenzind, op.cit., p 52.

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 GAS HEATING PHASE IN ELECTRICAL BREAKDOWN

Prepared by: Dr. Ernesto Barreto

Academic Rank: Senior Research Associate

Department and University: Atmospheric Sciences Research Center
State University of New York at Albany

Research Location: Flight Dynamics Laboratory, Vehicle Equipment
Division, Flight Vehicle Protection Branch,
Wright-Patterson AFB, Ohio

USAF Research: Capt. Pedro L. Rustan

Date: November 23, 1983

Contract No.: F49620-82-C-0035

The Gas Heating Phase in Electrical Breakdown^{*}

Ernesto Barreto
Atmospheric Sciences Research Center
State University of New York at Albany
Albany, NY 12222

Abstract

This report focuses attention on the neglected gas heating stage associated with a spark. A brief review is followed by summaries of recent experimental and theoretical investigations. It is shown that in all cases gas heating is preceded by the formation of a weakly ionized plasma. The transformation to a strongly ionized gas requires additional ionization that can be associated with several different physical processes: In small gaps (≤ 3.0 mm) with metal electrodes, the influx of electrons from a cathode spot with at least a single cell is responsible for gas heating. In longer gaps (1 to 3 cm) highly luminous ionization waves are produced. These can be reflected at the electrodes and increase the degree of ionization. They are well known but not clearly explained in longer discharges. This stage is followed by the formation of hot filaments either at the electrodes or in mid gap. These seem to be associated with the onset of effective ion-electron interactions. In large surface discharges (~ 1.0 m) the glow and hot channel behind it constitute a stable propagating unit that seems to be controlled by three body electron-ion recombination. The properties of these gliding discharges are similar to those of lightning leaders. In all cases there is strong evidence to support the concept of fluid dynamic effects associated with the hot electron population in the weakly ionized gas. Thus the propagation of stable discontinuities, either strong (shocks) or weak, is shown to be possible.

^{*}Work supported by the Office of Naval Research, the Wright Patterson Air Force Aeronautical Laboratories and the Air Force Office of Scientific Research.

I. Introduction

Traditional practical problems associated with the formation of sparks have been ignitions in coal mines and the damage produced by lightning. However, there has been recently an increase in the need to know the behavior and to improve our understanding of high pressure gas discharges as well as those through solids and liquids. The reasons are associated with a progressively increasing use of insulating plastic materials, the favorable economics of high voltage transmission lines, costly accidents on ships,¹ aircraft, spacecraft² and power lasers,³ the high efficiency and need for electrostatic precipitators, the increasing sensitivity of micro-electronic circuits to electromagnetic noise and, far more costly, the actual destruction of computer components by the ability of even small sparks to concentrate gigantic energy densities and destroy small regions in a microcircuit.⁴

The number of papers, books and conference proceedings regarding breakdown has, of course, also increased at a fast rate. This is not only because of the need to know but also because recording equipment with better than nanosecond time resolution and fast computers capable of handling complicated equations have become common place. Of necessity most of the work deals with specific problems for which a solution is desired: how to stop ignitions in a super tanker, punctures in a charged dielectric, the suppression of large voltage surges induced by lightning on power lines, the loss of satellites, etc. Each particular problem is interesting and contributes a little to the basic understanding of small and large sparks. Nevertheless, what has become clearly evident is our meager knowledge of basic processes in the latter stages of electrical breakdown. In particular, our lack of understanding of the change from weakly to strongly ionized gases that is usually accompanied by the observation of fast luminous waves and leads to gas heating and the collapse of high voltage differences. This report will first review some basic problems and then indicate how our work and that at several other laboratories contribute to the complicated experiment on gliding discharges at the Office National d'Etudes et de Recherches Aerospatiales (ONERA).

II. Some Basic Problems

From the start the concept of an electron avalanche was known to be deficient to explain a discharge that does not require externally produced electrons and is therefore self-sustained. Thus, Townsend himself introduced the need for an additional ionization process not directly related to the production of new electrons by their collision with neutral atoms or molecules. He thought that positive ions would ionize neutral molecules and introduced his secondary coefficient for ionization. This coefficient has remained in the literature but its physical significance has been shifted to more realistic ionization associated with ionized, excited or metastable molecules interacting and liberating electrons at a surface or in the gas. The current of a self-sustained discharge, I , can be expressed in terms of the distance it travels, x , and or primary α , and secondary γ , Townsend coefficients.

$$I = I_0 e^{\alpha x} [1 - \gamma(e^{\alpha x} - 1)]^{-1} = I_0 e^{\alpha x} [1 - \gamma e^{\alpha x}]^{-1} .$$

The denominator of this equation incorporates the factor $(1 - \gamma e^{\alpha x})$. Thus when $\gamma e^{\alpha x} \rightarrow 1$ the current becomes mathematically infinite. Townsend used this condition to define electrical breakdown. Experimentally it is associated with the upcurving in a straight line plot of the logarithm of the current discharge plotted vs distance. In air at atmospheric pressure and in a uniform field,⁵ the current at this stage is of the order of 10^{-10} A, very much smaller than currents of the order of 10 A at which gas heating actually begins and the discharge starts to change into a stable arc. Thus if gas heating and Townsend's breakdown are taken to be equivalent the same physical process is extrapolated by eleven orders of magnitude.

At low pressure the upcurving of the curve is associated with the onset of a steady cold glow discharge, and the condition $\gamma e^{\alpha x} \rightarrow 1$ merely indicates the start of the transition to a different steady discharge mode. Likewise at high pressures and low overvoltages the same condition seems to indicate that there is a transient glow that latter changes into an arc. However, the beauty of avalanches and the simplicity of the mathematical argument have carried Townsend's definition of breakdown into modern reviews.⁶ If by dielectric breakdown we understand the actual

collapse of the applied voltage and the transformation to a hot gaseous conductor across the electrodes the concept is not correct. This was evident from the start because avalanches and their associated secondary ionization were explicitly excluded from an explanation of the glow-to-arc transition.⁷ Nevertheless, glows were and are usually considered to be electrical breakdown.

Using cloud chambers it was soon realized that when avalanches reach a critical amplification (10^8 in N_2 at 300 Torr) the space charge field produced by a large difference in velocity between electrons and ions is more effective than that due to the applied voltage for the purpose of propagating ionization into the neutral gas. Thus, streamers develop that travel faster than the drift velocity of the electrons. Whether or not these streamers should be considered to be conducting or dielectric has been a source of controversy which until recently had little direct experimental support. For instance, one of the best established models (Dawson - Winn)⁸ assumes a significant separation of charge at the head of the streamer while actual recent measurements⁹ indicate that the medium is a weakly ionized gas and consequently the region of charge separation is limited to distances of the order of the Debye length. Initially it was believed that in gas mixtures impurities (e.g. O_2) could be ionized by radiation from the decay of excited molecules. Thus, strategically located seed electrons could be placed just ahead of the high field region at the tip of the propagating streamer. At the present moment this view has been challenged by a multitude of competitive processes with the result that there is no uncontested mechanism responsible for the seed electrons required for the propagation of positive streamers.¹⁰

The existence of negative (anode directed) streamers in corona discharges has never been clearly demonstrated. By contrast and, somewhat as a surprise, in uniform fields of the order of 2.0 cm in length it has been established that both positive and negative streamers are present, and that the negative streamer travels faster.¹¹ This is probably because it goes back into a region that is already ionized by the critical avalanche that produced the space charge which launched the streamer. This fact seems to support the concept that streamers should be considered as conductors. Accordingly, it should be noted that if a negative streamer reaches the

anode when the positive is still traveling, the latter accelerates.¹² This indicates an interaction between the cathode and the tip of the positive streamer at a speed very high compared to the streamer velocity ($10^4 - 10^6$ m/sec).

Regardless of small differences in interpretation, pioneering work by Wagner¹² and Tholl¹¹ in Germany using hydrogen in a uniform field and by Marode¹³ in France using positive point-to-plane geometry in air, has proven that, contrary to what was expected, the crossing of a gap by streamers does not immediately involve the formation of a hot channel. Instead as previously noted, a transient glow stage sets in. For a given configuration the duration of the glow stage decreases as the overvoltages increases and as we will see it may completely disappear. In molecular gases (N_2 , O_2 , H_2) the electron temperature is of the order of 1-4 eV limited by the onset of vibrational excitation of the molecules. The number density of electrons in a streamer is between $10^{12} - 10^{15}$ cm^{-3} . Consequently, the total energy stored in all electrons is a small fraction of the thermal energy of the neutrals and the gas remains cold. This is strikingly demonstrated by the fact that pure corona discharges do not ignite combustible hydrocarbon mixtures.¹⁴ In order to heat the neutrals, the number density of electrons has to increase to values about 10^{17} cm^{-3} . At this number the collision frequency between electrons becomes just as large as that between electrons and heavy particles which include a small (10^{-2}) fraction of ions. However, these have very large Coulomb cross-sections and a real thermodynamic equilibrium with identical particle temperatures is produced in times of the order of 5-10 nsec. The manner in which the electron density increases and whether or not the phenomenon is controlled by the electrodes is not always clear and will be discussed more in detail in this report. What is certain is that the weakly ionized plasma collapses, and that, if sufficient energy is available, the hot plasma produced increases its degree of ionization by at least a factor of ten to values as high as 10^{18} to 10^{19} electrons/ cm^{-3} in air. Metal from the cathode evaporates which may provide atoms with a lower ionization potential. Thus the discharge changes to a hot arc whose potential difference is only of the order of the ionization potential of the vapor atoms and whose current is limited by the power supply used.¹⁵ The

discharge sooner or later interacts with electrodes particularly the cathode and becomes affected by metallic properties.

In long positive point-to-plane sparks in air, leaders are produced that exhibit a bluish glowing fan traveling ahead of a bright white region that emits continuum radiation. By contrast, the radiation from the fan is made exclusively of discrete molecular lines thus indicating a low degree of ionization and a high electric field. This radiation is the same as that observed in a streamer or the positive column of a glow discharge. Thus it is assumed that they are all the same physical process, namely ionization by collision of electrons accelerated in an electric field. Clearly the transformation from a weakly to a strongly ionized gas in a leader takes place without the direct influence of nearby metal electrodes. Leaders have been analyzed in terms of two basic models. Kekez and Savic¹⁶ consider the head of the leader to be a very hot small region in complete local thermodynamic equilibrium. A shock wave is produced that overtakes the incoming high energy electrons in the glowing fan and the leader head regenerates itself by effective Coulomb interactions between the incoming electrons and the hot plasma. The other theory by Gallimberti¹⁷ considers that many negative ions are produced in the glowing region. Electrons are produced by detachment due to Joule heating in the convergent part of the fan. Thus detached electrons rapidly flow to the positive electrode and leave a net positive charge that starts a new corona fan and leads to the formation of a propagating stem whose temperature is of the order of the value required to detach electrons from negative ions ($\sim 2.0 \times 10^3$ °K). Regardless of the actual leader mechanism we want to point out that they have definitely thermally hot channels as evidenced by Schlieren photographs.¹⁸ Also, that they travel at speeds of the order of 10^4 m/sec which are smaller than the speed of streamers.

In steady glow discharges there is a large variety of waves that propagate from anode or cathode and produce standing or slowly moving luminous striations that exhibit potential differences of the order of the ionization potential of the gas. These waves may be self-excited in direct current discharges. They are more pronounced in molecular gases and are associated with spatial resonances of the electron gas in the

discharge. Typically they are not associated with gas heating and the collapse of the glow stage. They have been successfully analyzed using linearized models and have been called ionization waves. There is a large body of literature concerning these waves and we refer the interested reader to a review by Gascarden¹⁹ and a book by Franklin.²⁰ What should be noted is that these waves clearly exhibit the ability of a glow to sustain wave propagation associated with the electrons.

In all high pressure discharges there are also fast luminous pulses that travel along the discharge channel but, in contrast to the waves just mentioned, are definitely associated with increasing the degree of ionization in the discharge. The late Professor Loeb indicated that these waves are responsible for maintaining a stepped leader conducting for the time between steps and that they should share the same physics as the lightning leaders.²¹ He proposed that they should all be called ionization waves of potential gradient and indicated that even streamers corresponded to one form of these waves. His model of propagation was based on photo-ionization and avalanches. Quite in contrast R.G. Fowler²² has been, for a long time, proposing the idea that non-linear fluid waves can propagate through the electrons in a gas. Also that with sufficient overvoltage ionization waves may even propagate into a neutral gas as it is actually observed in long tubes at low pressure. In other words it is possible to have a strong discontinuity that ionizes the gas and can move at speeds that range between 10^5 and 10^7 m/sec.

III. Objective

The purpose of the simplistic review just presented is to note that gas heating by electrical discharges at high pressures, together with its associated flash of luminosity (the spark), is a transient phenomenon that occurs after the electrons in the discharge are sufficient in number to be strongly affected by Coulomb interactions. As recently noted in independently prepared reviews by Marode²³ and by the author,²⁴ it seems that at low overvoltages ($\lesssim 35\%$) there is always a transition to a weakly ionized plasma that is basically no different from a glow discharge. This fact seems to be independent of actual pressure, gas composition and electrode geometry. Thus before heating occurs the electrons in a discharge

are sufficient in number to act as an independent high temperature fluid that is not in equilibrium with neutrals or ions. Consequently, the concept of Townsend breakdown using a Paschen curve merely indicates a change to a new short-lived (10^{-6} - 10^{-7} sec) stage in the discharge that is nevertheless much longer than the time required for electrons to establish random properties by colliding with neutrals. ($\sim 10^{-12}$ sec). The change to a glow may irreversibly lead to gas heating but has little to do with the processes that produce it. Heating comes about as an intrinsic thermodynamic instability in the plasma which is known to occur at a well defined critical degree of ionization¹¹⁻¹²⁻²³⁻²⁴ ($\sim 10^{-2}$ in air corresponding to $\sim 10^{17}$ electron/cm³ at atmospheric pressure). What precisely brings about the increase in electron density to produce a spark is not clear and may be different under different experimental conditions. For instance, a small spark is definitely strongly affected by metallic properties of the electrodes but it is not clear that this interaction plays a major role in longer sparks and it must be ruled out in discharges that involve only dielectrics (e.g. lightning). A gas may heat up without evidence of space waves as the result of a subsonic expansion of electrons from a cathode spot into the pre-breakdown glow.²⁵⁻²⁶ However, if the electrons can be regarded as a fluid then it is possible for strong hydrodynamic discontinuities to exist. They will effectively increase the ionization in the gas. It is even conceivable that a strong shock wave may change the gas from neutral (instead of weakly ionized) to strongly ionized. This seems to be what happens at very high pressure or overvoltages. In all cases, it is clear that the final increase in ionization comes about by isotropic inelastic collisions of electrons with neutrals and not directly by the acceleration of individual electrons in the direction of the applied field. Note that this is basically the same criteria as that postulated by Loeb to explain the formation of ionizing waves of potential gradient. He pointed out that these waves are produced whenever the accumulation of electrons exceeds their ability to diffuse in the gas.²¹

IV. Some Recent Experiments

In the process of clarifying the physics of ignitions we have been able to show¹⁴ that a minimum ignition energy, that is to say, the minimum amount of energy required to heat a gaseous mixture to combustion temperature,

is associated with the ability of the spark to produce electron densities of 10^{17} cm^{-3} . The energies involved are $1-2 \times 10^{-4} \text{ J}$. Our small sparks are produced in gaps that are smaller than those required to produce an avalanche of critical size. Consequently, no streamers are produced. Instead, the discharge expands over the cathode by photoemission for a time period of $2-5 \times 10^{-7} \text{ sec}$. It fills a cylinder of diameter comparable to gap separation with a glow discharge. Then, very suddenly, in times of the order of 5 to 50 nsec practically all of the stored electrical energy goes into the gas, the current reaches a peak value of $\sim 10 \text{ A}$ and the voltage completely collapses. Thus, in a sense, we have isolated the simplest possible manner in which the glow-to-arc transition takes place: the available energy is just sufficient to heat the gas, there are no streamers and there is a glow stage that precedes the gas heating event.²⁴

It was soon established that gas heating was directly associated with the formation of a cathode spot as suggested by Cobine²⁷ as early as 1938 (see e.g. review by Lutz²⁸). A direct confirmation of this fact can be obtained by covering the cathode with a thin dielectric material (e.g. a Mylar sheet, or Polaroid coating). It is observed that the onset of ionization occurs at exactly the same electric field as without the dielectric. However, positive ions from the glow rapidly accumulate on the surface and reduce the field in the gap. No cathode spots are produced and no spark occurs. Only when the dielectric has accumulated sufficient charge to sustain a large surface streamer discharge does a spark take place. This, of course, requires energies much larger than the amount required to heat the gas with two metal electrodes. The surface streamer discharge seems to replace the cathode spot.¹⁴

Even a very carefully cleaned and polished electrode of any metal, including the noble metals, is always coated by a very thin oxide layer that may be patchy or continuous and varies in thickness from 5 to 500 nm.²⁹ The reason is that the equilibrium vapor pressure of O_2 over a metal surface is of the order of 10^{-10} Torr . Thus, unless very extraordinary care is taken a metal surface always incorporates a dielectric oxide layer. (A fact familiar to people that try to measure work functions). The fact is that the surface of a metal with its associated dielectric coating constitutes a very good capacitor. Some of the theories concerning the

formation of a cathode spot explicitly include changes in this oxide layer.³⁰ However, the physics of this process is probably in worse shape than that of gaseous breakdown because some of the leading scientists in the field have used convincing arguments to claim that, even conceptually, we simply do not comprehend what happens.³¹ Experimentally it is accepted that a cathode spot starts with a process of cold field emission and that it incorporates a characteristic smallest unit of surface known as a Kesaev cell (1-2 μm). We have been able to relate the gas heating event at minimum energy with the formation of at least a single Kesaev cell. Thus it is experimentally clear that gas heating in small gaps is controlled by events that take place at the cathode. If the electrons in the glow constitute a weakly ionized gas then it is possible for them to interact with those emitted from the cathode spot in such a way that the electron temperature increases and the ionization reaches the critical value for gas heating. We have shown that this is indeed the case and that the interaction is compatible with the ability of electrons to behave like an independent fluid.²⁵ Experiments in N_2 clearly exhibit a region of ionization that is shaped like a cone with its apex at the cathode spot. The angle of luminosity is of the order of 25° . This is about an order of magnitude larger than the cone of a typical avalanche,³⁴ but exactly of the right size to represent a region of turbulent mixing in a submerged electron jet.²⁶ (This angle is unique for all fluids).³² The picture that emerges is then that of electrons from the cathode spot being injected into a weakly ionized glow and producing electron turbulent mixing. This raises the electron temperature and produces further ionization. If this is the case there should exist a needle shaped inviscid core region in the jet that extends from the cathode into the gap. We have obtained photographs that clearly exhibit the existence of this inviscid core region. Not only that, but it has been shown that increasing the capacity results in periodic structures that can be calculated as the characteristic shock structure of an underexpanded supersonic jet.²⁶ The point being made is that we have good evidence to demonstrate subsonic and supersonic behavior with reference to the electron acoustic velocity ($\sim 7 \times 10^5$ m/sec in N_2 at atmospheric pressure assuming $T_e \sim 1.5$ ev). As the gap length is increased it becomes less evident that the discharge and final gas heating may be controlled by the electrodes.

A particularly clear experiment was reported by Stritzke, Sander and Raether in 1977.³³ (Contrary to most of the previous work by Wagner, Tholl and Koppitz in Professor Raether's group this paper is published in English). They consider breakdown in N_2 at 300 Torr with 20% overvoltage in a 2 cm uniform field gap, and take into account the fact that N_2 is not a hydrogen like atom as far as spectroscopic studies are concerned. About 100 electrons are released from the cathode by a UV flash and start an avalanche whose luminosity becomes recordable before it reaches its critical amplification of 10^8 at 14 mm from the cathode in 90 nsec. (See Fig. 1 taken from their publication). Then anode and cathode directed streamers propagate from the region where the avalanche becomes critical. They bridge the gap at 140 nsec.

Starting at 130 nsec three luminous waves are recorded. The first one follows the cathode directed streamer channel even before the latter reaches the metal surface. The speed of these waves is recorded as 0.5, 1.8 and 3.6×10^6 m/sec. The circuit current shows a rapid increase as soon as wave propagation starts. It is noted that the luminosity in the channel starts to increase rapidly as soon as the streamers reach the electrodes and that the rate at which it grows increases from $\exp(1.5)$ in 2.5 nsec to $\exp(3.0)$ in the same time interval but after the passage of one wave. Thus, the increasing speed of the waves may be associated with higher conductivity. The luminosity during the wave propagation stage is exclusively from molecular N_2 line emission. ($T_e \gg T_n$) and the current growth is accelerated much more than in the avalanche-streamer stage. When the current reaches about 10 A a region 2 mm behind the critical avalanche (at the "neck" of Allen and Phillips³⁴ cloud chamber experiments) becomes thermalized in a few nanoseconds as evidenced by the emission of continuum radiation. Then a hot thermalized plasma propagates towards anode and cathode. The current keeps growing to 1800 A when it becomes limited by the circuit. There are three things we must note about this outstanding paper. First, that gas heating starts at the same current (~ 10 A) as what we measure in smaller gaps without streamers. In fact this current value does not change by changing the overvoltage even though the whole current growth is accelerated. As noted, the time between the bridging of the electrodes by the discharge and thermalization decreases rapidly

with overvoltage. Secondly, it must be noted that even in a relatively small gap (2 cm) waves of ionization reflect from either electrode as was originally noted in much larger gaps. Note that they travel in weakly ionized channels, and are more effective for ionization than avalanches or streamers. Thirdly, that a thermalized region propagates from a midgap region toward both electrodes at speeds comparable to those of the ionizing waves in the weakly ionized gas stage. The paper makes no reference to metal lines or cathode evaporation although it probably does occur. For instance Wiese and Augis³⁵ start to see traces of metal lines at the surface of a gold cathode at 60 nsec in a discharge in Argon at 1 atm. Thus there is evidence of some cathode activity early in the discharge, but in all cases it is clear that the gas gets hot before ionization of metal vapor starts to become effective in the gap. In fact, Stritzke computes an increase in pressure by a factor of 25 in just a few nanoseconds, which is comparable with the time at which we measure a spark to become hot without streamers. It must be emphasized that only electron motion can account for these times, and that even though a streamer stage is produced, it is not clear that the electrodes do not play a significant role.

In positive point-to-plane gaps of 1-2 cm the streamer is produced in a region with a field much higher than that in a uniform geometry configuration. However, once produced the streamer travels in a region of comparatively negligible electric field. It has been shown that in air they can travel for a small distance even in a field free region.³⁶ Also that they propagate³⁷ for distances of the order of one meter with a field small compared to that required to start ionization in a uniform field (~ 7 vs 30 kV/cm). Once they cross a 1 to 2 cm gap streamers lead to a spark with a delay time that decreases with their amplitude.¹³ The circuit current during this delay stage is negligible compared to that when the streamer is traveling. After a pause, current again grows very rapidly, as in uniform fields. Bastien and Marode³⁸ have proposed an explanation for the breakdown process based on the assumption that excited molecules transfer their energy to the neutrals. This transfer together with some Joule heating leads to a small but rapid gas heating of the channel (a few 100's °K) with the net result that the density decreases

and the E/n_r value increases in a cylindrical region outlined by the primary streamer. Consequently, intensified ionization takes place and the spark issues. This model neglects the shielding of the field that would occur if the streamer can be considered to be a good conductor and any possible interaction with the electrodes. It assumes that there is sufficient energy on the excited states to be comparable to the neutral gas thermal energy.

In the past few years the problem has been intensively studied experimentally at the Universite de Pau-Adour by Dupuy, Gilbert and their associates.³⁹ They have incorporated a Schlieren system together with up-to-date instrumentation to record different stages in the transformation to a spark in positive point-to-plane geometries of the order of 1-2 cm. The Schlieren system does not record the density decrease predicted by Bastien and Marode. Instead, it seems that the channel gets very hot and at high pressure before it even starts to expand. This, of course, agrees with Stritzke's observation in uniform fields. They have also observed that before the collapse of the voltage there exists a stage in the discharge characterized by the formation of "filaments." These propagate from both anode and cathode into the gap and are thermalized. Thus the sequence of events is listed as follows: streamers, cathode activity and cathodic return front, anodic channel or secondary streamer, filaments and spark. We refer the reader to the publications listed as reference 39 for details. There are however some results of these studies that are particularly pertinent. The discharge does not become hot without activity at the electrodes. The anode directed wave has a speed between 0.7 and 5×10^6 m/sec. Its length of travel increases with capacity that determines whether or not the wave dies or crosses the whole gap. This is exactly the same as for our inviscid core jet lengths. It is important to emphasize an, as yet, unreported fact (Dupuy personal communication, 1983): A cathode directed secondary streamer and the wave from cathode to anode have luminous fronts that meet but pass through each other without any significant change in their propagation characteristics. This is a basic property of solitary waves in fluids. Finally, the spark channel is related to a critical charge density and consequently a critical energy which is determined to be 0.15×10^{-3} J and corresponds precisely to the

minimum energy required to heat the gas in our small sparks without streamers. The nature of the filaments is not clear. Dupuy, et al. suggest that they may be the same as leaders in long sparks. But they may also be associated with the thermalized channels observed by Stritzke to propagate toward the electrodes from the initial region of thermalization behind the critical avalanche as well as to the thermalized inviscid cores from our cathode spots.

The main purpose of reviewing all this experimental work is to demonstrate that in small sparks, just as in large sparks, the final ionization stages that lead to a collapse in the applied voltage are also always associated with the formation of luminous waves. Under many different circumstances these waves propagate at speeds of the order of 10^6 m/sec which is very close to the electron-acoustic velocity $(5 kT_e/3m_e)^{1/2}$. It makes no difference if the discharge is preceded by streamers or by successive avalanches. In fact, Koppitz⁴⁰ has shown that by enhancing photoemission using a CuI cathode coating the luminous wave fronts are no longer filamentary. Nevertheless he measures about the same velocity ($\sim 10^6$ m/sec).

In small gaps it is clear that the electrons that produce the voltage collapse are supplied from the outside of a region that is a weakly ionized plasma. The electrons have maxwellian distributions [$m_e/m_n \ll v_{ee}/v_{en} \ll 1$], obey an ideal gas law and can be described as collision dominated. ($\bar{t} \gg v_{ei}^{-1}$; $\bar{L} \gg \bar{c}_e v_{ei}^{-1}$ where \bar{t} and \bar{L} are typical time and length dimensions, \bar{c}_e is the average electron thermal speed and v_{ei} is the collision frequency with ions; $v_{ei} \sim v_{ee}$).

V. Theoretical Approaches

As indicated Fowler²² and his students have proposed that it is possible for non-linear effects to produce a strong shock-like wave in the electrons. This can propagate as a strong discontinuity into a neutral gas. Thus, he assumes the existence of a step like time independent solution for the system of equations of conservation of mass momentum and energy together with Poisson's equation. He assumes that the electron pressure gradient is the driving force. Consequently, the wave may propagate in or against the direction of the electric field. This has been

experimentally known for a long time (~ 1920). In contrast to this point of view Albright and Tidman⁴¹ and Klimbeil, et al.⁴² have proposed the propagation of a wave driven primarily by a strong electric field ($\sim 10^7$ v/m in air) into a neutral gas. Photoionization and constant electron temperature are considered to be important for the propagation of this wave.

Since our experiments with small sparks demonstrate that most of the electrons that heat the gas are injected into the spark gap by processes occurring at the electrodes, we can consider the possibility of wave propagation into a weakly ionized gas with the additional restriction that the source term in the mass conservation equation is small compared to all other terms.⁴³ Thus the weakly ionized gas provides the low pressure electron gas and those coming from the cathode are so numerous that their flux outweighs the ionization taken place. This is in accordance to the fact that inelastic collision frequencies are very small compared to their elastic counterpart ($\nu_{NE} \ll \nu_{ei} \ll \nu_{en}$). The Vlasov-Maxwell system of equations (collision term in Boltzman's equation equal to zero) can be written as a single non-linear matrix equation. When linearized about an initial steady state this equation yields, formally, a standard dispersion equation for the propagation of plane waves proportional to $\exp[\pm (kx - \omega t)]$ for each matrix component (electron density, fluid velocity, temperature and electric field).

$$\omega^2 = c^2 k^2 + \omega_p^2 = c^2 (2\pi/\lambda)^2 + n_e e^2 / \epsilon_0 m_e \text{ with } c^2 = (5kT_e / 3m_e)$$

The group velocity is given by $(\partial\omega/\partial k) = c^2(k/\omega)$. Thus plane waves propagate provided $[c^2 k^2 + \omega_p^2 - (2\pi \nu_{en})^2]^{1/2}$ is a real number. That is either $n > (\epsilon_0 m_e / e^2)(2\pi \nu_{en})$ or $\lambda < (c/\nu_{en})$. Consequently there is a critical density or wavelength for propagation that depends on the average electron neutral collision frequency. Using suitable values for air ($\nu_{en} \sim 10^{12} \text{ sec}^{-1}$) it can be shown that waves propagate at either a density lower than the critical thermalization value or for wavelengths of the order of the thickness of the oxide layer at the cathode. (This is, of course, what is experimentally confirmed by the shock wave pattern on the inviscid core structure in the electron turbulent jet experiments).²⁶ Asano⁴⁴ has shown that even for a thermally conducting viscous fluid with electron neutral collisions there are changes in attenuation of the waves but not in the

reference electron acoustic speed. Consequently, the whole Vlasov-Maxwell system of equations was, again, considered more in detail.⁴⁵ The importance of each force term in the momentum equation can be assessed by normalizing the equations and comparing the coefficients in front of each term. The equations are not linearized so that the formation of discontinuities due to the presence of small but finite non-linear terms remains in the mathematical description. It can be formally shown that for a characteristic propagation velocity of $\sim 10^6$ m/sec the driving force term is always, as expected, the electron pressure gradient. The values required for the characteristic time are compatible with the electron-neutral collision frequency that determines their distribution but not with the much smaller ion-electron interaction that determines their thermalization time. The problem needs further clarification because it is associated with the well known divergency of crosssections for Coulomb forces and the possibility of multiple interaction between electrons. Nevertheless, it is possible to show that the weakly ionized plasma is a suitable media for the propagation of fluid waves with an amplitude that is either stepped (shocks), S-shaped (satisfy Burger's equation) or even that of a soliton (satisfy non-linear Schrödinger or Korteweg-de-Vries equation). This is, again, as expected from the experimental section.

While we were concerned with analytic mathematical solutions to support the possibility of wave propagation in weakly ionized gases, I. Abbas and P. Bayle,⁴⁶ at the Université Paul Sabatier in Toulouse, put numbers into the same system of equations and prepared what the author considers the best hydrodynamic description of the spatio-temporal evolution of density momentum and energy in an electrical discharge. The mathematical formulation includes source terms, thermal and density diffusion, all interactions of electrons with heavy particles and the fact that electrons are not necessarily in equilibrium with the electric field.

By taking into account the finite life time of excited molecules it is shown that photoionization is not an important mechanism for the propagation of fast ionizing waves ($\sim 10^6$ m/sec) at high pressures. In dense nitrogen ($p > 50$ Torr) the ionization wave due to photoionization is shown to lag behind the electron collision ionization wave. Thus

irrespective of their actual ability to ionize, photons are emitted too late. A computer solution shows that some of the electrons in the shock front can actually become precursors traveling ahead of the actual front. This is due to thermal diffusion. It is shown that the shock zone cannot be the same as that in a regular shock in a neutral gas. Instead an electron shock is divided into two different regions separated by an electron pressure maximum. This separation is tied down to the absence of ambipolar diffusion that in turn limits the analysis to relatively small electron densities. Nevertheless, it is shown that the two regions comprising a shock are not transient in nature and can be maintained as a propagating structure. Thus, basically, our work is in agreement with that of Bayle and both are in support of the model proposed by Fowler. As far as space waves of ionization are concerned, the often quoted important role of photoionization does not seem to be justified for dense gases (This is not the situation for streamers in gas mixtures). The papers of Abbas and Bayle,⁴⁶ like the paper by Stritzke³³ in the experimental section, must be carefully read by all people interested in high pressure discharges.

The research at Toulouse is also tied to experimental results. Thus, a theory has been developed by P. Bayle, M. Bayle and E. Morales⁴⁷ to interpret streak camera records. It has been shown that the variations in optical density in a film can be used to evaluate the electric field and the electron density. G. Caumes⁴⁸ has prepared a thesis (Doctorate de 3 ème Cycle to appear in J. Appl. Phys., D. 1983), in which streak pictures from discharges in N₂-O₂ mixtures are systematically analyzed using a computer to store the data. The results provide a clear record of the propagation of space waves of ionization. Typically a small perturbation in electric field or electron number density is followed in steps of 0.04 nsec. It is clearly shown that in times of the order of 0.4 nsec and distances of about 0.2 cm an initial small disturbance grows to be a propagating electric field or electron density maximum. (Note, the speed is about 5×10^6 m/sec). The front is followed by oscillations that look very much like those in collisionless shocks (See Fig. 2 taken for Caumes thesis).

VI. The ONERA Experiment

A large Lucite plate, typically 2 mm thick, is placed in contact with a long (1.05 m), narrow (15 mm), flat (5 mm) grounded metal strip.

Both the Lucite plate and its attached long narrow electrode are supported perpendicular to the floor. Facing the plate on the side opposite to the grounded electrode there is a vertical row of metal points that constitute a comb-like structure (See Fig. 3). These points are not in contact with the dielectric but instead are attached to a motor and a high voltage supply. They can move back and forth horizontally in front of the dielectric plate. Thus corona discharges charge the surface of the dielectric until enough charge is accumulated to stop the air discharge. Above the level of the tip of the grounded electrode there is a metal half sphere (~ 3 cm diameter) that rests on the charged surface. It is supported by a rod that goes through the lucite and is connected to one end of a pressurized spark gap. The other side of this gap goes to the long grounded electrode through an induction current measuring device. After the surface is charged and the corona has stopped, the metal comb is removed to one side and the pressure lowered in the pressurized spark gap. A spark is produced that grounds the floating metal hemisphere located at the top of the system. Consequently, a highly overvolted discharge starts over the surface of the charged dielectric. If the surface voltage is high enough this discharge will reach the region directly over the metal grounded strip. Then a bright spark is produced that travels in a straight path following the region of high charge density determined by the metal strip. This is the gliding spark. Straight line propagation and a predetermined location provide ideal conditions for experimental studies. In addition, the speed of propagation varies between 0.6 and 2.8×10^6 m/sec and spectroscopic analysis shows that this is fast enough to guarantee that any changes in the dielectric due to the heat produced by the discharge occur after the front has passed over the region where such a change takes place. Thus the spark front is exclusively an air discharge that thermalizes and becomes very hot ($\sim 2.4 \times 10^4$ °K) without the influence of metallic properties or the control provided by the impedance of an external power supply. The currents produced are in the range of 20 to 120 A therefore, the processes taking place may be closely related to the leaders in lightning. We are happy to see a lower velocity limit that is very close to the reference electron acoustic speed. The discharge has been described in publications by Larigaldie, et al.,⁴⁹ who emphasize a preliminary model for the physical mechanisms, and by Borgade and Hartmann⁵⁰

who performed an initial spectroscopic study of a narrow region at half the length of the spark (1.05/2 m). Both Larigaldie and Hartmann are continuing their investigation beyond these preliminary studies. An internal ONERA report was prepared by Larigaldie⁵¹ and a new one is in preparation by Larigaldie and Lagage.

What can be described as an ideal experimental situation namely a sequential display of all processes leading to the formation of a high current hot channel, not affected by electrodes or external circuit has proven to be also a display of our ignorance concerning what is going on. It is therefore, appropriate to describe a little more in detail some of the experimental results. Discharges that propagate over a positively charged dielectric are called negative discharges and those over a negative dielectric positive discharges. One important finding is that these two types of discharges are completely different in character.⁴⁹ Positive discharges exhibit strong current oscillations that become stronger as the charge on the surface (corona voltage) increases; they also exhibit strong luminous bands (like striations) behind the traveling front. Streak pictures show that they travel always in steps. Negative discharges exhibit smooth current traces, and there are no repetitive luminous regions behind the traveling front. They are also initially stepped but above a critical voltage travel continuously over the dielectric. Their velocity decreases as they travel over the surface but this has not been carefully measured. The marked difference is obviously associated with the direction in which electrons travel and seems to support Hemmati⁵² that negative discharges (proforce waves) are always strong electron shocks. As noted, most of the ONERA effort has been concerned with negative discharges. Hence, unless specifically noted we will assume this is the case.

A plot of the current value at the start of propagation of the gliding spark, I , versus the voltage of the corona points, V_0 , provides a parabolic curve $I_0 \sim V_0^2$. Since V_0 is much higher than the corona threshold value (~ 5 kV) it can be assumed that it is also the starting surface voltage V_0 . It is then assumed that the voltage seen by the propagating tip at any other location is the same as that extrapolated from the I_0 vs V_0 plot using the current that is measured when the discharge goes over the point in question. With this assumption and a spark

velocity, v_s , measured to be linearly proportional to V_0 , it is possible to calculate the change in resistance and voltage per unit length along the discharge channel. The analysis⁴⁹ considers all the current flowing in the channel to be produced in a corona brush region ahead of the filamentary discharge. The channel has an ohmic resistance proportional to its length which is responsible for the decrease in current as the discharge elongates. The corona brush discharges the capacitor below its surface and feeds the current into a hot channel behind it. The experimental current traces can be reproduced through calculation using V_0 and v_s .

The same picture of a corona brush leading to a hot channel is supported by spectroscopic observations made by Borgade and Hartmann⁵⁰ along the axis of the discharge. They observe a region 23 mm long where only N_2 molecular line radiation is recorded. The rotational temperature in this region increases gradually to 1500 °K and then very rapidly within a few nanoseconds jumps to 24,000 °K. At this point atomic, N, O and H radiation is observed together with a strong continuum background. Stark broadening of the hydrogen lines is used to determine a maximum electron density of $7.2 \times 10^{17} \text{ cm}^{-3}$. The length of the front associated with molecular radiation (23 mm) seems to be independent of the applied voltage which for 125 kV corresponds to a time interval of 17 nsec. That is, the time changes with voltage but not the spatial dimension of the front. There seems to be some light before this front but since the spectroscopic equipment is about 2 m away the nature of this radiation is not clear. Further along the channel the density and temperature decrease monotonically to $1.5 \times 10^{16} \text{ cm}^{-3}$ and $\sim 15,000$ °K in a time interval of 800 nsec. It is pointed out that this is the same variation observed in lightning. During this last period the intensity of the continuum background decreases and the presence of very strong H and C lines together with CN lines evidence evaporation from the dielectric.

A ten nanosecond exposure frame⁵¹ from an image intensifier located as far away as the spectroscopic equipment (~ 2 m) exhibits a brush-like structure with several filaments that converge into a bright region that leads through a constriction into a bright stem ($\sim 100 \mu\text{m}$). This front has lateral dimensions also of the order of 2 cm (See Larigaldie, reference

49, Fig. 9 or 51, Fig. 1). Because of the molecular radiation the discharge in front of the hot region is considered to be a standard equilibrium glow discharge. It is not clear where in the converging luminous fan does the temperature actually increase. However, as noted, the jump occurs at 1500 °K. At this temperature O^- loses its electron in an average time of 2.5 nsec which is compatible with the observed rise time to 24,000 °K. Consequently, electron detachment is considered to be important as in Gallimberti's leaders (See reference 17, p. C7-212). (It is perhaps worthwhile to note that in times of the order of 10^{-8} sec O^- is probably the only negative ion that can be obtained because it is produced through dissociative attachment $O_2 + e \rightarrow O^- + O$ and it is unlikely that there is time for three body reactions to form other negative ions.) The glow discharge in front of the hot region is considered to have a uniform field $E \approx 1.3 \times 10^6$ V/m, a constant electron density $n_e = 10^{15}$ cm⁻³, and an electron temperature $T_e \approx 2 \times 10^4$ °K associated with the barrier produced by vibrational excitation. Thus a current is produced and Joule heating produces the gradual temperature increase. It can be argued that since the process is transient such a high field and electron density are compatible with each other. However, it must be noted that plasma shielding is not considered.²⁴ For instance, if the electron temperature is of the order of 2×10^4 °K, the shielding distance near a high voltage region at 125 kV is $\lambda_D (eV/kT_e)^{1/2} \approx 80$ μm and the time to produce the shielding is $(\omega_p/2\pi)^{-1} \approx 4 \times 10^{-12}$ sec. Both values are very small compared to experimental values of 2 cm and 2×10^{-8} sec. (λ_D is the Debye length and ω_p the plasma frequency).

It could be argued that since the actual ionization in the brush is only along the bright filaments, the actual effective field over the whole region is higher. At any rate, there seems to be a correlation between the computed values and those obtained from Toepler's laws.⁵¹ This basically indicates that a minimum energy must be deposited in a glow region before it becomes the hot region of a gliding discharge. The actual distribution of the electric field may not be important since only the potential difference and the total length are related to Toepler's laws.

If ionization takes place ahead of the hot channel it will have to occur either through avalanches or through interaction with an isotropic

electron distribution in the glow discharge region. An avalanche in air at a field of 4×10^6 V/m becomes critical ($n_e/n_{e0} \approx 10^8$) in a distance of 4 mm and a time of 20 nsec.⁵³ In that same time the spark front would have traveled $(2 \times 10^6 \text{ m/sec})(2 \times 10^{-8} \text{ sec}) = 4 \text{ cm}$, which is ten times the distance required to make a critical avalanche. There is no data to calculate avalanche growth in a much higher field, but it is known that at the critical stage they exhibit electron densities of $\sim 5 \times 10^{15} \text{ cm}^{-3}$ and that the positive ions left behind slow their propagation.^{34,54} The conclusion is that the glow region in front of the hot channel is not maintained by avalanches. Therefore if ionization takes place in the glow region it will have to be by interaction of the glow region itself with the field. This interaction has been postulated by Albright and Tidman.⁴¹ They have shown that in air there can be an interaction between a very high field (4×10^7 V/m) and a weakly ionized plasma ($n_e \sim 10^{15} \text{ cm}^{-3}$; $T_e \sim 1.7 \text{ eV}$). This tends to produce filamentary channels that propagate at speeds of the order of 10^6 m/sec. The front thickness is of the order of 10^{-5} cm, but since the analysis is based on an energy balance they do not consider any specific ionization mechanism or gas heating. A follow-on paper⁴² considers photoionization as the ionization process, but, as noted, this can be ruled out based on Bayle's calculations. Thus their work can account for the filamentary nature of the corona brush but does not consider the actual ionization process taking place.

Regardless of the actual ionization mechanism in the flow it has to occur in a time compatible with the observed distance and velocity: 2 cm and 10^6 m/sec. That is to say ionization has to occur in times of the order of 10^{-8} sec. Irrespective of how the front propagates it is necessary to consider if it is possible to have ionization within the glow in this short time. The ionization frequency, ν_i , for electrons with a maxwellian distribution (guaranteed by the condition $\nu_{ee}/\nu_{en} \gg m_e/m_n$) can be calculated. (See e.g. Abbas and Bayle⁴⁶ p. 654 or Larigaldie⁵¹ p.29). Assuming an ionization potential of 16 eV and the ions to be at room temperature (T_e in eV).

$$\nu_i = 1.66 \times 10^{11} T_e^{1/2} (1 + T_e/8) \exp(-16/T_e)$$

From this equation we can calculate the following table:

v_i^{-1}/sec	1.8×10^{-7}	1.0×10^{-8}	1.7×10^{-9}	5.2×10^{-10}
T_e/eV	1.5	2.0	2.5	3.0

It is clear then that within times of 10^{-8} sec electrons in a glow of energy larger than 2.0 eV can reproduce each other. It can even be shown that, as in Gallimberti's leaders, the glow may be in equilibrium with a dissociative recombination process $[e + (XY)^+ \rightarrow X + Y]$ provided $n_e \geq 2.1 \times 10^{15} \text{ cm}^{-3}$. This is the value used by Albright and Tidman, Larigaldie and ourselves and corresponds to those actually measured in streamers by Wagner and Tholl. It is concluded that a filamentary glow with $T_e \geq 2.0$ eV and $n_e \geq 10^{15} \text{ cm}^{-3}$ is compatible with the glow in the gliding discharge. But again, a maxwellian distribution implies effective Coulomb shielding and the formation of sheaths.

At this stage the ONERA experiment is necessarily suffering from a high ratio of interpretation to experimental data. Just the same as in other sparks, it is not clear how the electrons become sufficient in number to heat the gas. Bastien and Marode's mechanism of gradual heating can be ruled out on a time basis alone. There are no Dupuy's filaments or our cathode spots to provide the change. A process associated with an ion acoustic instability has been suggested but we fail to see how this can arise in a collision dominated gas ($v_{ee} \ll v_{en}$). Even if an ion acoustic wave can be produced, its speed is of the order of 3×10^3 m/sec roughly three orders of magnitude slower than the spark velocity. It could be argued that since the spark velocity has a lower limit that is comparable to the electron acoustic speed the whole process is associated with a supersonic electron shock. However, we are not aware of any description actually leading to gas heating and should point out that even though Fowler has assumed the existence of these shocks their true nature and indeed their existence is just becoming clear. We would strongly recommend that electron density measurements be made ahead of the hot region. (In the A and B periods of Borgade and Hartmann,⁵⁰ or the AB and BC of Larigaldie⁵¹). An important finding that will be described in forthcoming publications by Larigaldie is the possibility that gas heating may be

closely associated with superelastic collisions and three body electron-electron recombination ($e + e + X^+ \rightarrow e + X$). This process allows the neutrals to get hot while supplying energy to the electron gas. This seems to be the process that leads to a nearly completely ionized gas, and is compatible with experimental time and temperature changes seen by the spectroscopic study. Finally it should be noted that there has been recently a lot of speculation^{55,56} concerning the possibility of very fast (~ 500 eV) run-away electrons in highly overvolted gaps and in lightning. A gliding discharge could be easily adapted to look for them, for instance, by placing a sensor just below the Lucite surface.

VII. Conclusions

There is evidence that in all sparks, irrespective of pressure and geometry, gas heating is preceded by the formation of a glow discharge. This includes highly overvolted long gaps and surface discharges in which the glow travels ahead of a hot thermalized region. In all cases, the glow constitutes a weakly ionized channel with sufficient electron temperature and density to reduce electrical forces within the region where it exists. In small sparks the glow fills the whole inter-electrode region before any gas heating occurs. Then processes associated with the glow-electrode interface interact with a primarily isotropic distribution of electrons in the gap. The electron energy increases and the degree of ionization also increases to the stage where Coulomb interactions with heavy particles (ions) becomes important. The possibility that gradual heating reduces the gas density and increases the ionization rate is not supported by experimental evidence. On the other hand, the interaction of electrons from the cathode with those in the gap, the existence of waves of potential gradient that reflect at the electrodes and the formation of hot filaments associated with either electrode has been clearly demonstrated. In long laboratory sparks both the leading glow and the hot gas region that follow it, propagate at speeds of the order of 10^4 m/sec and involve currents of the order of 1 A. The transition from a glow to arc is not understood, as evidenced by mutually exclusive models based on either shock waves that overtake high energy electrons or, instead, comparatively low energy electrons undergoing attachment and detachment. The degree of ionization in the leader is small and full ionization becomes evident only with the return stroke.

Highly overvolted discharges over charged dielectrics travel at speeds of the order of 10^6 m/sec with currents of 100 A. They involve hot channels that seem to be fully ionized as indicated by initial spectroscopic studies. Superelastic collisions and three body electron-electron-ion recombination may play an important role in these discharges that exhibit characteristics similar to lightning leaders: electron densities, time variation, current magnitude and speed. This last one is of the order, or slightly higher than the electron-acoustic speed in the weakly ionized plasma ahead of the hot region. It suggests the existence of an electron shock. The actual transition from glow to hot channel is again not clear but the ability of these surface discharges to display all phenomena at a predetermined location makes them a powerful experimental tool. The ONERA sparks are produced using extremely high voltages (~ 125 kV). However, the critical parameter is the charge density and it is also possible to produce similar discharges using smaller voltages and thinner dielectrics. This has been demonstrated by Blythe and Carr.⁵⁷

It must be stressed that although in all discharges the transition from a glow to a hot channel is not clearly understood, it most certainly involves interactions at a degree of ionization where Coulomb forces between electrons are important. The concept of independent electrons interacting with an electric field is not appropriate at this stage and must be replaced by that of collective interactions in a plasma fluid.

From a practical point of view a streamer or glow discharge over a thin dielectric has the possibility of puncturing through and establishing a cathode or anode activity that may then heat the region above the dielectric. This is obviously related to the problem of attachment of discharges to painted, or epoxy covered metal surfaces. There are several practical examples that can be listed to indicate the importance of these processes. First is the fact that aircraft seems to be more vulnerable to lightning discharges when the plane travels in a positively charged cloud.⁵⁸ This may be related to the fact that weakly ionized discharges produce a cathode region from which a strong wave can be launched. (Only strong electron discontinuities, shocks, are possible with negative discharges). Again there is evidence that the process of dissociative attachment ($O_2 + e \rightarrow O^- + O$) produces discontinuities in lasers

and that at atmospheric pressure the disturbance travels at about the electron-acoustic velocity.³ There is the problem of surface discharges over thin dielectrics associated with thermal blankets in satellites. Balamin, et al.⁵⁹ have reproduced satellite conditions and bombarded with electrons (20 keV) a thin (50 μm) Kapton dielectric with an aluminum backing. Very strong discharges are produced that follow surface disturbances associated with brushing or rubbing. Discharges can then be made to follow straight paths. Their average speed is 7.1×10^5 m/sec. Even a very "clean" surface in the presence of a hot plasma can lead to the production of cathode spots that contaminate the plasma and interfere with the operation of a reactor.⁶⁰

VIII. Acknowledgements

This work was made possible by G.A. DuBro of WPAFB, J. Hughes of ONR and J. Taillet of ONERA. These three gentlemen provided me with a unique opportunity to learn and to discuss my work with many people in Europe. It is a pleasure to acknowledge their help. Likewise, everyone at ONERA should be thanked for being kind and helpful during my visit. In particular, J.L. Bouley who is a model host and S. Larigaldie and P. Lagage who started endless constructive arguments that I hope will continue for a long time.

P. Strizke, I. Sander and H. Raether

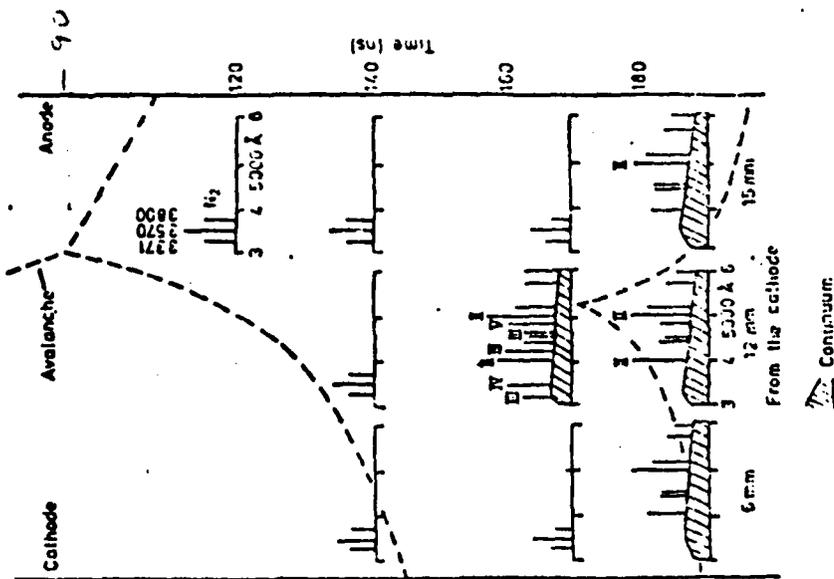


Figure 1. Schematic spectra in N_2 emitted from the discharge at different times and at different distances from the cathode (6, 12, 15 mm). A streak picture drawn by the dashed lines gives a survey of the different stages of the discharge.

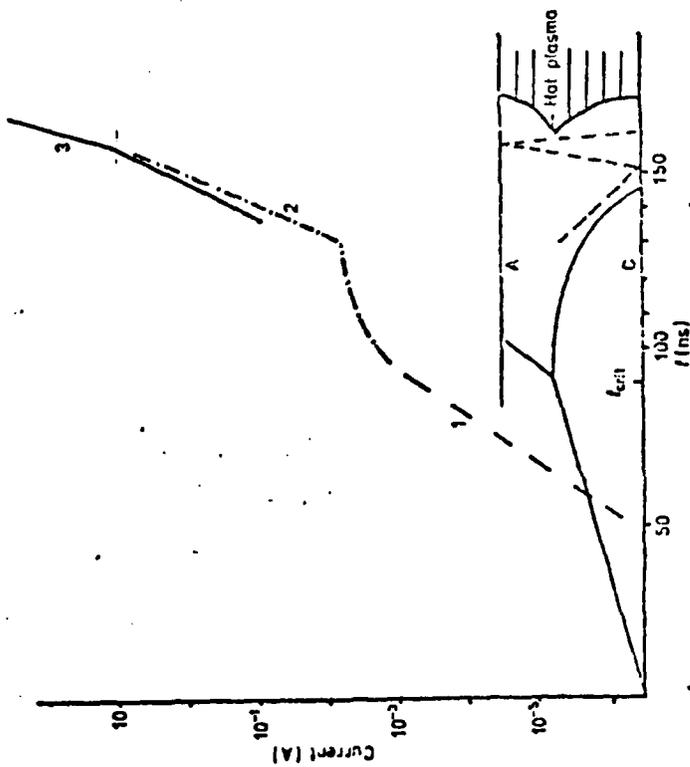


Figure 5. Current as function of time: (1) theoretical electron current during the avalanche stage, calculated by equation (5); (2) current calculated by integrating measured electron densities (Equation 6); (3) measured current. Besides the avalanche and the streamers the observed ionization waves are drawn (dashed lines in the schematic streak picture).

Figure 1. Reproduced from reference 33

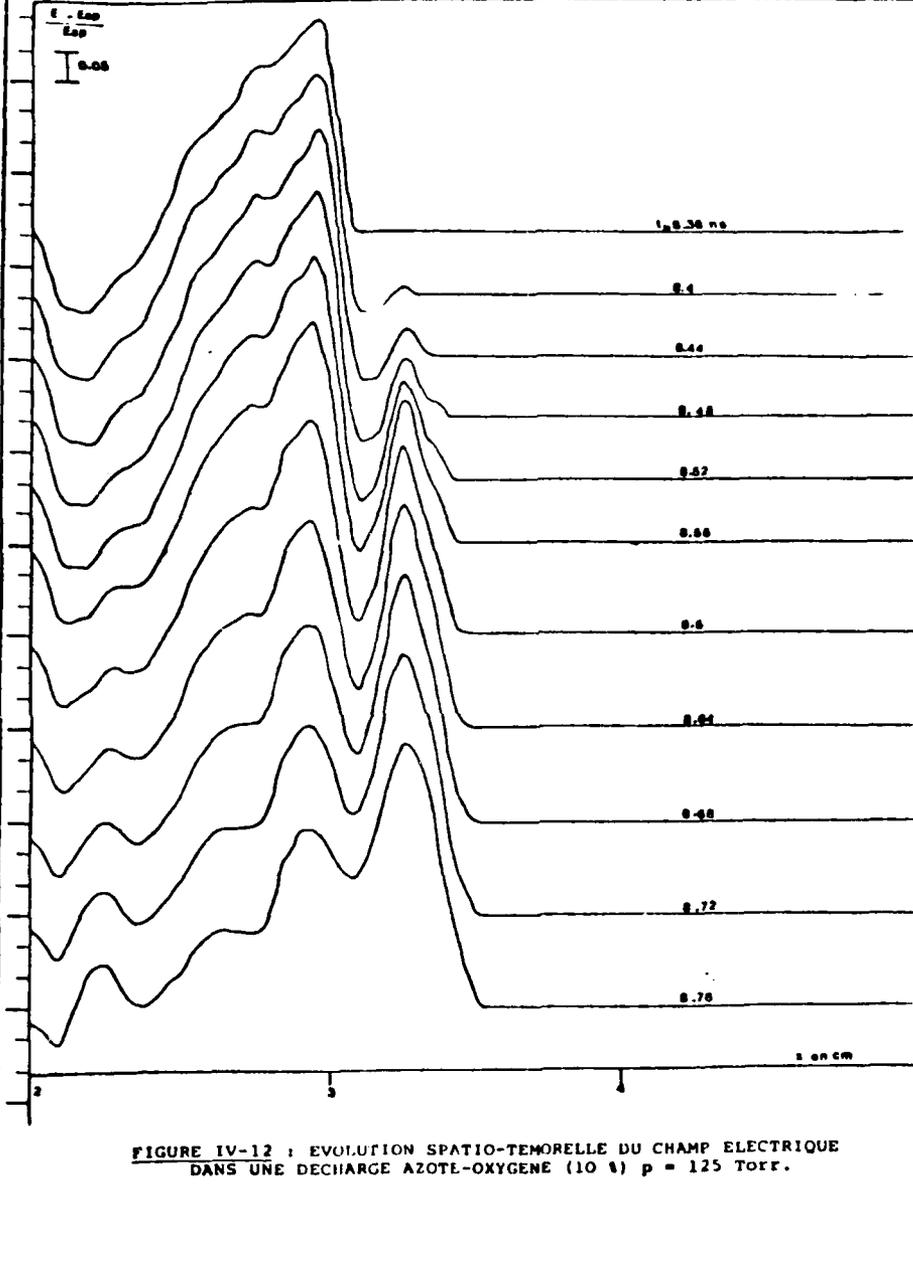


FIGURE IV-12 : EVOLUTION SPATIO-TEMORELLE DU CHAMP ELECTRIQUE
DANS UNE DECHARGE AZOTE-OXYGENE (10 %) p = 125 Torr.

Figure 2. Reproduced from reference 48

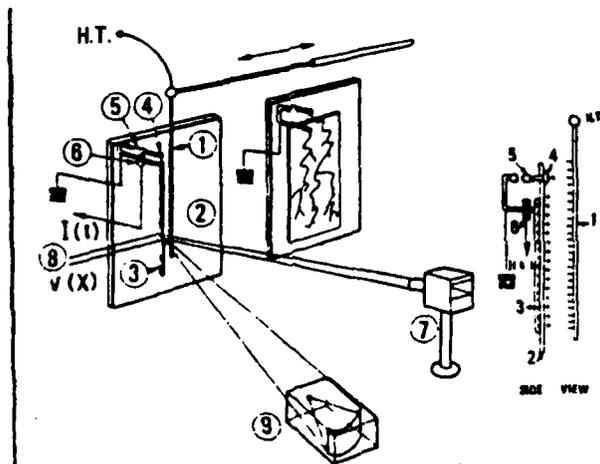


FIG. 1. Experimental setup: (1) Metallic comb (high voltage). (2) Dielectric slab. (3) Metallic strip (grounded). (4) Floating electrode. (5) Spark gap. (6) Current probe. (7) Electro-optical image converter. (8) Optical fibers. (9) Spectroscopy.

27114-07S01.10

© 1981 American Institute of Physics

7114

Figure 3. Reproduced from reference 49

IX. REFERENCES

1. Bass, R.L., T.E. Owen and F.T. Dodge. Bulk Carrier Operations Safety Enhancement Project. U.S. Dept. of Commerce, Natl. Technical Information Service, PB-256 188, June 7, 1976.
2. Space Charging Technology. NASA Conference Publication 2182 (AFGL-TR-81-0270). Conf. U.S. A.F. Acad., Colorado Springs, Nov. 12-14, 1982.
3. Douglas-Hamilton, D.H. and S.A. Mani. J. Appl. Phys., 45, 4406 (1974).
4. Chase, E.W. Electrostatic Discharge Damage to Integrated Circuits. Invited Paper, Electrostatics 83. Inst. of Physics Conference, Series number 66, St. Catherine College, Oxford, UK, 13-15 April 1983.
5. Loeb, L.B. Basic Processes of Gaseous Electronics. Univ. of Cal. Press, Berkeley, California, 1955.
6. Dutton, J. in Electrical Breakdown and Discharges in Gases. E.E. Kunhardt and L.H. Luessen, Eds. NATO ASI Series, Vol. 89a, Plenum Press, NY, 1983.
7. Loeb, L.B. Fundamental Processes of Electrical Discharge in Gases. Wiley, NY, 1939.
8. Dawson, G.A. and W.P. Winn. Z. Phys., 183, 159 (1965).
9. Geary, J.M. and G.W. Penney. Phys. Rev. A., 17, 1483 (1978).
10. Sigmond, R.S. and M. Goldman. Electrostatics 83. Inst. of Physics, Conf. Series Number 66, 81(1983).
11. Tholl, H., Z. Naturf. 19a, 346 (1964); 19a, 704 (1964); 22a, 1068 (1967); 25a, 420 (1970).
12. Wagner, K.H. Z Phys. 189, 465 (1966).
13. Marode, E. J. Appl. Phys., 46, 2005 (1975); 46, 2016 (1975).
14. Barreto, E., S.I. Reynolds and H. Jurenka. J. Appl. Phys., 45, 3317 (1974).
15. Pfender, E. in Gaseous Electronics, Vol. I, Electrical Discharges. M.U. Hirsh and H.J. Oskam, Ed., Acad. Press, New York, 1978.
16. Kekez, M.M. and P. Savic. in Electrical Breakdown and Discharges in Gases. E.E. Kunhardt and L.H. Luessen, Eds. NATO ASI Series, Vol. 89a, Plenum Press, New York, 1983.
17. Gallimberti, I. J. de Physique, Colloque C7, Supplément au No. 7, Tome 40, Juillet 1979, page C7-193.

18. Dupuy, J., A. Gibert, P. Domens, B. Hutzler, J.P. Riu, F. Edlinger. 399/RGE-5/81-Mai, 1981.
19. Gascarden, A. in Gaseous Electronics, Vol. I. Electrical Discharges, M.N. Hirsh and H.J. Oskam, Eds., Acad. Press, NY, 1978.
20. Franklin, R.N. Plasma Phenomena in Gas Discharges. Clarendon Press, Oxford, 1976.
21. Loeb, L.B. *Science* 146, 1417 (1965); *J. Geophys. Res.*, 71, 4711 (1966); *J. Geophys. Res.*, 73, 5813 (1968).
22. Fowler, R.G. *Adv. in Electronic and Electron Physics*. 20, 1 (1964); 35, 1 (1974); 41, 1 (1976).
23. Marode, E. in Electrical Breakdown and Discharges in Gases. E.E. Kunhardt and L.H. Luessen, Eds. NATO ASI Series, Vol. 89b, Plenum Press, New York (1983).
24. Barreto, E. in Encyclopedia of Chemical Processing and Design. J.J. McKetta and W.A. Cunningham, Eds., 17, 426, Marcel Dekker, Inc., New York (1983).
25. Barreto, E. Electrostatics 1979 (Inst. of Physics, Univ. of Reading, Berkshire London 1979). p. 135.
26. Jurenka, H. *J. Appl. Phys.*, 53, 6115 (1982).
27. Cobine, J.D. *Phys. Rev.*, 53, 911 (1938).
28. Lutz, M.A. *IEEE Trans. Plasma Sci.*, PS-2, 1 (1974).
29. Roberts, M.W. and C.S. McKee. Chemistry of the Metal-Gas Interface. Clarendon Press, Oxford, Chpt. 11 (1978).
30. Guile, A.E. and B. Jüttner. *IEEE Trans. Plasma Science*, PS-8, 259 (1980).
31. Ecker, G. in Vacuum Arcs. J.M. Lafferty, Ed., John Wiley, New York (1980).
32. Landau, L.D. and E.M. Lifshitz. Fluid Mechanics. Addison-Wesley Pub. Co., Reading, Mass. (1959), p. 133.
33. Stritzke, P., I. Sander and H. Raether. *J. Phys. D: Appl. Phys.*, 10, 2285 (1979).
34. Allen, K.R. and K. Phillips. *Proc. Roy. Soc. A*, 274, 163 (1963).
35. Wiese, L.L. and J.A. Augis. *J. Appl. Phys.*, 48, 4528 (1977).
36. Winn, W.P. *J. Geoph. Res.*, 70, 3265 (1965); *J. Appl. Phys.*, 38, 783 (1967).
37. Phelps, C.T. *J. Geophys. Res.*, 70, 5799 (1971); *J. Appl. Phys.*, 47, 2929 (1976).

38. Bastien, F. and E. Marode. Gaseous Dielectrics III. Pergamon Press, 1982. See also. J.P. Boeuf and E. Marode. J. Phys. D: Appl. Phys., 15, 2169 (1982).
39. Bertault, P., J. Dupuy and A. Gibert. J. Phys. D: Appl. Phys, 10, L219 (1977); Dubois, J., J. Phys. D: Appl. Phys. 13, 405 (1980); Dupuy, J. and A. Gibert, J. Phys. D: Appl. Phys., 15, 655 (1982); 16, 379 (1983); Gibert, A., J. Dupuy, M. Bayle and P. Bayle, J. Phys. D: Appl. Physics, 16 (1983). (in print)
40. Koppitz, J. J. Phys. D: Appl. Physics, 6, 1494 (1973).
41. Albright, N.W. and D.A. Tidman. Phys. Fluids, 15, 86 (1972).
42. Klingbeil, R., D.A. Tidman and R.F. Fersnler. Phys. Fluids, 15, (1969) (1972).
43. Barreto, E., H. Jurenka and S.I. Reynolds. J. Appl. Phys., 48, 4510 (1977).
44. Asano, N. J. Phys. Soc., Japan, 36, 861 (1974).
45. Jurenka, H. and E. Barreto. J. Appl. Phys., 53, 3581 (1982).
46. Abbas, I. and P. Bayle. J. Phys. D: Appl. Phys., 14, 649 (1981); 14, 661 (1981).
47. Bayle, P., M. Bayle and E. Morales. J. Phys. D: Appl. Phys., 13, (1980).
48. Caumes, G. "Etude des Decharges Transitoires dans les gaz par traitement d'images des cameragrammes de cinematographie ultra-rapide." esis Universite Paul Sabatier de Toulouse. N° D'ordre 2589. To appear in J. Phys. D: Appl. Phys. (1983).
49. Larigaldie, S., G. Labaune and J.P. Moreau. J. Appl. Phys., 52, 7114 (1981); Larigaldie, S., J. de Physique. Collque No. 7, C7-429 (1979).
50. Borgade, M.C. and G. Hartmann. J. Appl. Phys., 53, 8568 (1982).
51. Larigaldie, S. "Simulation D'un Foudroiement-Analyse des Mecanismes Associes a la Decharge." ONERA, RF 5/7236 PY, 25 Août, 1982.
52. Hemmati, M. "The Exact Solution of the Electron-Fluid Dynamical Equations." Ph.D. Dissertation. Univ. of Oklahoma, 1983.
53. Wagner, K.H. Z. Phys., 241, 258 (1971).
54. Raether, H. "Electron Avalanches and Breakdown in Gases." Butterworths, Washington, D.C., 1964.

55. Kunhardt, E.E. in Electrical Breakdown and Discharges in Gases. E.E. Kunhardt and L.H. Luessen, Eds. NATO ASI Series, Vol. 89a, Plenum Press, New York, 1983.
56. Parks, G.K., R.J. Spiger, B.H. Mank and J. Chin. EOS Trans. Amer. Geophysical Union, 61, 978 (1980).
57. Blythe, A.R. and G.E. Carr. J. of Electrostatics, 10, 321 (1981).
58. Pitts, F.L. AIAA 19th Aerospace Sciences Meeting. Paper No. 81-0083. (Conference held in St. Louis, MO, January 12-15, 1981).
59. Balamin, K.G., M. Gossland and R.D. Reeves. Proc. Intl. Symp. on Spacecraft Materials in Space Environment. CNES, ESA and CERT. Toulouse, France, 8 to 12, June 1982 (ESA-SP-178).
60. Johnson, C.T., A. Lee and G.S. Smeltzer. J. Appl. Phys., 52, 6543 (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

Laser-Induced Helium Plasma

Prepared by: Stanley Bashkin
Academic Rank: Professor
Department and University: Physics, University of Arizona
Research Location: Air Force Geophysics Laboratory, OPR Branch, Hanscom Air Force Base, Massachusetts
USAF Research: Dr. Russell Armstrong
Date: August 4, 1983
Contract No: F49620-82-C-0035

Laser-Induced Helium Plasma

by

Stanley Bashkin

ABSTRACT

A spectroscopic study was made of a laser-induced helium plasma. The intensities of several spectral lines were studied as a function of gas pressure and time subsequent to the initiating laser pulse. The widths of some lines were measured as a function of time. The information was used to determine the electron temperature and density as a function of time, and also to identify certain restrictions on the kinds of transitions which can occur in the plasma. From the latter, some problems which appear in the literature have been solved. The essential feature of the solution is that angular momentum must be conserved, a requirement which limits the final states which can be formed.

ACKNOWLEDGMENTS

It gives me great pleasure to acknowledge my indebtedness to Drs. Russell Armstrong, William Blumberg, and Jonathan Lurie, with whom I was closely associated during the fellowship period, and with whom I hope to have further research interaction. Needless to say, those activities could not have taken place without the sponsorship of the Air Force Systems Command, Air Force Office of Scientific Research, and the good offices of the Air Force Geophysics Laboratory, OPR Branch, at Hanscom Air Force Base.

I. INTRODUCTION

One of the important areas of study by the Optical Physics Research (OPR) branch of the Air Force Geophysics Laboratory (AFGL) is the emission of optical and infrared radiation from the upper atmosphere. The radiation arises from atoms, molecules, and ions of the atmospheric gases, stimulated by the passage of particles, photons, and man-made objects through that region.

In order to evaluate the kinds of optical/infrared transitions which occur in the upper atmosphere, OPR has developed apparatus intended to simulate or otherwise represent the appropriate events. This apparatus includes gas cells into which electrons or laser light can be injected, and spectrometers/interferometers for making measurements on the radiations which the electrons or lasers generate. It is possible to make measurements as a function of type of gas, gas pressure, injected power, electron energy, and time. This last is especially useful when the initiating energy source is pulsed.

My own work in the past twenty years or so has been in accelerator-based atomic physics, with particular emphasis on the light which is radiated when fast atomic ions are sent through thin foils. This so-called method of "beam-foil spectroscopy" is my major individual contribution to modern atomic research. Among other things, it has made possible the measurement of the radiative lifetimes of excited states in multiply-ionized atoms, and similar measurements when the emissions lie in the vacuum ultraviolet. The seventh international conference on this subject will be held at Oxford in the summer of 1984. I organized the first and third such conferences, and have served on the program committees of all the others.

In addition, a colleague (Professor John O. Stoner, Jr.) and I have published four large volumes of energy-level and Grotrian diagrams for all optical emissions from all stages of ionization for the first 25 chemical elements.

It was the combination of my laboratory and collative experience in atomic spectroscopy which led OPR to feel that I could be of assistance in their program. At the same time, the problems on which they were working had a certain appeal to me because there were difficulties in interpreting the spectroscopic information which had been obtained at OPR or had been published in the open literature. Those difficulties posed a challenge which I thought was intriguing.

The foregoing comity of interests and backgrounds made it reasonable that I meet the responsibilities of my fellowship at AFGL/OPR.

II. OBJECTIVES OF THE RESEARCH EFFORT

In the course of my pre-summer visit, I was shown both published and unpublished data which Dr. Russell Armstrong of OPR and some external colleagues had taken on the time-dependence of the emissions from laser-induced plasmas in nitrogen, oxygen, and argon. Those data looked to be rather complicated, but the unpublished results suggested to me that there might be some interesting multiplicity-dependent effects which could be used to interpret the results. However, because of the complexity of the systems which had been investigated, I proposed that we try a similar experiment, but on the atomically simpler gas, helium.

This suggestion was accepted, and Dr. Jonathan Lurie was assigned to work with me. By the time I arrived for the summer's activities, he had already taken some preliminary data, which enabled us to refine the experimental technique for the problem at hand.

On the basis of the early studies, we were able to establish these goals:

1. Measure the spectral emissions in the wavelength range from 350 to 600 nm as a function of gas pressure in the range from 50 to 400 torr.
2. Observe the intensities of certain spectral lines as a function of time subsequent to the laser pulse.

3. Determine the electron temperature and density in the plasma as a function of time and various physical parameters.
4. Study the influence of laser power.
5. Survey the literature on helium plasmas, and interpret our results in the light of what was already known.
6. Apply our results and the analysis thereof to the development of further experiments and, of course, to the original work on the atmospheric gases.

III. APPROACH TO THE OBJECTIVES

Light from a Nd:YAG laser at 1.06 microns was focused into a cell which was filled with helium gas to the desired pressure. This light caused the helium to enter the plasma phase and emit light of its own. That light was examined in two different ways. In one, an optical multi-channel analyzer was used in intervals of 16 nm in order to obtain the emission spectrum. In the other, the light was brought into a Spex Czerny-Turner scanning monochromator so that the behavior of individual spectral lines could be examined.

The laser was pulsed with a pulse width of 10 ns and a repetition rate of 100 per sec. This made it possible to follow the intensity of selected spectral lines as a function of time. These measurements constituted the main thrust of the work, and were made as a function of gas pressure. The most useful data were obtained in the pressure range between 100 and 400 torr.

In addition, several spectral lines, especially 468.6 nm from the $n = 4$ to $n = 3$ transition in He II and 587.5 nm from $2p \ ^3P^0 - 3d \ ^3D$ in He I, were scanned, so as to determine the electron temperature and density as a function of time.

It quickly developed that this approach was entirely satisfactory, and excellent data were obtained, permitting us to satisfy the first four goals mentioned in II. At the same time, goal 5 was met, although the literature on

He plasmas includes several hundred papers. We were able to study perhaps 50 of the most significant papers. In doing so, it appeared that our experiment was virtually unique, in that only one or two other authors had reported on the phenomena we found to be interesting. Moreover, the previous work along these lines was either neglected, as regards fundamental analysis, or confused, because of inadequate experimental breadth. Hence, we were in the position of making a significant contribution to the physics of the helium plasma.

The literature also included some puzzles, in that certain cross sections were found to be zero, the rates of some other reactions unusually small, and there were apparent contradictions in a number of papers. We were then in the position of being able to use our findings to clarify all of the foregoing, so that goal 5 was more than met.

Goal 6 remains to be met, but it is our hope that at least part of it will be achieved in the course of preparing the results of the summer's work for publication in a research journal. That writing is now in progress.

IV. NEW OBJECTIVES

The first new objective is to complete a paper describing our work and submit it to an appropriate scientific journal for publication. The first draft of such a paper has now been finished, and the second draft is in progress.

The second objective is to take advantage of our experiment so as to improve the quality and extend the quantity of the data. This can be done, for example, by modifying the optics, particularly the match of f /numbers between the collimating lenses and the spectrometer. Such a modification would permit working at lower pressures and at later times, while still giving good signal/background data. Another modification that would be helpful is installing electronics to give digital rather than just analogue output from the detector.

A third objective is to repeat the experiment with hydrogen, rather than helium. It is not clear whether hydrogen or helium is the simpler system as regards our plasma conditions, but, nonetheless, the extension of the measurements to hydrogen would give important information regarding the model we have developed to explain the helium results.

The fourth and final objective is to repeat the experiment with atmospheric gases. Those gases are, of course, of primary interest to AFGL/OPR, but it seemed sensible to defer their study until the work on helium and hydrogen was finished. That viewpoint prevails in the light of the summer's experience, so the investigation of nitrogen and oxygen remains to be carried out.

V. RECOMMENDATIONS

a. The most important recommendation is that the new objectives of Section IV be given high priority. The physics is of great interest and the scientific feasibility of the measurements has been clearly demonstrated. The main handicap is a shortage of manpower that can be assigned to such work. My colleague, Dr. Jonathan Lurie of AFGL/OPR, has been given additional responsibilities, and it is not clear that he could devote adequate time to the plasma research. On the other hand, the apparatus is relatively simple to use, and my suggestion is that two skilled experimenters be brought into the laboratory as temporary personnel for the purpose of doing this work. I estimate that a month's hard work by two people could meet at least the second and third of the new objectives, so it would not be particularly expensive. If SCEEE could finance such a project, I would be willing to assist in finding the proper people to deal with this problem.

b. In discussions with Dr. Armstrong and his colleague, Dr. William Blumberg, it developed that there would be considerable AFGL/OPR interest in supporting a new project in my home laboratory at the University of Arizona, and it is my intention to request support from a mini-grant so that

such a program could be started. This new work would deal with the possibility that heavy particles - i.e., protons and elements of higher charge - could generate radiations in the infrared region of the spectrum when passing through atmospheric gases. Since protons and helium nuclei with energies in the neighborhood of 1 MeV impinge upon the upper atmosphere and participate in auroral events, there is an obvious application to understanding processes that take place in that part of the earth.

In order to look at these possibilities, particles from the smaller (2 MV) of the two Van de Graaff accelerators in my laboratory would be energized to about 1 MeV and sent into suitable gas targets. Radiations would be analyzed with a Fabry-Perot interferometer. The accelerator, with all necessary auxiliary apparatus, is available, and Drs. Armstrong and Blumberg have agreed to lend me an interferometer adequate for exploratory studies. I would be joined in this endeavor by two Arizona colleagues - Professor Laurence C. McIntyre, who is highly experienced in both accelerator and atomic experimental work, and Professor John O. Stoner, Jr., who is a fine spectroscopist with excellent experience with interferometers as well as with particle accelerators. Thus we believe that we would be in a good position as regards the proposed investigations.

c. Finally, I would urge that the SCEEE program be expanded so as to enable vigorous prosecution of those researches which appear to be especially promising. Ten weeks are, after all, rather a short time for completing a serious task. In those instances where at least a good start has been made, it would be valuable if SCEEE were to be able to support a major effort to ensure full exploitation of the potential.

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 LIQUID SLOSHING IN SPIN-STABILIZED SATELLITES

Prepared by: Joseph R. Baumgarten, Professor
Daniel E. Hill, Research Engineer

Department & University: Mechanical Engineering Department
Iowa State University

Research Location: Arnold Engineering Development Center
Engine Test Facility
Arnold AFS, TN 37389

USAF Research: Dr. John T. Miller
Aeropropulsion Programs Department
ETF/AEDC

Date: July 29, 1983

Contract No: F49620-82-C-0035

INVESTIGATION OF LIQUID SLOSHING IN SPIN-STABILIZED SATELLITES

By

Joseph R. Baumgarten

and

Daniel E. Hill

ABSTRACT

Certain configurations of spin-stabilized spacecraft consistently develop a coning or nutating motion during the perigee burn. This motion consists of sinusoidal oscillations about the pitch and yaw axes at the same frequency, but with a 90° phase difference. The sloshing of liquid fuel stores is suspected as a source of these nutations. The moving liquid in its spherical containers has been modeled as an equivalent pendulum, pivoted with the main body of the payload, and moving relative to it with rotating constraint. The equations of motion of the spacecraft with a pendulum system have been derived. Numerical solution is accomplished on the digital computer. Comparison is made to flight test data of actual spacecraft.

ACKNOWLEDGMENTS

The authors wish to acknowledge the consultation and assistance of Dr. John T. Miller, Senior Lead Engineer, the AEDC Research Colleague, during this investigation. Mr. Joseph R. Parker, Jr., Program Manager, Analysis and Evaluation, Aeropropulsion Programs Department, is acknowledged for his logistic help and support in the completion of this study. This research was supported by the Air Force Office of Scientific Research/Air Force Systems Command, United States Air Force, under contract F49620-82-C-0035. This investigation was completed at AEDC, Arnold AFS, Tennessee.

I INTRODUCTION

Launchings of several of the STAR 48 Communication Satellites from the Space Shuttle have consistently resulted in a nutating motion of the spacecraft. Flight data from roll, spin, and yaw axis rate gyros indicate a constant frequency, equal amplitude, sinusoidal oscillation about the yaw and pitch axis. The vector combination of these two components of vibration results in a coning motion of the satellite about its spin axis. The vehicle is spin stabilized at launch, having a 1 rev/sec spin velocity imparted to it.

After launching from the Shuttle, in the perigee phase of its orbit, the satellites power assist module (PAM) fires its thruster to establish a geosynchronous earth orbit. It is this axial thrust that gives rise to the coning which predominates after PAM-motor burnout. Consistently, flight data from rate gyros indicates the steady-state coning and a 0.5 cps small amplitude disturbance superimposed on the 1 rev/sec spin velocity.

Combustion instabilities in the PAM rocket motor were thought to be the source of a side force which would induce the coning motion. In order to investigate the presence of any such combustion instabilities, a STAR 48 motor was fired at the Engine Test Facility, AEDC, Arnold Air Force Station. A test rig having lateral and axial load cells was utilized, and the rig allowed the PAM to be spun at 1 rev/sec during the firing. A spectral analysis was completed of the resulting load cell records obtained during firing. A spectral analysis was completed of the resulting load cell records obtained during firing. The test results indicated no significant forces at the required frequency (one half cycle per second) and it was concluded that combustion instabilities could not be the source of moments about the principal axes of the spacecraft which caused coning motion.

Dr. John T. Miller, Sverdrup Technology, Inc., AEDC Group, had followed the static test of the STAR 48 motor. He completed an analysis of the payload (the communication satellite), and found that a 55 ft-lb external moment at one half cycle/sec would sustain the coning motion.

He suspected that sloshing motion of liquid stores in the vehicle is the mechanism for creating and sustaining the nutation of the spacecraft. The authors, having had some experience in modeling of sloshing fluids {1,2}* made application through SCEEE for summer support in order to join in this investigation with Dr. Miller, the AEDC Research Colleague.

II OBJECTIVES

Certain configurations of spin stabilized spacecraft consistently develop a coning or nutational motion during the launch phase. This motion consists of sinusoidal oscillations about the pitch and yaw axes at the same frequency, but with 90° phase difference. The sloshing of fluid fuel stores is a probable source of these nutations. The object of this study is to develop the dynamic response equations of an equivalent spherical pendulum model of the liquid and container. The pendulum model would be refined by comparing the numerical solution of the response equations with flight data from roll, pitch, and yaw gyros. With a satisfactory model of the structure and its stored liquid, a parametric study could be completed of the important design parameters which affect the amplitude and frequency of the nutation.

III THEORY

This investigation has been initiated in order to study the general problem of the dynamic effects of moving parts on the motion of a spin-stabilized spacecraft. The problem has been formulated from various points of view by Roberson {3}, Grubin {4}, Kane and Sobala {5}, and Edwards and Kaplan {6}. Robinson modeled a rigid main body with an arbitrary number of moving components. He chose the composite center of mass of the system as the reference point. This formulation resulted in time varying moments of inertia and a moving reference point. Grubin avoided this problem by choosing the vehicle center of mass as the reference point. He could easily identify the instantaneous position of the moving mass with reference to the vehicle. Kane and Sobala investigated the problem of attitude control through controlled motion of an internal mass

*Numbers in brackets designate references

in the spacecraft. Edwards and Kaplan studied the detumbling of a spacecraft by the programmed motion of a movable internal mass. The equations of motion of the spacecraft were derived with the origin of a coordinate set fixed in the principal coordinates of the main mass of the spacecraft. Thus, a fixed reference point was defined at the body center of mass and the instantaneous position of the movable mass was defined relative to the main mass. This is the approach taken here in this present study.

Let a satellite with some even number of semi-spherical fuel tanks be represented as shown in Figure 1a. The sloshing fuel can be represented in this first approximation as a single degree of freedom pendulum with some equivalent mass oscillating at some identifiable radius {1,2}. Let the instantaneous angular positions of two diametrically opposite pendulums be given by α and β . Referring to Figure 1b, let the roll axis spin velocity, ω_x , control the instantaneous orientation of the fuel tanks in the y-z plane, and $\sigma = \omega_x t$. If diametrically opposite tanks have instantaneous relative positions which are 90° out of phase then

$$\alpha = a \sin \omega_f t, \quad \beta = b \cos \omega_f t \quad (1)$$

These vectors combine to one unbalanced pendulum having an oscillation given by

$$\delta = (a^2 + b^2)^{1/2} \sin(\omega_f t + \psi) \quad (2)$$

with the forcing frequency given by ω_f . Then, referring to Figure 1c, a single equivalent sloshing tank with mass m_e results, having a sloshing amplitude γ which oscillates at a forcing frequency ω_f . The instantaneous position of the mass m_e is given by the coordinates

$$x = d + l \cos \delta \quad (3)$$

$$y = (e + l \sin \delta) \sin \sigma \quad (4)$$

$$z = (e + l \sin \delta) \cos \sigma \quad (5)$$

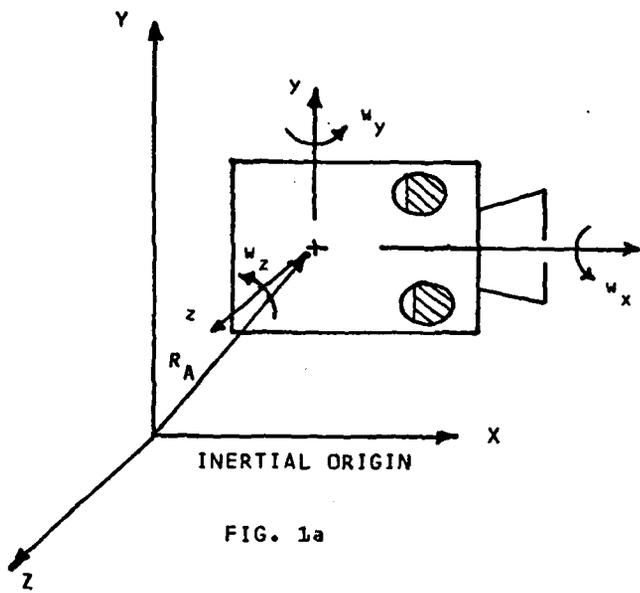


FIG. 1a

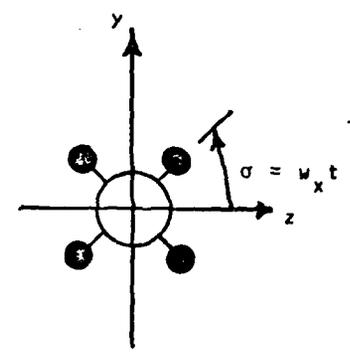


FIG. 1b

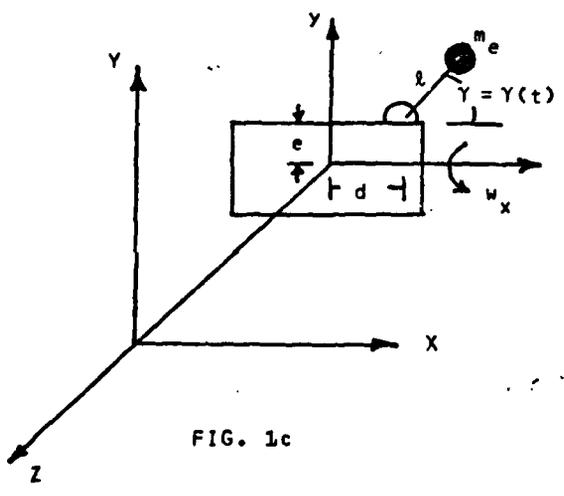


FIG. 1c

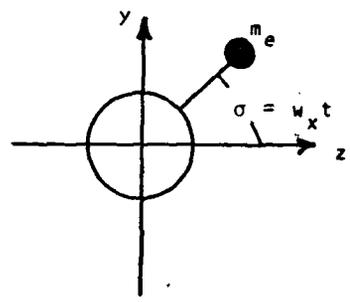


FIG. 1d

The oscillating amplitude δ is assumed to be small and the horizontal position x essentially constant. Thus

$$\dot{x} = d + l$$

$$\ddot{x} = 0$$

$$\ddot{y} = 0 \quad (6)$$

$$y = (e + l\delta) \sin \omega_n t \quad (7)$$

$$\dot{y} = l\dot{\delta} \sin \omega_n t + (e + l\delta) \omega_n \cos \omega_n t$$

$$\ddot{y} = (l\ddot{\delta} - (e + l\delta) \omega_n^2) \sin \omega_n t + 2l\dot{\delta} \omega_n \cos \omega_n t \quad (8)$$

$$z = (e + l\delta) \cos \omega_n t \quad (9)$$

$$\dot{z} = l\dot{\delta} \cos \omega_n t - (e + l\delta) \omega_n \sin \omega_n t \quad (10)$$

$$\ddot{z} = (l\ddot{\delta} - (e + l\delta) \omega_n^2) \cos \omega_n t - 2l\dot{\delta} \omega_n \sin \omega_n t \quad (11)$$

With this definition of the kinematics of the equivalent pendulum, one can now proceed to define the dynamic response of the spacecraft system. The work of Edwards and Kaplan {6} serves as basis for this analysis.

The summation of torques with respect to the inertial origin is given by

$$\Sigma \bar{M} = \dot{\bar{H}} + s \times \bar{a} \quad (12)$$

where H is the angular momentum, s the static moment of the body with respect to some arbitrary point, and a is the acceleration of that point. When the reference point is the system center of mass, the second term of equation (12) disappears and the torque summation becomes

$$\Sigma \bar{M} = \dot{\bar{H}} \quad (13)$$

We define the mass ratio of the combination of vehicle mass M and equivalent fluid pendulum mass m_e as

$$\mu = \frac{m_e M}{M + m_e} \quad (14)$$

and the moments of inertia of the spacecraft about the principal body fixed axes as I_x , I_y , and I_z . Then, equation (13) yields a set of three coupled nonlinear differential equations in terms of the vehicle angular velocities and the pendulum mass position, thus (recalling that the x position is essentially constant)

$$\begin{aligned} & (I_x + \mu(y^2 + z^2))\dot{\omega}_x + (I_z - I_y + \mu(y^2 - z^2))\omega_y \omega_z \\ & + \mu(-x\dot{y}\dot{\omega}_y + (2y\dot{y} + 2z\dot{z})\omega_x + yz(\omega_z^2 - \omega_y^2) - xz(\dot{\omega}_z + \omega_x \omega_y) \\ & + x y \omega_x \omega_z + y\ddot{z} - z\ddot{y}) = 0 \end{aligned} \quad (15)$$

$$\begin{aligned} & (I_y + \mu(x^2 + z^2))\dot{\omega}_y + (I_x - I_z + \mu(z^2 - x^2))\omega_z \omega_x \\ & + \mu(-y z \dot{\omega}_z + x y (\dot{\omega}_x + \omega_y \omega_z) + 2z(\dot{z}\omega_y - z\dot{y}\omega_z) - 2\dot{y}x\omega_x \\ & + x z (\omega_x^2 - \omega_z^2) + y z \omega_x \omega_y - x \ddot{z}) = 0 \end{aligned} \quad (16)$$

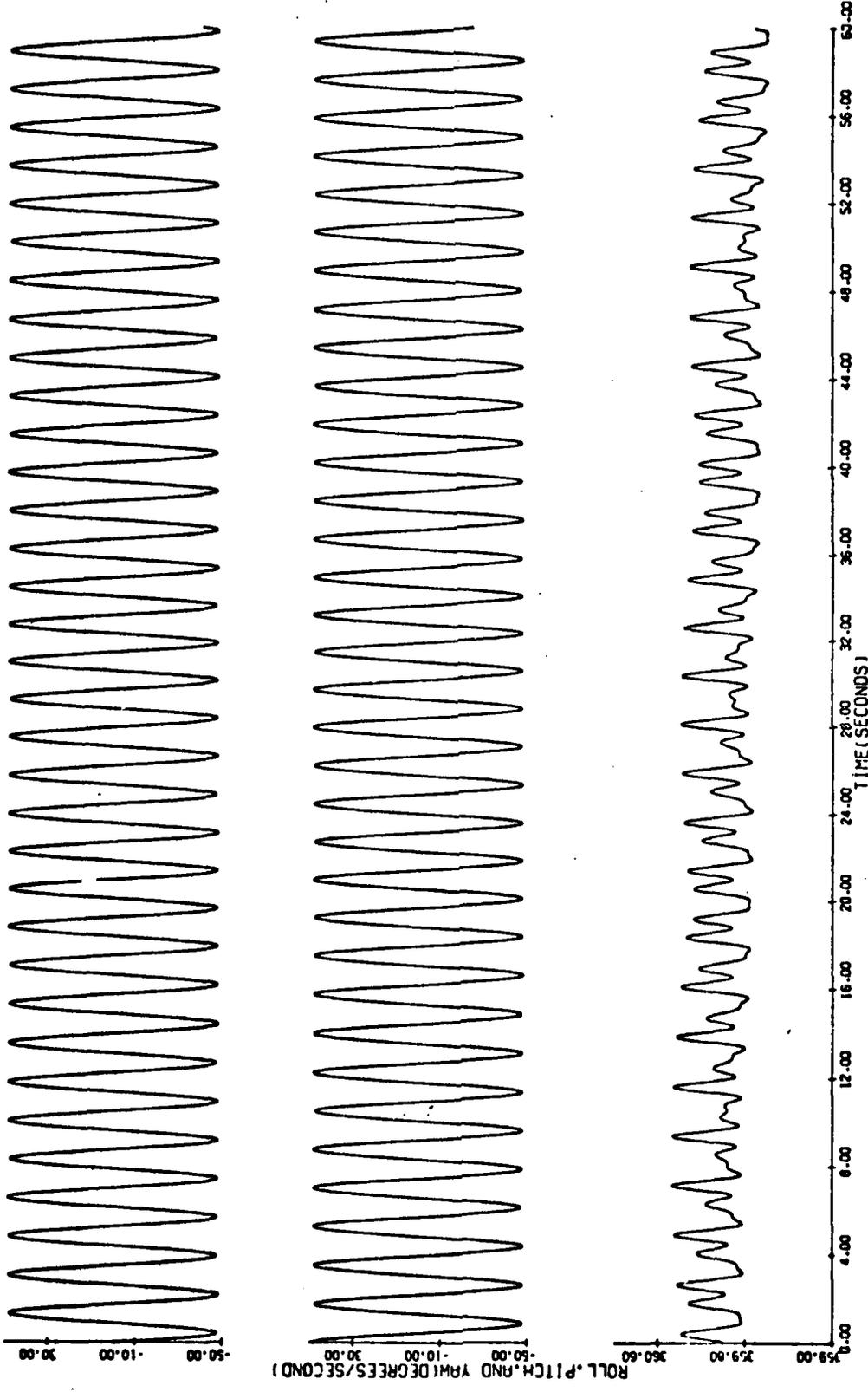
$$\begin{aligned}
& (I_z + \mu(x^2 + y^2))\dot{\omega}_z + (I_y - I_x + \mu(x^2 - y^2))\omega_x\omega_y \\
& + \mu(-x\dot{z}\dot{\omega}_x - y\dot{z}\dot{\omega}_y + xy(\dot{\omega}_y^2 - \dot{\omega}_x^2) + 2y(\dot{y}\dot{\omega}_z - \dot{z}\dot{\omega}_y) \\
& - 2x\dot{z}\dot{\omega}_x - y\dot{z}\dot{\omega}_x\dot{\omega}_z + x\dot{z}\dot{\omega}_y\dot{\omega}_z + x\dot{y}) = 0 \quad (17)
\end{aligned}$$

IV RESULTS

The system is given its initial spin velocity, ω_x , and the fluid sloshing is given a small amplitude of oscillation γ at a forcing frequency ω_f . The dynamic response equations of the vehicle have been programmed for solution on the digital computer and a parametric study of the physical constants of the vehicle has been initiated. The phase of the satellite motion being studied here is that which occurs after the completion of the booster rocket burn. In this phase, the principal inertias of the spacecraft are constant, the axial thrust is now zero, and a steady state sloshing can be applied as the sole dynamic effect on the vehicle.

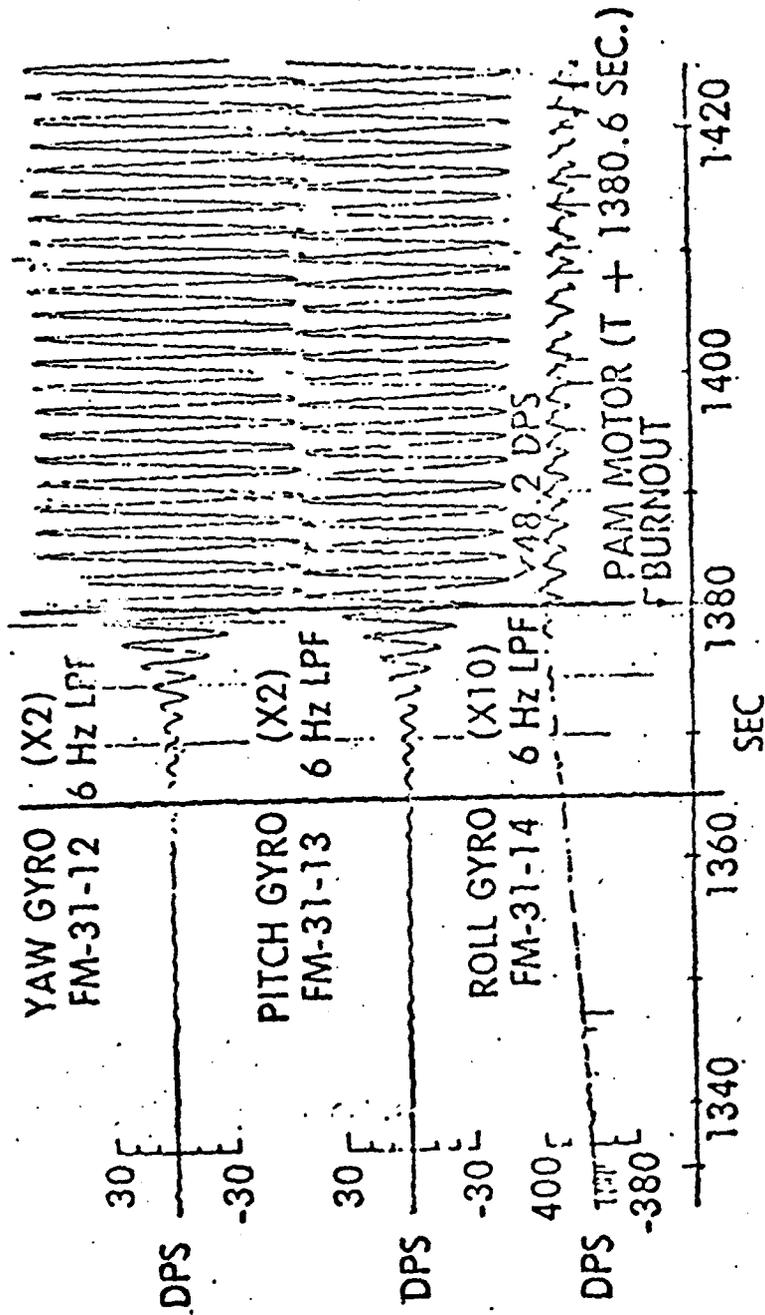
The digital computer solution of equations (15), (16), and (17) is shown plotted in Figure 2. The roll axis velocity ω_x is seen to be initialized at its 1 rev/sec spin velocity. As time advances, a small amplitude periodic variation in spin velocity of 0.5 Hz is seen to attain. The pitch and yaw axis velocities ω_z and ω_y are seen to develop a steady state amplitude of approximately 50 deg/sec and the periodic oscillations are seen to be out of phase by 90°.

These results agree reasonably well with Figure 3, the flight test data from the roll, pitch, and yaw gyros of the RCA-C vehicle launch. This simple mechanism, an equivalent planar pendulum, satisfactorily models the effect of the sloshing fuel stores on board the spacecraft.



CONING SIMULATION
FIGURE 2

DEFINITION OF CONING PROBLEM



FLIGHT DATA - RCA - C'

FIG. 3

V RECOMMENDATIONS

It is recommended that the study of variation of design parameters be expanded to include the effect of varying fuel levels in the storage tanks. The assumption of a planar pendulum representation of the sloshing fluid can be refined; a two degree of freedom spherical pendulum model can be constructed for the fuel stores.

Having a working mathematical model of the sloshing fluid in the spacecraft, many design considerations can now be pursued. The authors plan to propose to AFOSR the investigation of passive control of the sloshing fluid by means of tank interconnection and by valving. Models of a control system for amplitude and phase control of tank sloshing are deemed possible in view of the success achieved with the modeling results presented in Figure 2.

These results suggest that further pursuit of the causes for coning need not be made. Investigation of combustion instabilities are no longer necessary and redesign of nozzle pins is not indicated to be necessary. Schemes of baffeling of the fuel tanks and of slosh control by porting of tanks should be pursued.

REFERENCES

1. Sayar, B. A., and Baumgarten, J. R. "Pendulum Analogy for Nonlinear Fluid Oscillations in Spherical Containers", ASME TRANS, Journal of Applied Mechanics, December 1981, Vol. 48, pp 769-772.
2. Sayar, B. A. and Baumgarten, J. R. "Linear and Nonlinear Analysis of Fluid Slosh Dampers", AIAA Journal, Vol. 20, No. 11, Nov. 1982, pp 1534-1538.
3. Roberson, R. E., "Torques on a Satellite Vehicle from Internal Moving Ports", Trans. ASME, Journal of Applied Mechanics, Vol. 25, June 1958, pp 196-200.
4. Grubin, C., "Dynamics of a Vehicle Containing Moving Parts", Trans. ASME, Journal of Applied Mechanics, Vol. 29, September 1962, pp 486-488.
5. Kane, T. R., and Sobala, D., "A New Method for Attitude Stabilization", AIAA Journal, Vol. 1, No. 6, June 1963, pp 1365-1367.
6. Edwards, T. L., and Kaplan, M. H., "Automatic Spacecraft Detumbling by Internal Mass Motion", AIAA Journal, Vol. 12, No. 4, April 1974, pp 496-502.

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

Searching for Precursors to Laser-Induced Damage

Prepared Jointly by: Dr Michael F Becker

Academic Rank: Associate Professor

**Department and University: Department of Electrical Engineering
University of Texas at Austin**

and

Dr Fred E Domann

Academic Rank: Associate Professor

**Department and University: Department of Physics
University of Wisconsin - Platteville**

**Research Location: Optical Coating Damage Laboratory
Air Force Weapons Laboratory
Kirtland AFB, New Mexico**

USAF Research: Dr Alan Stewart

Date: 1 August 1983

Contract No.: F49620-82-C-0035

Searching for Precursors to Laser-Induced Damage

by

Michael F. Becker

and

Fred E. Domann

Abstract

Several experimental techniques are investigated to try to find a precursor to permanent damage caused by laser radiation. These techniques include charge emission, neutral particle emission, surface photoconductivity, and surface potential measurements. The techniques are applied to a wide variety of optical materials: diamond-turned copper, silicon, and various dielectric films on fused silica substrates. Charge emission was found to be a useful technique, and its extension to multi-pulse experiments is recommended. Neutral particle emission was too insensitive to be useful. Photoconductivity data were gathered, but an ultrahigh vacuum system in conjunction with photon cleaning is probably necessary to gather meaningful data. Suggestions for further investigations are offered, including a surface potential probe.

Acknowledgements

The authors hereby acknowledge the Southeastern Center for Electrical Engineering Education; the Air Force Office of Scientific Research; the Air Force Weapons Laboratory, Kirtland Air Force Base, New Mexico; and Dr Arthur Guenther for providing the opportunity of making this an interesting and productive summer.

We would also like to thank Dr Alan Stewart for his helpful advice and generous contribution of his time and effort. We thank the rest of the staff of the Optical Coating Damage Laboratory for their assistance whenever it was needed.

Finally, one of us, F.E.D., would like to thank his wife, Lou Anna, and daughter, Kate, for relieving him of his usual summer duties.

MICHAEL F BECKER
Associate Professor
Department of Electrical
Engineering
University of Texas at Austin

FRED E DOMANN
Associate Professor
Department of Physics
University of
Wisconsin - Platteville

I. Introduction.

In the years that laser-induced damage has been studied, two important, fundamental questions still remain unanswered.¹ (1) What microscopic mechanism or mechanisms lead to laser damage? (2) Is there a reliable (non-destructive) precursor to laser damage? These two questions are clearly related, and an understanding of one will aid in understanding the other. In this summer research project, we will concentrate on the second question and seek precursor information in laser damage using a variety of new techniques. The study is to be carried out using a wide variety of samples, and the results will be correlated with other commonly used diagnostics for laser damage.

II. Objectives.

The problem of finding a precursor to laser damage will be addressed using the following experimental techniques: (1) charge emission, (2) neutral particle emission (detected by rapid pressure fluctuations), (3) surface photoconductivity, and (4) surface potential changes. The apparatus for each of these diagnostics is to be constructed, tested, and applied to as wide a range

of samples as possible. Ideally, samples should include metals, semiconductors, insulators, and dielectric thin films. A description of the apparatus and the experimental results for each of the four diagnostics will be found in each of the four subsequent sections of this report.

III. Charge Emission.

The emission of negative charge into vacuum during laser irradiation was the first diagnostic to be studied in this project. However, we will first describe the characteristics of the laser system which are common to all of the experiments which were conducted.

The laser was a Quantel Nd:YAG laser oscillator amplifier system. The 1.06 micron pulses were 5.07 ± 1.0 ns long and were produced by passive Q-switching. The pulse profiles were measured with a fast vacuum photodiode and were always smooth Gaussians. After amplification, the pulses passed through an attenuator consisting of a rotating half-wave plate and a polarizer. At the sample location, the laser beam was brought to a focus by a 2 m focal length lens. The focal plane of the lens was scanned with a narrow (<50 micron) slit in order to

determine the beam spot size on the target. By this method, the beam radius at the focus was found to be 180 microns. In addition, the beam profiles were reasonably Gaussian. The laser energy in the focal plane was measured by an energy meter which sampled part of the incident beam. This meter was calibrated against a second meter placed behind the focal plane. Using all the preceding information, the energy, the energy density (fluence), and the intensity (power/area) of the laser pulse incident on the sample could be determined for every laser shot.

A variety of samples were used in these experiments: metals, semiconductors, and dielectrics. Diamond-turned copper mirrors were tested as being representative of the best metal optics. Single-crystal silicon wafers polished on both sides represented the semiconducting materials, and fused silica substrates coated with half-wave thick dielectric coatings represented the wide band-gap insulators. Two high quality coating materials were tested in these experiments: Ta_2O_5 and Al_2O_3 . These dielectric films were also tested for photoconductivity, and these experiments will be described in Section V.

During the course of the work looking for precursors to damage, the thresholds for single-pulse damage were measured for all of the optics which we tested. These results are shown in Table I.

Sample	Single-Pulse Damage Threshold	Extremes of Damage-No-Damage
Silicon Wafer	1.6 J/cm ²	1.1-2.18 J/cm ²
Copper, Diamond- Turned	2.1 J/cm ²	2.0-2.4 J/cm ²
Cr-26 Film	10.4 J/cm ²	9.8-12.0 J/cm ²
Ta ₂ O ₅ Film	~ 10 J/cm ²	9.4-10.8 J/cm ²
Al ₂ O ₃ Film	Uncertain	50-125 J/cm ²

Table I. Single-pulse damage threshold for all the samples which were tested (except for the photoconductivity samples, see Section V.)

The main emphasis of this section is on the relation between charged particle emission and laser-induced damage. The data were collected by placing the samples in a vacuum of about 10⁻⁶ Torr. A charge collecting wire was placed about 2 cm above and slightly to the side of the sample surface. The wire was charged to +1,500 volts (sufficient to collect all emitted negative charge), and the negative charge pulses were amplified by an Amptek A-

203 and A-206 charge sensitive amplifier chip set. The sensitivity was calibrated as 1.75×10^{13} volts/coulomb. In practice, the ultimate sensitivity was limited by induced noise to about 10^{-15} coulomb, and the dynamic range was slightly more than two orders of magnitude. The charge emission waveform was then acquired by the Tektronics transient recorder and computer. We should point out that judging whether a sample is damaged or not is a controversial subject. The criterion we used was that damage occurred if any visible change could be observed under an optical Normarski microscope at up to 400x.

In general, the results of the charge emission experiments fell into two categories:

(1) No charge emission is observed until the material surface is damaged, or

(2) The surface emits charge prior to damage but there is no obvious correlation between charge emission and subsequent damage.

Silicon and the Cr-26 dielectric film fell into category 1, while copper and the remaining dielectrics fell into

category 2. For category 1, the quantity of charge emitted was always so great as to saturate the charge collection electronics.

Some experiments were conducted in which more than one pulse was incident on the same site. In this case, all the materials in category 2 showed a "cleaning" effect where the charge emission eventually became undetectable after a sufficient number of pulses of the same energy were incident on the site. Presumably, this effect is due to laser cleaning of surface contaminants by evaporation and ionization. The quantity of charge emitted showed no correlation with laser energy in all cases except one. For copper, the charge emission prior to damage was noisy but shows a monotonic increase with increasing laser fluence. This result is plotted in Figure 1. Because of the scatter in the data, there is no way to discriminate between a power law dependence of emitted charge on fluence or an inverse exponential dependence. Theoretically, the former dependence is associated with multiphoton processes, and the latter is associated with thermal emission over an energy barrier.

The morphologies of the damaged samples were studied using Normarski microscopy. In general, there was more

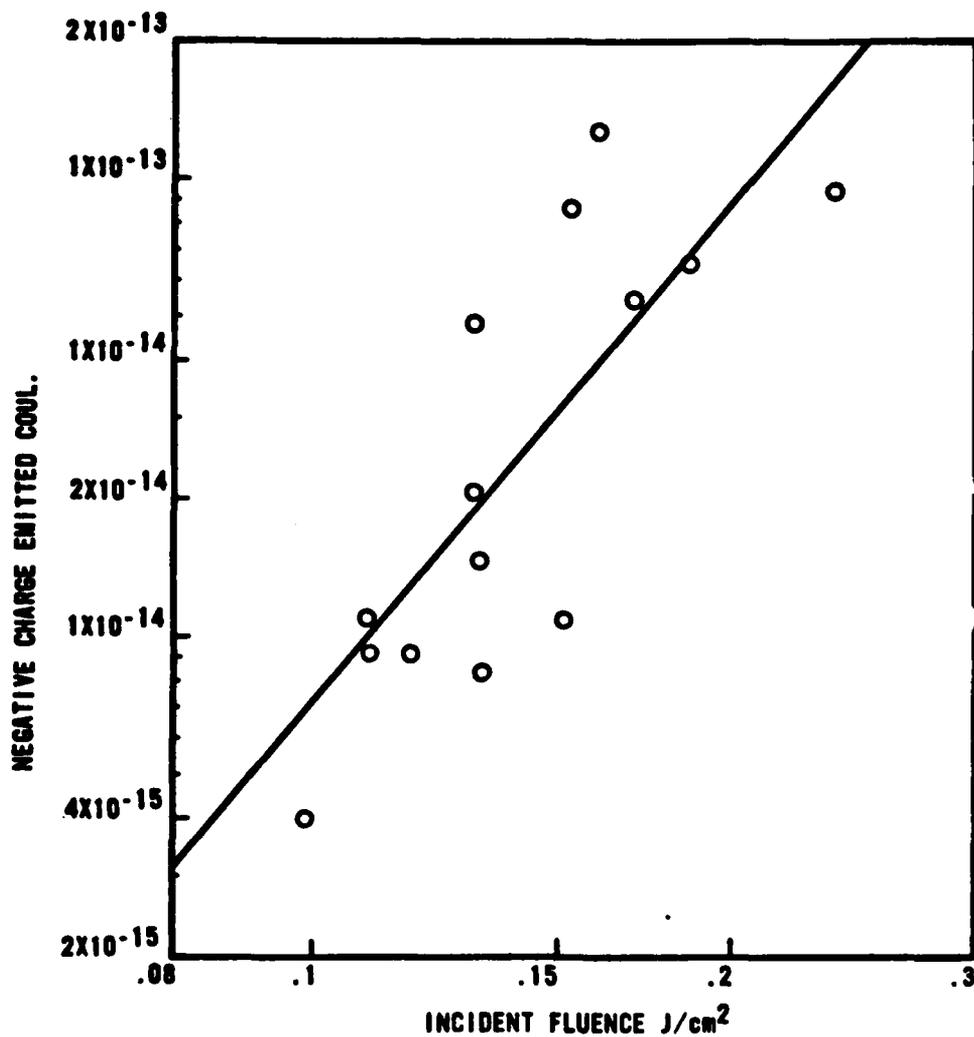
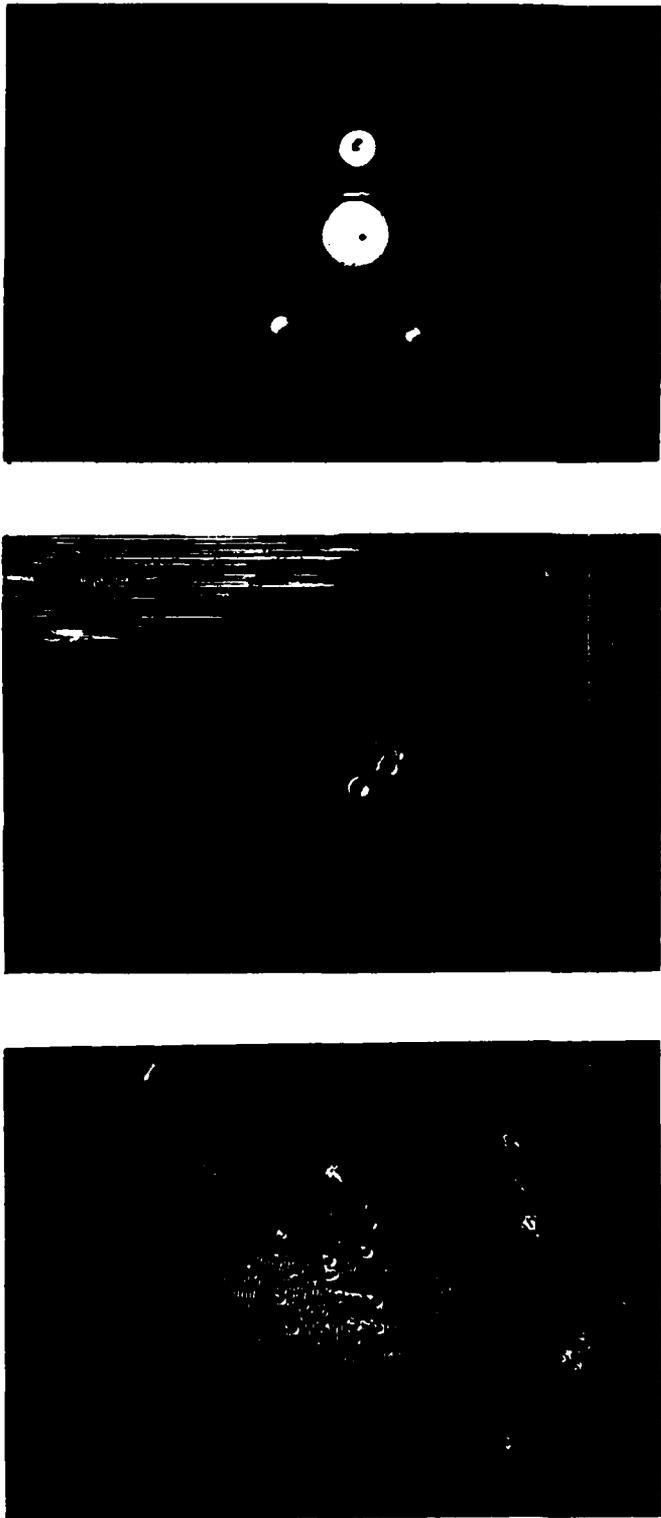


Figure 1. Negative charge emitted from a diamond-turned copper mirror versus laser fluence. Each point represents a single laser shot on a fresh site at an energy well below damage.

information to be gained from the more pure and ordered materials; i.e., more structure was visible on single-crystal silicon than on polycrystalline copper or on glassy dielectrics. Emphasis was placed on near damage threshold events only.

Typical results showed that the copper formed small melt pits on damage with no apparent correlation between the pits. At higher fluences, more uniform melting resulted in ripple and diffraction ring formation in addition to melt pits. Silicon showed a wealth of information and gave distinctly different results for single-pulse and multi-pulse damage. For single-pulse damage, shallow features are formed: grooves, ripples, and diffraction rings. But for multi-pulse damage, only correlated chains of pits are formed, and each chain always contains more than one pit. This damage morphology has been observed and models for it discussed for the case of picosecond laser pulses.² The high quality dielectric films damaged in what appear to be circular regions. These regions are of high contrast under the Normarski. A limited number of examples of the damage morphologies we have observed are shown in Figure 2.



a.

b.

c.

Figure 2. Typical single shot near threshold damage morphologies, Normarski micrographs.

a. Copper mirror, 524x, 5.60 J/cm²;

b. Silicon, 262x, 2.62 J/cm²;

c. Ta₂O₅ film on fused silica, photoconductivity sample, 262x, 21.6 J/cm².

In conjunction with the experiments on silicon, the nonlinear absorption of the 1-mm-thick wafers was measured at fluences up to the single-pulse damage threshold. The data are plotted in Figure 3 such that if two-photon absorption is present, a linear plot will result. This effect shows clearly in the figure. Unfortunately, the sample had very parallel faces, and 1.06 microns fell on an etalon resonance. If this were not the case, the nonlinear absorption coefficient could be easily determined from the slope of the curve.

In summary, we find charge emission to be an indicator of surface cleanliness only for the first laser pulse. In some materials, it is a reliable indicator of the first damage event. Work remains on the significance of charge emission in the multi-pulse regime.

IV. Neutral Particle Emission.

Neutral particle emission data were collected from one sample each of copper and silicon. Neutral particles were detected with a fast ionization gauge Model FIG-1, manufactured by Beam Dynamics, Inc. The operating principles of the fast ionization gauge are identical with those of a standard Bayard-Alpert ionization gauge, but the dimensions

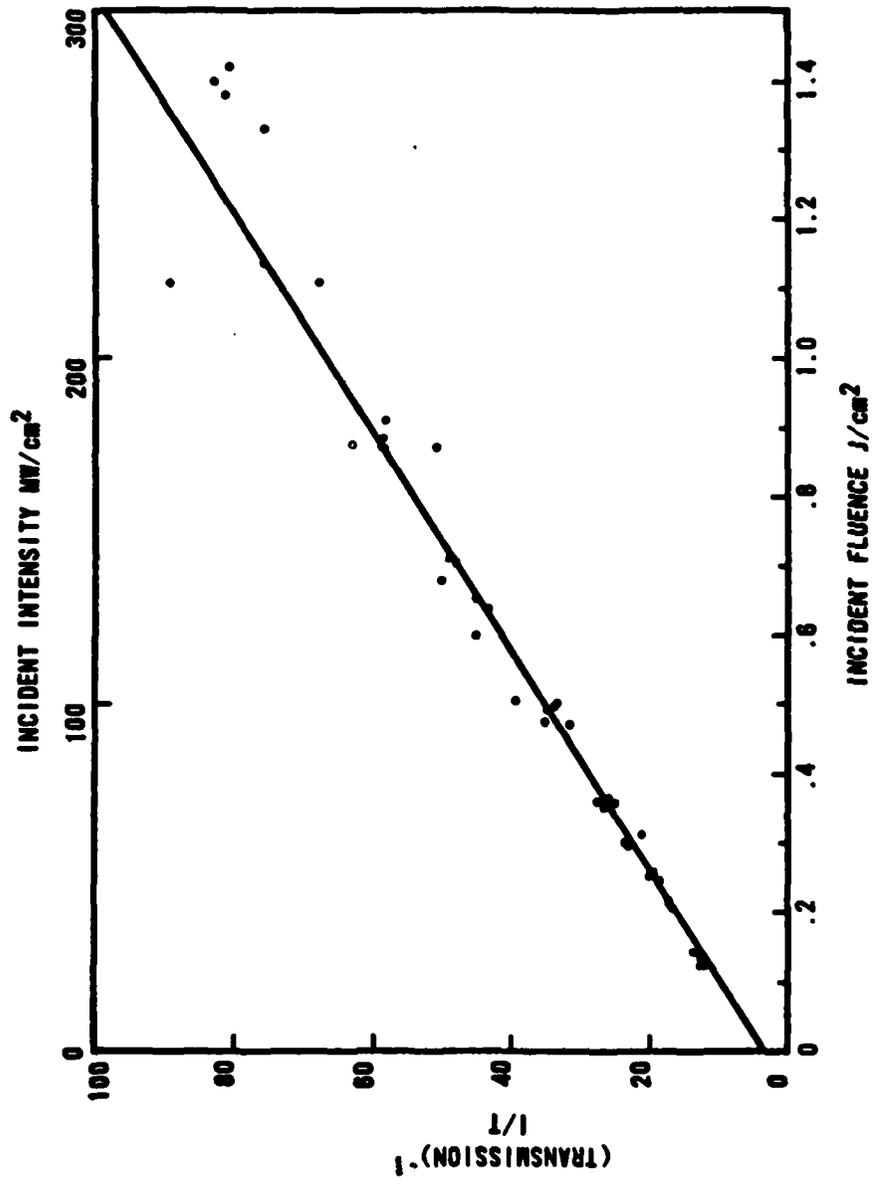


Figure 3. Inverse of sample transmission versus laser fluence for a 1-mm-thick silicon sample. Each point represents a single laser pulse, and the damage threshold is at the right edge of the figure. The laser pulse length was 5 nsec.

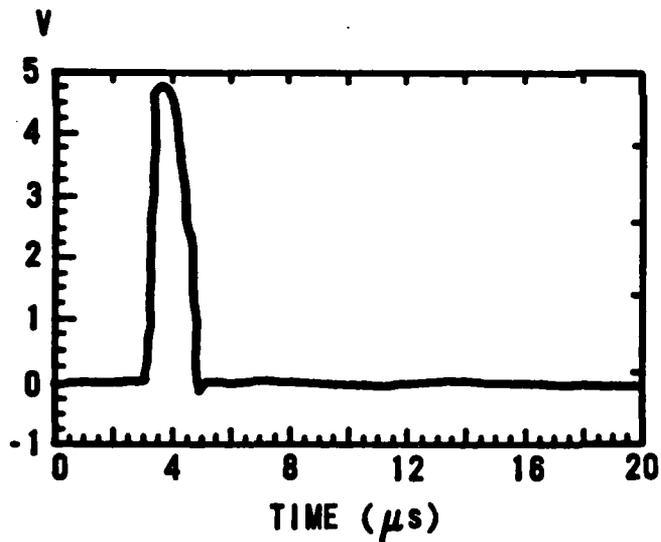
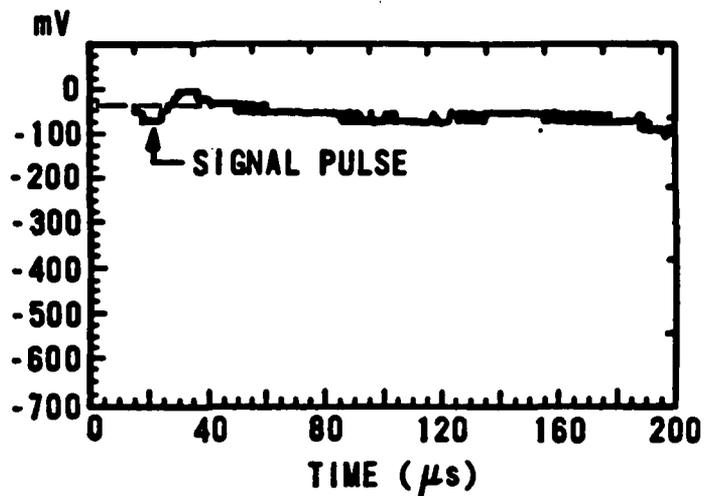
of the fast ionization gauge are significantly smaller, thus reducing the ion travel times from the grid region (where the ionization occurs) to the collector. The fast ionization gauge was controlled with a Model FC-1 Control Unit which regulated the emission current (continuously adjustable, 5 μ A - 3MA) and provided a DC voltage output signal (4mV - 10V). The response time of the device was 3 μ s, and the sensitivity was 1×10^5 volt Torr⁻¹ per mA of emission current. During these measurements, the emission current was nominally adjusted to 0.5 mA.

Figure 4 shows a typical output that was obtained when a copper sample was irradiated with a 3.1 J/cm² pulse, which is well above the single-shot damage threshold for copper at 2.1 J/cm². The small neutral particle emission in Figure 4a can be compared to the saturated ion emission signal shown in Figure 4b.

Similar results were obtained for neutral particle emission from silicon.

The reason for the small signal was no doubt due to the small solid angle subtended by the fast ionization gauge, together with its small ionization volume (330 mm³). This

(a) Neutral particle emission signal obtained with the FIG.



(b) Ion emission signal.

Figure 4. Neutral particle and ion emission from copper.

diagnostic was deemed too insensitive, based on the above results, and was not used in subsequent experiments.

V. Photoconductivity.

Photoconductivity data were collected for three samples of Al_2O_3 and one of Ta_2O_5 . Gold on chromium conducting strips 0.045 inches wide and $0.1\mu m$ thick were vapor-deposited on top of the dielectric coatings. The sample was mounted in a ceramic (MACOR by Corning) holder shown in Figure 5, and electrical connections to the conducting strips were made with gold-plated, electromechanical relay contacts. The conducting strips were deposited such that four different spacings between strips, called channels, resulted. These channel widths were 0.08, 1.21, 1.65, and 2.03 mm. Voltages were applied to the conducting strips such as to produce electric fields from about 5.0×10^4 V/m to 1.7×10^6 V/m. Charge flowing into the positive electrode was sensed with the charge amplifier described in Section III (Charge Emission). The output of the charge amplifier was sent to the transient digitizer.

Since the noise level of the charge amplifier output was approximately 0.020 mV and the digitizer saturated at approximately 2.2 V, the dynamic range of charge detecting

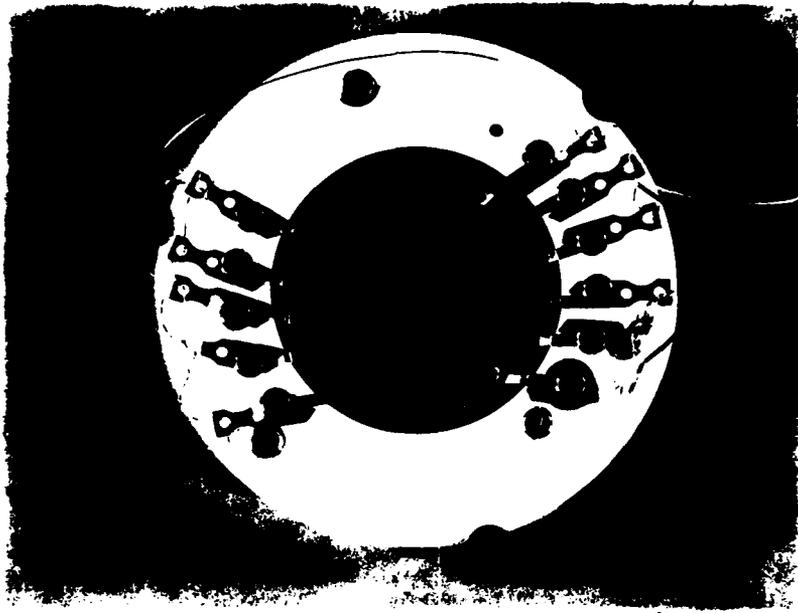


Figure 5. A photograph of a sample mounted in the ceramic sample holder. The small, evenly spaced dots across one of the conducting strips are laser-damage spots used as markers.

system covered about two orders of magnitude. It turns out that this range was not adequate to measure the amount of charge collected as a function of laser radiation for laser fluences adequate to cause damage. For laser fluences that resulted in an unsaturated charge collection signal, an attempt was made to find a functional relationship between the charge collected and laser fluence. The results for several samples (Al_2O_3 and Ta_2O_5) are shown in Table II. It was found that these data were uncorrelated on a linear, log-log, or exponential graph.

Laser Fluence (J/cm^2)	1.97	6.68	7.19	6.12	5.29	4.77	6.61	5.95	5.46	5.76	4.8
Charge Collected ($\times 10^{15}$ coulombs)	8.7	37.7	100	49.0	12.0	13.0	130	77.0	14.3	77.0	44.0

(a)

Laser Fluence (J/cm^2)	0.58	0.56	0.55	0.54	0.56	0.58	0.64	0.62
Charge Collected ($\times 10^{14}$ coulombs)	8.84	2.0	4.13	1.57	7.4	8.8	9.69	0.713

(b)

Table II. (a) Charge collected versus laser fluence
for Al_2O_3
(b) Charge collected versus laser fluence
for Ta_2O_5

The data in Table II show that for a given laser fluence, Ta_2O_5 conducts charge much more readily than Al_2O_3 , and that for either type of sample the charge conduction varied widely from site to site for a given radiation fluence.

In fact, for one Al_2O_3 sample (which was irradiated at 1.0 Atm), there was no apparent predamage charge conduction.

It was also found that for those samples of Al_2O_3 that were irradiated in a vacuum chamber ($P = 2 \times 10^{-6}$ Torr), predamage charge collection was readily apparent. This raises the question of whether the charge collected was actually due to the photoconductivity or charge emission. Since no provision was made in the experiment to distinguish between the two, one must conclude that either could have produced the observed signal.

To summarize the photoconductivity experiment then, several key items should be pointed out: (1) The dynamic range of the charge collecting instrumentation was too small to determine if a relationship between collected charge and laser fluence existed; (2) the data were very

site-specific, probably due to surface contaminants; and (3) some experimental provision should be made to determine whether the collected charge is due to photoconductivity or charge emission (see Recommendations).

VI. Electrostatic Probe.

It is possible that the accumulation of surface charge or surface dipoles or the change in surface potential might be a precursor to laser damage. An instrument to measure surface potential was researched, designed, and constructed during the course of this work. It utilizes a vibrating capacitor probe to measure surface potential. A recent method to do this is described in reference 3, and this type of device is reviewed in reference 4. Surface potential is sensitive to surface charge, surface dipoles, and adsorbed impurities. For this reason, it might be useful in the study of laser damage.

The probe utilizes an electromagnetic drive coil with a probe tip 1 mm in diameter. The tip may easily be vibrated from .1 to .25 mm excursion for voltages between 5 and 12 Vrms. A plan view of the probe is shown in Figure 6.

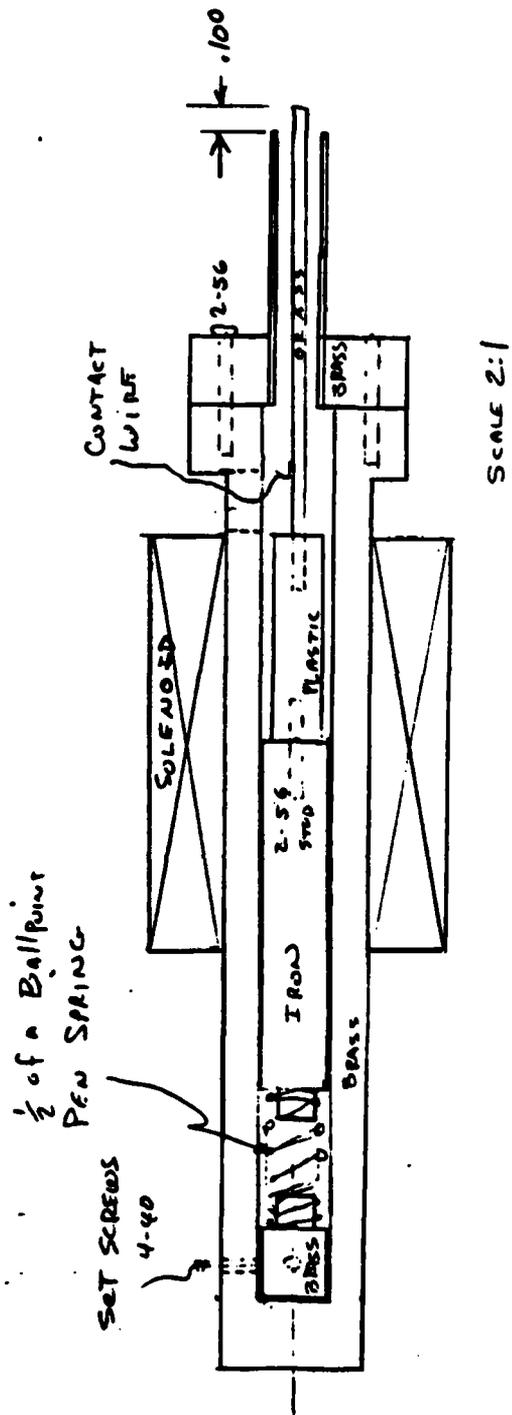


Figure 6. Plan view of the electrostatic vibrator probe for measuring surface potential.

At the end of the project, the probe had not been tested with a laser-damaged sample. The continuation of this investigation will be treated in Section VII.

VII. Recommendations.

Charge emission has been shown to be a useful diagnostic in laser damage studies. Some improvements can still be made, and we make the following recommendations:

a. It is recommended, of course, to raise the saturation level of the charge amplifier so that the dynamic range well exceeds the level at damage threshold.

b. It is recommended that multi-pulse experiments be conducted to separate cleaning effects from damage and pre-damage phenomena.

c. It is recommended that an energy analysis system be added to the system to facilitate photoelectron or photoion spectroscopy. Data obtained from such experiments may be very useful in determining the energetics of laser damage and lead to useful microscopic models.

d. It is recommended that a sensitive mass spectrometer be added to the vacuum chamber to help identify the liberated species. The mass spectrometer would also be useful to determine the cleanliness of the surface.

It is still uncertain whether surface conductivity will be a useful diagnostic for laser damage. Because conductivity is significantly effected by surface contaminants⁵ it is recommended that conductivity measurements similar to those described in Section V be carried out in a vacuum chamber capable of reaching pressures well below those achieved with the present system. To estimate a reasonable background pressure, we may use the fact that t_1 is approximately $10^{-6}/P$, where t_1 is the time in seconds for one monolayer of a contaminant to adsorb on an atomically clean surface and P is the background pressure in Torr. Thus, for example, if one assumes a perfectly clean surface in a background pressure of 10^{-9} Torr, one monolayer of contaminants will be adsorbed in approximately 10^3 seconds. At these pressures, a sample could be cleaned via photon desorption while in the vacuum chamber and afford the experimenter ample time to collect data.

Finally, we recommend that tests using the surface potential measuring device be conducted as this technique still looks promising.

References

1. The field is well documented by the proceedings of the 1st through 14th Annual Boulder Damage Symposia which are printed as NBS Special Publications.
2. See for example: M. F. Becker, R. M. Walser, Y-K. Jhee, and D. Y. Sheng, "Picosecond Laser Damage Mechanisms at Semiconductor Surfaces," SPIE Vol. 322, Picosecond Lasers and Applications, p. 93, Bellingham, Wash., SPIE (1982) and the references contained therein.
3. J. C. Campuzano and R. G. Greenler, "Instrument for Combining Reflection-Absorption Infrared Spectroscopy with Other Surface-Sensitive Techniques," Rev. Sci. Instrum. 52, May 1981, pp. 678-683.
4. N. A. Surplice and R. J. D'Arcy, "A Critique of the Kelvin Method of Measuring Work Functions," Journal of Physics E: Scientific Instruments 3, 1970, pp. 477-482.
5. K. L. Mittal, ed., Surface Contamination Vol. 2, New York, Plenum Press, 1979, p. 967.

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

PROPAGATION LOSS IN ELECTROSTATICALLY

VARIABLE SAW DELAY LINES

Prepared by: Henry L. Bertoni

Academic Rank: Professor of Electrophysics

Department and University: Department of Electrical Engineering and Computer Science, Polytechnic Institute of New York

Research Location: Rome Air Development Center, Electromagnetic Sciences Division, Antennas and RF Components Branch, RF and SAW Components Section

USAF Research Alan J. Budrean

Date: August 26, 1983

Contract No: F49620-82-C-0035

PROPAGATION LOSS IN ELECTROSTATICALLY VARIABLE

SAW DELAY LINES

by

Henry L. Bertoni

ABSTRACT

Variations in SAW time delay due to an applied D.C. electric field has been suggested as a means for achieving the phase shift needed for each radiating element in phased-array antennas. Important device parameters are: 1) insertion loss, and; 2) sensitivity of delay change to D.C. terminal voltage. These parameters can be in competition, in that device modifications intended to improve one parameter may make the other worse. These parameters are considered for the normal-field configuration, and for the in-plane configuration in which both D.C. electrodes are on the same surface as the SAW. The in-plane configuration can have high voltage sensitivity if the electrodes are close together, but diffraction loss for this case is found to be large for long path lengths. Methods for limiting diffraction loss without reducing voltage sensitivity are investigated.

ACKNOWLEDGMENTS

The author wishes to thank Alan J. Burdrear and Paul H. Carr for suggesting the topic of this report and for sharing their ideas and experimental results. The author also wishes to thank Andrew J. Slobodnik, Jr. and Jose H. Silva for their support and technical discussions. The financial support of the Air Force Systems Command, Air Force Office of Scientific Research and the Rome Air Development Center is gratefully acknowledged.

I. Introduction

The phase shifters currently used in phased array antennas are typically of the latching ferrite type. The cost of these phase shifters constitutes a significant fraction of the cost of the antenna. Alternative means of achieving phase shift that is less expensive is therefore highly desirable. It has been found that strong D.C. electric fields change the time delay in surface acoustic wave (SAW) devices [1-3]. This effect has the potential for providing phase shift from inexpensive devices fabricated by photolithographic techniques.

The first configuration of D.C. electrodes to be studied [1,2] employed an electric field normal to surface on which the SAW propagates. The field was generated by a voltage applied to a metallic film deposited on the surface. Because of the thickness of the piezoelectric crystal plate needed for mechanical strength, rather high terminal voltages were needed to achieve significant velocity change.

Budreau, et.al. [3] proposed the in-plane configuration, in which D.C. voltage is applied between two electrodes, both deposited on the SAW surface and lying on either side of the propagation region. Because the electrode gap can be made much less than the plate thickness, this configuration can have higher velocity sensitivity to D.C. terminal voltage. For convenience in fabrication the electrodes were in the form of bars approximately 50 SAW wavelengths wide.

As a result of attenuation in the metal film, the electrodes affect the SAW propagation as if they were semi-infinite. Shorting of the R.F. electric field by the electrodes slows the SAW propagating beneath them. Thus, the SAW propagates from input to output transducer in a high velocity channel or gap between two semi-infinite regions of low velocity. This arrangement is opposite to that known to act as a waveguide [4,5]. For the long paths needed to achieve voltage sensitivity, the high velocity gap may have large diffraction loss. This loss is expected to increase with decreasing gap width.

II. Objectives of the Research Effort

The purpose of this study is to investigate the tradeoffs between propagation loss and voltage sensitivity, and to seek improvements that can be achieved by modification of the electrode geometry. Transducer bandwidth requirements strongly favor the high coupling cuts of LiNbO_3 , for which the relation between velocity change and electrostatic field has been obtained experimentally. For this reason, we do not consider improvements in voltage sensitivity that might result from using other materials.

Consideration is first given to the normal field configuration. It is shown that raising the electrode slightly off the surface reduces SAW attenuation. However, voltage sensitivity is decreased.

Diffraction loss for the in-plane configuration is then investigated. Scalar wave theory is used to represent the two dimension SAW propagation. An expression is derived for the diffraction loss due to wide electrodes. This expression confirms the expectation of high loss for long, narrow gaps.

Two methods for reducing diffraction loss are considered. One method employs electrodes whose width is one half of an appropriately defined wavelength. To first order, these electrodes are invisible to the SAW, so that the diffraction loss is nearly that of a free surface.

Another method considered to reduce diffraction loss employs an overlay of acoustically fast material deposited on top of the electrodes. For sufficiently thick overlay, the SAW under the electrodes becomes fast as compared to the gap region, so that the configuration acts as a slot waveguide [4,5]. This case involves a mode coupling loss that is independent of gap length, rather than diffraction loss.

III. Basic Device Considerations

The performance of electrostatically varied delay lines involves the three factors, dispersion, insertion loss and voltage sensitivity. Of the three, dispersion appears to be the least significant for application to phased array radar, since serious degradation of the antenna performance occurs only for velocity changes greater than about 10% over the operating band. Insertion loss and voltage sensitivity are critical factors, and changing device parameters to improve one factor frequently makes the other worse.

The size of the three factors, and some suggestions for improvement are discussed below for both the normal field and in-plane configurations.

A. Normal-Field Configuration

The normal field configuration is shown in Fig. 1, including a modified electrode configuration. This modification employs a top electrode separated from the surface of the piezoelectric crystal by a distance H. In the usual device, the top electrode consists of a plating of thickness h deposited directly on the crystal surface (H=0).

1. Dispersion

For thin platings, dispersion is proportional to ωh , where ω is the radian frequency ($\omega = 2\pi f$), as discussed in Ref. [6]. A conducting plating can slow the Rayleigh wave as a result of mass loading, over and above the slowing due to shorting of the piezoelectric field. The slowing due to mass loading by a 600 Å thick Al plating on LiNbO_3 is given by

$$\frac{\Delta V}{V_R} = \begin{cases} \left(\frac{f}{f_0}\right) 2.03 \times 10^{-3} & (16 \text{ } 1/2^\circ \text{ doubly rotated}) \\ \left(\frac{f}{f_0}\right) 3.142 \times 10^{-3} & (37^\circ - X) \end{cases} \quad (1)$$

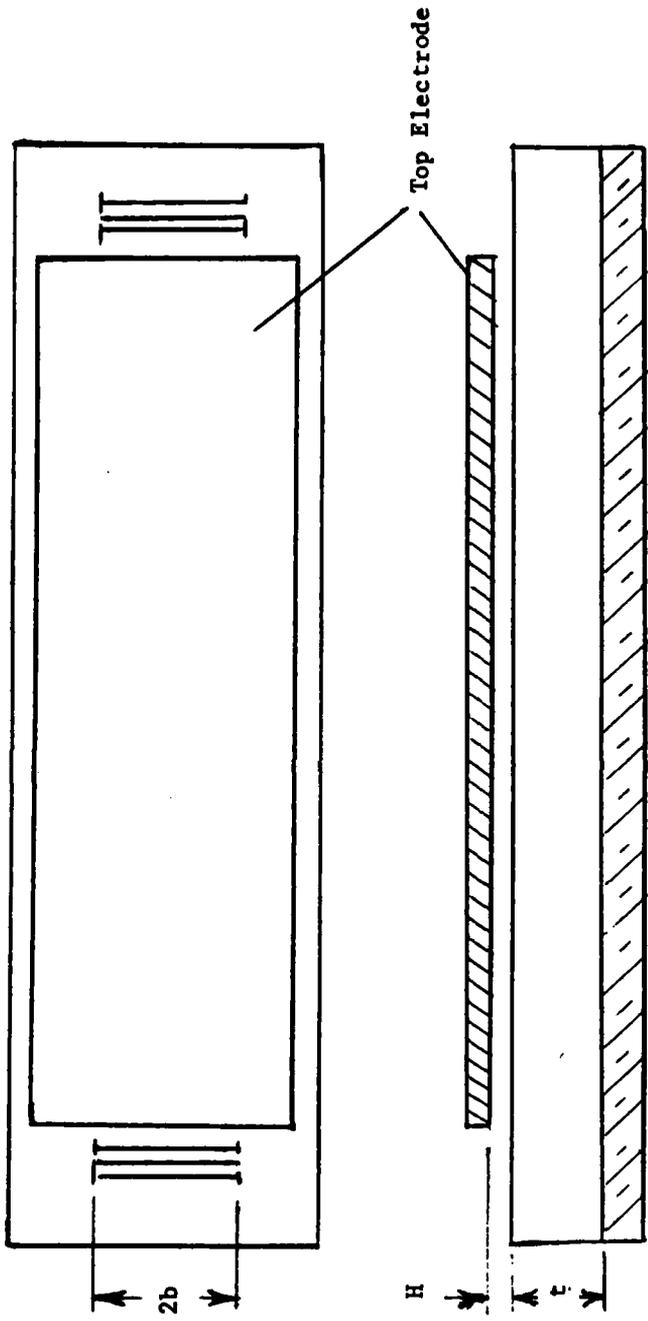


Fig. 1 Top and side views of the normal-field electrode configuration

where $f_0 = 1$ GHz. It is seen that even for a 1 GHz bandwidth the dispersion is only a few tenths of a percent, and hence not significant for radar applications. For the raised electrode configuration ($H \neq 0$), no dispersion will occur.

2. Insertion Loss

Insertion loss for the normal field configuration consists of the transducer and diffraction loss found in any delay line, plus the loss in the plating for the case $H=0$. This configuration places no special limit on the aperture $2b$ of the interdigital transducers, so that diffraction loss can be made small by choosing the aperture large enough.

The presence of a film deposited on the crystal will cause loss that is approximately proportional to f^2 . Using a laser probe Davis and Weller [7] measured attenuation for frequencies in the range 270-730 MHz for Al films 500, 1000 and 2000 Å thick on YZ L_1NbO_3 . Their data for the best films approximately fits the expression

$$\alpha = 9.5 \left(\frac{f}{f_0} \right)^{2.2} \left(\frac{h}{600} \right)^{.4} \text{ db/cm} \quad (2)$$

where h is in Angstroms and $f_0 = 1$ GHz. For comparison, the free surface attenuation on L_1NbO_3 is given in Ref. [8] as

$$\alpha = 2.5 \left(\frac{f}{f_0} \right)^2 + 0.54 \left(\frac{f}{f_0} \right) \text{ db/cm} \quad (3)$$

Thus at $f = 1$ GHz, a good quality film of thickness 600 Å increases the attenuation by 6.5 db/cm.

Using the concept of viscosity of the election gas in metal films, Snider, et. al. [9], have computed the attenuation due to thin Al films on L_1NbO_3 . For kh small, the attenuation is proportional to $\omega^2 h$. Scaling their calculations at 2 GHz leads to the expression

$$\alpha = 13.2 \left(\frac{f}{f_0} \right)^2 \left(\frac{h}{600} \right) \text{ db/cm} \quad (4)$$

This attenuation is somewhat greater than the measurements [7] and has a different power dependence on h . Their calculations do not include loss in the $L_1\text{NbO}_3$ crystal or dissipation in the film due currents generated by the electric field. In spite of these differences, (2) and (4) give roughly the same attenuation for 600 Å thick films.

Lifting the top electrode off the surface ($H \neq 0$) eliminates the loss due to mechanical vibration of the electrode material. For H greater than about a Rayleigh wavelength, the currents induced in the electrode, and hence dissipation, will be negligible. However, this reduction in loss will be accompanied by a reduction in sensitivity, as discussed below.

3. Voltage Sensitivity

Assume that the change in SAW velocity is due to the electrostatic bias field E_0 present within the crystal. Then, lifting the electrode off the surface will reduce E_0 , and hence reduce the change in velocity per applied volt. Since $E_0 = V_{DC}/(t + \epsilon H)$, where V_{DC} is the applied voltage and ϵ is the relative dielectric constant normal to the plate, lifting the electrode off the surface reduces the voltage sensitivity by the factor

$$\frac{t}{t + \epsilon H} = \frac{1}{1 + \epsilon H/t} \quad (5)$$

If, $H = 5 \mu\text{m}$ and $t = 10 \text{ mill} = 250 \mu\text{m}$, then using $\epsilon = \epsilon_{xx}^T = 84$ for $L_1\text{NbO}_3$ gives the value 0.37 for the factor (5), while a gap $H = 10 \mu\text{m}$ gives the value 0.23.

B. In-Plane Field Configuration

The in-plane configuration shown in Fig. 2 allows the D.C. bias electrode to be moved closer together in order to improve voltage sensitivity and to avoid the dispersion and loss associated with the deposited film of normal-field configuration [3]. If $2b < 2a$, then the film does not lie in the SAW path, and hence does not cause dispersion. However, the film does increase diffraction loss, which may become large as the electrode separation $2b$ is made small. Diffraction loss is considered after a preliminary discussion of the effects of geometry on the voltage sensitivity.

1. Geometrical Effects

The electric field between the electrodes is shown end on Fig. 3. Because the D.C. electric field is transverse to the saggital plane; it will affect the SAW velocity through a different combination of crystal constants than in the normal field configuration. The significance of this change in orientation has not yet been studied.

In addition to the orientation effect, the geometry of Fig. 3 will influence the voltage sensitivity through the variation of E_x over the range $-a < x < a$. A simple expression for E_x along the crystal surface is obtained from the solution for zero thickness electrodes that are semi-infinite in x , and assuming the principal axes of the crystal to be oriented along the coordinate axes of Fig. 3. With these approximations one has [10]

$$E_x = \left(\frac{V_{DC}}{\pi a} \right) \frac{1}{\sqrt{1 - (x/a)^2}} \quad (6)$$

For electrodes of finite thickness h , (6) is not valid for $a-x$ less than about h , and E_x remains finite as $x \rightarrow a$.

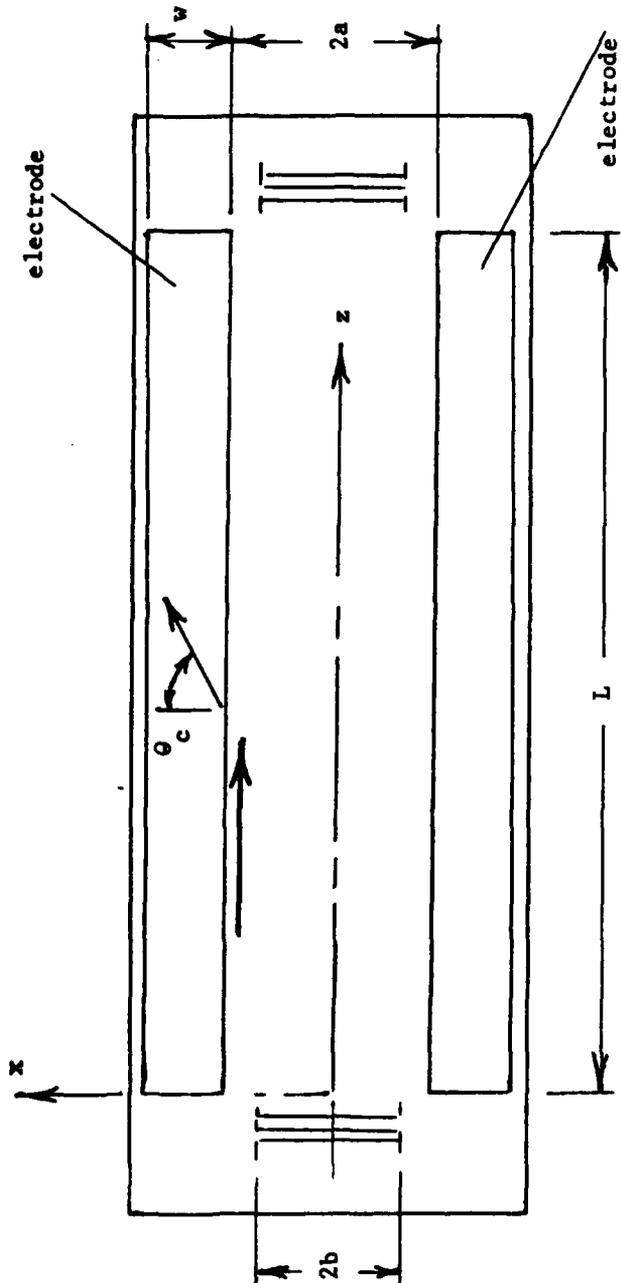


Fig. 2. Top view of the electrodes for the in-plane field configuration.

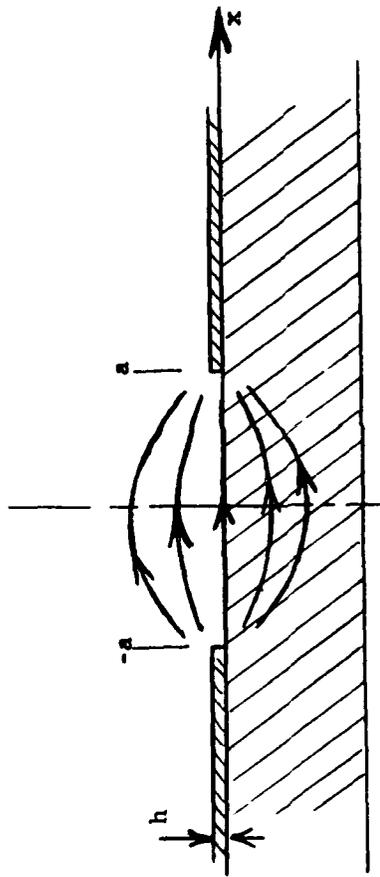


Fig. 3. End view of the in-plane field configuration showing the fringing static electric field.

The effect of the spatial variation of \underline{E} can be understood from Fig. 2. Suppose that the transducer width $2b$ is sufficiently large (or L is sufficiently small) so that SAW propagation may be taken as being parallel to the z axis. For each value of x , the SAW velocity $V_R(x)$ will be different due to the non-uniformity of E_x . Thus the phase of the SAW at the output transducer will depend on x . For small changes in velocity, the phase variation with x is given by

$$\phi(x) = \frac{\omega L}{V_R(x)} - \left(\frac{\omega L}{V_R} \right) \left[1 - \frac{\Delta V(x)}{V_R} \right]$$

where V_R is the free surface Rayleigh velocity. If $A(x)$ is the SAW amplitude distribution at the output transducer, the output voltage V_{RF} will be proportional to

$$V_{RF} \propto e^{ik_1 L} \int_{-b}^b A(x) e^{-ik_1 L \Delta V(x)/V_R} dx \quad (7)$$

where

$$k_1 = \omega/V_R. \quad (8)$$

Assuming that $\Delta V(x)$ is proportional to E_x of (6), then $\Delta V/V$ can be written

$$\frac{\Delta V(x)}{V} = \gamma \frac{V_{DC}}{2a} \frac{2}{\pi} \frac{1}{\sqrt{1-(x/a)^2}} \quad (9)$$

where γ is a phenomenological sensitivity parameter. Expression (9) is symmetric in x , so that assuming $A(x)$ is also symmetric, (7) reduces to

$$V_{RF} \propto 2e^{ik_1 L} \int_0^b A(x) \exp \left[-ik_1 L \frac{\Delta V(0)}{V_R} \frac{1}{\sqrt{1-(a/x)^2}} \right] dx \quad (10)$$

where

$$\frac{\Delta V(0)}{V_R} = \gamma \frac{V_{DC}}{2a} \frac{2}{\pi} \quad (11)$$

is the change of velocity along the centerline of the device.

The change in velocity is obtained experimentally by measuring the change in phase of the output voltage with applied D.C. bias voltage. As seen from (10), the change of output phase in fact represents a weighted average of the change in velocity. Two aspects of the integrand make the central portion of gap between the D.C. electrodes more important in the weighting than the region near the electrodes ($x \rightarrow a$). The first aspect relates to the variation of $A(x)$, while the second relates to the rapid phase variation as $x \rightarrow a$.

In general, $A(x)$ will decrease as the ends of the transducer are approached. For $a > b$, as in Fig. 2, the integration does not even approach the edges $x = a$ of the D.C. electrode. Even if $a = b$ and $A(a)$ is not small, the rapid variation of phase in the exponential as $x \rightarrow a$ will cause the integration over neighboring half periods to cancel. The primary contribution to (10) comes from a range of $x < a$ for which $1/\sqrt{1-(x/a)^2} \approx 1$. Thus we may approximate (10) as

$$V_{RF} \propto 2e^{ik_1 L} e^{-ik_1 \Delta V(0)/V_R} \int_0^b A(x) dx. \quad (12)$$

Expression (12) indicates that the apparent velocity change is approximately the value $\Delta V(0)$ along the center line. Thus, the electrode geometry reduces the voltage sensitivity by $2/\pi$ as compared to that of parallel plane electrodes having the same spacing.

To test the foregoing model, we have plotted the variation of $\Delta V(0)/V_R$ in (11) with gap width $2a$ in Fig. 4 superimposed on measured values of sensitivity ($\Delta V/V$ per volt) reported by Budreau, et. al. [3]. The amplitude of the computed curve has been normalized so as to pass through the measured point at $2a = 40 \mu\text{m}$. The agreement with the measured variation is quite good considering the simplicity of the model.

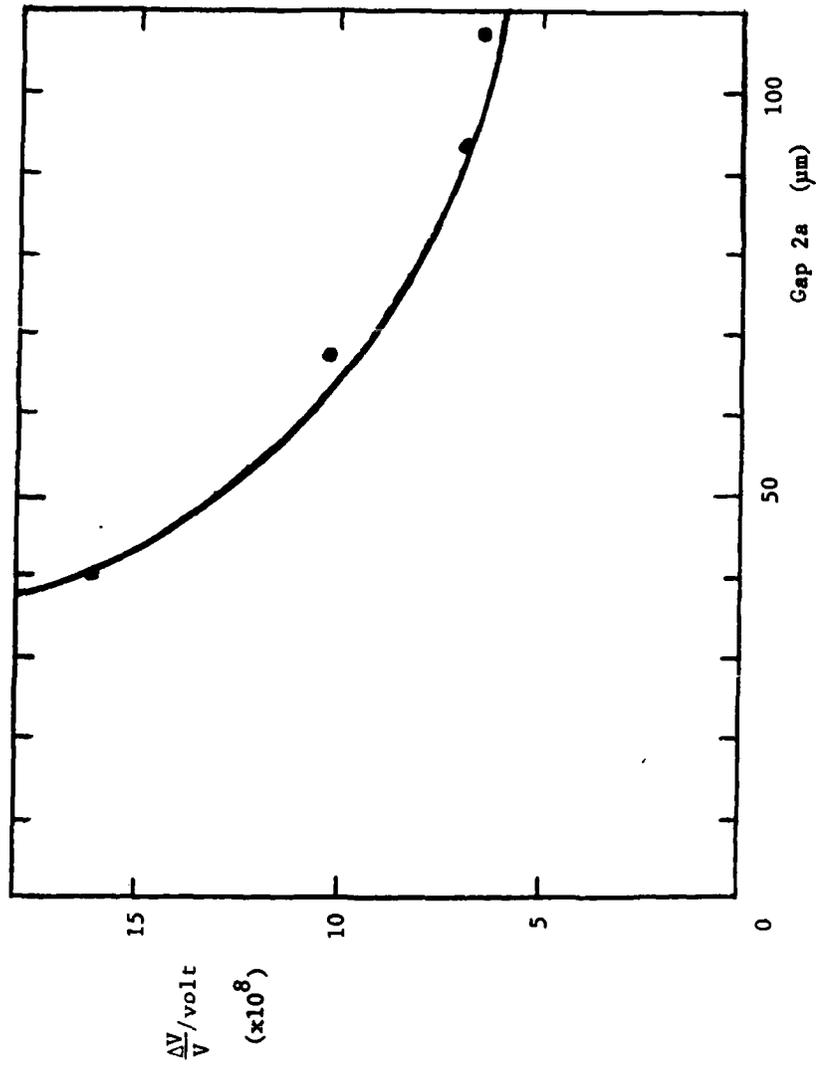


Fig. 4. Variation of voltage sensitivity with electrode gap for the in-plane field configuration. Measured points are taken from Ref. [6] for the 37-X cut of LiNbO_3 .

In the actual devices measured, the transducer aperture $2b$ was greater than the separation $2a$ between the electrodes. The SAW fields propagating beneath the electrodes would not be affected by the D.C. field in the gap. If these SAW fields reached the output transducer they would affect the phase of the output voltage, and hence, the apparent SAW velocity. The accuracy of the simple model discussed above suggests that these fields do not reach the output transducer, probably because of attenuation due to the film. Consistent with SAW attenuation under the film, the normal component of D.C. field between the electrodes and the ground plane at the bottom of the crystal plate does not seem to have significantly influenced the measured SAW velocity.

The largest changes in velocity reported in [3] for the in-plane configuration was about $\Delta V/V_R \approx 2 \times 10^{-5}$ for $L = 14$ mm at a frequency of 900 MHz. For these values, the phase changes along the center line due to applied D.C. voltage is $k_1 L \Delta V/V_R = (0.006)2\pi$. Because the phase change is so small, the variation of phase with x across the aperture will be less than π , except for $x \rightarrow a$. As a result, little phase cancellation will occur in the integral of (10). Had the phase change along the center line been close to 2π , then significant phase cancellation would result in the integration, and one would expect to see marked variations in the amplitude of the R.F. output voltage with applied D.C. voltage, and possibly phase reversals.

In the next section we discuss insertion loss due to diffraction. It is shown that this loss can be reduced by using narrow D.C. electrodes. The model described above indicates that width of the electrodes does not affect sensitivity, so that some improvement in loss is possible without sacrificing sensitivity.

2. Diffraction Loss

The D.C. electrodes shown in Fig. 2 for the in-plane configuration slow the SAW propagating beneath them due to piezoelectric shorting and mass loading.

Because of the slowing, waves propagating parallel to the edge of the electrode in the free surface region will be refracted at a finite angle θ_c into the electrode region, as shown in Fig. 2. As a result of this refraction, less energy will reach the output transducer, and hence path loss will increase.

An integral expression for the diffraction loss due to wide D.C. electrodes is derived in Appendix A. Because the fields propagating beneath the electrodes for the case when the transducer aperture (2b) is greater than electrode spacing (2a) do not seem to contribute to the output, we have treated the simpler case $2a > 2b$. Using Fourier transform techniques and deforming the path of integration, it is shown that for electrode spacing wide compared to SAW wavelength, the diffraction loss D_L between the transducers is given by the factor

$$D_L = \frac{2b}{\pi} \int_0^{\infty} \left(\frac{\sin \eta_1 b}{\eta_1 b} \right)^2 e^{i(\chi - k_1)L} f(\eta_1) d\eta_1 \quad (13)$$

where

$$f(\eta_1) = \frac{\eta_1 \eta_2}{\eta_1^2 + (\eta_2^2 - \eta_1^2) \cos^2 \eta_1 \alpha} \quad (14)$$

In these expressions η_1 and κ are the wave numbers along x and z, respectively, of a SAW propagating on the free surface of the piezoelectric crystal. For the same wavenumber κ along z, a SAW propagating in the electroded region has wavenumber η_2 . Thus κ and η_2 may be regarded as functions of η_1 . Finally, k_1 is the free surface SAW wavenumber for a wave propagating in the z direction. In the parabolic approximation,

$$\left. \begin{aligned} \chi &= \sqrt{k_1^2 - \alpha_1^2 \eta_1^2} \\ \eta_2 &= \frac{1}{\alpha_2} \sqrt{(k_2^2 - k_1^2) + \alpha_1^2 \eta_1^2} \end{aligned} \right\} \quad (15)$$

Here k_2 is the plated surface SAW wavenumber for a wave propagating along z and α_i ($i=1,2$) is related to the anisotropy parameter $\gamma = d^2V(\theta)/d\theta^2$ of Ref. [11] via

$$\alpha_i^2 = 1 + \gamma_i. \quad (16)$$

In the case when $\eta_1 = \eta_2$, so that SAW propagation on the plated surface is the same as the free surface, it is seen from (14) that $f(\eta_1) = 1$. For this condition (13) reduces to the ordinary diffraction loss between two uniform IDT's. If $\eta_2^2 - \eta_1^2 \ll \eta_1^2$, it is argued in Appendix A that for $a \gg b$, (13) reduces to the value obtain for a free surface, as is to be expected. The integration in (13) cannot be expressed in terms of tabulated functions, and must be evaluated numerically.

While numerical integration is in general required to evaluate (13), simple analytic approximations can be obtained in the free surface case for large L . Using the approximation

$$\chi \approx k_1 - \frac{\alpha_1^2}{2k_1} \eta_1^2 \quad (17)$$

it is seen that the exponential will go through rapid variations outside the range

$$0 \leq \eta_1 \leq \frac{2\pi}{\sqrt{\alpha_1^2} \lambda_1 L}.$$

The rapid variations in turn result in cancellation for the portion of the integration interval outside of this range. In this range of η_1 lies within the central lobe $0 < \eta_1 < \pi/b$ of $(\sin \eta_1 b / \eta_1 b)^2$, then this function may be approximated by unity. The requirement that the upper limit of the range be less than π/b is equivalent to the far field condition

$$F \equiv \left(\frac{L \alpha_1^2}{\lambda_1} \right) / \left(\frac{2b}{\lambda_1} \right)^2 > 1 \quad (18)$$

for transducers on a free surface [1].

For a free surface $f(\eta_1) = 1$ so that using (17), D_L of (13) becomes

$$D_L \approx \frac{2b}{\pi} \int_0^{\infty} \exp\left[-i\left(\frac{\alpha_1^2}{2k_1} L\right) \eta_1^2\right] d\eta_1 = \frac{e^{-i\pi/4}}{\sqrt{F}}. \quad (19)$$

This expression gives the well known fact that diffraction loss varies as $1/\sqrt{L}$ for large L .

When the D.C. electrodes are present, it is not simple to approximate D_L . In Appendix C an approximation is derived for case $a=b$ when $k_1 b$ is large and $F \gg 3$. The approximation derived there gives

$$D_L \approx 2\left(\frac{2}{\pi}\right)^2 e^{-i\pi F/4} e^{-bF/B} + 4\left(\frac{2}{\pi}\right)^4 \frac{b}{B\sqrt{F}} e^{-i\pi/4} \quad (20)$$

where

$$B = \frac{2}{\pi} b^2 \sqrt{k_2^2 - k_1^2} / \alpha_2.$$

As an example, ignoring mass loading effects of the electrodes, for 37-X LiNbO_3 , $B/b = 1.23 b/\lambda_1$ where λ_1 is the wavelength in the direction of propagation on the free surface. Then

$$D_L \approx 0.811 e^{-i\pi F/4} e^{-0.812(F\lambda_1/b)} + \frac{0.534 e^{-i\pi/4}}{\sqrt{F} (b/\lambda_1)}$$

If $b = 10\lambda_1$ and $F \lesssim 50$, the term containing the exponential is larger than the algebraic term, and can even be larger than D_L for a free surface.

In order to understand the variation of D_L with the parameters L and b , we have numerically evaluated the integral in (13). The results are plotted in Fig. 5, together with the diffraction loss for a free surface. For a narrow gap between the electrodes, the diffraction loss is greater than for a free surface, and it increases rapidly for F greater than about 5. For wide gaps, the diffraction

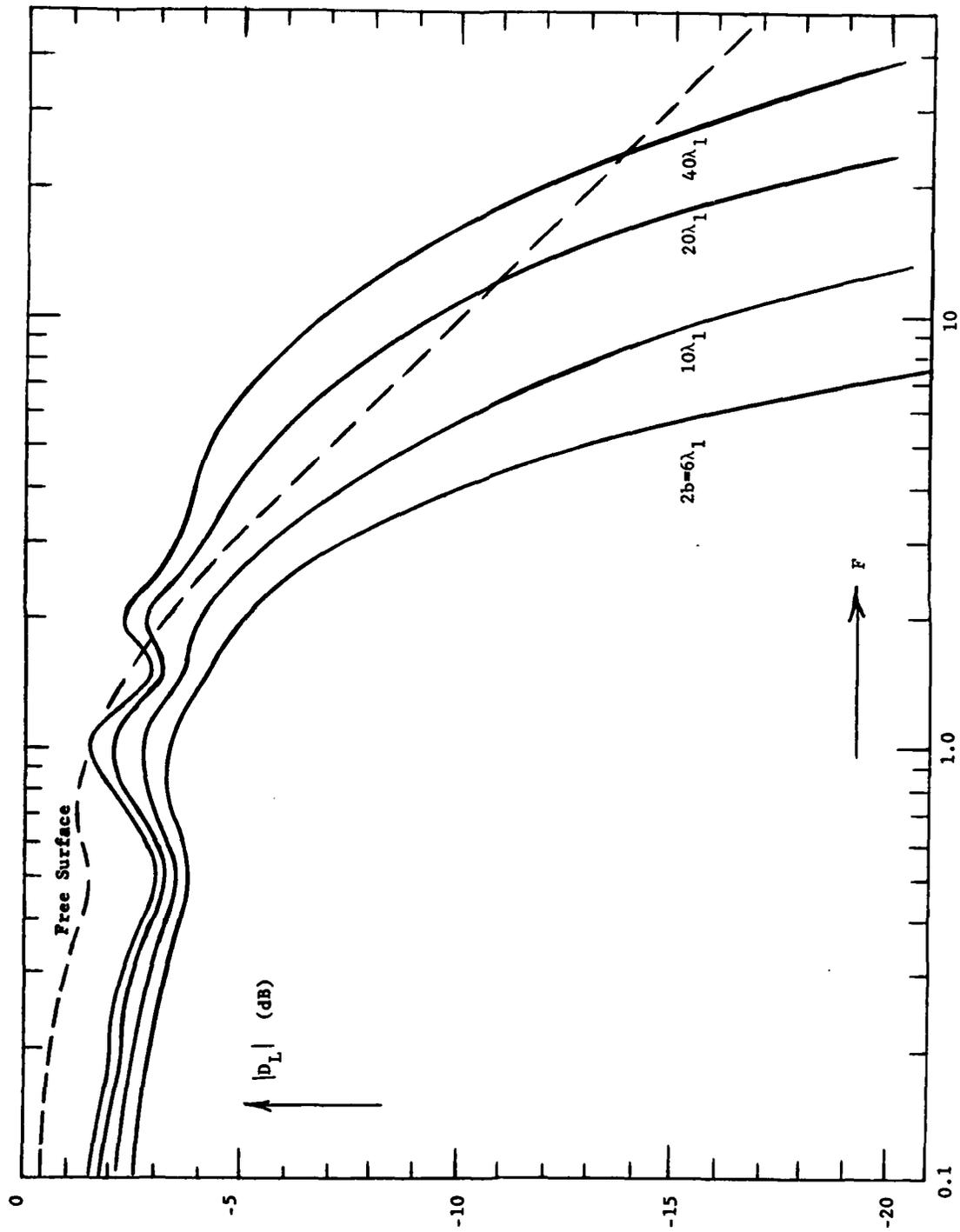


Fig. 5. Variation of diffraction loss with distance parameter F for an electrode gap $2b$.

loss can be less than the free surface loss over some range of F . However, for large F the diffraction loss increases rapidly and becomes greater than the free surface loss.

Application of the in-plane field configuration requires narrow gaps $2b$ for voltage sensitivity and long paths of propagation L to achieve large changes in time delay. For this choice of parameters, the use of wide electrodes leads to an increase in diffraction loss over that of a free surface. In the next section we consider methods for reducing diffraction loss.

IV. Reduction of Diffraction Loss for the In-Plane Configuration

The measurements of Budreau, et.al.[3] indicate that for the high coupling cuts of LiNbO_3 , the in-plane configuration suffers from low sensitivity of the velocity to applied electric field (volts/m), as compared to the normal field configuration. Since the D.C. bias electrodes can be made much closer for the in-plane configuration than for the normal field case, the sensitivity to applied terminal voltage can be made higher. However, bringing the D.C. electrodes closer together has been seen to increase diffraction loss.

In this section we discuss modifications of the electrodes to reduce this loss. The first type of modification employs electrodes of narrow width w . The second employs surface platings in the region $|x| > b$ that stiffen the surface thereby creating a slot waveguide in the region between the D.C. electrodes.

A. Narrow Electrodes

Wide electrodes, whose attenuation makes them appear to be semi-infinite, conduct energy away from the fields in the gap and thereby increase diffraction loss. One possible approach to reducing this loss is to use narrow electrodes to prevent the energy from being lost. Electrodes of finite width w act as $\Delta V/V$ waveguide. Thus, the field reaching the output transducer will be composed of

modes guided by the electrodes, in addition to the unguided fields that are similar to the fields on the free surface of a crystal.

The reduction D_L in the output voltage due to diffraction loss can be separated into a component D_f representing the effect of the unguided field, and D_g representing the effect of the guided waves. Thus

$$D_L = D_f + D_g \quad (21)$$

Expressions for D_f and D_g are derived in Appendix B. Neglecting dissipation of the guided waves by loss in the electrode platings, $|D_g|$ is independent of the separation L between the transducers, but decreases rapidly with separation $2a$ between the electrodes. On the other hand, D_f decreases with the ratio L/b^2 .

The loss D_f due to the unguided fields is given by (13), with $f(\eta_1)$ replaced by (B-7). Further manipulation of (B-7) yields

$$f(\eta_1) = \left\{ \left[\cos \eta_2 w \sin \eta_1 a + (\eta_2/\eta_1) \sin \eta_2 w \cos \eta_1 a \right]^2 + \left[\cos \eta_2 w \cos \eta_1 a - (\eta_1/\eta_2) \sin \eta_2 w \sin \eta_1 a \right]^2 \right\}^{-1} \quad (22)$$

The functional dependence of $f(\eta_1)$ is sufficiently complex that it is not easily seen how to choose w to optimize D_f . For the case when $w = 0$, or the case when $k_1 = k_2$ and $\alpha_1 = \alpha_2$ so that $\eta_2 = \eta_1$, it is easily established that $f(\eta_1) = 1$, which is the value for a free surface. The free surface condition may be approximated for finite w if it is chosen such that

$$w \eta_2(0) = w \sqrt{k_2^2 - k_1^2} / \alpha_1 = \pi \quad (23)$$

Using such half-wave electrodes, $f(0) = 1$.

1. Half-Wave Electrodes

To evaluate the diffraction loss for electrodes of width w satisfying

(23), we observe that the principle contribution to (13) comes from the interval $0 < \eta < \pi/b$, where $\pi/b \ll \sqrt{k_2^2 - k_1^2}/\alpha_1$. Thus we may use an approximation for $f(\eta_1)$. To develop the approximation we make use of

$$\eta_2 \approx \frac{1}{\alpha_2} \sqrt{k_2^2 - k_1^2} + \frac{(\alpha_1^2/\alpha_2^2)\eta_1^2}{2\sqrt{k_2^2 - k_1^2}} = \eta_2(0) + \frac{\alpha_1^2/\alpha_2^2}{2\eta_2(0)} \eta_1^2 \quad (24)$$

For w satisfying (23), to order η_1^2

$$\begin{aligned} \cos \eta_2 w &\approx 1 \\ \sin \eta_2 w &\approx -w \frac{\alpha_1^2/\alpha_2^2}{2\eta_2(0)} \eta_1^2 \end{aligned} \quad (25)$$

Using (24) and (25) in (22), we find that

$$\begin{aligned} f(\eta_1) = \left\{ \left[\sin \eta_1 a - \frac{w}{2} (\alpha_1^2/\alpha_2^2) \eta_1 \cos \eta_1 a + O(\eta_1^3) \right]^2 \right. \\ \left. + \left[\cos \eta_1 a + O(\eta_1^3) \right]^2 \right\}^{-1} \end{aligned} \quad (26)$$

In (26) we have not approximated $\sin \eta_1 a$ or $\cos \eta_1 a$ since for $a \gg \lambda_1$, they can vary significantly over the range of η_1 giving the primary contribution to (13).

Neglecting powers of η_1 greater than the first

$$f(\eta_1) \approx 1 + w (\alpha_1^2/\alpha_2^2) \eta_1 \sin \eta_1 a \cos \eta_1 a \quad (27)$$

When this expression for $f(\eta_1)$ is substituted into (13), the first term in (27) gives the diffraction loss for a pair of transducers on a free surface. The second term in (27) gives a correction that is small since the integrand varies as η_1^2 for η_1 small. Thus the presence of the electrodes will result in only a small change in the diffraction loss over that of a free surface.

When the far-field condition (18) is satisfied, (13) can be approximated in closed form. In this case, the range of η_1 giving the primary contribution to the integral is determined by the exponent, rather than $(\sin \eta_1 b / \eta_1 b)^2$.

Thus we may approximate $\sin \eta_1 a \approx \eta_1 a$ and $\cos \eta_1 a \approx 1$. With these approximations and the help of (17) we have for $a = b$

$$D_f = \frac{e^{-i\pi/4}}{\sqrt{F}} \left[1 - i \frac{W}{b} \frac{\alpha_1^2}{\alpha_2^2} \frac{\pi}{2F} \right]. \quad (28)$$

For b large compared to w , the correction term in (28) will have only a small effect on $|D_f|$.

When w satisfies (23), the $n = 1$ mode is at cutoff, and hence gives vanishing contribution to (B-20). Thus we need keep only the $n = 0$ mode in (B-20) and obtain

$$D_g = \frac{4bk_1}{\alpha} \left(\frac{\sinh y_0 b}{y_0 b} \right)^2 \frac{y_0 e^{-2y_0 a} e^{i(x-k_1)L}}{\left(1 + \frac{\alpha_1^2}{\alpha_2^2} \frac{y_0^2}{\eta_2^2} \right) - \frac{y_0 W}{2} \frac{\alpha_1^2}{\alpha_2^2} \left(1 + \frac{y_0^2}{\eta_2^2} \right)} \quad (29)$$

where y_0 is decay constant along x of the guided wave fields in the free surface region, and η_2 and κ are found from (15) with $\eta_1 = iy_0$.

To estimate the size of $|D_g|$ relative to $|D_f|$, we make use of the computations of Ref. [8] for waveguides on $Y - Z$ LiNbO_3 . Assuming $b = a$ and that $y_0 b \gg 2$, it is found that

$$|D_g| \approx 1.22 (\lambda_1/b) = 0.869/(y_0 b).$$

If wide transducers are used so that $F \sim 1$, then $\lambda_1 \ll b$ and $|D_g| \ll |D_f|$. For this case the guided wave give only a small contribution to the output voltage. The guided wave may result in some amplitude and phase ripple with frequency since its propagation constant $\kappa(iy_0)$ is different from k_1 .

2. Other Electrode Widths

Electrodes that are an integer multiple of the width given by (23) will also lead to approximations for $f(\eta_1)$ similar to (27). The presence of such electrodes will therefore only slightly perturb D_f from its value for a free

surface. However, such electrodes will have more modes above cutoff, thereby complicating the phase and amplitude dependence of D_g with frequency.

The half-wave electrodes result in less diffraction loss because they are to first order invisible to a SAW propagating along the z axis. One is led to speculate that an improvement in the loss could be achieved if the electrodes reflected the SAW back into the gap between the electrodes. Quarter-wave electrodes, having width half that given by (23), would have the largest reflection coefficient. Unfortunately, the reflection coefficient is negative, as can be seen from (B-3) for $\eta_1 = 0$ and $\eta_2 w = \pi/2$ so that reflected waves add destructively with those still propagating in the gap. This result can also be seen from $f(\eta_1)$ in (22) if $\eta_2(0)w = \pi/2$, then $\cos \eta_2 w = 0$, $\sin \eta_2 w = 1$ and for η_1 small

$$f(\eta_1) \approx \frac{\eta_1}{\eta_2(0) \cos \eta_1 a}$$

Because $f(\eta_1)$ depends linearly on η_1 , the integral in (13) is smaller than it would be for a free surface.

B. Slot Waveguides

One way to achieve a narrow gap between electrodes, without increasing the diffraction loss is to plate the electrodes of Fig. 2 with a material such as AlN that stiffens the surface [12]. If the SAW velocity is thereby increased over the free surface SAW velocity, the gap region will act as a waveguide [4,5].

If the electrode consists of a very thin metalization, then the mechanical perturbation of the surface will be due only to the film used to stiffen the surface. To first order for thin platings, the electrical shorting and mechanical loadings are additive effects. Thus, for thin platings of hexagonal material, such as AlN, with C axis normal to the surface, the wave number k_2 under the plating is [6]

$$\frac{k_2 - k_1}{k_1} = \frac{\Delta V}{V_1} + \frac{V_1 h}{4P} \left[\left(\rho' - \frac{1}{V_1^2 s_{66}} \right) |v_x|^2 + \rho' |v_y|^2 + \left(\rho' - \frac{1}{V_1^2} \frac{s_{11}}{s_{11}^2 - s_{12}^2} \right) |v_z|^2 \right] \quad (30)$$

Here, V_1 is the SAW velocity along z on a free surface, $\Delta V/V_1$ is the velocity change due to shorting of the piezoelectric fields. The plating material has thickness h , mass density ρ' and elastic stiffness constants s_{ij} in the coordinate system where Z is along the C axis. Finally, v_x , v_y and v_z are the particle velocity components at the surface $y=0$ of a SAW carrying power along z of P watts per meter.

Reference [14] lists the compliances C_{ij} for A2N. If these are used to obtain the s_{ij} elements as if the material were non-piezoelectric, then $1/s_{66} = 1.10 \times 10^{11}$ (n^2/N) and $s_{11}/(s_{11}^2 - s_{12}^2) = 1.35 \times 10^{11}$ (n^2/N). Also, $\rho = 3.26 \times 10^3 \text{ kg/m}^3$ so that for the 37-X cut of LiNbO_3 expression (30) becomes

$$\frac{k_2 - k_1}{k_1} = .0268 - 0.222 h/\lambda_1 \quad (31)$$

It is seen that a layer of thickness $h \approx 0.12 \lambda_1$ is necessary to overcome the slowing resulting from shorting of the electric field. The first order perturbation theory leading to (31) may not be accurate for films of this thickness, and a more accurate computer calculation is required to establish the actual value of h needed to overcome the slowing due to the conducting film.

For thicker platings, the SAW velocity $V_2 = \omega/k_2$ will be greater than the free surface velocity V_1 . The highest SAW velocity possible is limited by the velocity of the shear wave propagating parallel to the surface. Platings that are thick enough to raise the SAW velocity above the shear wave velocity will cause the SAW to leak energy into the bulk of the crystal. The slowest of the two shear waves propagating along x in LiNbO_3 has velocity that is 1.022 times

Rayleigh velocity on the 37-X cut. Therefore $(k_2 - k_1)/k_1 > 0.022$ for the SAW under the plating to be a bound wave.

The condition necessary for guidance of a wave by a slot waveguide is that k_2 lie between the shear wavenumber and k_1 . When this condition is satisfied, the wavenumbers κ_n of the propagating waveguide modes lie between k_1 and k_2 , so that η_2 is imaginary and η_1 is real. The dispersion relation for the guided waves of even symmetry about the center line $x = 0$ is

$$\eta_2 = i\eta_1 \tan \eta_1 b \quad (32)$$

where η_1 and η_2 are related to κ_n via (A-6). Equation (32) may be solved graphically for η_1 using the construction of Fig. B-1 with subscripts 1 and 2 interchanged and $2b$ used for w . Because the transducers are symmetric about the centerline, only the symmetric ($\eta = 0, 2, 4, \dots$) modes need be considered.

The dispersion of slot waveguide modes is discussed in Ref. [4] for the case of dielectric layers on anisotropic substrates. The presence of the shorting layer will modify somewhat the dispersion. In any case, the phase velocity $V_n = \omega/\kappa_n$ of the guided waves must lie between V_1 and V_2 . Since V_2 must be less than the shear wave velocity, the difference between V_n and V_1 can be no more than a few percent. As a result, mode dispersion will not be a significant factor in the application of gap guides to phase shifters for phased array radar. If two or more modes can propagate, differences in their phase velocities may however produce undesirable amplitude variations with frequency.

The output voltage in the presence of the guide can be obtained by modifying the analysis of Appendix A. In expression (A-14), the poles of the integrand are those of the factor

$$\frac{1 + \Gamma e^{i2\eta_1 a}}{1 - \Gamma e^{i2\eta_2 a}} = \frac{\eta_1 - i\eta_2 \tan \eta_1 a}{\eta_2 - i\eta_1 \tan \eta_1 a} \quad (33)$$

It is seen from (33) that the poles occur at the roots of the dispersion relation (32) for the guided waves. Referring to Fig. A-2, the poles associated with the modes propagating in the positive z direction are located along the real κ axis to the right of the branch point k_2 .

When the integration path is deformed around the branch cut through k_2 , these poles are captured. The output voltage due to the poles is therefore

$$V_0 = 2Q \frac{b^2}{\alpha_1^2} \sum_n \left(\frac{\sin \eta_1 b}{\eta_1 b} \right)^2 \frac{\eta_1 - i\eta_2 \tan \eta_1 a}{\frac{d}{d\kappa}(\eta_2 - i\eta_1 \tan \eta_1 a)} \frac{e^{i\kappa_n L}}{\eta_1} \quad (34)$$

where the sum is taken over the propagating modes and η_1 and η_2 are evaluated at the modal wavenumber κ_n . Making use of (15) and (32), and normalizing (34) by the factor (B-9), the contribution to the diffraction loss from the guided waves is found to be

$$D_g = 2 \sum_n \left(\frac{\sin \eta_1 b}{\eta_1 b} \right)^2 \frac{k_1}{\kappa_n} \frac{\eta_2 b e^{i(\kappa_n - k_1)L}}{\eta_2 a + i \left(\frac{\alpha_1^2}{\alpha_2^2} \cos^2 \eta_1 a + \sin^2 \eta_1 a \right)} \quad (35)$$

The contribution to the output voltage from the continuous spectrum is represented by the integration around the branch cut originating from k_2 in the κ plane. Because of the factor $(\sin \eta_1 b / \eta_1 b)^2$, the contribution from the continuous spectrum will be small if at $\kappa = k_2$, $\eta_1 b > \pi$. With the help of (A-6) this condition is

$$b \sqrt{k_1^2 - k_2^2} / \alpha_1 > \pi \quad (36)$$

Alternatively, for F large, phase cancellation due to the exponential term will make the contribution from the continuous spectrum small. We do not consider this contribution further.

Condition (36) also implies that two or more modes of the slot guide are propagating. As an example, suppose that k_2 is 1% smaller than k_1 . Using $\alpha_1 = 0.892$ for 37-X LiNbO₃,

$$b\sqrt{k_1^2 - k_2^2}/\alpha_1 = (0.317)\pi b/\lambda_1$$

A gap $2b$ slightly larger than $6\lambda_1$ will satisfy (36).

To estimate D_g of (35) let $a=b$, $\alpha_1=\alpha_2$ and assume that $b\sqrt{k_1^2 - k_2^2}/\alpha_1 = \pi$. Under these conditions the second mode will be at cutoff with $\eta_2 = 0$, and hence will give a zero contribution to (35). With the approximation $\eta_1 b \approx \pi/2$ for the lowest mode gives with the help of (15)

$$\gamma_2 b = i \frac{\alpha_1}{\alpha_2} \sqrt{\frac{b^2}{\alpha_1^2} (k_1^2 - k_2^2) - \eta_1^2 b^2} = i\pi \frac{\sqrt{3}}{2}$$

The $n = 0$ term in (35) then gives

$$|D_g| \approx 2 \left(\frac{2}{\pi} \right)^2 \frac{\pi \sqrt{3}/2}{\pi \sqrt{3}/2 + 1} = 0.59$$

corresponding to -4.5 dB loss independent of L . For comparison, consider electrodes that are not stiffened. If $2b = 6\lambda_1$ and $\alpha_1^2 L/\lambda_1$ is 250 or greater, then $F \geq 7$ and the diffraction loss is more than 25 dB. Thus, a substantial reduction in diffraction loss can be achieved for narrow D.C. electrode spacings by plating over the electrodes with a material that stiffens the surface.

V. Recommendations

In order to compare the various electrode configuration, we consider 1 GHz SAW devices on 38-X cut LiNbO_3 . This cut was found to give the highest voltage sensitivity for both the normal field and in-plane configurations [3]. For the normal field case, an applied field of 1 MV/m gives a velocity change $\Delta V/V = 1.5 \times 10^{-4}$ in an $L=1.4$ cm path. For the in-plane case $\Delta V/V = 6.2 \times 10^{-6}$ for $V_{DC}/2a = 1$ MV/m.

Based on the foregoing values, we have computed the D.C. voltage necessary to produce a change in delay of 1/2 an R.F. cycle (180° phase shift) at the output transducer. These values are listed in Table 1 for a 1 GHz SAW. For the raised electrode, the factor (5) was used in computing V_{DC} .

Propagation loss is also listed in the table. For plated electrode in the normal field configuration, the loss over free surface attenuation is due to loss in the film. The transducers are assumed wide enough to make diffraction loss negligible. For the in-plane configuration, loss represents the asymptotic forms of D_L valid for transducers that are short enough so that $F > 1$. In the case of slot waveguide, the loss represents mode coupling, which depends on $2b$, but is independent of L . The value -5 dB is for $2b \approx 6\lambda_1$.

It is seen from Table 1 that relatively long paths are required to achieve 1/2 cycle delay changes with D.C. voltages under 100V. Propagation loss will be significant in all configurations. Using slot waveguides with $2b = 10 \mu\text{m} \approx 2.5\lambda_1$ gives the lowest attenuation.

The foregoing conclusion suggests that slot guide deserves further consideration as a basis for phase shift devices. It is important to determine the effect of losses in the metal film and AuN plating on the guided wave attenuation. A more rigorous theory must be employed to compute dispersion for SAW propagation under the plating, including the effects of finite metal thickness

and the piezoelectric properties of AlN. When computing the dispersion of modes in the slot waveguide, it may also be necessary to account for effects at the edge of the plating, which are known to be important for gratings. Finally the use of narrow electrodes may reduce loss and improve confinement. Criteria for determining the width of the narrow electrodes, and their location with respect to the edge of the AlN plating, should be developed.

TABLE 1
ATTENUATION AND VOLTAGE SENSITIVITY FOR 1 GHz SAW
ON THE 38-X CUT OF $L_1\text{NbO}_3$

Normal Field Configuration* [250 μm thick plate, wide transducers]	Propagation Loss Over Free Surface Attenuation of $-2.5L$ dB	VDC for 1/2 Cycle Delay Change
1. Plated Electrode ($H=0$)	-7 dB	466/L
2. Raised Electrode ($H=5\mu\text{m}$)	0	1,260/L
B. In-Plane Configuration* [$2a=2b$ several wavelengths]*	-1.8-1.79x10 L/(2b) dB	$45 \left(\frac{2b}{L} \right)$
1. Wide Electrodes	-45+20 $\log_{10} (2b)$	
2. Half-Wave Electrodes	-10 $\log_{10} L$ dB	
3. Lowest Mode of Slot Waveguide	~ -5 dB	

*path length L is measured in cm,
Gap width 2b is measured in μm

APPENDIX A: Diffraction Loss Due to D.C. Electrodes

In order to compute the diffraction loss due to the D.C. electrodes in Fig. 2 we assume them to be semi-infinite in x and infinite in z . For wide electrodes that are slightly lossy, as is typically the case, the foregoing assumptions allow us to accurately model the diffraction loss in the simplest possible structure. The resulting geometry is shown in Fig. A-1. The electrode separation $2a$ is assumed to be greater than or equal to the transducer aperture $2b$. Finally, the SAW velocity in the plated region 2 is assumed to be slightly smaller than that of the free surface region 1.

Previous studies [4] have shown that the SAW amplitude can be described by a potential function $\psi(x,z)$. For crystal cuts, such as 37 - X LiNbO₃, that satisfy the parabolic approximation, away from the sources the potential function in regions $j = 1, 2$ satisfies the wave equation

$$\left(\alpha_j^2 \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2} + k_j^2 \right) \psi(x,z) = 0. \quad (\text{A-1})$$

Here k_j is the wavenumber of the SAW propagating in the z direction and α_j , is related to the anisotropy parameter $\gamma = d^2v(\theta)/d\theta^2$ of Ref. [8] via

$$\alpha_j^2 = 1 + \gamma_j. \quad (\text{A-2})$$

Note that we are concerned with waves propagating primarily along z , with only a small x variation, so that the parabolic approximation is appropriate.

To find the SAW generated by the transducer, we represent the transducer, we represent the transducer by a collection of point sources $\delta(x-x')\delta(z)$ along

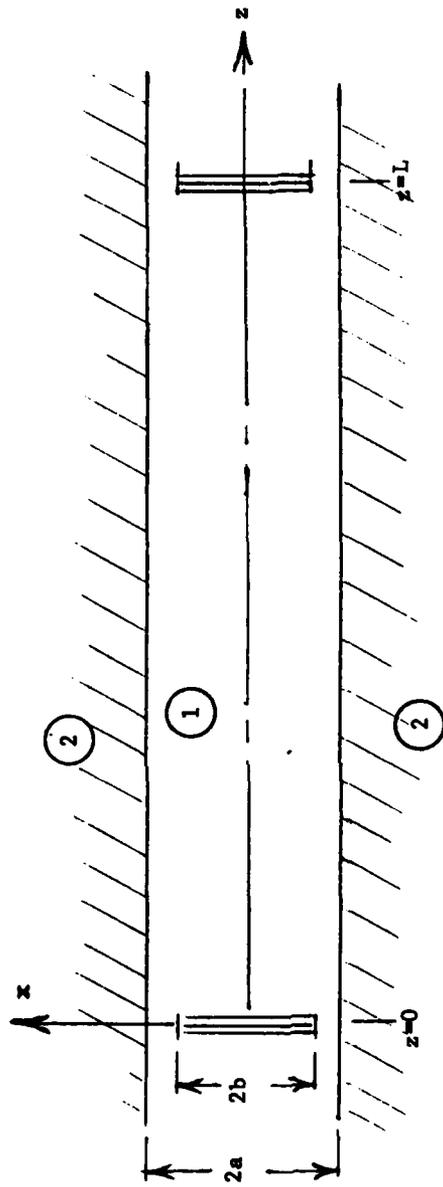


Fig. A-1. Configuration for computing diffraction loss due to electrodes

the interval $-b < x' < b$. For each point source the potential $\Psi_g(x, z, x')$ must satisfy

$$\left(\alpha_j^2 \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2} + k_j \right) \Psi_g = \delta(x-x') \delta(z). \quad (\text{A-3})$$

If we further employ a Fourier integral representation along z , then for time dependence $\exp(-i\omega t)$

$$\Psi_g(x, z; x') = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi(x, x', \kappa) e^{i\kappa z} d\kappa. \quad (\text{A-4})$$

The transform $\Phi(x, x', \kappa)$ satisfies the equation

$$\left[\alpha_j^2 \frac{\partial^2}{\partial x^2} + (k_j^2 - \kappa^2) \right] \Phi = \delta(x-x'). \quad (\text{A-5})$$

The solution to (A-5) can be constructed in terms of exponential solutions $\exp(\pm i\eta_j x)$, where

$$\eta_j = \frac{1}{\alpha_j} \sqrt{k_j^2 - \kappa^2}. \quad (\text{A-6})$$

At the boundaries $x = \pm a$ we require that Φ and its normal derivative be continuous [4]. As a result, the boundaries can be described by the reflection coefficient

$$\Gamma = \frac{\eta_1 - \eta_2}{\eta_1 + \eta_2}. \quad (\text{A-7})$$

Thus for $a > x > x'$

$$\Phi = A \left[e^{i\eta_1(x-a)} + \Gamma e^{-i\eta_1(x-a)} \right] \quad (\text{A-8})$$

while for $-a < x < x'$

$$\Phi = B \left[e^{-i\eta_1(x+a)} + \Gamma e^{i\eta_1(x+a)} \right]. \quad (\text{A-9})$$

To find the coefficients A and B we return to (A-5), which implies that Φ is continuous at $x = x'$, but that $\partial\Phi/\partial x$ jumps by $1/\alpha_1^2$. Using these conditions and solving for A and B gives

$$\Phi = \frac{1}{i2\alpha_1^2 \eta_1 D} \left\{ e^{i\eta_1|x-x'|} e^{-i\eta_1 a} + 2\Gamma \cos[\eta_1(x+x')] + \Gamma^2 e^{-i\eta_1|x-x'|} e^{i2\eta_1 a} \right\} \quad (\text{A-10})$$

where

$$D = e^{-i2\eta_1 a} - \Gamma^2 e^{i2\eta_1 a}. \quad (\text{A-11})$$

Substituting this expression into (A-4) gives $\psi_g(x, z; x')$.

To find the field at the output transducer it is necessary to integrate $\psi_g(x, z; x')$ over the aperture $-b < x' < b$. The output voltage is then found by integrating the field over the output transducer $-b < x < b$. Thus the output voltage is

$$V_0 = Q \int_{-b}^b \int_{-b}^b \psi_g(x, L; x') dx dx' \quad (\text{A-12})$$

where Q is some constant of proportionality. Interchanging the order of integration over x', x with that of the integration over κ allows us to write

$$V_0 = \frac{Q}{2\pi} \int_{-\infty}^{\infty} \left[\int_{-b}^b \int_{-b}^b \Phi(x, x', \kappa) dx dx' \right] e^{i\kappa L} d\kappa. \quad (\text{A-13})$$

Substituting from (A-10) for $\Phi(x, x'; \kappa)$ and carrying out the integrations over x' and x , after some manipulation one obtains

$$V_0 = \frac{Q}{i\pi \alpha_1^2} \int_{-\infty}^{\infty} \left\{ \left[\left(\frac{\sin \eta_1 b}{\eta_1 b} \right)^2 \frac{1 + \Gamma e^{i2\eta_1 a}}{1 - \Gamma e^{i2\eta_1 a}} - i \left[\frac{\sin 2\eta_1 b}{2\eta_1^2} - \frac{b}{\eta_1} \right] \right\} \frac{e^{i\kappa L}}{\eta_1} d\kappa. \quad (\text{A-14})$$

Since for η_1 small

$$i \left[\frac{\sin 2\eta_1 b}{2\eta_1^2} - \frac{b}{\eta_1} \right] \approx i \frac{2}{3} \eta_1 b^3 \quad (\text{A-15})$$

the only pole singularities of the integrand occur at the roots of

$$(1 - \Gamma e^{i2\eta_1 a})_{0-36} = 0. \quad (\text{A-16})$$

Using (A-7), it can be seen that the integrand in (A-14) is independent of the sign of η_1 , so that its only branch points occur at $\kappa = \pm k_2$, where $\eta_2 = 0$.

The branch points $\pm k_2$ in the complex κ - plane are indicated in Fig. A-2, together with the associated branch cuts. The integration path implied in (A-14) is also indicated. With these cuts, $\text{Im } \eta_2 > 0$ on the top Riemann sheet, and $\text{Re } \eta_2 > 0$ along the portion of the integration path in the interval $-k_2 < \kappa < k_2$. In addition, the branch points $\kappa = \pm k_1$ of the function η_1 are indicated in Fig. A-2, along with the associated branch cuts (shown dashed). These cuts are not needed to evaluate (A-14), but will clarify the subsequent discussion.

Taken together, the terms in the curved brackets of (A-14) vanish for $|\kappa| \rightarrow \infty$. Thus for $L > 0$, the integration path may be deformed into the upper half plane, as shown in Fig. A-2. For $k_2 > k_1$, the pole singularities of (A-14) will be of the leaky wave type, and hence lie on the Riemann sheet for which $\text{Im } \eta_2 < 0$. Such poles are not intercepted during the deformation indicated in Fig. A-2.

Along the new path, η_1 is real and the terms in the square brackets of (A-14) individually vanish as $\kappa \rightarrow i\infty$. Furthermore, the term

$$\left[\frac{\sin 2\eta_1 b}{2\eta_1^2} - \frac{b}{\eta_1} \right] \frac{e^{i\chi L}}{\eta_1}$$

has the same value on both sides of the branch cut, so that its integration over the deformed path vanishes. Hence we are left with only the first term in the square bracket of (A-14).

Now make the change of variable to η_1 in the integration. Thus

$$\begin{aligned} \chi &= \sqrt{k_1^2 - \alpha_1^2 \eta_1^2} \\ \eta_2 &= \frac{1}{\alpha_2} \sqrt{(k_2^2 - k_1^2) + \alpha_1^2 \eta_1^2} . \end{aligned} \quad (\text{A-17})$$

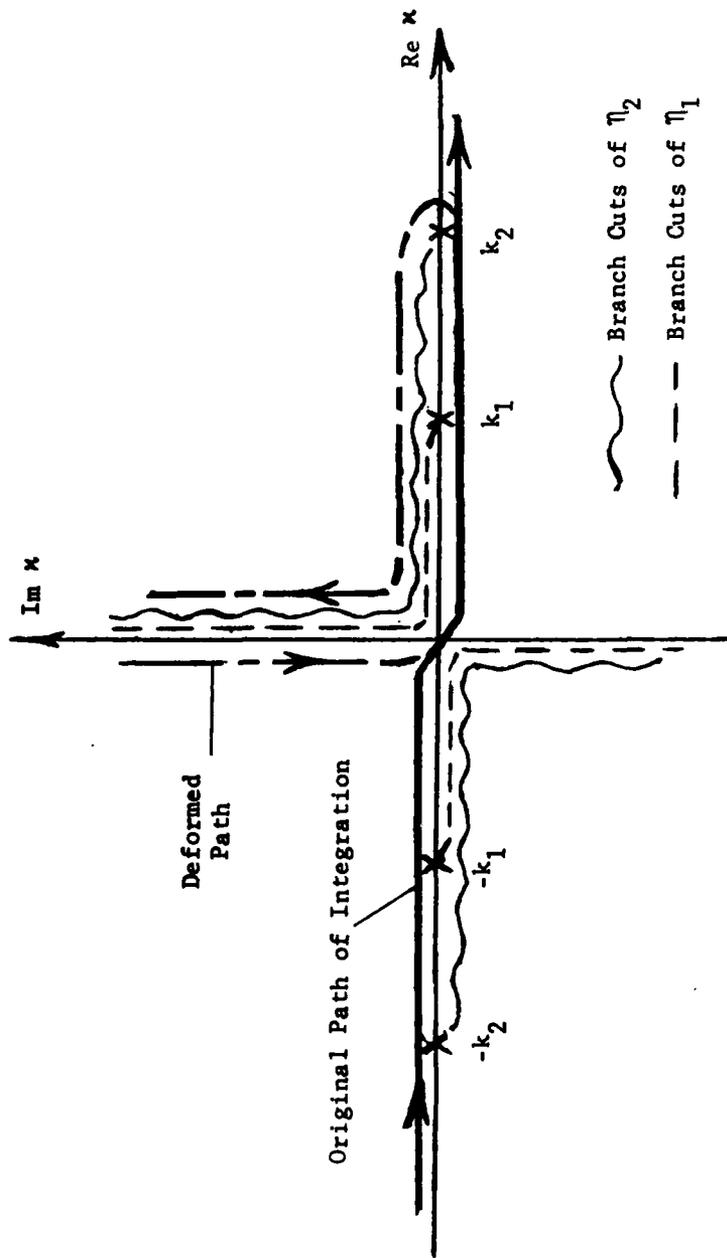


Fig. A-2. Integration path in the complex κ plane.

With this change of variable, (A-14) becomes

$$V_0 = \frac{Q}{i\pi} \int_{P_1} \left(\frac{\sin \eta_1 b}{\eta_1} \right)^2 \frac{1 + e^{i2\eta_1 a}}{1 + e^{i2\eta_1 a}} \frac{e^{i\kappa L}}{\kappa} d\eta_1 \quad (\text{A-18})$$

where the mapping of the integration path P_1 in the complex η_1 -plane is shown in Fig. A-3. The branch points $\pm k_1/\alpha_1$ of κ are also shown, together with the branch points $\pm (1/\alpha_1)\sqrt{k_2^2 - k_1^2}$ of the function η_2 in (A-17). With the choice of branch cut shown in Fig. A-3, η_2 has the same sign as η_1 along the real η_1 axis, which is consistent with the sign of η_2 along the deformed integration path in Fig. A-2. Note that for $k_2 = k_1$, this branch cut disappears. Also, $\Gamma = 0$ for this case and (A-18) reduces to the diffraction integral for two uniform transducers on a free surface, as is to be expected.

To further simplify (A-18), we use (A-7) to show that

$$\frac{1 + \Gamma e^{i2\eta_1 a}}{1 + \Gamma e^{i2\eta_1 a}} = \frac{\eta_1 \eta_2 - i(\sin \eta_1 a)(\cos \eta_1 a)(\eta_2^2 - \eta_1^2)}{\eta_2^2 - (\eta_2^2 - \eta_1^2)(\sin \eta_1 a)^2} \quad (\text{A-19})$$

Along the real η_1 axis, the real part of (A-19) is an even function while the imaginary part is odd. Thus, the contribution V_{OR} to (A-18) from the real axis portion of the path P_1 is given by

$$V_{OR} = \frac{2Qb^2}{i\pi} \int_0^{\infty} \left\{ \left(\frac{\sin \eta_1 b}{\eta_1 b} \right)^2 \left[\frac{\eta_1 \eta_2}{\eta_2^2 - (\eta_2^2 - \eta_1^2)(\sin \eta_1 a)^2} \right] \right\} \frac{e^{i\kappa L}}{\kappa} d\eta_1 \quad (\text{A-20})$$

For the portion of the integration path P_1 along the imaginary η_1 axis, let $\eta_1 = iy$ where $0 < y < (1/\alpha_1)\sqrt{k_2^2 - k_1^2}$. On the right-hand side of the branch cut η_2 is positive real, while on the left hand side it is negative

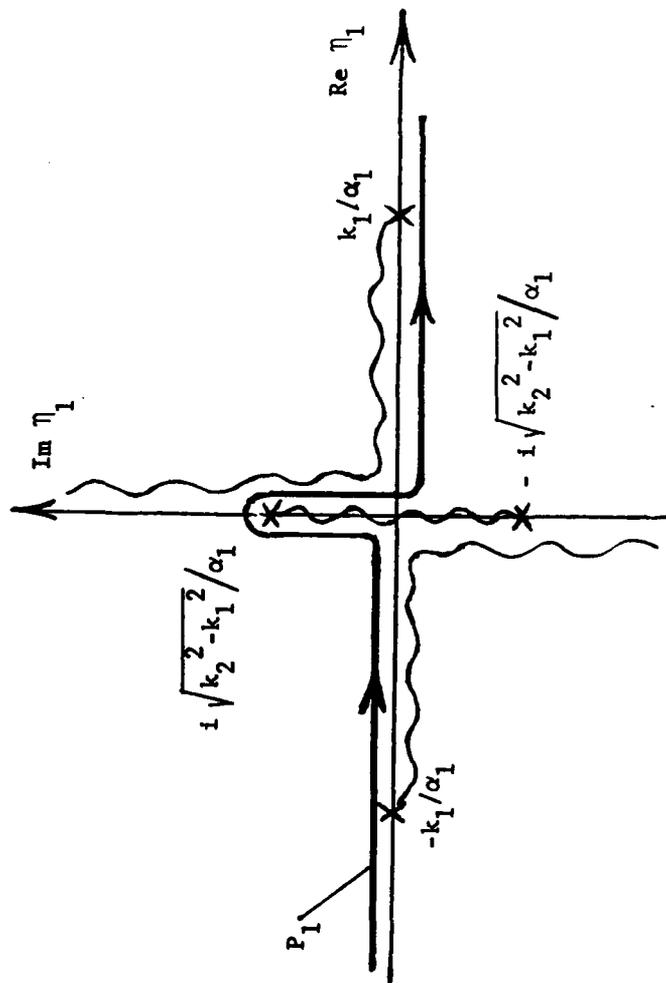


Fig. A-3. Integration path in the complex η plane.

real. Substituting $\eta_1 = iy$ into (A-19), and making use of the symmetry of η_2 , the contribution V_{OI} to (A-18) from the portion of the path along the imaginary axis can be written

$$V_{OI} = \frac{2Qb^2}{i\pi} \int_0^{\sqrt{k_2^2 - k_1^2}/\alpha_1} \left\{ \left(\frac{\sinh yb}{yb} \right)^2 \frac{y\eta_2}{\eta_2^2 + (\eta_2^2 + y^2)(\sinh ya)^2} \right\} \frac{e^{i\alpha_1 L}}{\alpha_1} dy \quad (A-21)$$

where from (A-17)

$$\begin{aligned} \alpha_1 &= \sqrt{k_1^2 + \alpha_1^2 y^2} \\ \eta_2 &= \frac{1}{\alpha_2} \sqrt{(k_2^2 - k_1^2) - \alpha_1^2 y^2} \end{aligned} \quad (A-22)$$

In order to understand the dependence of V_{OR} and V_{OI} on the parameters a , b we substitute expressions (A-17) and (A-22) for η_2 and assume that $k_2^2 - k_1^2 \ll k_2^2$. The two factors in the curved brackets of (A-20) are plotted in Fig. A-4 for the case $a, b \gg 2\pi\alpha_2/\sqrt{k_2^2 - k_1^2}$. The plot of $(\sin \eta_1 b / \eta_1 b)^2$ is self explanatory. The plot of the second factor is obtained by approximating η_2 by $\sqrt{k_2^2 - k_1^2}/\alpha_2$ for η_1 small. Then the square bracket term in (A-20) for $\eta_1 \ll \sqrt{k_2^2 - k_1^2}/\alpha_1$ becomes

$$\left[\right] \approx \frac{\eta_1 \sqrt{k_2^2 - k_1^2} / \alpha_2}{\frac{k_2^2 - k_1^2}{\alpha_2^2} (\cos \eta_1 a)^2 + \eta_1^2 (\sin \eta_1 a)^2} \quad (A-23)$$

which has maxima when $\eta_1 a$ is an odd multiple of $\pi/2$, and minima when $\eta_1 a$ is a multiple of π .

For $\eta_1 \gg \sqrt{k_2^2 - k_1^2}/\alpha_1$, it is seen that $\eta_2 \approx \alpha_1 \eta_1 / \alpha_2$ and the square bracket term in (A-20) can be approximated as

$$\left[\right] \approx \frac{\alpha_2}{\alpha_1} \left\{ 1 - \frac{1}{2} \left[\left(\frac{\alpha_1^2}{\alpha_2^2} - 1 \right) + \frac{k_2^2 - k_1^2}{\alpha_2^2 \eta_1^2} \right] (1 - \cos 2\eta_1 a) \right\} \quad (A-24)$$

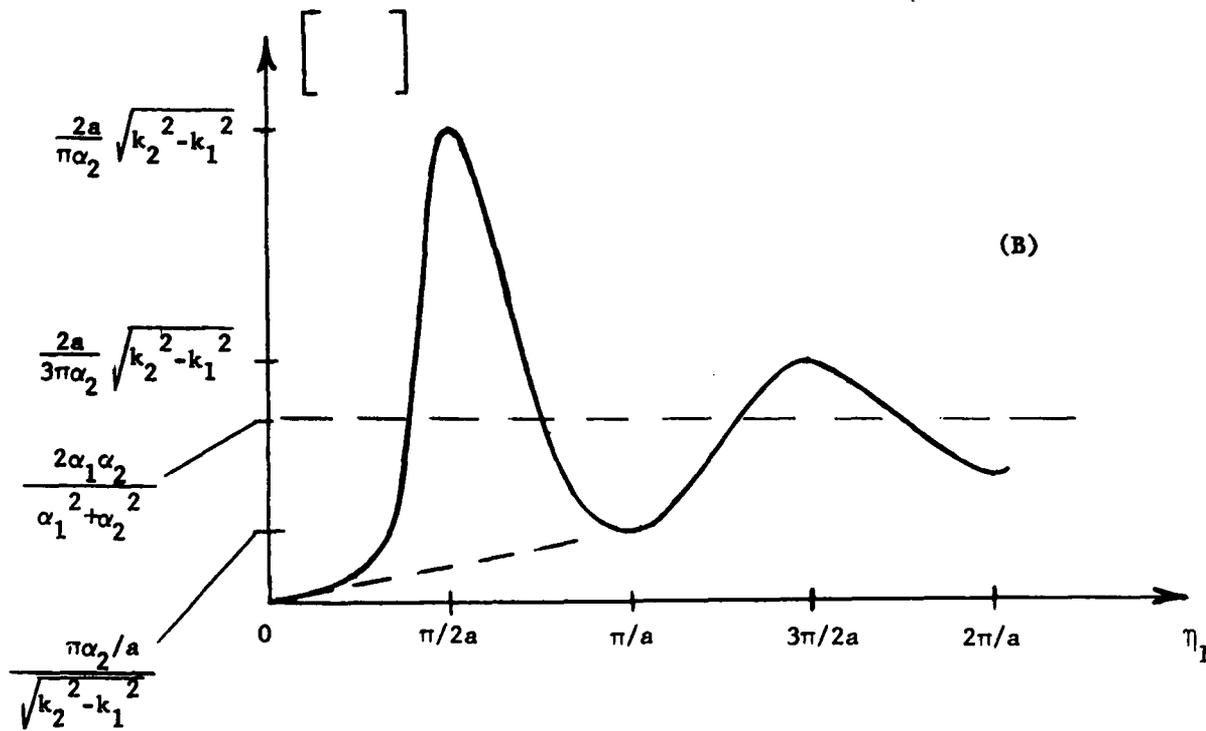
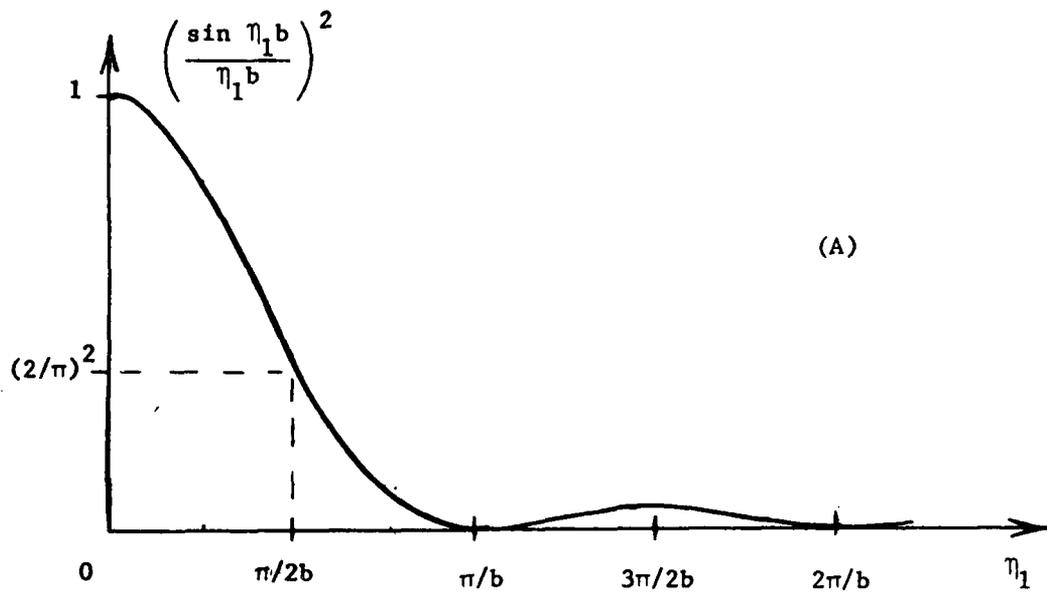


Fig. A-4. Variation of the factors in the integrand of V_{OR}

Because α_1 is close to α_2 , this factor is seen to have small oscillations about

$$\frac{\alpha_2}{\alpha_1} \left[1 - \frac{1}{2} \left(\frac{\alpha_1^2}{\alpha_2^2} - 1 \right) \right]. \quad (\text{A-25})$$

For example, the 37-X cut of LiNbO_3 has $\alpha_1^2 = 0.795$ and $\alpha_2^2 = 0.687$ so that $1/2(\alpha_1^2/\alpha_2^2 - 1) = 0.079$ and (A-25) has value 0.857.

When the electrode spacing $2a$ is large compared to the transducer aperture $2b$, the period of oscillation π/a of the factor in (A-23) or (A-24) is small compared to the period of oscillation π/b of $(\sin \eta_1 b / \eta_1 b)^2$. These rapid oscillations will cancel in the integration of (A-20), with the result that the integral is close to its value for $k_2 = k_1$. In other words, when $a \gg b$ the output voltage is close to that for transducers on a free surface.

For the case $a=b$, the product of the two factors of Fig. A-4 is sketched in Fig. A-5. For comparison, we have also sketched dashed in Fig. A-5 the term in the curved bracket of (A-21) for the same case $a=b$. The area under dashed the curve will be smaller than that under the solid curve by a factor proportional to the ratio of amplitudes of the two curves. This ratio is proportional to $\alpha_2^2 / [\alpha^2(k_2^2 - k_1^2)]$. For the 37-X cut of LiNbO_3 , this factor is $(0.33)(\lambda_1/\alpha)^2$. Thus for a wide aperture ($a \gg \lambda_1$), the contribution V_{OI} of (A-21) to the total output voltage V_o is small compared to V_{OR} of (A-20).

Because the principal contribution to the integration in (A-20) comes from $\eta_1 < 2\pi/b \ll k_1$, the factor $1/\kappa$ can be replaced by $1/k_1$. With the foregoing approximations, the diffraction loss D_L due to the electrodes and/or due to path length L is given by V_{OR} normalized to its value for $L=0$ and $k_1=k_2$. Thus

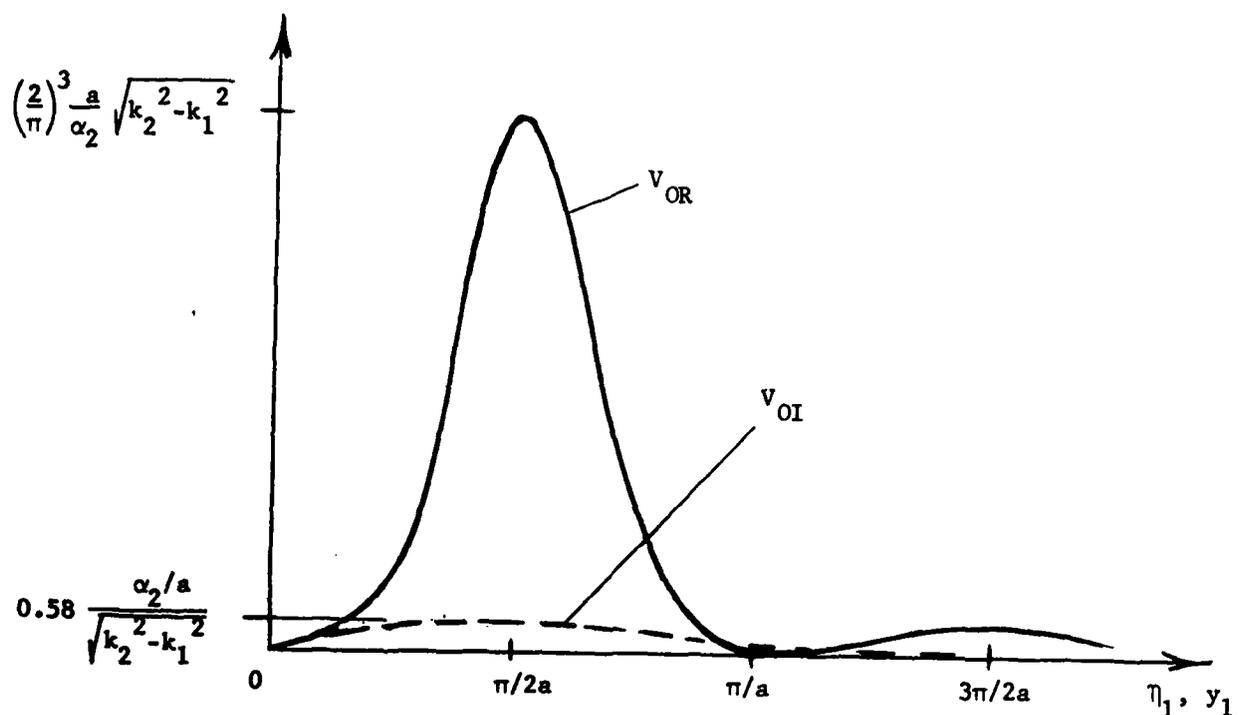


Fig. A-5. Comparison of the integrands for V_{OR} and V_{OI} .

$$D_L = \frac{2b}{\pi} \int_0^{\infty} \left(\frac{\sin \eta_1 b}{\eta_1 b} \right)^2 \left[\frac{\eta_1 \eta_2}{\eta_2^2 - (\eta_2^2 - \eta_1^2) (\sin \eta_1 a)^2} \right] e^{i(x-k_1)L} d\eta_1 \quad (\text{A-26})$$

where we have made use of

$$\int_0^{\infty} \left(\frac{\sin \eta_1 b}{\eta_1 b} \right)^2 d(\eta_1 b) = \frac{\pi}{2} . \quad (\text{A-27})$$

While (A-26) was derived using the quadratic approximation for the dispersion curve, the approximation served primarily to make clear the location of branch points and branch cuts during the deformation of the integration path. However, the branch points occur explicitly in (A-26) only via the definition of κ and η_2 , and not in integration path. This result suggests that if numerically generated dispersion curves were available for κ and η_2 , (A-26) could be used for materials for which the quadratic approximation is not valid.

To gain insight into (A-26), we approximate D_L for the case of large L and large $a = b$. When L is large, the exponential term undergoes rapid phase variation for $\eta_1 > \pi/b$, with the result that the contributions to the integral from neighboring segments of the η_1 axis cancel each other. However, the large first maximum of the function

$$f(\eta_1) = \frac{\eta_1 \eta_2}{\eta_2^2 - (\eta_2^2 - \eta_1^2) (\sin \eta_1 a)^2} \quad (\text{A-28})$$

evident in Fig. A-4(b) will significantly influence the value of the integral.

The shape of the maximum suggests an approximation of the form

$$f(\eta_1) \approx \eta_1 \frac{(2a/\pi)^2 \sqrt{k_2^2 - k_1^2/\alpha_2}}{1 + B^2 (\eta_1^2 - \pi/2a)^2} \quad (\text{A-29})$$

which gives the correct value of $f(\pi/2a)$. The coefficient B is found by requiring that the second derivative of the approximation (A-29) evaluated at $\pi/2a$ give the correct value for large a . This requirement leads to

$$B = \frac{2}{\pi} a^2 \sqrt{k_2^2 - k_1^2/\alpha_2} . \quad (\text{A-30})$$

The factor η_1 in (A-29) cancels one power of η_1 in the factor $(\sin \eta_1 b / \eta_1)^2$ in (A-26). The first two terms in the series expansion of $(\sin \eta_1 b / \eta_1)$ about $\pi/2b$ are

$$\frac{(\sin \eta_1 b)^2}{\eta_1} \approx b \frac{2}{\pi} \left[1 - \frac{2b}{\pi} \left(\eta_1 - \frac{\pi}{2b} \right) \right]. \quad (\text{A-31})$$

With the help of (A-29) and (A-31), the integral in (A-26) for $a=b$ can be written

$$D_L = \frac{2}{\pi} b^2 \frac{k_2^2 - k_1^2}{\alpha_2} \int_0^{\infty} \frac{(2 - \eta_1 2b/\pi) e^{i(\alpha - k_1)L}}{1 + B^2(\eta_1 - \pi/2b)^2} d\eta_1. \quad (\text{A-32})$$

Because of the large factor B in the denominator of (A-32), the integrand will be small for $\eta_1 < 0$. Hence we may extend the lower limit to $\eta_1 = -\infty$.

If we now use the approximation (17) for κ , and carry out a partial fraction expansion of the integrand, (A-32) becomes

$$D_L = \frac{b}{i2} \left(\frac{2}{\pi} \right)^4 \left\{ \left(\frac{\pi}{2} - i \frac{b}{B} \right) \int_{-\infty}^{\infty} \frac{\exp[-i(F/\pi)(\eta_1 b)^2]}{\eta_1 b - (\pi/2 + ib/B)} d\eta_1 - \left(\frac{\pi}{2} - i \frac{b}{B} \right) \int_{-\infty}^{\infty} \frac{\exp[-i(F/\pi)(\eta_1 b)^2]}{\eta_1 b - (\pi/2 - ib/B)} d\eta_1 \right\} \quad (\text{A-33})$$

where F is defined in (18). The integrals in (A-33) can now be evaluated in closed form.

To evaluate the integrals in (A-33), the paths of integration along the real η_1 axis are deformed into paths at 45° to the real axis starting in the second quadrant of the complex η_1 plane, and ending in the fourth quadrant. During the deformation, the pole at $(\pi/2 - ib/B)$ in the second integral is captured, giving rise to a residue contribution. The integrations along the slant path can be expressed in terms of the complementary error function. As a result of this evaluation one obtains [19]

$$D_L = \frac{1}{i2} \left(\frac{2}{\pi} \right)^4 \left\{ 2\pi i \left(\frac{\pi}{2} + i \frac{b}{B} \right) \exp \left[-i \frac{F}{\pi} \left(\frac{\pi}{2} - i \frac{b}{B} \right)^2 \right] + \left(\frac{\pi}{2} - i \frac{b}{B} \right) (i\pi) e^{-T_+^2} \operatorname{erfc}(-iT_+) - \left(\frac{\pi}{2} + i \frac{b}{B} \right) (i\pi) e^{-T_-^2} \operatorname{erfc}(-iT_-) \right\} \quad (\text{A-34})$$

where

$$T_{\pm} = \sqrt{\frac{F}{\pi}} \left(\frac{\pi}{2} \pm i \frac{b}{B} \right) e^{i\pi/4}. \quad (\text{A-35})$$

From (A-30), for $a=b$ it is seen that b/B is proportional to $1/b$, and hence for b large $b/B \ll \pi/2$. Also, for F large, $|T_{\pm}|$ will be large so that [13]

$$e^{-T^2} \operatorname{erfc}(-iT) \sim \frac{i}{\sqrt{\pi} T}. \quad (\text{A-36})$$

Using (A-36) in (A-34) and, after combining terms containing the complementary error function, neglecting b/B compared to $\pi/2$ in amplitude terms

$$D_L = 2 \left(\frac{2}{\pi} \right)^2 e^{-i\pi F/4} e^{-bF/B} + 4 \left(\frac{2}{\pi} \right)^4 \frac{b}{B\sqrt{F}} e^{-i\pi/4}. \quad (\text{A-37})$$

The approximations leading to (A-37) are valid for F greater than about three. Since $b/B \ll \pi/2$, the first term in (A-37) will be larger than the second for F less than about $(B/b)^2$.

APPENDIX B: Effects of Finite Width Electrodes

Electrodes of finite width w are shown in Fig. 2. Consistent with the representation of the SAW amplitude in terms of a potential function, as discussed in Appendix A, the impedance looking in the x -direction for waves having variation $\exp(ikz)$ along z is $1/\eta_1$ in the free surface region, and $1/\eta_2$ in the plated region. Thus the impedance seen at $z = a$ looking into the finite width plating region can be found from the transmission line analogue for $\exp(-i\omega t)$ time dependence as

$$Z_{in} = \frac{1}{\eta_2} \frac{(1/\eta_1) - i(1/\eta_2) \tan \eta_2 w}{(1/\eta_2) - i(1/\eta_1) \tan \eta_2 w} \quad (B-1)$$

The impedance (B-1) reflects waves incident along x with reflection coefficient

$$\Gamma = \frac{Z_{in} - (1/\eta_1)}{Z_{in} + (1/\eta_1)} \quad (B-2)$$

Substituting from (B-1) we obtain

$$\Gamma = \frac{i(\eta_2^2 - \eta_1^2) \tan \eta_2 w}{2\eta_1 \eta_2 - i(\eta_2^2 + \eta_1^2) \tan \eta_2 w} \quad (B-3)$$

This expression for Γ may now be used in the integral expression (A-14) for the output voltage V_o .

The reflection coefficient appears in (A-14) through the function

$$\hat{f}(\eta_1) = \frac{1 + \Gamma e^{i2\eta_1 a}}{1 + \Gamma e^{i2\eta_2 a}} \quad (B-4)$$

Substituting for Γ from (B-3), and after suitable manipulation, it is found that

$$\hat{f}(\eta_1) = \frac{(1 - i \frac{\eta_1}{\eta_2} \tan \eta_2 \frac{w}{2})(1 - i \frac{\eta_2}{\eta_1} \tan \eta_2 \frac{w}{2}) - i(\frac{\eta_1}{\eta_2} - \frac{\eta_2}{\eta_1}) \tan \eta_2 \frac{w}{2} e^{i2\eta_1 a}}{(1 - i \frac{\eta_1}{\eta_2} \tan \eta_2 \frac{w}{2})(1 - i \frac{\eta_2}{\eta_1} \tan \eta_2 \frac{w}{2}) + i(\frac{\eta_1}{\eta_2} - \frac{\eta_2}{\eta_1}) \tan \eta_2 \frac{w}{2} e^{i2\eta_1 a}} \quad (B-5)$$

It is seen from (B-5) that \hat{f} is unchanged by changing the sign of η_2 . Hence the branch points at $\pm k_2$ in Fig. A-2 are not present in the integrand. However, the branch points at $\pm k_1$ are present in \hat{f} , reflecting the fact that the free-surface region extends to $|x| \rightarrow \infty$.

The function \hat{f} will have one or more poles in the segment of the real κ axis between k_1 and k_2 . In this segment η_2 is real and η_1 is positive imaginary. If the original path of integration is deformed as shown in Fig. A-2, then the portion of the path from k_1 to k_2 and back again to k_1 will enclose these poles. Changing the variable of integration to η_1 results in the path P_1 of Fig. A-3 for the integral of (A-18). In the η_1 plane, the integration along the real axis gives the contribution from the continuous spectrum to V_0 . The portion of P_1 from the origin to the point $i\sqrt{k_2^2 - k_1^2}/\alpha_1$ and back to the origin encompasses the poles of \hat{f} , and gives the contribution to V_0 from the guided waves.

1. Continuous Spectrum Contribution

To compute the contribution from the continuous spectrum we note that $(\sin \eta_1 b / \eta_1 b)^2$ and κ are even functions of η_1 . Thus only the symmetric portion of \hat{f} will survive the integration along the real η_1 axis. Define symmetric part $f(\eta_1)$ as

$$f(\eta_1) = \frac{1}{2} [\hat{f}(\eta_1) + \hat{f}(\eta_1)] . \quad (B-6)$$

Recalling that η_2 is odd function of η_1 along the $\text{Re} \eta_1$ axis, after suitable manipulation it can be shown that

$$f(\eta_1) = \left\{ (\cos \eta_2 w)^2 - \left(\frac{\eta_1}{\eta_2} - \frac{\eta_2}{\eta_1} \right) (\sin \eta_2 w) (\cos \eta_2 w) (\sin 2\eta_1 a) \right. \\ \left. + \frac{1}{2} (\sin \eta_2 w)^2 \left[\left(\frac{\eta_1^2}{\eta_2^2} + \frac{\eta_2^2}{\eta_1^2} \right) - \left(\frac{\eta_1^2}{\eta_2^2} - \frac{\eta_2^2}{\eta_1^2} \right) \cos 2\eta_1 a \right] \right\}^{-1} \quad (B-7)$$

The output voltage V_{OR} due to the continuous spectrum is now given by (A-20) with the term in the square bracket replaced by $f(\eta_1)$. Thus

$$V_{OR} = \frac{2Qb^2}{i\pi} \int_0^{\infty} \left(\frac{\sin \eta_1 b}{\eta_1 b} \right)^2 f(\eta_1) \frac{e^{ik_1 L}}{k_1} d\eta_1. \quad (B-8)$$

Dividing this expression by

$$\frac{2Qb^2}{i\pi} \left[\int_0^{\infty} \left(\frac{\sin \eta_1 b}{\eta_1 b} \right)^2 d\eta_1 \right] \frac{e^{ik_1 L}}{k_1} = \frac{Qb}{ik_1} e^{ik_1 L} \quad (B-9)$$

normalizes V_{OR} to the output voltage when the two transducers are close together on a free surface. The normalized voltage is represented by the symbol D_f . Note that k in the denominator of (B-8) can be replaced by k_1 with little error for $b \gg \lambda_1$.

2. Guided Wave Contribution

The guided waves decrease exponentially in amplitude for $|x| \rightarrow \infty$. This delay results from the fact that η_1 is imaginary at the pole. Substituting $\eta_1 = iy$ into the denominator of \hat{F} in (B-5), and equating it to zero, gives the dispersion equation of the guided waves as

$$\left(1 + \frac{y}{\eta_2} \tan \eta_2 \frac{W}{2} \right) \left(1 - \frac{\eta_2}{y} \tan \eta_2 \frac{W}{2} \right) - \left(\frac{y}{\eta_2} - \frac{\eta_2}{y} \right) \tan \eta_2 \frac{W}{2} e^{-2ya} = 0. \quad (B-10)$$

The first term in (B-10) gives the dispersion equation for an isolated strip, while the second term represents the perturbation due to coupling between the strips [4]. If the strips are separated by several wavelengths, the coupling is small, and hence the guided wave solutions for the pair of strips will be close to the solutions for an isolated strip.

As a result of the coupling between strips, each mode of the isolated strip will split into two modes. One mode is symmetric with respect to the centerline

$x = 0$, and the other is antisymmetric. For symmetric transducers, as is assumed here, only the symmetric mode is excited. Because $f(\eta_1)$ describes interactions in a symmetric structure, its poles are only those of guided waves having even symmetry about the centerline.

The modes of an isolated strip have either even or odd symmetry about the center of the strip. Even modes, including the lowest mode, are given mode index ($n = 0, 2, 4, \dots$), and are solutions of the dispersion equation

$$y = \eta_2 \tan \eta_2 \frac{w}{2} . \quad (\text{B-11})$$

The dispersion equation of the odd modes, having mode index ($n = 1, 3, 5, \dots$), is

$$y = - \frac{\eta_2}{\tan \eta_2 \frac{w}{2}} . \quad (\text{B-12})$$

Insight into these solutions is gained by solving (15) for $y = i\eta_1$ in terms of η_2 real. Thus

$$y = \frac{1}{\alpha_1} \sqrt{(k_2^2 - k_1^2) - \alpha_2^2 \eta_2^2} . \quad (\text{B-13})$$

The solution to (B-11) or (B-12) is then given by the intersection of the plot of the right-hand side of these equations with the plot of the right hand side of (B-13). Such plots are indicated in Fig. B-1. It is seen that one or more real solutions for η_2 exist depending on the size of $\sqrt{k_2^2 - k_1^2} / \alpha_2$ relative to π/w . The cutoff frequency for the n th mode is found from the condition

$$\frac{1}{\alpha_2} \sqrt{k_1^2 - k_2^2} = \frac{n\pi}{w} . \quad (\text{B-14})$$

To find the voltage V_{OI} due to the guided modes let $\hat{N}(\eta_1)$ be the numerator of $\hat{f}(\eta_1)$ in (B-5), and let $\hat{D}(\eta_1)$ be its denominator. Since the poles are encircled in the clockwise direction, then the residue contribution is

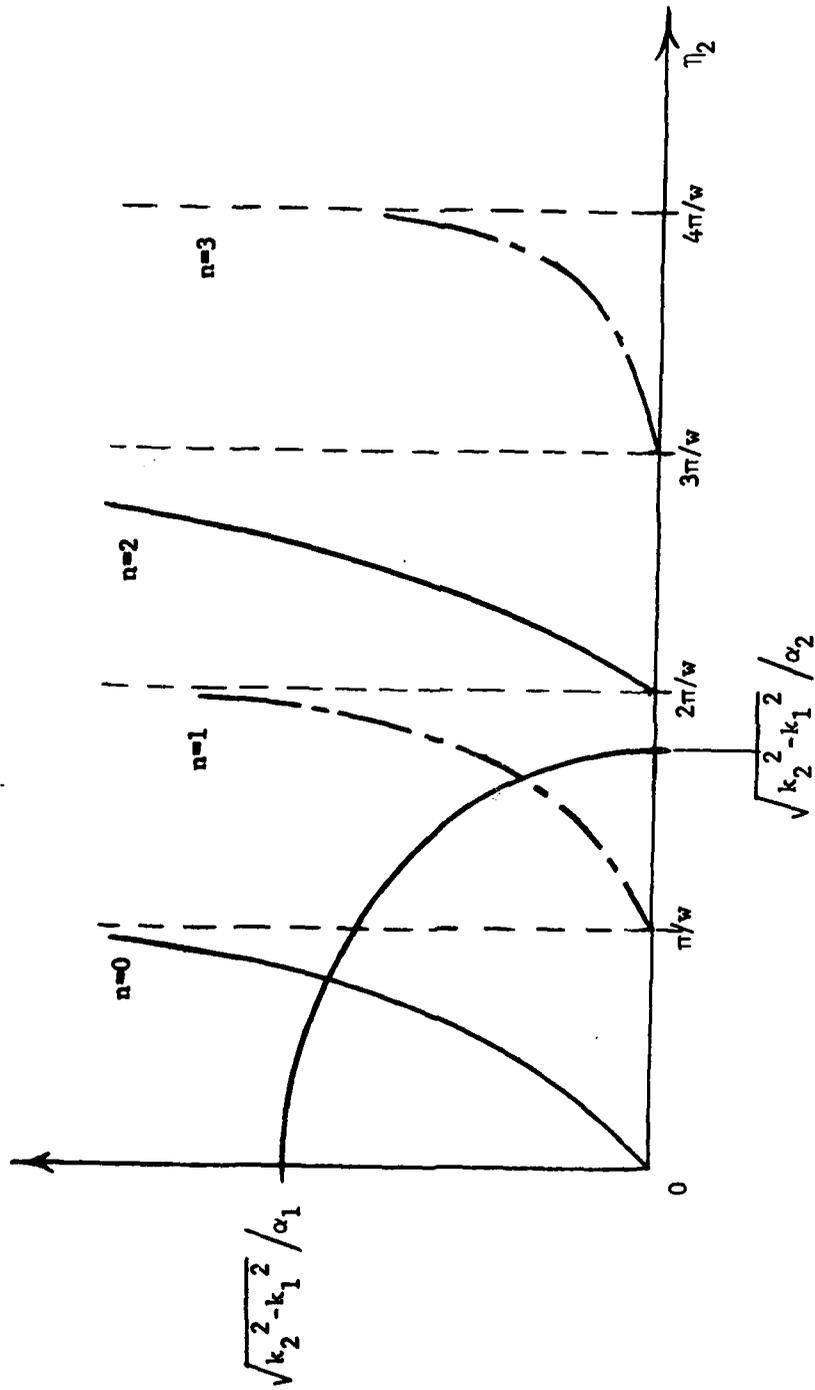


Fig. B-1. Construction for finding the solutions to the dispersion relation for an isolated strip.

$$V_{OI} = -2Qb^2 \sum_{n=0,1,\dots} \left[\left(\frac{\sin \eta_1 b}{\eta_1 b} \right)^2 \frac{\hat{N}(\eta_1)}{d\hat{D}(\eta_1)/d\eta_1} \frac{e^{i\alpha(\eta_1)L}}{\alpha(\eta_1)} \right]_{\eta_1 = iy_n} \quad (B-15)$$

where iy_n is the location of the n th pole, and the sum is taken over the poles for which η_1 is pure imaginary, corresponding to propagating guided waves.

In order to evaluate (B-15), we approximate the pole location by the solution of (B-11) or (B-12) for an isolated strip. For the even modes we use (B-10) and (B-11) to obtain

$$\begin{aligned} \hat{N}(iy_n) &= 2 \left(\frac{y_n}{\eta_2} + \frac{\eta_2}{y_n} \right) \tan \eta_2 \frac{w}{2} e^{-2y_n a} \\ &= 2 \left(1 + y_n^2 / \eta_2^2 \right) e^{-2y_n a}. \end{aligned} \quad (B-16)$$

For odd modes, (B-10) and (B-12) give

$$\hat{N}(iy_n) = -2 \left(1 + \eta_2^2 / y_n^2 \right) e^{-2y_n a}. \quad (B-17)$$

For $2y_n a$ large, the derivative of the term containing the exponential in $D(\eta_1)$ will be small, when evaluated at $\eta_1 = iy_n$, as compared to the derivatives of the other term. Neglecting the exponential, it can be shown that at the even numbered poles

$$\frac{d\hat{D}}{d\eta_1} = \left(1 - i \frac{\eta_1}{\eta_2} \tan \eta_2 \frac{w}{2} \right) \left[\frac{d}{d\eta_1} \left(-i \frac{\eta_2}{\eta_1} \tan \eta_2 \frac{w}{2} \right) \right] \quad (B-18)$$

while at the odd numbered poles

$$\frac{d\hat{D}}{d\eta_1} = \left[\frac{d}{d\eta_1} \left(-i \frac{\eta_1}{\eta_2} \tan \eta_2 \frac{w}{2} \right) \right] \left(1 - i \frac{\eta_2}{\eta_1} \tan \eta_2 \frac{w}{2} \right) \quad (B-19)$$

Using the parabolic approximation (15) for $\eta_2(\eta_1)$ and (B-11) or

(B-12), the various terms in the residue series (B-15) can be simplified. After some manipulation one finds that

$$V_{OI} = -4iQb^2 \sum_n \left(\frac{\sinh y_n b}{y_n b} \right)^2 \frac{y_n e^{-2y_n a} e^{i\alpha(iy_n)L} / \alpha(iy_n)}{\left(1 + \frac{\alpha_1^2}{\alpha_2^2} \frac{y_n^2}{\eta_{2n}^2} \right) + \frac{y_n w}{2} \frac{\alpha_1^2}{\alpha_2^2} \left(1 + \frac{y_n^2}{\eta_{2n}^2} \right)} \quad (B-20)$$

where the sum is taken over the propagating modes and η_{2n} is the value of η_2 at $\eta_1 = iy_n$. At cutoff for each mode $y_n = 0$, so that the contribution to (B-20) from each mode initially increases from zero as frequency increases above its cutoff. Far enough from cutoff so that $2y_{na} > 4$, $\sinh y_{nb} \approx (1/2)e^{y_{nb}}$, which simplifies the evaluation of (B-20). Dividing (B-20) by the factor (B-9) gives the normalized voltage that has been labeled D_g in Section II.

References

1. S.G. Joshi and B.B. Dasgupta, "Electronically Variable Surface Acoustic Wave Time Delay Using a Biasing Electric Field," Proc. 1981 IEEE Ultrasonics Symp., pp. 319-323.
2. S.G. Joshi, "A temperature Compensated High Voltage Probe Using Surface Acoustic Waves," Proc. 1982 IEEE Ultrasonics Symp., pp. 317-320.
3. A.J. Budreau, P.H. Carr and J.H. Silva, "New Configuration for Electronically Variable SAW Delay Line," Proc. 1982 IEEE Ultrasonics Symp., pp. 399-400.
4. R.V. Schmidt and L.A. Coldren, "Thin Film Acoustic Surface Waveguides on Anisotropic Media," IEEE Trans., SU-22 (1975), pp. 115-122.
5. H.F. Tiersten, "Elastic Surface Waves Guided by Thin Films," J. Appl. Phys., 40 (1969), pp. 770-789.
6. B.A. Auld, Acoustic Fields and Waves in Solids, Vol. II, Wiley and Sons, New York (1973), pp. 274-278.
7. K.L. Davis and J.F. Weller, "SAW Attenuation in Metal Film - Coated Delay Lines," Proc. 1979 IEEE Ultrasonics Symp., pp. 659-662.
8. A.J. Slobodnik, Jr., P.H. Carr and A.J. Budreau, "Microwave Frequency Acoustic Surface-Wave Loss Mechanisms on LiNbO_3 ", JAP, 41 (1970), pp. 4380-4387.
9. D.R. Snider, H.P. Fredricksen and S.C. Schneider, "Surface Acoustic Wave Attenuation by a Thin Film," JAP, 52 (1981), pp. 3215-3222.
10. E. Durand, Electrostatique at Magnetostatique, Masson, Paris (1953), pp. 302-303.
11. T.L. Szabo and A.J. Slobodnik, Jr., "The Effect of Diffraction on the Design of Acoustic Surface Wave Devices," IEEE Trans., SU-20 (1973), pp. 240-251.
12. P. Carr, private conversation.

13. M. Abramowitz and I.A. Stegun, Handbook of Mathematical Functions, National Bureau of Standards, Washington, DC (1964), pp. 297-298.

14. K. Tsubouchi, K. Sugai and N. Mikoshiba, "Zero Temperature Coefficient Surface-Acoustic-Wave Devices Using Epitaxial AlN Films," Proceedings of the 1982 Ultrasonics Symposium, pp. 340-345.

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

DILEMMAS OF COMBAT PSYCHIATRY: WORLD WAR II AND VIETNAM

Prepared by:	Dr. Willie A. Bragg	Cynthia B. Bragg
Academic Rank:	Assistant Professor	Graduate Student
Department and University:	Department of Early Childhood and Special Education University of Cincinnati	Social and Philosophical Foundations of Education Rutgers, The State University
Research Location:	Neuropsychiatry Branch, Clinical Sciences Division, USAF School of Aerospace Medicine, Brooks Air Force Base, Texas	
USAF Research Colleague:	Drs. David Jones and Bryce Hartman	
Date:	August 10, 1983	
Contract No:	F49620-82-C-0035	

DILEMMAS OF COMBAT PSYCHIATRY: WORLD WAR II

AND VIETNAM

by

Willie A. Bragg
and
Cynthia B. Bragg

ABSTRACT

A review was conducted to investigate issues surrounding combat psychiatry. The focus of this investigation was to examine the development of combat psychiatry from the Civil War to Vietnam. Special attention was given to World War II and Vietnam conflict relative to incidence, diagnosis and treatment of psychiatric casualties. The results demonstrate the significant gains made in preventive psychiatry. Recommendations for further study are outlined.

ACKNOWLEDGEMENT

This research project could not have been conducted without the sponsorship of the Air Force Systems Command, the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education.

I would like to thank the Air Force personnel at the Clinical Sciences Division, especially the Neuropsychiatry Branch for their hospitality and generosity. I would also like to thank Dr. David Jones for his assistance and advice during this project.

Appreciation is extended to Mr. Lloyd Foster for his encouragement and support during this research.

Finally, special thanks is extended to Dr. Bryce Hartman for his guidance throughout this endeavor.

I. INTRODUCTION

What affects teacher performance and why have they chosen other professions? These are just a few of the many questions asked by educators. Stress, related to teacher burn-out, has been the focus of much discussion. Attention has been given to the underlying causes of teacher burn-out. Preventive measures including programs and workshops have been initiated to address this syndrome. Educators, more specifically, teachers have complained about heavy teaching assignments, large class enrollments, insufficient materials, apathetic administration and lack of parental support as possible reasons for teacher burn-out. These frustrating situations can, and in many instances, do affect teacher performance in the classroom. Thus, the teacher's inability to adequately carry out classroom responsibilities has been linked to "teacher burn-out" syndrome.

Similar to the educational community, combat psychiatry has addressed the issue of stress during war. Simply stated, "soldier burn-out" syndrome can be defined as the soldier's inability to perform military duties during combat situations. Therefore, soldier burn-out is analogous to "teacher burn-out" in many respects. Several researchers have suggested that environmental and personality variables affect combat stress breakdown while others infer that duration and nature of combat significantly influence the disorder. Although unanswered questions remain, this paper has addressed some issues surrounding treatment of psychiatric casualties during wartime.

II. OBJECTIVES

The primary purpose of this project was to examine the causes of psychiatric casualties, symptoms exhibited and treatment procedures utilized during combat. Special attention was given to World War II and the Vietnam conflict since these wars differed along several dimensions. No attempt was made to isolate techniques relative to department (i.e., Army, Navy, Air Force) but to outline approaches and principles of combat psychiatry. The specific objectives were as follows:

- (1) To present a historical review of combat psychiatry as it relates to terminology and treatment of psychiatric casualties.
- (2) To examine etiology of psychiatric casualties during World War II and Vietnam conflict.
- (3) To isolate symptoms exhibited by servicemen during World War II and Vietnam conflict.
- (4) To review treatment procedures used by military psychiatrists during World War II and Vietnam conflict.

III. APPROACH

To accomplish the above objectives, a comprehensive review of the literature was conducted with the assistance of Strughold Aeromedical and Base Libraries. Documents concerning combat psychiatry from the American Civil War to Vietnam were collected.

Second, an annotated bibliography was prepared. This involved article review and summary.

Third, material relevant to World War II and Vietnam conflict was isolated. This aided in the completion of the final report.

Fourth, all articles concerning combat psychiatry were categorized and key terms were selected. This procedure was necessary in preparing a computerized storage and retrieval program on combat psychiatry.

A. Historical Review

The Civil War marked the beginning of combat psychiatry. It was during this time that the Union and Confederate Armies attempted to use hospital ships to evacuate the wounded. Medical attention was given to nonpsychotic emotional disorders of war, which were classified as "nostalgia." This diagnostic term reflected the belief that soldiers were homesick and that they lacked sufficient character to adjust to the combat situation (1, 2).

Hausman and Rioch (3) also suggested that the term nostalgia adequately described this disorder. However, they implied that little official concern was given to psychiatric casualties until William Hammond, who was Surgeon General of the Union Army, described the syndrome. They reported the incidence of nostalgia during the Civil War as 2.34 cases per 1,000 troops in the first year and 3.3 per 1,000 troops in the second year.

Arthur (4) and Johnson (5) defined combat breakdown during World War I. They stated that men were constantly exposed to danger, disease, fatigue and unknown horrors. Thus, the terms "shell shock" and "war neuroses" were used to describe the large numbers of psychiatric casualties. Implicit in the terminology was the association of artillery fire with some part of the brain. Miller (6) reported the incidence of psychiatric casualties as 34 per 1,000 troops.

Salmon emphasized the importance of treatment during World War I. Principles of treatment included (1) proximity - as close to the battle scene as possible, (2) immediacy - as soon as possible and (3) expectancy - the understanding that the psychiatric casualty would return to duty.

The beginning of World War II signaled an abrupt change in policy concerning the problem of psychiatric casualties. Military personnel felt that psychiatric problems could be prevented by exclusion. Thus, a massive screening program was developed to identify persons who were "at risk" or who showed signs of poor physical stamina and low intelligence. Because the incidence of psychiatric casualties drastically increased during the war, the screening program was abandoned and a return to the cardinal principles of proximity, immediacy and expectancy were seen (7). The important diagnostic development during World War II was the change in terminology for psychiatric casualties. "Exhaustion" rather than war neuroses was used to denote prolonged or heavy combat exposure. Modifications of exhaustion were used such as "combat fatigue" and "flying fatigue" (8).

The psychiatric literature on the Korean War demonstrated the continuation of psychiatric principles utilized during World War II. In addition, neuropsychiatric teams were established. These treatment procedures were effectively used until the Vietnam conflict. "Combat exhaustion" was the term used to describe psychiatric casualties (9).

During the Vietnam conflict, psychiatrists continued to investigate the stresses of combat. "Combat fatigue" was applied to Vietnam casualties.

term implied that soldiers with healthy personalities developed this condition due to extreme stress.

On the other hand, "combat neurosis" was applied to those individuals with long-standing neurotic problems aggravated by combat. Thus, the emergence of "pseudocombat fatigue" was seen which described men who had a history of social maladaptiveness. Combat stress was not seen as the primary cause of breakdown (10). Consequently, combat fatigue and pseudocombat fatigue were terms used to denote different underlying causes. Although the terms differed, psychiatric casualties in all major conflicts starting with the Civil War represented a small percentage of individuals treated. According to De Fazio (11), the psychiatric casualty rate for all departments was lower for the Vietnam conflict. Overall, it was 12 per 1,000 as compared to 37 per 1,000 in the Korean War and 101 per 1,000 in World War II (See Table 1).

The causes of psychiatric casualties have made a complete circle over the past years. Terminology has changed and treatment has been modified to reflect these changes (See Table 2).

Based on this review, contradictions exist between Vietnam and World War II. Therefore, the following sections examine the issues surrounding these two wars with emphasis on incidence, symptoms and treatment of psychiatric casualties.

B. World War II

1. Nature

World War II began in Europe on September 1, 1939. Honoring their pledge to Poland, Britain and France declared war on Germany September 3, 1939. Psychiatric casualties were numerous following the Tunisia fighting, the first large scale land combat of the U.S. Army.

TABLE 1
 PSYCHIATRIC CASUALTIES (1861-1975)

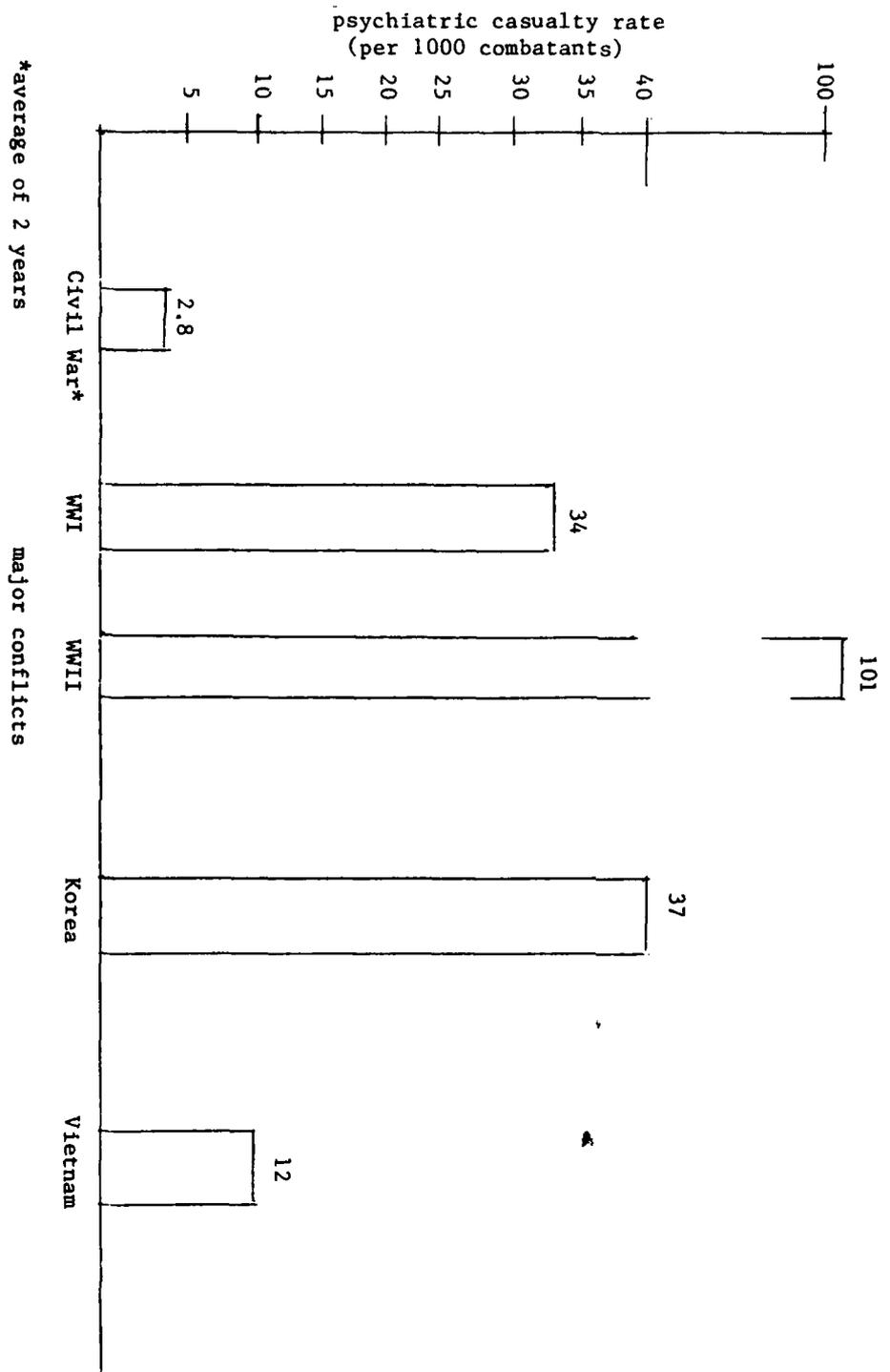


TABLE 2
DEVELOPMENT OF MILITARY PSYCHIATRY

WAR	YEAR	TERMINOLOGY	TREATMENT
American Civil War	1861-1865	Nostalgia	Evacuation/Hospital
World War I	1914-1918	War Neuroses Shell Shock Battle Shock	Principles of Psychiatry (Salmon) Proximity Immediacy Expectancy
World War II	1939-1945	Exhaustion Combat Fatigue Flying Fatigue Operational Fatigue Old Sergeants Syndrome Psychoneuroses	Screening Program Reinstatement of Psychiatric Principles
Korean War	1950-1953	Combat Exhaustion	Continuation of Psychiatric Principles Establishment of Neuropsychiatric Teams Close to Battalion
Vietnam War	1957-1975	Combat Fatigue Combat Neurosis Pseudocombat Fatigue	Varied Approaches Drug Therapy Hypnotherapy, Psychotherapy Psychiatric Principles Proximity Immediacy Expectancy Simplicity Centrality

World War II was a war of movement, with distant objects to be attained in continual combat. Soldiers fought on land and carried weapons, ammunition and other supplies on foot. Physical fatigue and emotional stress was attributed to heavy and prolonged battle. During the latter part of the war, psychiatric casualties were no longer viewed as "psychoneurosis" but as "exhaustion." Consequently, the nature of the war affected symptoms and dictated treatment (12).

3. Symptoms

Swank and Marchand (13) described American soldiers' reactions before, during and after combat. In addition to self-inflicted wounds, palpitations, abdominal pains, vomiting and backaches were common complaints of soldiers prior to combat. During the early days of combat, urinary frequency, intense thirst, sweating, anorexia, fear of eating and being left alone were indicators of early signs of exhaustion. The authors also pointed out that the onset of exhaustion began to appear between 25 and 30 days of combat. This condition was marked by more frequent occurrences of fear. The soldier began to lose confidence in himself and became overcautious, staying close to his trench. Irritability, memory lapse and sleeplessness became apparent during mounting states of exhaustion.

Similar to Swank and Marchand's observations, Grinker and Spiegel (14) examined combat disorders. They grouped cases of psychiatric casualties into five categories: (1) free anxiety states, (2) phobic states, (3) conversion states, (4) psychosomatic reactions and (5) depressions.

The desire to escape and fear of death was seen by the Germans as secondary psychological causes of war neuroses. They also reported cases of extreme depression, apathy and psychosomatic complaints. However, psychiatric casualties were rare in the German Army.

Cline and Rath (15) viewed the breakdown of the French Army in its entirety as a psychiatric casualty. The devastating defeat of the French by the German Army caused complete disorganization including lack of group cohesion, leadership and morale.

3. Treatment

As previously mentioned, the development of neuropsychiatry division was a major advancement for American combat psychiatry. During World War II, psychiatric teams were established and placed in close proximity to the fighting. Psychiatric personnel also trained general medical officers.

Mental hygiene consultation services were established in the continental U.S. placing mental health personnel close to the soldiers being evacuated and treated. The units which were established at training camps, concentrated on personnel classification, psychiatric services and consulting and preventive psychiatry. Implementation of preventive psychiatric principles included (1) early detection of psychiatric problems of individuals or groups, (2) early treatment of individuals in area closest to combat situation, (3) avoidance of hospital atmosphere and (4) early return to duty.

The German and French medical personnel also used similar treatment techniques. Psychiatric casualties were given brief periods of rest away from the hospital.

C. Vietnam Conflict

1. Nature

The involvement of the United States in Vietnam was a different kind of experience. U.S. Participation in previous wars had been clearly understood and supported. However, antiwar demonstrations and political unrest surrounded the Vietnam conflict. Soldiers began to question the purpose of the war and concentrated on surviving the experience. The uncertainty of U.S. involvement in Vietnam had deleterious effects on leadership, cohesion and morale of the soldiers.

The guerrilla-like warfare made it difficult to distinguish friend and foe. The possibility of being killed by unseen or unknown enemies generated mistrust among the soldiers. In addition, vast differences (i.e., race, language, culture) existed between the Vietnamese people and American soldiers (16).

2. Symptoms

According to Jones (17) servicemen moved through three phases during their Vietnam experiences. The arrival period, occurring between three weeks and three months was categorized as apprehensive enthusiasm. Jones related the apprehension to fear of death, separation from family and expectations of discomfort. Sleepwalking, enuresis and nightmares were signs exhibited by soldiers.

Resignation described the middle stage which lasted until one month before evacuation. Although the soldiers had adjusted to the combat environment, signs of loneliness, loss of appetite and weight and psychosomatic complaints became evident.

Greenberg, Pearlman and Gampel (18) also observed symptoms of Vietnam servicemen. They investigated the sleep patterns of nine combat fatigued soldiers and concluded that short REM latency and variability of REM time were directly related to the disorder.

Similar results were shown in the Schosberg's study. Schosberg and Benjamin (19) examined sleep patterns of combat fatigued soldiers. Their findings showed that the individuals exhibited REM deficiency, elimination of Stage IV and shorter sleep periods.

Much has been written about the incidence and cause of psychiatric disorders. While Walker and Cavenar (20) and Renner (21) attribute the inability to detect psychiatric casualties, Moskos (22) and Tiffany and Allerton (23) point out the 12-month rotation policy as the primary reason for few psychiatric casualties during Vietnam. Because symptoms and diagnosis varied, treatment modifications were made during Vietnam conflict.

3. Treatment

During the early phases of the conflict, combat casualties were treated by the use of phenothiazines and other tranquilizing drugs. Hypnotherapy and psychotherapy have also been reported as treatment measures. However, implementation of psychiatric principles (i.e., proximity, immediacy, expectancy, simplicity, centrality) have proved more beneficial in treating psychiatric casualties.

IV. RESULTS

The findings of this investigation show that the development of combat psychiatry is reflected in previous wars. Modifications relative to treatment

have had a significant impact on military psychiatry. As diagnosis and treatment changed, terminology changed.

Various signs of battle breakdown were observed in servicemen during World War II and Vietnam. They included anorexia, nausea, bedwetting and sleep disorders. These symptoms were viewed by psychiatrists as overt manifestations of battle fatigue and were treated in several ways.

Extreme stress due to combat was seen as the primary cause of breakdown during World War II while uncertainty and adjustment were viewed as reasons for breakdown during the Vietnam conflict. It is unclear as to the effects of other variables such as personality and military conditions (e.g., leadership, unit cohesion) on psychiatric casualties.

Striking differences were noticed in the number of psychiatric casualties during World War II and Vietnam. Reasons for these differences are speculative.

V. RECOMMENDATIONS

The following recommendations are based on the review of material concerning combat psychiatry.

1. It is suggested that military psychiatrists scrutinize the nature of future wars. This will enable projection of possible symptoms and facilitate treatment measures. The results of the information gathered may also be applicable to civilian psychiatry.
2. It is suggested that service personnel be trained in the use of "coping" mechanisms for extreme isolation. Future wars may not foster aspects of comradery such as group cohesion and support. Therefore, servicemen and

women must be able to adjust to the war setting in order to effectively carry out military duties.

3. It is suggested that military personnel be trained to identify and treat psychiatric casualties. If psychiatric teams are located a distance from the war zones, "buddy care" may be required to provide temporary treatment.

REFERENCES

1. Kormos, H. R. The Nature of Combat Stress. In C. R. Figley (Ed.).
Stress Disorders Among Vietnam Veterans. New York: Brunner/Mazel, 1978.
2. Baker, S. L. Traumatic War Disorders. In H. I. Kaplan and A. M. Friedman
(Eds.), Comprehensive Textbook of Psychiatry. Baltimore: Williams &
Williams, 1980.
3. Hausman, W. and Rioch, D. Military Psychiatry. Arch. Gen. Psychology,
16:722-739, 1967.
4. Arthur, R. J. Reflections on Military Psychiatry. American Journal of
Psychiatry, 135:2-7, 1978.
5. Johnson, A. W. Combat Psychiatry: A Historical Review. Medical Bulletin,
26:305-308, 1969.
6. Miller, E. The Neuroses in War. New York: The MacMillan Company, 1944.
7. Balson, P. M. and Dempster, C. R. Treatment of War Neuroses from Vietnam.
Comprehensive Psychiatry, 21:167-175, 1980.
8. Glass, A. J. Combat Exhaustion. U.S. Armed Forces Medical Journal,
2:1471-1478, 1951.
9. Hales, R. E. and Jones, F. D. Teaching Principles of Combat Psychiatry
to Army Psychiatry Residents. Military Medicine, 148:24-27, 1983.
10. Boydston, J. A. and Perry, C. J. Military Psychiatry. In H. I. Kaplan
and A. M. Friedman (Eds.), Comprehensive Textbook of Psychiatry.
Baltimore: Williams & Williams, 1980.

11. De Fazio, V. J. Dynamic Perspectives on the Nature and Effects of Combat Stress. In C. R. Figley (Ed.), Stress Disorders Among Vietnam Veterans. New York: Brunner/Mazel, 1978.
12. Stokesbury, J. L. A Short History of World War II. New York: William Morrow & Company, Inc., 1980.
13. Swank, R. L. and Marchand, W. E. Combat Neuroses: Development of Combat Exhaustion. Archives of Neurology and Psychiatry, 55:236-247, 1946.
14. Grinker, R. R. and Spiegel, J. P. Men Under Stress. New York: McGraw Hill, 1945.
15. Cline, R. W. and Rath, F. The Concept of an Army as a Psychiatric Casualty. Journal of the Royal Army Medical Corps, 128:79-88, 1982.
16. Howard, S. The Vietnam Warrior: His Experience and Implications for Psychotherapy. American Journal of Psychotherapy, 30:121-135, 1976.
17. Watson, P. War on the Mind: The Military Uses and Abuses. Chapter II, Combat Psychiatry. New York: Basic Books, Inc., Publishers, 1978.
18. Greenberg, R., Pearlman, C. and Gampel, D. War Neuroses and the Adaptive Function of REM Sleep. British Journal of Medical Psychology, 45:27-33, 1971.
19. Schosberg, A. and Benjamin, M. Sleep Patterns in Three Acute Combat Fatigue Cases. The Journal of Clinical Psychiatry, 39:546-549, 1978.
20. Walker, J. I. and Cavenar, J. O. Their Problems Continue. The Journal of Nervous and Mental Disease, 170:174-180, 1980.
21. Renner, J. A. The Changing Patterns of Psychiatric Problems in Vietnam. Comprehensive Psychiatry, 14:169-181, 1973.

22. Moskos, C. C. The American Combat Soldier in Vietnam. *Journal of Social Issues*, 31:25-37, 1975.
23. Tiffany, W. J. and Allerton, W. S. Army Psychiatry in the Mid-60's. *American Journal of Psychiatry*, 123:810-818, 1967.

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

EFFECTS OF RADIATIVE HEAT FEEDBACK

ON SOLID ROCKET PROPELLANT COMBUSTION

Prepared by:	M. Quinn Brewster
Academic Rank:	Assistant Professor
Department and University:	Department of Mechanical and Industrial Engineering, University of Utah
Research Location:	Air Force Rocket Propulsion Laboratory (DYC)
USAF Research:	Jay N. Levine
Date:	September 9, 1983
Contract No.:	F49620-82-C-0035

EFFECTS OF RADIATIVE HEAT FEEDBACK
ON SOLID ROCKET PROPELLANT COMBUSTION

by

M. Quinn Brewster

ABSTRACT

The effects of particle-enhanced radiative heat feedback on steady and transient combustion of solid propellants is investigated. Radiative influence on the steady-state burn rate, the linearized pressure-coupled response function and ignition/extinction phenomena are considered. It is shown that under typical motor conditions although the radiative flux is usually a small fraction of the total heat flux, the radiative flux is often preferentially absorbed by particle additives in the propellant. This results in pre-heating of the particles which can produce a "glo-plug" effect at the burning propellant surface wherein hot particles (inert or reactive) act as ignition sources for gas-phase chemical reactions. This effect may be profitably exploited in controlling high-frequency instability. It is also shown that at low pressures and low burn rates particle radiation can be a controlling influence in ignition/extinction phenomena. Particularly, the variation of the lower pressure deflagration limit with particle loading can be completely accounted for with proper consideration of particle radiation. Many such observations, often arbitrarily attributed to catalytic and other effects, may be actually radiative in origin.

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 the opportunity to pursue research at the Air Force Rocket Propulsion Laboratory. Assistance from the staff and personnel of the laboratory was greatly appreciated. In particular stimulating discussions, helpful suggestions and the cordial friendship of Jay Levine and Wayne Roe are gratefully acknowledged.

I. INTRODUCTION AND OBJECTIVES

Essentially every effort directed at modelling of solid rocket propellant combustion, both steady and unsteady, has included the assumption that radiative heat feedback to the propellant surface may be neglected. This assumption certainly appears to be justified at first glance. A simple energy balance made around a small control volume at the burning propellant surface (Section II.A) indicates that indeed, under typical motor operating conditions, even the maximum theoretical black-body radiation is usually a very small fraction of the total required heat flux. However several physically observed combustion phenomena which can be linked with the presence or absence of particles in the propellant have thus far successfully defied completely accurate explanation. Foremost among these are the problems of combustion instability and ignition/extinction. With a realization of the profound influence the presence or absence of solid or liquid particles can have on the heat transfer in other combustion systems (e.g. sooty hydrocarbon flames and metal droplet combustion), this study was undertaken to investigate the possible role radiative heat transfer may play in the complex processes controlling solid propellant combustion. An extensive literature search was performed looking for common threads and forgotten or perhaps overlooked observations and as a result two potentially significant conclusions were drawn. One involves particle heat transfer, radiation and the role of particles in suppressing combustion instability, called the "glo-plug" effect and the other concerns the role of particles and radiation in observed extinction and ignition of laboratory propellant strands.

I.A. Radiative Feedback Versus Conductive

The natures of radiative and conductive heat transport are fundamentally so different that a brief mention of these differences will serve to guide interpretation of further observations and results.

I.A.1. Spatial Influence: Volume Versus Local Thermal radiation is described as both energy transported by electromagnetic waves and by photons. It can propagate without the presence of a medium (i.e., in a vacuum). Its effect is a volume effect. That is, spatially remote elemental volumes of matter may exchange energy directly by radiation but only indirectly by conduction (molecular diffusion) through the intervening medium. The mathematical consequence of this is that integro-differential equations describe radiative transport and differential equations describe conductive transport (convection is considered to be merely fluid motion coupled with conduction). The physical consequence of this for combustion of solid propellants is that a synchronizing mechanism exists for correlating otherwise random events occurring at spatially remote locations. For example a random increase in the gas temperature directly above the burning propellant surface, as might be caused by one or more aluminum particles igniting and ejecting from the surface, would not be expected to influence the burn rate significantly except near that one location if conduction alone were considered. Thus, remotely occurring events would remain unsynchronized (in the absence of any other synchronizing influence). Radiative transport, however, may serve as a synchronizing mechanism to correlate random events (see Section III.C).

I.A.2. Temporal Influence: Speed of Light Versus Molecular Diffusion Since radiation travels at the speed of light the usual time constant associated with transient conduction in a material with finite thermal capacitance is not present. One consequence of this for burning propellants is that the increase in the real component of the pressure-coupled response function associated with the thermal inertia of the solid at low frequencies and the gas at high frequencies may be diminished by increasing the fraction of total heat feedback due to radiation.

I.A.3. Spectrally Selective Absorption And Emission The propagation of radiation is governed by the optical constants of the medium (here taken to be the complex refractive index $\tilde{n} = n - ik$). The real part n governs the speed of propagation and the imaginary part k (which can be related to the absorption coefficient k_a by $k_a = 4\pi k/\lambda$, λ = wavelength in vacuum) governs extinction of the waves by absorption. Solid materials generally emit and absorb in a continuous fashion with respect to wavelength with relatively minor variations in the optical constants. Gases emit and absorb in bands, if at all. One consequence of this in solid propellant combustion is that a significant portion of the radiation flux incident on the propellant surface can be absorbed preferentially by particles in the fuel/oxidizer matrix which itself is comparatively transparent to radiation in most cases. This can result in a preferential pre-heating of the particles due to the typically low thermal conductivity of the surrounding fuel/oxidizer matrix. This effect can be significant even if the radiant heat feedback is much smaller in magnitude than the total (conductive) heat flux because the

particles usually constitute such a small fraction of the total propellant composition. In the case of aluminized propellants the preferentially absorbed radiant flux could likely be sufficient to initiate combustion of aluminum somewhere below the propellant surface whereupon further increase in particle temperature would be the result of aluminum combustion. In either case the appearance of preferentially heated particles at the propellant surface forms the basis of the "glo-plug" postulate of suppressing combustion instability (Section III.B.2) which is felt to be the primary significant result of this study.

I.B. Gas-phase Versus Condensed-phase Characteristics

Accurate description of both gas-phase and condensed-phase is necessary for complete radiative transport modelling. It will be assumed here that radiative transport in the gas-phase is dominated by the continuous (black-body) radiation of the solid or liquid particles present. This is a reasonable first-approximation, although the influence of gaseous radiation may certainly be non-negligible at elevated pressures.

The situation in the solid or condensed phase may be considerably more complicated. Double-base propellants are largely transparent to radiation. In fact, early studies [1] on effects of radiation in double-base propellant combustion were motivated by the problem of premature ignition below the grain boundary due to absorption of radiation by particulate impurities. Ammonium perchlorate (AP) is also transmissive below about 2 microns [2]. Polymeric fuel binders tend to be more absorptive than AP but the absorption occurs in many narrow spectral regions. Furthermore, in composite propellants, mis-

matches in the refractive index at the boundaries between oxidizer particles and binder will cause scattering of radiation. Thus the problem is one of heat transport by combined conduction and radiation in an inhomogeneous, scattering, absorbing (and possibly emitting) medium.

The direction of net radiant flux will, in most practical situations, be from gas-phase to solid. Such will be the case in motor environments and in laboratory strand experiments when an external radiant flux is imposed. Thus the effective absorptivity of the solid and emissivity of the gas will be of primary importance. In the case of ignition/extinction studies involving propellant strand, however, the net radiative flux may be either toward or away from the solid. In either of the latter two cases the usually unimportant effective emissivity of the propellant must then be considered.

II. STEADY-STATE BURNING

II.A. Simple Energy Balance-Relative Magnitude of Radiative Flux

In this section a simple energy balance will be carried out to indicate the approximate magnitude of the radiative flux compared with the total heat flux. Consider the model pictured in Fig. 1 with the solid propellant initially at temperature T_0 approaching the origin ($x = 0$) at the gas-interface with linear mass regression rate m . The possibility of sensible heat release by condensed-phase chemical reactions, surface phase change, and gas-phase chemical reactions is indicated by the source/sink terms Q_s , L_s , and Q_f , respectively ($Q_{s,f} > 0$: exothermic; $L_s > 0$: endothermic). Neglecting relatively minor

enthalpy changes due to the appearance or disappearance of chemical species (other than the obvious contributions Q_s , L_s and Q_f) the steady-state first-law energy balances on control volumes 1, 2 and 3 take the following forms:

$$\text{Gas (cv1): } kg (dT/dx)_{s+} = m[Q_f - C_p(T_f - T_s)] \quad (1)$$

$$\text{Surface (cv2): } mL_s + k_\infty (dT/dx)_{s-} = kg (dT/dx)_{s+} \quad (2)$$

$$\text{Solid (cv3): } k_\infty (dT/dx)_{s-} = m[C_s(T_s - T_0) - Q_s] \quad (3)$$

The resulting total energy balance is

$$\text{Total: } L_s + C_s(T_s - T_0) = Q_s + Q_f - C_p(T_f - T_s) \quad (4)$$

where the thermodynamic relation

$$dL_s = (C_p - C_s)dT_s \quad (5)$$

ensures that T_f is independent of T_s . Radiation, if present, would simply augment the conductive heat flux terms, $k(dT/dx)$. (For more details including simple chemistry see [3].)

The magnitude of the heat feedback can be estimated in the following way. T_0 , C_s , C_p and $Q_s + Q_f - L_s$ are known or can be estimated from thermochemical data. Thus T_f can be estimated using Equation (4). Assuming reasonable values, as follows, the total heat feedback q can be estimated from either Equation (2) or (3).

$$L_s = Q_s = 0 \quad Q_f = 800 \text{ cal/g} \quad c_s = C_p = 0.33 \text{ cal/gK}$$

$$T_s = 900 \text{ K} \quad T_0 = 300 \text{ K} \quad \rho_\infty = 1.95 \text{ g/cm}^3$$

$$m = \rho_\infty r = 0.195, 1.95, 7.8 \text{ g/cm}^2\text{s}$$

The resulting flame temperature is $T_f = 2724 \text{ K}$ and the heat fluxes corresponding to the given burning rates are $q = k(dT/dx)_s = 20, 200, 800 \text{ cal/cm}^2\text{s}$. The maximum theoretical black-body radiation at T_f is $q_r = \sigma T_f^4 = 75 \text{ cal/cm}^2\text{s}$ which is in the line with more detailed radia-

tive transfer calculations [4]. Thus the influence of radiation will be strongest at low pressures and low regression rates. Examination of equations (1) through (4) reveals the sensitivity of total heat flux to changes in various parameters. In particular, exothermic sub-surface chemical reaction ($Q_s > 0$) leads to reduced total heat flux and therefore increased relative importance of radiative feedback. In connection with this it is interesting to note that Beckstead [5] has estimated that 75% of the total chemical heat release in AP monopropellant flames occurs in the condensed (liquid) phase due to reaction of ammonia with perchloric acid.

In the case of aluminized propellants the preceding calculation using the gas temperature T_f to compute the radiant flux may result in a significant underestimate of the true value. Particle temperatures of 2000-2200 K have been indicated as typical of aluminum during combustion in solid propellant oxidizing environments at 1 atm [6-8]. At higher pressures (50 atm) 4800 K has been reported for the oxide-laden flame and 3200 K for the particle itself [9]. While the question of how to determine the effective emissivity of a cloud of burning aluminum particles is still one of considerable uncertainty it would not be totally unreasonable to suspect an effective temperature of 3500 K ($q_r = 200 \text{ cal/cm}^2\text{s}$) at a pressure of 50 atm. (At a vapor pressure of 50 atm the equilibrium temperature of Al is 3800 K and, if liquid hydrocarbon droplet combustion studies are any indication, the actual particle temperature will not be far below that value.)

II.B. Steady Burn Rate

Theoretical prediction of the steady linear regression rate of

homogeneous solid propellants without radiative influence is still a subject of controversy [3]. While most of the burning rate formulae have evolved from laminar flame theory [10] some disagreement exists over the influence of the solid-phase [11, 12]. Nevertheless various formulae often have surprisingly similar functional forms. Burn rate modelling for composite propellants is considerably more complex [13] and therefore not amenable to yielding simple analytical burn rate expressions.

Several studies have attempted to incorporate radiative transfer into conventional laminar flame theory [14-16]. Rather than dwell on the differences and similarities of these various models, a simple representative form, similar to the Ibiricu-Williams [15] and Miller [3] models, will be discussed here for the purpose of clarifying radiative influence. The semi-analytical-empirical formula

$$m^2 = CP^v / (A + B/m) \quad (5)$$

where

$$A = C_s(T_s - T_0) - Q_s \quad (6)$$

(total heat feedback to solid per unit mass flow)

$$B = q_{r,loss} \quad (7)$$

(net radiative heat loss from solid)

$$C, v = \text{constants} \quad (8)$$

has been found to predict qualitatively, at least, the burning rate behavior of most homogeneous and composite propellants. For zero radiative flux ($B = 0$) Equation (5) reduces to the usual burn rate expression. For $B > 0$ there is a net loss of radiation leading to minimum deflagration pressure P_{DL} (Section IV.A.). For $B < 0$ there is a net gain of radiant energy (as from an external source or emission

from the flame). The form in which the radiation term appears (divided by m) is an indication of the fact the radiant flux is independent of the mass burning rate whereas conductive feedback is intimately related to mass flux. Conductive heat must, in effect, "swim upstream" against the convective flow to reach the surface and pyrolyze (or vaporize) reactants which in turn produce the convective flow. The greater the rate of heat conduction the greater the convective flow which causes greater resistance to heat conduction, until a steady equilibrium state is reached. Radiant energy, on the other hand, "travels for free" so to speak. Enhancement of the radiant transport rate does not produce a self-opposing or inhibiting effect, as in conduction. Extension of this idea from the gas-phase into the solid explains many of the dynamic burning characteristics of solid propellants exposed to intense radiation fields. It explains, for example, why a propellant burning stably can be extinguished by sufficiently sudden and strong burst of radiation exposure. The manner in which radiation appears in (5) (additive with initial temperature T_0) also explains the interpretation of radiation as merely augmenting T_0 and [15] discusses this point.

Several experimental studies [17-20] have been aimed at investigating the influence of external radiation ($B > 0$) on regression rates of pure AP [17-18], and composite propellants [19-20]. A general result of these studies is that for large external fluxes the burn rate increases linearly with incident flux. For pure AP the magnitude of m as well as the constant of proportionality increased as the concentration of absorbing particles in the sample increased [17]. This observation is in agreement with Equation (5). Solving Equation (5) for m yields

(9).
$$m = [- B + (B^2 + 4A CP^v)^{1/2}]/2A \quad (9)$$

The asymptotic limit of (9) for $B^2/4A CP^v \gg 1$ yields,

$$m = - B/A = (\alpha_{eff}/A)q_{incident} \quad (10)$$

which correctly predicts that m increases linearly with $q_{incident}$ for large incidence with the slope depending on propellant absorptivity.

In [19], for composite propellants, m increased linearly with $q_{incident}$ but the slope appeared to be independent of absorptivity (although m did increase significantly as α_{eff} did). This is an indication of the limited usefulness of Equation (5) for composite propellants.

In summary, the effect of incident radiation is to increase burn rate, whether the source of the radiant energy is external, or chemical energy released in the flame zone. The case of net radiant loss from the propellant will be discussed in Section IV.A.

III. UNSTEADY (OSCILLATORY) BURNING

Many aspects of oscillatory burning and combustion instability are still poorly understood, at best. About all that can be said with certainty is that in many situations the inclusion of particles in the propellant acts to suppress pressure-coupled oscillatory burning. The acoustic theory of viscous damping by the particles cannot account completely for this behavior [21]. One possible mechanism, which has not been considered before, for explaining this behavior is enhanced radiative feedback due to increased solids loading in the gas phase. The underlying principle being that radiative feedback may to an extent be de-coupled from pressure oscillations whereas conductive feedback is

always pressure-coupled. Enhancement of radiative feedback would then decrease the magnitude of the real part of the pressure-coupled response function. This decrease in acoustic gains combined with the associated increase in acoustic losses due to viscous particle damping may be a more accurate picture of the true influence that adding particles has on unsteady combustion behavior. In any event, a complete description of possible radiative transfer effects must include both gas and solid phase considerations.

III.A. Gas-Phase Considerations

One of the most significant experimental observations regarding the ability of particles to suppress unstable (oscillatory) burning is that to be effective the particles must be introduced into the motor chamber through the solid propellant [21]. That the effectiveness of the particles is associated with their presence near the burning surface was demonstrated by two experiments involving aluminized propellants and aluminum oxide particles. In the first test it was found that a motor which could be stabilized by including a certain amount of aluminum in the propellant could not be stabilized by introducing an equivalent amount of aluminum oxide particles at the head end [21]. In the second test, a double-length motor was burned with aluminum being added to only half of the propellant. With aluminum in the rearward section that section burned stably and the forward section burned unstably. With aluminum in the forward section, again that section burned stably and the rear section burned unstably, in spite of the presence of the oxide particles in the gas [21].

In connection with this another very important observation made

with window bomb studies is that those particles which are effective in suppressing instability also tend to burn (if reactive) or be incandescent (if inert) very near the propellant surface [21,22]. Various explanations have been offered for this. Workers at Rohm and Haas [23] proposed a mechanism based on alteration of the heat feedback by the energy source/sink effect of the particles. Cheung et al. [21] proposed that catalysis of the combustion reactions by the particles was responsible. Rudy and Bain [23] suggested a "pilot-light" theory with physically plausible arguments but no analytical support. Another possible explanation which was considered in this study is what may be called the "cold particle layer" effect. The idea is that in those cases where the particles failed to control instability a layer of relatively cold particles was radiatively insulating the propellant surface from the flame, thereby increasing the relative magnitude of pressure-coupled conductive feedback. Momentum, mass and energy conservation indicate that indeed a high concentration of particles with low velocity and low temperature populate the region close to surface. However an analysis of the effective emissivity of such a particle cloud indicated that for typical particle loadings the magnitude of the effect was too small. The "glo-plug" concept which was finally adopted and is presented in Section III.B.2. is essentially a solid-phase phenomenon. It extends and adds analytical support to Rudy's Pilot-Light theory.

III.B. Solid-Phase Considerations

III.B.1. Quasi-Steady Gas, Homogeneous Solid Early attempts to model instability pointed out that, at least at low frequencies, the

thermal inertia of the solid-phase was the prominent influence in determining the pressure-coupled acoustic response (or admittance) function [24-29]. A key result of this type of quasi-steady gas approach for determining propellant response was that the dependence on frequency ω and burn rate m entered the problem only through the non-dimensional frequency parameter $\Omega = \omega \rho_{\infty} k_{\infty} / m^2 C_s$. This means that for a given set of propellant properties and neglecting radiation the response function is a unique function of Ω (curve 1, Fig. 2). When radiation is included burn rate will enter as another independent parameter (as discussed in Section II.8) and will cause either an increase (curve 4) or decrease (curves 2, 3; Fig. 2) in the peak of the response function according to whether the radiative feedback is pressure-coupled or de-coupled. The maximum will, however, still occur at the same value for Ω for the same propellant properties. Cantrell et al. [28,29] obtained results similar to those of Fig. 2 considering only gaseous radiation, which is by definition pressure-coupled. The effect of in-depth absorption on response function is similar to that of burn rate. Thus, for equal amounts of absorbed energy, the case of surface absorption would correspond to, say, curve 1 (Fig. 2) and that of in-depth absorption to curve 2.

III.B.2. Preferential Particle Heating - "Glo-Plug" Effect

Mihlfeith, Baer and Ryan [19] conducted an interesting series of tests which led to a glaring discrepancy when experimental results were compared with predictions by the homogeneous solid theory of the previous section. They subjected AP composite propellants to a periodic external radiant flux and measured the radiant-flux normalized response function

using combustion recoil. The result was that a frequency shift in the response function was observed between propellants which were identical except for inclusion of a very small (1 wt.%) amount of carbon black in one sample, as indicated in Fig. 3. The amplitude shift between curves 1 and 2 (Fig. 3) can be accounted for theoretically by the difference in depth or location of absorption (if not in the amount of absorption itself). But the frequency shift of the peak cannot be accounted for. In searching for an answer, the observation noted in Section III.A. regarding incandescent particles near the burning surface suggested that the assumption of a homogeneous solid be examined. That frequency enters only through the parameter Ω and even the form of Ω itself can be traced to the unsteady term in the energy equation for the solid, $\rho_s C_s (\partial T_s / \partial t)$. Unless the assumption of a homogeneous solid is relaxed, and another unsteady term is introduced to account for preferential absorption of radiation by the particles, $\rho_p C_p (\partial T_p / \partial t)$, frequency will always enter through the same parameter Ω and the theory will predict no frequency shift. Allowing for separate particle and matrix (oxidizer and binder) energy equations introduces a new time constant, the magnitude of which will depend on how efficiently the radiant energy absorbed by the particles can be distributed by conduction to the matrix. This would account for the frequency shift in the response function at low frequencies which was observed in [19].

At high frequencies the effect of preferentially preheated particles would likely be much different. At high frequencies the thermal lag of the solid phase is no longer considered to be a controlling influence. The combustion response must be determined by processes

with much smaller time constants, such as gas-phase thermal inertia or chemical kinetics. While the mechanisms involved in dynamic combustion response at high frequencies are still very poorly understood the observation remains that particles which are most effective in controlling high frequency instability are those which are either incandescent or burn easily near the surface [21, 22]. The "glo-plug" concept is that inert particles, preferentially preheated by radiant energy may act as ignition sources near the surface to stabilize combustion. This is essentially the same "pilot-light" theory earlier proposed by Rudy [22] for reactive particles. An analysis of the "glo-plug" concept, which involves simultaneous solution of the energy and radiative transport equations in the solid, was carried out. Typical results are presented in Fig. 4. For 10 μm particles differential temperatures of 300 K can easily be reached at the propellant surface. The arguments made by Cheung [21] regarding sufficient particle concentrations near the surface to justify the chemical catalysis postulate can also be made for the "glo-plug" postulate.

The complete expression for particle temperature at the surface is somewhat complicated and dependence on the key parameters obscure. It is therefore useful to consider the expression valid in the limit $fv \ll 1$.

$$T_p(x=0) = T_\infty(x=0) + (A_1 Q_a q_r D_p / 2k_\infty) \quad (11)$$

A_1 is a constant related to the reflectivity of the surface, Q_a is the absorption efficiency of the particles, D_p the particle diameter and k_∞ the thermal conductivity of the matrix. Equation (11) would seem to indicate the bigger the particles the better for instability control. However, larger particles means having fewer also, for a given mass

loading. Therefore there must be an optimum D_p corresponding to a condition somewhere between having a few very hot particles and many relatively cool particles. The ideal would be to have as many as possible with the minimum required ignition temperature.

III.C. Non-Acoustic Oscillatory Burning

The observation has been made by several investigators [30-34] that certain (usually metallized) propellants can burn with periodic oscillations independent of acoustic coupling with a cavity. The oscillations are accompanied always by periodic bright flashes of luminosity and also by pressure oscillations if burned in a confined cavity. It has also been observed that the oscillations correlate with the appearance and disappearance of metal drops on the surface [31]. Undoubtedly the periodicity is related to the time constants of such processes as liquid metal agglomeration and ejection from the surface. The puzzling point, however, is how such random, spatially remote events could become synchronized to produce well-defined, synchronous oscillations. Certainly pressure oscillations, if burning takes place in a cavity, could be a synchronizing influence [30]. But the phenomenon can be observed at atmospheric pressure with a hand-held glass-blowing torch held to aluminized AP strands [31]. In such a case it is hard to think that pressure oscillations could be strong enough to be a synchronizing mechanism. Radiative feedback, however, could be. The sudden radiant emission caused by molten aluminum flowing into the hot flame zone would be transmitted beyond the immediate vicinity to remote surface locations. Moreover that emission would be preferentially absorbed by solid aluminum particles, enhancing agglomeration and ejection and a

feedback loop necessary for self-sustained oscillations is established. Close examination of simultaneous traces of pressure and flame luminosity oscillations also indicates that, although the appearance is generally simultaneous, radiation oscillations may precede pressure oscillations by approximately one cycle. The above sequence of events could also conceivably occur in a full-scale motor environment, to the detriment of stability, but it is probable that the magnitude of the effect would be less because of the more intense and more uniform radiation field.

IV. IGNITION AND EXTINCTION

Ignition and extinction of solid propellants has been an active and diverse area of research [35-42]. Many topics could be included but only one having particularly strong relation to particles and radiation will be discussed. That is the low pressure deflagration limit P_{DL} of certain propellants (e.g. pure AP) and the effect of particle additives on P_{DL} .

IV.A. Low Pressure Deflagration Limit

The low pressure deflagration limit of certain materials such as AP had been well-observed and successfully explained theoretically in 1960 by Spalding [41] as being due to radiative energy loss from the hot surface. Levy and Friedman [17] studied experimentally the influence of adding "catalyst" particles of copper chromite Cu-0202 to the AP. Those results are reproduced in Fig. 5. At low particle concentrations P_{DL} increases as the volume fraction f_v (or weight fraction) of particles is increased. Probably this is due to enhanced radiative loss caused by increasing the effective emissivity of the AP. At an intermediate

value of particle concentration, P_{DL} begins to decrease and Levy and Friedman have suggested [17] that this was caused by increasing importance of catalysis in the gas-phase, out-weighing the heat loss effect. Another explanation, however, which can be analytically justified (at least approximately) was developed during the course of this study and follows.

IV.B. A Simple Model

The particles which serve to increase the effective emissivity of the surface at wt. % $< 10^{-3}$ also increase the effective emissivity of the hot gas above the surface. The second effect offsets the first at higher particle loadings. A straight-forward radiative heat transfer analysis treating the gas as gray and isothermal, the surface as gray and diffuse and the surroundings as at zero degrees readily yielded the net radiative heat loss from the surface as shown in Fig. 6. As particle loading increases so does the heat loss from the surface. Eventually emission by the flame becomes significant enough to reduce the surface heat loss. Differentiating implicitly the approximate burn rate formula (5) and setting equal to zero yields

$$P_{DL}^v = (-B^2/4AC) \quad (12)$$

or

$$P_{DL} \sim q_{loss}^{2/v} \quad (13)$$

which correctly predicts the behavior of P_{DL} at larger particle loadings ($v \sim 1$). The first conclusion to be drawn from this is not a new one: ignition/extinction tests of strands must be interpreted carefully. The second one is new: particle radiation may be a controlling influence.

V. SUMMARY AND RECOMMENDATIONS

Two potentially significant conclusions have been drawn in this

study regarding the role particle radiation may play as a controlling mechanism in solid propellant combustion. The first, the "glo-plug" concept, applies at high pressures and for conditions typical of actual motor environments. Even though, under these conditions, it is questionable that radiative transport is any more than a small fraction of the total heat feedback to the propellant surface, it has been shown how preferential absorption of that radiant flux by particles near the surface may serve to suppress unstable, pressure-coupled oscillatory burning by preheated particles acting as sources of ignition for the gas-phase reactions.

The second conclusion applies at low pressures and in laboratory strand burning experiments. Under these conditions it is very possible for radiant fluxes to be a large fraction of the propellant surface heat flux. It has been demonstrated that particle radiation can be a controlling influence in solid propellant ignition and extinction. Both emission of radiation by particles in the hot flame zone and absorption by particles and matrix in the solid-phase can play determining roles. Therefore careful consideration should be given to particle-radiation interaction in attempting to interpret results of laboratory ignition/extinction experiments.

Further studies should be devoted to re-examining present ideas concerning both steady and transient solid propellant combustion modeling with proper incorporation of particle-radiative heat transfer effects. Thermal emission from the hot gas zone is significantly affected by the presence of particles. Selective absorption by particles in the solid propellant may be a very important effect which

has been largely ignored. In summary, explanations attributing certain observations to either catalytic or viscous damping effects by particles may be incorrect or only partially accurate in their descriptive models. Further study should be directed toward elucidating the controlling influence(s) under any given set of conditions, be it chemical, mechanical, or radiative in nature.

REFERENCES

1. Huggett, C., Bartley, C. E., and Mills, M. M., Solid Propellant Rockets, Princeton University Press, 1960, p. 38.
2. Ohlemiller, T. J. and Summerfield, M., "A Critical Analysis of Arc Image Ignition of Solid Propellants," Report of work performed under contract AF 49(638) 1268, Propulsion Division, AFOSR.
3. Miller, M. S., "In Search of an Idealized Model of Homogeneous Solid Propellant Combustion," Technical Report ARBRL-TR-02383, U.S. Army Armament Research and Development Command, Ballistic Research Laboratory, Aberdeen Proving Ground, Maryland, Dec., 1981.
4. Pearce, B. E., "Radiative Heat Transfer within a Solid-Propellant Rocket Motor," J. Spacecraft and Rockets, Vol. 15, No. 2, March-April, 1978, pp. 125-128.
5. Beckstead, M. W., "Combustion Mechanisms of Composite Solid Propellants," 19th JANNAF Combustion Meeting, Vol. II, Chemical Propulsion Information Agency (CPIA) publication 366, pp. 93-100.
6. Ermakov, V. A., Razdobreev, A. A., Skorik, A. I., Pozdeev, V. V. and Smolyakov, S. S., "Temperatures of Aluminum Particles at the Time of Ignition and Combustion," translated from Fizika Goreniya i Vzryva, Vol. 18, No. 2, March-April, 1982, pp. 141-143.
7. Babuk, V. A., Belov, V. P. and Shelukhim, G. G., "Combustion of Aluminum Particles in Composite Condensed Systems under Low and High Pressures," translated from Fizika Goreniya i Vzryva, Vol. 17, No. 3, May-June, 1981, pp. 26-31.
8. Grigor'ev, V. E., Zarko, V. E., and Kustenoyii, K. P., "Experimental Investigation of the Agglomeration of Aluminum Particles in Burning Condensed Systems," translated from Fizika Goreniya i Vzryva, Vol. 17, No. 3, May-June, 1981, pp. 3-1-.
9. Micheli, P. L. and Schmidt, W. G., "Behavior of Aluminum in Solid Rocket Motors," AFRPL-TR-77-29, Dec., 1977.
10. von Karman, T., "The Present Status of the Theory of Laminar Flame Propagation," Sixth International Symposium on Combustion, Reinhold, New York, 1957, pp. 1-11.
11. Culick, F. E. C., "An Elementary Calculation of the Combustion of Solid Propellants," Astronautica Acta, Vol. 14, Pergamon Press, 1969, pp. 171-181.
12. Williams, F. A., "Quasi-Steady Gas-Phase Flame Theory of Unsteady Burning of a Homogeneous Solid Propellant," AIAA J., Vol. 11, No. 9, Sept. 1973, pp. 1328-1330.

13. Cohen, N. S., "Review of Composite Propellant Burn Rate Modelling," AIAA J., Vol. 18, No. 3, March, 1980, pp. 277-293.
14. Coates, R. L. and Kwak, S., "Effect of External Radiation on the Burning Rates of Solid Propellants," J. Spacecraft and Rockets, Vol. 9, No. 10, Oct. 1972, pp. 742-745.
15. Ibiricu, M. M. and Williams, F. A., "Influence of Externally Applied Thermal Radiation on the Burning Rates of Homogeneous Solid Propellants," Comb. and Flame, Vol. 24, 1975, pp. 185-198.
16. Penner, S. S. and Oife, D. B., "The Influence of Radiant-Energy Transfer on Propellant Burning Rates and Ablative Rates Controlled by an Intense Radiation Field," Research Paper P-118, Institute for Defense Analyses, Research and Engineering Support Division, April 1964.
17. Levy, J. B. and Friedman, R., "Further Studies of Pure Ammonium Perchlorate Deflagration," Eighth Symposium (International) on Combustion, The Williams and Wilkens Co., Baltimore, 1962, pp. 663-672.
18. Hertzberg, M., "The Laser-Induced Combustion of Pure Ammonium Perchlorate and the Structure of its Composite Propellant Flames," NASA CR 66919.
19. Mihlfeith, C. M., Baer, A. D. and Ryan, N. W., "Propellant Combustion Instability as Measured by Combustion Recoil," AIAA J., Vol. 10, No. 10, Oct. 1972, pp. 1280-1285.
20. Horton, M. D. and Youngberg, L. Z., "Effect of Radiant Energy on the Burning Rate of a Composite Solid Propellant," AIAA J., Vol. 8, No. 10, Oct. 1970, pp. 1738-1741.
21. Cheung, H., Lou, R. L., and Fioravanti, S., "Investigation of Unstable Burning in Composite Propellants," Report No. 0181-01F, Aerojet General Corporation, Sacramento, California, July 1959.
22. Rudy, T. P. and Bain, L. S., "Chemical Control of Propellant Properties," Report AFRPL-TR-81-53, Aug. 1981.
23. Woodward Waesche, R. H., Ballistic Section Progress Report No. 78, Rohm and Haas Co., Jan., 1959.
24. Denison, M. R. and Baum, E., "A Simplified Model of Unstable Burning in Solid Propellants," ARS J., Vol. 31, 1961, pp. 1112-1122.
25. Culick, F. E. C., "A Review of Calculations for Unsteady Burning of a Solid Propellant," AIAA J., Vol. 6, No. 12, Dec., 1968, pp. 2241-2255.

26. Culick, F. E. C., "Calculation of the Admittance Function for a Burning Surface," Astronautica Acta, Vol. 13, Pergamon Press, 1967, pp. 221-237.
27. T'ien, J. S., "Oscillatory Burning of Solid Propellants including Gas Phase Time Lag," Comb. Sci. and Tech., Vol. 5, 1972, pp. 47-54.
28. Cantrell, R. H., Hart, R. W., and McClure, F. T., "Linear Acoustic Gains and Losses in Solid Propellant Rocket Motors," AIAA J., Vol. 2, June 1964, pp. 1100-1105.
29. Cantrell, R. H., McClure, F. T., and Hart, R. W., "Effects of Thermal Radiation on the Acoustic Response of Solid Propellants," AIAA J., Vol. 8, No. 8, March 1965, pp. 418-426.
30. Eisel, J. L., Horton, M.D., Price, E. W. and Rice, D. W., "Preferred Frequency Oscillatory Combustion of Solid Propellants," AIAA J., Vol. 2, No. 7, July 1964, pp. 1319-23.
31. Inami, Y. H. and Shanfield, H., "Nonacoustic Combustion Pulsations of Ammonium Perchlorate Containing Aluminum," AIAA J., Vol. 2, No. 7, July 1964, pp. 1314-1319.
32. Beckstead, M. W., Ryan, N. W., and Baer, A. D., "Non-acoustic Instability of Composite Propellant Combustion," AIAA J., Vol. 4, No. 9, Sept. 1966, pp. 1622-1628.
33. Shelukhin, V. F., Buldakov, V. F. and Belov, V. P., "Experimental Investigation of the Combustion Process in Heterogeneous Condensed Systems," translated from Fizika Goreniya i Vzryva, Vol. 5, No. 1, 1969, pp. 42-51.
34. Price, E. W., "Review of the Combustion Instability Characteristics of Solid Propellants," Chapt. 5, Part 1, Advances in Tactical Rocket Propulsion, Agard Conference Proceedings, Nov. 1, 1965.
35. Baer, A. D. and Ryan, N. W., "Ignition of Composite Propellants by Low Radiant Fluxes," AIAA J., Vol. 3, No. 5, May 1965, pp. 884-889.
36. Kulkarni, A. K., Kumar, M., and Kuo, K. K., "Review of Solid-Propellant Ignition Studies," AIAA J., Vol. 20, No. 2, Feb. 1982, pp. 243-244.
37. Kumar, R. K., "Gas Phase Ignition of a Composite Solid Propellant Subjected to Radiant Heating," Comb. Sci. and Tech., Vol. 30, 1983, pp. 273-288.
38. Bradley, H. H. and Williams, F. A., "Theory of Radiant and Hypergolic Ignition of Solid Propellants," Comb. Sci. and Tech., Vol. 2, 1970, pp. 41-52.

39. Baer, A. D. and Ryan, N. W., "An Approximate but Complete Model for the Ignition Response of Solid Propellants," work under grant AFOSR 40-66 and 40-67.
40. Ohlemiller, T. J., Caveny, L. H., DeLuca, L., and Summerfield, M., "Dynamic Effects on Ignitability Limits of Solid Propellants Subjected to Radiative Heating," Fourteenth Symposium (International) on Combustion, pp. 1297-1307.
41. Spalding, D. B., "The Theory of Burning of Solid and Liquid Propellants," Comb. and Flame, Vol. 4, March 1960, pp. 59-76.
42. Johnson, W. E. and Nachbar, W., "Deflagration Limits of the Steady Linear Burning of a Monopropellant with Application to Ammonium Perchlorate," Eighth Symposium (International) on Combustion, pp. 678-689.

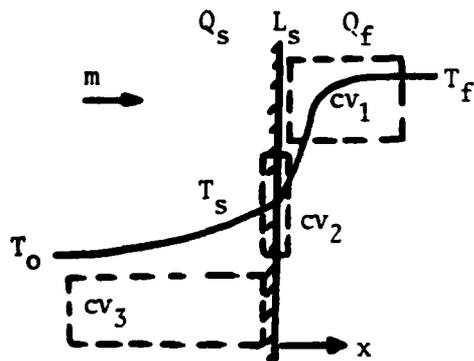


Fig. 1. Simple Energy Balance

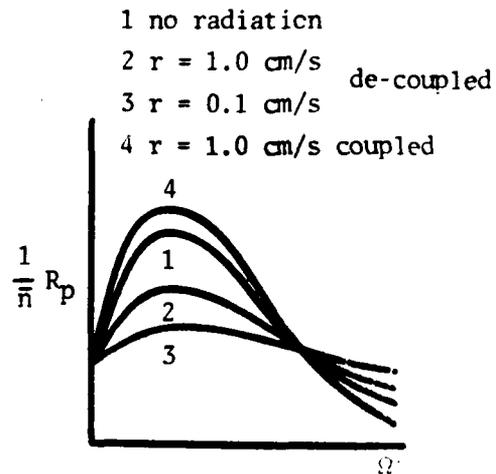


Fig. 2. Effect of Radiation on Quasi-Steady Gas-Phase Response Function

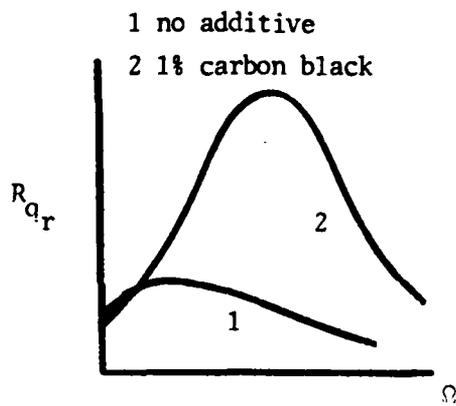


Fig. 3. Effect of Additives on Response Function [19]

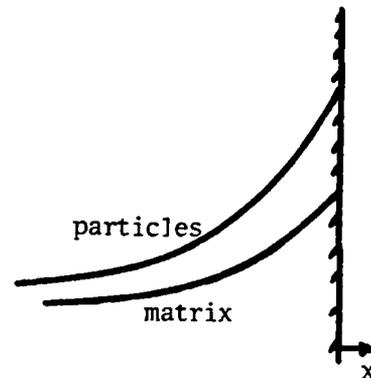


Fig. 4. Particle and Matrix Temperature Profiles in Solid Propellant

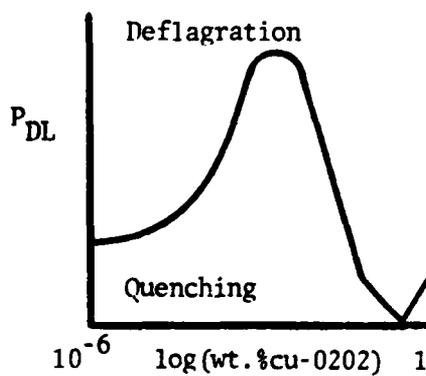


Fig. 5. Pressure Deflagration Limit

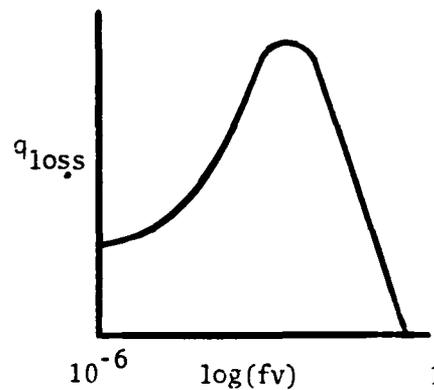


Fig. 6. Radiative Loss

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICER OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

AIRCREW-RELEVANT MAN-MONKEY ANALOGS FOR

EVALUATION OF CD AGENTS: PITCH AND ALCOHOL

Prepared by: L. W. Buckalew

Academic Rank: Associate Professor

Department and University: Department of Psychology
Alabama A & M University

Research Location: Air Force School of Aerospace Medicine,
Crew Technology Division, Aerospace Research
Branch

USAF Research Colleague: Dr. William F. Storm

Date: July 29, 1983

Contract No: F49620-82-C-0035

AIRCREW-RELEVANT MAN-MONKEY ANALOGS FOR
EVALUATION OF CD AGENTS: PITCH AND ALCOHOL

by

L. W. Buckalew

ABSTRACT

With the growing possibility of military personnel encountering chemical environments in any future conflicts, a major research thrust exists for the evaluation of CD agents to protect the viability of military operations. USAF interests are primarily concerned with flight line operations. The major problem in testing CD agents is their toxic nature which, except for a very restrictive range of doses, often precludes use of human subjects. One solution to this is the use of non-human primates and extrapolation to humans. However, it becomes necessary to develop and validate task/performance man-monkey analogs. This project was devoted to exploring the feasibility of an analog of relevance to aircrew operations. Within task and testing parameters imposed by previous USAF CD research using monkeys, a potential task analog was developed involving the Primate Equilibrium Platform (PEP), GAT-1, and a compensatory tracking task. This task is equilibrium-based and uses the pitch plane. Alcohol is suggested as the agent, as it may readily be administered to both humans and monkeys thus facilitating provision of comparative data. Recommendations reflect on major considerations for continuing efforts to develop and validate this task analog: standardization of physiological (sensory) modality, use of a roll plane in combination with pitch, and redefinition of 'equivalent' drug doses.

Acknowledgment

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, educational, professionally developmental, and productive summer at the Air Force School of Aerospace Medicine, Brooks AFB, Texas. He would like to acknowledge the Crew Technology Division, and in particular the Aerospace Medical Research Branch, for its hospitality, support, and excellent working conditions.

A particular expression of appreciation and thanks is due Dr. William F. Storm for his collaboration, guidance, and supportive encouragement in this research effort. Additionally, and of large impact on the conduct and completion of this project, Drs. Burton and Miller deserve appreciable recognition for their supportive interface and contributions in providing a highly professional working environment. Of particularly outstanding and commendable nature was the cooperative and team-orientation atmosphere which permeates this laboratory and speaks to the highest standards of professionalism and service.

I. INTRODUCTION:

Chemical agents, loosely united under the rubric of 'drugs,' have long been known to man, particularly in medical and recreational contexts. The written story of opium is believed to have begun some 6,000 years ago and alcohol has been dated to about 8000 BC. ¹ Salicylates, as internal analgesics, may be traced back some 2,400 years to the Greeks, and the hallucinogenic Datura plant containing such agents as atropine, scopolamine, and hyoscyamine was chronicled in ancient Chinese texts and used by both Greeks and Aztecs. Cannabis has a formal history dating from 2737 BC. Hence, concern with and use of drugs is rooted in ancient history.

The modern era of drug history may be tied to several scenarios of the mid 1800s: invention of the hypodermic, importation of Chinese railroad workers and their opium habits, morphine addiction during the Civil War, and the widespread dissemination of patent medicines. These conditions contributed a sensitizing effect to drug use, abuse, and needs for regulation. More recently, there appeared three major contributors to the popularity and interest in drugs: introduction of psychotherapeutic agents, attention given drug use by youth as a social phenomenon, and sensitization to drugs used during pregnancy and carcinogens. With the increasing importance attached to drugs came the careful regulation of drug development and dissemination.

It has been advanced that a drug, per se, is neither 'good' nor 'bad.' ¹ It is the pattern of a drug's use within a society which determines its value. Some drugs, respecting restrictions on use, serve medicinal needs, though these same drugs can be abused and thus present a quite different perspective. Similarly, many 'social' drugs used in moderation serve a recreational function, though too may be abused and result in socially and personally detrimental effects. While this perspective of a drug as a "two-sided coin" seems appropriate for most drugs, there are chemical agents which defy this analogy. Two such types are hallucinogens, such as LSD and PCP, and chemical warfare agents, the mass dissemination of which would be deleterious to the physiological and/or psychological integrity of an organism. The threat of chemical (drug) weapons appears very real, accentuated by recent

reports of field deployment in Central and Southeast Asia. There appears a growing possibility of military personnel encountering chemical environments in any future conflict, a realization which has stimulated concerted research interest. A major defensive/protective thrust now exists for the evaluation of potential CD agents in order to preserve the viability of military operations, with USAF particular concerns for flight line operations.

The discovery of German nerve agent (anticholinesterase) stockpiles after World War II rapidly stimulated interest in developing, testing, and making available to field forces appropriate pre-treatment or antidote CD agents. ² These agents (anticholinergics), such as atropine, were first studied in organized fashion in the 1950s, with particular concern for determination of therapeutic doses. ² Since then, considerable research has been conducted, primarily by the British, on atropine drug effects in humans. Relatively comprehensive overviews of the pharmacology and pharmacodynamics of anticholinesterases and anticholinergics are available. ^{3,4,5} For two anticholinergic CD agents, atropine and benactyzine, extensive investigations have been conducted on their physiological effects, ^{2,6,7,8,9,10} with particular emphasis on dose effects. Considerably fewer efforts have been devoted to determining performance effects, both psychomotor and cognitive, of CD agents in humans, though some research exists. ^{2,7,11} A review of atropine's physiological and performance effects in humans, with specific concern for USAF interests, was recently contributed. ¹²

Due to the toxic nature of CD agents, in vivo human testing is often precluded, thus necessitating the use of animals to estimate human performance decrements. Examples of this orientation are reflected in previous USAF efforts. ^{13,14,15} However, while advantageous if not necessary, there are several important problems associated with the use of animal subjects in CD research. One problem is the development and validation of animal models for drug effects. A second and related problem is validating animal performance measures/tasks as analogs to human behaviors, a consideration further defined by the necessity to approximate flight line, aircrew operations. Both problems must be dealt with, in sequence, to legitimize animal-man ex-

trapolations, and each primary problem has several subroutines. For example, what animal model is most appropriate for CNS drug effects and what performance tasks used with an animal model can readily be generalized to human behavior, particularly that related to aircrew operations?

Animals are frequently used in drug studies because they allow a degree of control not possible in retrospective or prospective human studies, and animal investigations allow a greater range of tests than possible with humans.¹⁶ Hence, pre-clinical drug development research and much medical and psychological research uses animals. The efficacy of such research is predicated on the validity of specific animal models, the most prevalent of which involve rodents and primates. As alcohol is one of the most researched CNS-acting drugs, much of the debate and discussion about animal models for drug effects has centered around this chemical agent.^{16,17,18,19,20,21,22} It would appear that, respecting differences in metabolism and dose equivalency, a rodent model for alcohol effects seems tenable. Considering an animal model for CD agent effects, data are accumulating which suggest the efficacy of non-human primates.^{13,14,15,23} A major contribution towards establishing the validity of a monkey model for CD drug effects was made by Mattsson and colleagues²⁴ who reviewed, integrated, and attempted to equivocate findings of 14 studies using human or monkey subjects, CD agents, and related tasks. These authors concluded the profitability of pursuing a non-human primate model for CD agent effects and encouraged development of appropriate analogs. While it seems justifiable to assume the adequacy of both rodent and monkey models for appropriate drug effects, preference for the latter is encouraged by the physiological, anatomical, and cognitive similarity of monkeys and humans.²⁵

The problem of performance function similarity between non-humans and humans is not so easily dispensed, though consideration of analogs has been encouraged.²⁴ Even the selection of tasks differentially responsive to drug effects is complicated, with important theoretical and operational considerations.^{26,27} However, based on the extensive review by Mattsson and colleagues,²⁴ there appear several potential task/performance man-monkey analogs: timing, equilibrium maintenance,

and reaction time. Of these tasks, equilibrium maintenance is the most researched for non-human primates, particularly within the context of CD agent challenge. ^{13,14,15,28} However, the empirical validation of man-monkey analog tasks in equilibrium maintenance necessitates administration of atropine or benactyzine to humans over a range of doses, a situation generally precluded by the toxicity of these agents. Hence, to establish and validate a task/performance man-monkey analog for evaluation of CD agents, it is necessary to use a less noxious drug, though one known to offer an appropriate dose-response effect for performance decrement. Alcohol (ethanol) appears to satisfy these requirements.

II. OBJECTIVES:

A need exists for a performance measuring task that is valid with both humans and animals, yet offers functional implications for flight line operations and is sensitive to CNS-acting drugs. Based on the review of previous testing of CD agents in non-human primates, ²⁴ equilibrium maintenance appears to offer such a task. The comparative pharmacology of the vestibular system has extended virtually throughout the complete vertebrate subphylum. ¹³ Of the three tasks suggested to hold promise as potential man-monkey analogs, ²⁴ equilibrium appears most relevant to USAF interests. The sense of balance, so important in aviation operations, ^{15,29,30} involves multiple sensory inputs with several neurotransmitters implicated as pharmacological substrates. ¹³ The importance of the vestibular sense to aircrew operations may be judged by considering associated neural pathways and functions. Vestibular connections involve the cerebellum and spinal cord (motor control and reflexes), medulla (arousal and vital functions such as nausea and vomiting), the temporal cortex (assumed related to dizziness), nuclei controlling neck muscles (head position), and cranial nerve nuclei (eye muscles). ³¹ Given that (a) the neurochemical and neuroanatomical systems of equilibrium are similar among primates, (b) the vestibular sense is important to aircrew operations, (c) a non-human primate model for drug effects is acceptable, and (d) alco-

hol challenge is appropriate to CD interests, exploration of a man-monkey equilibrium task analog using ethanol as a CD-related challenge appears plausible.

Previous USAF research 13,14,15 with monkeys indicated equilibrium processes as sensitive to CD agent (drug) effects. While few research efforts have been devoted to alcohol effects on equilibrium in non-human primates, previous studies 29,32,33,34,35 have found equilibrium-related tasks sensitive to alcohol effects in humans. Based on these studies of both monkeys and humans, a correlatable task/performance involving a Primate Equilibrium Platform (PEP) for monkeys and modified General Aviation Trainer (GAT) pitch control function appears viable. Both the PEP and modified GAT function offer pitch control performance data directly related to aircrew/flight line operations, and both tasks have been shown sensitive to CNS-acting drugs.

It was the objective of this study to empirically determine the responsiveness of a modified GAT operation involving pitch control to ethanol challenge, i.e. differential dose effects reflected in performance decrements. If the GAT pitch task is found adequately sensitive to ethanol in humans, and respecting the similarity of this task to monkey PEP functions and aircrew operational relevance, the final step in validating a man-monkey analog for testing CD agents would be ethanol challenge to monkeys using the PEP task. Determination of the relationship between human (GAT) and monkey (PEP) performance under ethanol challenge could support an analog allowing extrapolation from monkeys under CD agent (atropine, benactyzine) challenge to humans in an aircrew-relevant function. What follows is a suggested experimental paradigm for an initial empirical investigation. The Recommendations section of this report will be devoted to denoting appropriate comparative data and discussions of important methodological considerations for future studies.

III. METHODS AND MATERIALS:

a. Subjects

The subjects will be 10-12 male volunteers between 20 and 35 years of age who meet the body size requirements for a Class 1 flying physical (AF Regulation 161-43). Basic anthropometric characteristics

of the sample will be presented in tabular form. All subjects will be non-rated and all will have prior experience with alcohol. No subjects who, by interview, acknowledge excessive alcohol histories (more than three cocktails or the equivalent per day) will be used and volunteers will be excluded from service if taking any prescription or OTC drug. Subjects will not be paid for participation.

b. Apparatus

The pitch control task, assumed analogous to the PEP (see Figure 1) task "flown" by monkeys, ¹⁴ will involve a Link General Aviation Trainer (see Figure 2) modified to offer a single dimension (pitch) compensatory tracking task. Input forcing functions and control stick dynamics will be matched as closely as possible for the two systems (PEP and GAT), and the modified GAT-1 placed in an enclosure as with similar relative dimensions to the PEP enclosures employed in previous studies. ^{13,14,15} All GAT-1 instruments will be masked from vision and windows covered to minimize cues to pitch. An auditory alarm will be used to alert subjects when GAT attitude is beyond five degrees of horizontal, a situation functionally similar to the mild electric shock applied to monkeys when the PEP deviated more than 15 degrees. It must be noted that, while the PEP operationally allows a deviation of 30 degrees from true horizontal, the analogous GAT-1 function is restricted to approximately seven degrees deviation. Time-sampled deviations from individually determined horizontals, defined during training to obtain a stable performance, will be loaded into a DEC PDP11/34 computer. Alcohol solutions, mixed with orange juice, will involve 190 proof ethanol, and blood-alcohol levels (BALs) will be measured using a breathalyzer.

c. Procedure

Each subject will undergo four daily 5-minute training sessions in the modified GAT-1 apparatus until a stable performance is obtained. Such performance is defined as when six consecutive slopes of adjusted root mean squares (ARMS), regressed against time, are not significantly different as determined by ANOVA. The derivation of ARMS as performance metrics has been previously described. ³⁶

Under a double-blind condition, prepared ethanol solutions inclusive of a placebo treatment will be administered to individual subjects and BALs measured to determine the point at which testing is to begin. Given an appropriate BAL, a single 5-minute GAT-1 performance will be recorded for each subject. A treatment condition design, counterbalanced for order, will be used so that each subject is tested under each BAL: 0.00%, 0.03%, 0.06%, and 0.09%. The testing of a subject under each BAL will be separated by 24 hours to minimize drug carryover effects. During testing, the BAFB Clinic will be notified as to the time of testing and BAL condition anticipated, and no subject will be permitted to drive a motor vehicle with a BAL greater than .05%, as transportation will be furnished if needed.

In both the performance stabilization and drug-testing phases, subjects will be instructed to "keep the auditory alarm off" by appropriately manipulating the GAT-1 'stick' to adjust pitch. The modified GAT-1 will be driven from the horizontal by a program of input signals determined as appropriate and analogous as possible to that used in PEP testing. The same drive program will be used for all performance stabilization trials, drug trials, and over all subjects. Dynamics for this GAT control program are currently being designed under USAF contract.

IV. RESULTS:

The nature of the data anticipated should be similar to that generated by PEP studies. ^{13,14,15} ARMS (adjusted root mean squares) reflecting individual performance effects will represent the error of the GAT around its average position, measured in degrees. For each subject, this average position constitutes mean deviation from horizontal, in degrees, during stabilization trials. Hence, for each dose level of alcohol, a subject's performance will be compared with that subject's data under control (stabilization) conditions. The inclusion of a .00% BAL condition test serves as a supplementary control data point potentially reflective of placebo effects. Data analysis is expected to utilize previous reported techniques. ³⁶

V. RECOMMENDATIONS:

Prior to initiating PEP trials with alcohol challenge to complete the analog testing, it would be operationally beneficial to compare the human performance data generated with that of related studies. This comparison should reflect on the reliability of present data. The following summary table of appropriate comparative studies involving alcohol effects on vestibular functions is provided to such end.

<u>STUDY</u>	<u>FINDINGS</u>
Thyssen 37	Postural imbalance at BAL .6 mg/ml
Savolainen 38	Dose dependent effects at .4 and .8 g/kg
Gilson 33	Effect on tracking performance during whole-body movement at BAL .027%
Fregly 39	Ataxia effect and performance decrement at BAC 20 mg/100 ml
Evans 40	Impairment in stance stability, linear effect from 25 ml/70 kg
Schroeder 41	Effect on tracking during pitch plane motion at 2 ml/kg

There are a number of factors reflected in the literature which should be carefully considered in future efforts to develop this performance/task analog. These factors have, within limits imposed by operation under an approved protocol, been taken into account in the basic design suggested in this report for human testing, and they should receive appropriate attention relative to the PEP scenario.

1. Previous PEP studies 13,14,15 have utilized an equilibrium-based compensatory tracking task involving only the pitch plane. The proposed methodology for human testing duplicates this condition, though there are indications that lateral motion may be more complex than anterior-posterior sway and hence be more sensitive to alcohol effects. 42 Alcohol has been noted to produce no effect on tracking while in the pitch plane in the absence of motion, though the yaw position did reflect an effect. 41 During motion, alcohol did affect tracking in both pitch and yaw planes, though absolute error was greater for pitch plane stimulation. Given findings of human research,

it may be advantageous in future studies involving both the GAT and PEP to utilize both anterior-posterior (pitch) and lateral (roll) plane stimulation.

2. From a physiological standpoint, the vestibular system in primates incorporates two components: subjective vertigo and nystagmus.³¹ These two subsystems operate independently, as the former responds to eye movement stimulation while the latter utilizes fluid movement stimulation. Vertigo is often characterized as subjective, as the organism must learn to interpret sensory input and determine, on an individual basis, what constitutes a horizontal plane. Hence, there are two subsystems working collectively to contribute equilibrium maintenance, and each subsystem may be differentially responsive to drug challenge. The proposed methodology involving human GAT compensatory tracking attempts to distinguish between the two subsystems and control for nystagmus by minimizing visual cues as to pitch plane. Thusly, any resulting alcohol effect, if any, may be directly related to subjective vertigo. Given the flying mission and the absence of visual cues associated with nystagmus during high altitude flights, initial exploration of vertigo functions and drug effects seem most appropriate. Previous PEP studies^{13,14,15} have allowed both vestibular subsystems to operationally contribute to performance data. It is important in defining the clarity and reliability of any drug effect on an equilibrium-based performance task to respect the fact that the presence or absence of visual stimuli is likely related to any alcohol effect.⁴³ Evidence would suggest that visual information (nystagmic) is very important in counteracting alcohol effects.⁴⁴

3. Given standardized task parameters and distinction of sensory system parameters, the problem emerges as to the equivalence of drug metabolism rates among primates. It has been noted that data from monkeys and humans are very different.²⁴ Hence, proportionately equal drug doses are not necessarily equal in the sense of blood levels of an agent, time to peak effect, or drug deactivation. The human dose response curve is much steeper than in monkeys and humans are much more sensitive to performance degradation, on a mg/kg basis, than are monkeys.¹⁵ Respecting this important difference, care must be exercised in deriving any man-monkey extrapolation. It would seem appropriate to

pursue basic comparative studies as to onset time, length, and deactivation time for agents of interest based on common physiological indices mediated by CNS systems. This data could be used, on an agent-by-agent basis, to derive mathematical constants to be applied to any man-monkey extrapolation. On an interim basis, and for alcohol studies, it is suggested that instead of relying on administering proportionately similar doses based on body weight, BAL/BAC levels could be time sampled in monkeys and humans and performances keyed to specific BALs. One caution to this practice is the appreciation that a breathalyzer gives only indirect and approximate measures of BAL. 35

In final analysis, pursuit of man-monkey analogs for the evaluation of CD agents is a necessity. There is a requirement, in consideration of data extrapolations, to deal with the problems of task/physiological function standardization and systemic chemical agent measurement. Alcohol does appear to offer an opportunity, as an initial model agent, though measurement of atropine and similar CD agents will necessitate such procedures as time-based blood analysis. Also, assuming demonstrable drug effects from CD agents of interest, it is suggested that such effects be investigated relative to state dependency. Were this phenomenon substantiated, training for critical aircrew operations could be expanded to include mild exposure to specific CD agents. Given such training, it is possible that performance decrements associated with prophylactic doses of a CD agent could be reduced/minimized and thus the operational integrity of a mission be preserved.

REFERENCES

1. Ray, O., Drugs, Society, and Human Behavior (3rd ed.), St. Louis, C. V. Mosby, 1983.
2. Headley, D. B., "Effects of Atropine Sulfate and Pralidoxime Chloride on Visual, Physiological, Performance, Subjective, and Cognitive Variables in Man: A Review," Military Medicine, Vol. 147, pp. 122-132, 1982
3. Julien, R. M., A Primer of Drug Action (3rd ed.), San Francisco, W. H. Freeman, 1981.
4. Kalser, S. C., and McLain, P. L., "Atropine Metabolism in Man," Clinical Pharmacology and Therapeutics, Vol. 11, pp. 214-227, 1970.
5. Ketchum, J. S., Sidell, F. R., Crowell, E. B., Aghajanian, G. K., and Hayes, A. H., "Atropine, Scopolamine, and Ditran: Comparative Pharmacology and Antagonists in Man," Psychopharmacologia, Vol. 28, pp. 121-145, 1973.
6. Brown, B., Haegerstrom-Portnoy, G., Adams, A. J., Jones, R. T., and Jampolsky, A., "Investigation of Visual Performance After Administration of Cholinergic Blocking Agents," Report No. 801, US Army Medical Research and Development Command, Ft. Detrick, Maryland, 1980.
7. Brown, B., Baker, R., Adams, A., Haegerstrom-Portnoy, G., Jones, R., and Jampolsky, A., "Investigation of Visual Performance After Administration of Cholinergic Blocking Agents, II. Atropine," US Army Medical Research and Development Command, Ft. Detrick, Maryland, 1982.
8. Cullumbine, H., McKee, W. H. E., and Creasey, N. H., "The Effects of Atropine Sulphate upon Healthy Male Subjects," Quarterly Journal of Experimental Physiology, Vol. 40, pp. 309-319, 1955.
9. Mirakhur, R. K., "Comparative Study of the Effects of Oral and I.M. Atropine and Hyoscine in Volunteers," British Journal of Anaesthetics, Vol. 50, pp. 591-598, 1978.
10. Perry, D. J., and Mount, E., "A Comparison of the Effect of Atropine and Placebo on the Galvanic Skin Response," Journal of Investigative Dermatology, Vol. 22, pp. 497-501, 1954.

11. Moylan-Jones, R., "The Effect of a Large Dose of Atropine Upon the Performance of Routine Tasks," British Journal of Pharmacology, Vol. 37, pp. 301-305, 1969.
12. Lobb, M. L., Phillips, J. D., and Winter, A. S., "Aircrew Performance Effects of Atropine Sulfate with and without an Anticholinesterase Challenge: Review and Interpretation," USAF Contract No. F33615-83-R0611, USAF School of Aerospace Medicine, Brooks AFB, Texas, 1983.
13. Bennett, C. T., Lof, N. E., Farrer, D. N., and Mattsson, J. L., "Equilibrium Performance Changes Produced by Atropine in M. Mulatta and M. Fascicularis," SAM-TR-81-29, USAF School of Aerospace Medicine, Brooks AFB, Texas, 1981.
14. Farrer, D. N., Yochmowitz, M. G., Mattsson, J. L., Barnes, D. J., Lof, N. E., Bachman, J. A., and Bennett, C. T., "Behavioral Effects of Benactyzine on Equilibrium Maintenance and a Multiple Response Task," SAM-TR-79-19, USAF School of Aerospace Medicine, Brooks AFB, Texas, 1979.
15. Farrer, D. N., Yochmowitz, M. G., Mattsson, J. L., Lof, N. E., and Bennett, C. T., "Effects of Benactyzine on an Equilibrium and Multiple Response Task in Rhesus Monkeys," Pharmacology Biochemistry and Behavior, Vol. 16, pp. 605-609, 1982.
16. Riley, E., "Ethanol as a Behavioral Teratogen: Animal Models," in Biomedical Processes and Consequences of Alcohol Use, DHHS Publication No. (ADM) 82-1191, 1982.
17. Abel, E. L., "Effects of Ethanol and Pentobarbital in Mice of Different Ages," Physiological Psychology, Vol. 6, pp. 366-368, 1978.
18. Chernoff, G. F., "A Mouse Model of the Fetal Alcohol Syndrome," Teratology, Vol. 11, pp. 14A, 1975.
19. Kolata, G. B., "Fetal Alcohol Advisory Debated," Science, Vol. 214, pp. 642-645, 1981.
20. Myers, R. D., "Tetrahydroisoquinolines in the Brain: the basis of an Animal Model of Alcoholism," Alcoholism: Clinical and experimental Research, Vol. 2, pp. 145-152, 1978.

21. Streissguth, A. P., Landesman-Dwyer, S., Martin, J. C., and Smith, D. W., "Teratogenic Effects of Alcohol in Humans and Laboratory Animals," Science, Vol. 209, pp. 353-361, 1980.
22. Wittl, F. P., "Alcohol and Birth Defects," FDA Consumer, (May) 1978.
23. Campbell, M. E., Blick, D. W., Lanum, J., Wheeler, T. G., and Callin, G. D., "Blood Cholinesterase as a Function of Physostigmine," SAM-TR-81-19, USAF School of Aerospace Medicine, Brooks AFB, Texas, 1981.
24. Mattsson, J. L., Bennett, C. T., and Farrer, D. N., "Behavioral Effects of Atropine and Benactyzine: Man-Monkey Comparisons," SAM-TR-81-16, USAF School of Aerospace Medicine, Brooks AFB, Texas, 1981.
25. Sands, S. F., and Wright, A. A., "Monkey and Human Pictorial Memory Scanning," Science, Vol. 216, pp. 1333-1334, 1982.
26. Allen, R. W., Stein, A. C., and Jex, H. R., "Detecting Human Operator Impairment with a Psychomotor Task," Paper No. 286, Systems Technology, Inc., Hawthorne, California, June 1981.
27. Jex, H. R., McRuer, D. T., Allen, R. W., and Klein, R. H., "Dynamic Characteristics of Alcohol-Impaired Human Controllers," paper presented at the 6th Triennial IFAC World Congress, Boston, August 1975.
28. Longo, V. G., "Behavioral and Electroencephalographic Effects of Atropine and Related Compounds," Pharmacology Review, Vol. 18, pp. 965-996, 1966.
29. Chiles, W. D., and Jennings, A. E., "Effects of Alcohol on a Problem Solving Task," FAA-AM-72-11, Civil Aeromedical Institute, Oklahoma City, Oklahoma, 1972.
30. Guedry, F. E., Gilson, R. D., Schroeder, D. J., and Collins, W. E., "Some Effects of Alcohol on Various Aspects of Oculomotor Control," NAMRL-1206, Naval Aerospace Medical Research Lab, Pensacola, Florida, 1974.
31. Carlson, N. R., Physiology of Behavior (2nd ed.), Boston, Allyn and Bacon, 1980.

32. Dowd, P., "Influence of Alcoholic Liquors on the Vestibulo-Ocular Responses to Coriolis Stimulation," SAM-TR-73-253, USAF School of Aerospace Medicine, Brooks AFB, Texas, 1974.
33. Gilson, R. D., Schroeder, D. J., Collins, W. E., and Guedry, F. E., "Effects of Different Alcohol Dosages and Display Illumination on Tracking Performance During Vestibular Stimulation," Aerospace Medicine, Vol. 43, pp. 656-660, 1972.
34. Henry, P. H., Davis, T. Q., Engelken, E. J., McNee, R. C., Keiser, H. N., Triebwasser, J. H., and Lancaster, M. C., "Evaluation of Two Link GAT-1 Trainer Tasks by Experienced Pilots at Three Alcohol Dose Levels," SAM-TR-74-53, USAF School of Aerospace Medicine, Brooks AFB, Texas, 1974.
35. Henry, P. H., Flueck, J. A., and Lancaster, M. C., "Laboratory Assessment of Pilot Performance Using Nonrated Subjects at Three Alcohol Dose Levels," SAM-TR-74-27, USAF School of Aerospace Medicine, Brooks AFB, Texas, 1974.
36. Yochmowitz, M. G., Patrick, R. P., Jaeger, R. J., and Barnes, D. J., "New Metrics for the Primate Equilibrium Platform," Perceptual and Motor Skills, Vol. 45, pp. 227-234, 1977.
37. Thyssen, H. H., Brynskov, J., and Jansen, E. C., "Alcohol and Postural Imbalance. A Force Plate Study," Z. Rechtsmed., Vol. 87, pp. 257-260, 1981.
38. Savolainen, K., Riihimaki, V., Vaheri, E., and Linnoila, M., "Effects of Xylene and Alcohol on Vestibular and Visual Functions in Man," Scandinavian Journal of Work and Environmental Health, Vol. 6, pp. 94-103, 1980.
39. Fregly, A. R., Bergstedt, M., and Graybiel, A., "Relationships Between Blood Alcohol, Positional Alcohol Nystagmus and Postural Equilibrium," Quarterly Journal of Studies on Alcohol, Vol. 28, pp. 11-21, 1967.
40. Evans, M. A., Martz, R., Rodda, B. E., Kiplinger, G., and Forney, R. B., "Quantitative Relationship Between Blood Alcohol Concentration and Psychomotor Performance," Clinical Pharmacology and Therapeutics, Vol. 15, pp. 253-260, 1974.

41. Schroeder, D. J., Gilson, R. D., Guedry, F. E., and Collins, W. E., "Alcohol and Disorientation-Related Responses. VI. Effects of Alcohol on Eye Movements and Tracking Performance During Laboratory Angular Accelerations about the Yaw and Pitch Axes," FAA-AM-72-34, Civil Aeromedical Institute, Oklahoma City, Oklahoma, 1972.
42. Woollacott, M. H., "Effects of Ethanol on Postural Adjustment in Humans," Experimental Neurology, Vol. 80, pp. 55-68, 1983.
43. Schroeder, D. J., "Influence of Alcohol on Vestibular Responses to Angular Accelerations," Aerospace Medicine, Vol. 42, pp. 959-970, 1971.
44. Begbie, G. H., "The Effects of Alcohol and of Varying Amounts of Visual Information on a Balancing Test," Ergonomics, Vol. 9, pp. 325-333, 1966.

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 BLAST AND FRAGMENT LOADING ON
REINFORCED CONCRETE

Prepared by: DR. CHESTER E. CANADA / DOUGLAS J. YOVAISH
Academic Rank: ASSOCIATE PROFESSOR / GRADUATE STUDENT
Department and University: SCHOOL OF CIVIL ENGINEERING
OKLAHOMA STATE UNIVERSITY, STILLWATER, OK.
UNIVERSITY OF FLORIDA, GAINESVILLE, FL.
Research Location: AIR FORCE ENGINEERING AND SERVICE CENTER
ENGINEERING RESEARCH DIVISION
TYNDALL AFB, FL.
USAF Research colleague: LT. TOM HILFERTY
Date: JULY 22, 1983
Contract No: F49620-82-C-0035

COMBINED BLAST AND FRAGMENT LOADING
ON REINFORCED CONCRETE

by

Dr. Chester E. Canada and Douglas J. Yovaish

ABSTRACT

A preliminary computational procedure is developed to predict the dynamic response of reinforced concrete to the combined blastwave and fragment loading arising from the near-by detonation of a metal-encased high explosive charge. For the physical models and scaled distances used here, the predicted shockwave profiles in concrete for cased and uncased charges are significantly different. The spatial gradient at the shockfront for a cased charge is relatively steeper yielding conditions more conducive for spalling.

If spalling occurs, less concrete is available to resist local shear failure; so, the probability of perforation is increased. This rationale appears to be consistent with recent experimental programs which show concrete spalling and perforation to be significantly greater for a cased charge than for an uncased charge.

ACKNOWLEDGEMENTS

This research effort, which was conducted at Tyndall AFB, was sponsored by the Air Force Systems Command and the Air Force Office of Scientific Research through the Southeastern Center For Electrical Engineering Education.

The authors wish to thank all of their colleagues and associates at the Engineering and Services Center of Tyndall Air Force Base for their hospitality and helpfulness over the ten week period. Specifically they wish to thank their technical contact, Lt. Tom Hilferty for his readily available guidance and encouragement. Other personnel at Tyndall who provided timely information and help in a commendable professional manner included Major John Centrone and Capt. Paul Rosengren.

SYMBOLS

A	AREA (IN ²)
B	FRAGMENTATION PARAMETER (LB ^{1/2} IN ^{-7/6})
E	MODULUS OF ELASTICITY (PSI)
H	WALL THICKNESS (IN)
L	WAVELENGTH (IN)
M	MASS DENSITY (LB-S ² /IN ⁴)
N	NUMBER OF FRAGMENTS
P	MAXIMUM OVERPRESSURE (PSI)
Q	GRUNEISEN'S RATIO
R	STANDOFF DISTANCE (FT)
S	STRESS (PSI)
T	WAVE DURATION (S)
V	SPECIFIC VOLUME (IN ⁴ / (LB-S ²))
W	WEIGHT (LB)
Z	SCALED DISTANCE (FT/LB ^{1/3})
a	EXPONENTIAL DECAY PARAMETER, RADIUS ()
c	SPEED OF SOUND (IN/S)
d	DIAMETER (IN)
e	SPECIFIC ENERGY
i	IMPULSE (PSI-S)
r	PRESSURE INTERCEPT (PSI)
k	HARDNESS PARAMETER
l	LENGTH (IN)
m	MASS (LB-S ² /IN)
n	EXPONENT
p	VARIABLE OVERPRESSURE (PSI)
r	VARIABLE DISTANCE (IN; FT)
t	TIME (s); THICKNESS (IN)
u _p	PARTICLE VELOCITY (IN/S)
v	VELOCITY (IN/S)
x,y	DISTANCES ALONG WALL (IN)
z	DISTANCE INTO WALL (IN)

I. INTRODUCTION: This effort addresses the response of aboveground, blast resistant, reinforced concrete facilities caused by near-miss detonations of air-delivered bombs. The structural loading profile that results from a near-miss detonation generally has a peak pressure substantially higher and a duration substantially shorter than the corresponding material strength and natural period of most structures.

A structure located near the source will generally be subjected to both missile impacts and a highly directional, transient blast wave. In the high pressure (near-field) region, blast pressures decay rapidly with distance from the source. As a result, the blast loading on a specific area of a structure is highly dependent on the separation of that element from the source. Often that part of the structure nearest the source will be subjected to a loading magnitude capable of causing the localized material failures of spalling or breaching. Such structures must be examined for localized shearing responses in addition to the more conventional overall flexure and shearing modes of response.

H. Hader (1) recently reported an extensive investigation concerning the effects of bare and steel-cased explosive charges on reinforced concrete walls. The main conclusion was that cased charges were considerably more effective in causing perforation than bare charges. In fact, the paper reported that cased charges caused perforations in reinforced concrete walls at distances up to ten times larger than bare charges of the same weight. Further, the wall thickness required to prevent perforation is up to three times larger for cased than for bare charges. Empirical curves to predict spalling and perforation for the two situations were given in the reference. The curves are reproduced here as Fig 1. Detailed discussions of the results

of Ref 1 are contained in Ref 2-4. The increased effectiveness of a cased charge in causing localized concrete failure is supported by each of these references.

Various modifications were investigated in Ref 2-4, to address both shearing and flexural modes of behavior. For flexure, main reinforcement percentages as low as 0.25% were found effective. For spalling, all the successful fixes involved some mechanism to decouple the blast wave and fragments from an unimpeded, direct line-of-sight interaction with the wall.

The emphasis of this effort will be to develop an understanding of material and structural responses to combined fragment and blast loadings. This enhanced understanding of mechanisms involved should prove useful for choosing among the possible structural modifications.

II. OBJECTIVE: The objective of this study is to develop a rationale to predict the combined missile and blast wave induced loading on a reinforced concrete structure from the detonation of a nearby, steel-cased, high explosive charge. Structural and material responses predicted due to this combined loading will be compared with applicable experimental results. Subject areas that need refining and improving will be noted.

III. DISCUSSION: The analysis procedure adopted here is to separately determine the pressure-time loading profiles for blast and fragments on the reinforced concrete element. Each profile has the form

$$p = p_0(1-t/T)e^{-at/T}. \quad (1)$$

If $a = 0$, which is the case most often used, the profile is triangular. For the problem under consideration, simultaneous loading is assumed. The two pressure-time profiles are simply added linearly to obtain a single wave, of rather involved geometry for element loading. Response of the structural element, including spall, local shear, and flexural modes is determined.

Spalling and local shearing are across-the-plate phenomena; so, characteristic times will be small relative to flexural response. A detailed knowledge of the wave profile within the concrete is required for the spall analysis. Since the peak pressures obviously decay with distance, an attenuation analysis is necessary. The main steps in the analysis procedure are featured in the following narrative.

a. LOADING DUE TO BLAST WAVE: The determination of blast wave loading involves few assumptions and is relatively straightforward. Air burst curves for reflected over-pressure and impulse in Ref. 6. are used in conjunction with a preliminary modification (7), of a computer program "REICON" (8) to predict loading. The program "REICON" was initially written to investigate shear failure in soil-covered reinforced concrete beams and slabs. The code contains shear failure criteria based on a critical impulse and an initial, critical velocity to cause both edge and localized shear failures. Only the localized shear failure mode for a slab is considered in this effort.

A free-field pressure distribution from a spherical charge of the form

$$p = p(x,y) f(t) \quad (2a)$$

where $p(x,y) = kz^{-n}$ (2b)

and $f(t) = (1 - t/T) \exp(-at/T)$ (2c)

is used in REICON. Free-field pressures from cylindrical charges are predicted by segmenting the cylinder into individual charges applying Eq 2 to each segment, and linearly summing the results. Reflected pressures resulting from the interaction of the underground, free-field wave with a concrete wall is approximated in REICON by

$$p(x,y) = (kz^{-n}) (1 + R/r) \quad (3a)$$

where $z = r/W_{HE}^{1/3}$. (3b)

The curves of Ref 6 are converted herein to predict shocks at ground level by using an equivalent charge weight. An equivalent weight factor of 1.8 is recommended in the reference. The converted curves are both shifted to the right, and the scaled impulse curve is also shifted upward. These resulting curves were fit to the form of Eq 3. The resulting relations for reflected shockwaves at ground level are

$$p(x,y) = 5210(z^{-1.81}) (1 + R/r) \quad (4a)$$

and $i(x,y) = 0.184 W_{HE}^{1/3}(z^{-1.81}) (1 + R/r)$ (4b)

with an effective duration

$$t(x,y) = (a^2/(e^{-a} + a - 1)) [(3.53E -5) W_{HE}^{1/3}]. \quad (4c)$$

These equations agree with the reference curves for scaled distances less than about 1.5, but are increasingly in error for larger scaled distances.

b. LOADING DUE TO FRAGMENTS: For this study, the cylindrical axis of the cased charge is oriented vertically on detonation. A symmetric, cylindrically expanding fragment distribution of uniform velocity limited to the horizontal

direction is assumed. The accuracy of this assumption will obviously depend on the initial velocity distribution caused by the detonation geometry. The initial uniform velocity of the expanding fragments and the resulting decay in velocity as a function of distance are predicted in Ref 5 as

$$v_0 = (2E')^{1/2} [2W_{HE}/(2W_C + W_{HE})]^{1/2} \quad (5a)$$

and $v = v_0 \exp((-1.59 E^{-3})r/W_f^{1/3}). \quad (5b)$

A view of the assumed situation is shown in Fig 2. For purposes of this study the wall will be divided into sectors of constant height (l) width (t) (as shown in Fig. 3). The average velocity normal to the surface of the wall in sector j is approximately

$$v_{nj} = v_j [\text{SIN } \Theta_j - \text{SIN } \Theta_{j-1}] / (\Theta_j - \Theta_{j-1}) \quad (6)$$

where v_j is obtained from Eq 5. The weight of the median fragment and the number of fragments striking sector j are predicted from the Mott Theory (6) as

$$W_f = [Bt_c^{5/6} d_c^{1/3} (1 + t_c/d_c) (\ln .5)]^2 \quad (7)$$

and $N_j = N_T \cdot (\Theta_j - \Theta_{j-1}) / 2\pi \quad (8a)$

with $N_T = ((\ln 0.5)^2 / 2) W_C / W_f. \quad (8b)$

Unless otherwise specified, parameters in REICON/SPALL default to these for TNT.

The impulse imparted to the wall in sector j equals:

$$i_{fj} = m_j v_{nj} / A_j \quad (9a)$$

where $m_j = (W_C/g) (\Theta_j - \Theta_{j-1}) / 2\pi \quad (9b)$

and $A_j = T(1)$ (9c)

The average spacing between median fragments in sector j is

$$x_{fj} = (A_j/N_j)^{1/2}. \quad (10)$$

The depth of penetration of a fragment with median weight in sector j is (6)

$$z_{fj} = (5.61E-8)k'(5000/f'_c)^{1/2}W_f^{0.4}(v_{nj})^{1.8} \quad (11)$$

where $k' = 0.7$ for mild steel, and $k' = 1.0$ for armor piercing steel. The impulse specified in Eq 9 is deposited in the wall over this depth z_{fj} . If the fragment velocity is almost constant over most of its penetration into the wall, then the interaction time for deposition of the impulse is

$$t_{fj} = z_{fj}/v_{nj} \quad (12)$$

This interaction time is the assumed duration of the resulting shock wave of magnitude and form

$$p_{fj} = (i_{fj}/t_{fj})[a^2/(e^{-a} + a^{-1})]. \quad (13)$$

The shock wave emanating from a single fragment impact would exhibit a spherical expansion. The peak pressure at the shock front would decay approximately as the inverse of the distance from a reference volume of radius a (9). For the situation of several fragments striking an area, the individual shock waves will interfere to both spall concrete from the front face and eventually form a planar shock front propagating into the wall. The depth of spall for a given region on the front face is assumed equal to the average depth of fragment penetration for that region.

The model used to predict pressures propagating into the wall assumes that the energy of striking fragment is deposited in a characteristic volume equal to the equivalent spherical volume of a median fragment. This reference volume is situated at the terminal location of the fragment Z_f . A spherically expanding wave emanates from this characteristic volume. The pressure in this wave decays inversely with propagation distance until it has traveled a distance equal to the fragment spacing at which time a plane wave is formed. The pressure when a plane wave finally forms is

$$P_{fj} = (a/x_{fj}) P_{fj} \text{ (INITIAL)} \quad (14)$$

where $P_{fj} \text{ (INITIAL)}$ is the value in Eq 13.

Further attenuation of this wave, since it is now considered a plane wave, proceeds by hydrodynamic attenuation as discussed below.

c. HYDRODYNAMIC ATTENUATION: Obviously, the peak pressure at the shock front reduces in magnitude as a function of propagation distance. Quite often this decay process is ignored for design considerations and the wave applied at the back face is simply equated with the wave applied to the front face. This approximation appears adequate for plane waves, possessing a long wavelength relative to the propagation distance of interest. However, for non-planar waves or waves with short lengths, decay processes become significant. Attenuation as a function of geometrical divergence is considered for non-planar waves. Pressure decay due to hydrodynamic attenuation is considered for planar shocks of short wavelength. Obviously, as discussed below, further work is needed in this area to improve prediction accuracies.

Hydrodynamic attenuation arises because the sum of the particle velocity and local speed of sound behind a shock front exceed the velocity of the front. Consequently, disturbances behind the front overtake it, resulting in pressure reductions as a function of propagation distance. Fig 4 shows the physical situation. The decrease in pressure as a function of distance for the assumed elastic behavior is

$$Dp/Dz = -dp/dz [u_p + c]/c_0 - 1] \quad (15)$$

with the parameters defined in Fig 4. An expansion equation of state is required to evaluate the local speed of sound c , which is highly dependent on the physical expansion process along the release isentrope. Many isentropic equations-of-state have been proposed to describe characteristics peculiar to a material. In the high pressure region, concrete obviously undergoes some irreversible crushing. A completely accurate description of the attenuation process would dictate the use of an isentrope sensitive to this irreversible crushing (10). Significant attentuations would be predicted for those regions involving severe concrete crushing. For the first order approximation intended here, it will be assumed, however, that most of the concrete material remains in the elastic phase. A Mie-Gruneisen equation-of-state is assumed, namely

$$e = (p + E)V/Q \quad (16)$$

where Q is the Gruneisen ratio. Data for the Gruneisen ratio of concrete are quite limited. A constant value of 2.3, as reported in Ref 10, will be used here.

Evaluation of Eq 16 for an isentropic expansion leads to

$$P_S = E/(Q+1) \cdot [(V_0/V)^{Q+1} - 1]. \quad (17)$$

The local speed of sound behind the shock is

$$c^2 = (-V^2 dp/dV)_s \quad (18a)$$

leading to

$$c = c_0 (V_0/V)^{Q/2}. \quad (18b)$$

The resulting expression for $(u_p + c)$ is

$$(u_p + c) = c_0 [(V_0/V)^{Q/2} + (1-V/V_0)]. \quad (19)$$

Finally from Eq 15,

$$DP/DZ = (dp/dZ) [(V_0/V)^{Q/2} - V/V_0] \quad (20a)$$

$$\text{where } \rho_H = \rho_0 c_0^2 (1-V/V_0). \quad (20b)$$

Determination of the locus of peak pressures for the first transit of the structural element depends on the initial blast wave profile and the propagation distance. The locus curve is

$$p_{j+1} = p_j + (DP/DZ) \Delta Z_j. \quad (21)$$

The value used for ΔZ must be small enough to assure resolution of the pressure attenuation.

d. CONCRETE SPALLING: Several SPALLING models (11) have been proposed to predict the behavior of various classes of

material. Unfortunately, insufficient data are available to utilize most of these models. Perhaps the simplest known model is used here; where, concrete spall occurs at the instant a tensile stress greater than the tensile strength S_u exists at a plane in the concrete. The phenomena is thus treated as a brittle process. The process under consideration is illustrated in Fig 5. Fig 5a is a spatial plot showing the incident compression wave within the concrete of peak pressure P , and wavelength L at the time of first contact with the back face of the concrete element. Fig 5b portrays the situation a short time later when the interaction between the rarefaction tail of the incident compressive wave and the peak of the reflected rarefaction wave has created a tensile stress at the plane shown. If p is greater than the tensile strength, then spall occurs at the plane noted by ΔZ . Material to the right of this fracture plane moves faster than the material to the left. In the absence of any additional restraining forces, material on the two sides of the fracture plane will increasingly separate with time.

If the spall plane is located within the concrete cover of the back face, no additional restraining forces arise to prevent the relative motion of a spalled segment. If, on the other hand, the spall plane lies between negative and positive reinforcement, some restraint to the relative motion on the two sides of the plane will eventually be provided as stresses develop in the reinforcing steel. Even more restraint to the relative separation would be provided in those elements containing web reinforcement. Generally blast-resistant concrete contains, as a minimum, both negative and compressive reinforcement, and probably temperature steel as well.

For the preliminary calculations conducted here, specified regions on the back face of the wall are evaluated for spall only down to that depth corresponding to the innermost layer of negative reinforcing steel. The containment provided by the reinforcing bars is assumed sufficient to prevent relative separations at possible fracture planes formed between the negative and positive bars. Possible breaching is considered to be the most significant shear failure mode for this area between the layers of steel.

After a specified region of the wall is first checked for concrete removed due to fragment impacts on the front face and spalling on the back face, that region is then evaluated for possible breaching failure. This analysis includes reductions in wall thickness as well as any "trapped impulse" removed in a spalled layer.

If a spall plane is produced at ΔZ in Fig. 5b, then the impulse trapped in the spalled segment is

$$i = (pL/a^2c_0) [(1-a + 2a \Delta Z/L)(e^{-2a\Delta Z/L}) - (1-a)] \quad (22)$$

and the velocity of the segment is

$$v = i/M\Delta Z. \quad (23)$$

The prediction of spalling for the model used here is directly dependent on the values of concrete tensile strength and spatial slope of the pressure wave profile. The slope is sensitive to the value selected for 'a'. Quite often insufficient data is available for a judicious selection of 'a' so a value of zero is simply used. This yields a triangular wave shape.

Concrete spalling is the result of dynamic behavior. The process has no static counterpart. The value for dynamic tensile strength is certainly higher than the static strength, but questions of absolute value remain. Kot(12) reports evidence of a tensile strength equal to half of the static compressive strength, whereas, other investigators use values as low as 13% of the compressive strength. The effects of variations in the decay constant and the tensile strength are discussed in the results section.

KOT also found that spalling is dependent on the angle of incidence of a compressive wave at a free surface. He predicts that for typical concretes, spalling will not occur for angles exceeding 45 degrees. This limiting angle will be used in this analysis.

V. RESULTS: Only one test geometry is examined here with the computational scheme developed in this report. Obviously, a more complete parameter study is required to judge the credibility of the assumed models but time was simply not available. The test situation examined involves a 12-inch thick reinforced concrete wall located near a cylindrical explosive charge such as shown in Figure 2. The charge weighs 100 pounds and has a length-to-diameter ratio of three. Both bare and mild steel cased charges are considered. The steel casing is 1/4-inch thick and contoured to the shape of the charge. The case weighs about 56.1 pounds.

The concrete wall is made of 5,000 psi concrete and has both 60,000 psi flexural and shear reinforcing bars at densities of 0.5% and 0.25% respectively. Symmetric distributions are assumed for the negative and positive bars. The distance from the back face to the inside edge of the negative reinforcing steel is about three inches. This is the depth at which concrete spalling is examined.

Results featured by the computational exercise are for those standoff distances at which spalling and breaching are imminent. Parameter variations include the cased vs the bare charge and the exponential decay constant. These results are given in Table I. The standoff distances for imminent concrete spalling are functionally dependent on the value used for tensile strength. A value of 2,000 psi was used. The average loading due to the blast wave over the 24-inch diameter area was determined using an analytical solution for the average pressure predicted by Eq. 4a and the corresponding impulse predicted by Eq. 4b. Critical impulse is determined from the methods of Ref 8 with the wall thickness remaining after possible concrete spalling.

The results of Table I are consistent with recent experimental data in that greater wall damage is predicted for cased charges than for bare charges. Breaching is indicated for a cased charge at about 1.6 times the distance for a bare charge. This factor of 1.6 changes little with variation of the exponential decay constant. The distance ratio changes from about 2.2 to 6.2 as the decay constant varies from zero to five.

VI. RECOMMENDATIONS: Several assumptions were necessary in the analysis described. Some of these assumptions were reasonable and require no further consideration. Unfortunately, there were also relatively unsupported assumptions made that were critical parameters for the predicted results. These assumptions were made in the fragment induced loading, the shock wave attenuation and concrete spalling analyses. Additional analytical and experimental work is recommended for these areas. The work would include a parameter study of shockwave profiles and

attenuation processes in concrete. Correlation of observed spalling with measured shockwave geometries would be necessary. The analytical part of the study would include the development of physical models for attenuation and spalling consistent with the experimental observations.

The physical models used here permit limited reductions in effective wall thicknesses due to fragment penetration and concrete spalling. This reduction in thickness directly affects the remaining load capacity of the the wall. Additional work is necessary in this area to better estimate reserve capacity after initial spall and fragement induced damage.

REFERENCES

1. H. Hader. "Effects Of Bare and Cased Explosive Charges on Reinforced Concrete Walls," Symposium proceedings of the interaction of non-nuclear munitions with structures, U.S. Air Force Academy, Colorado (May 10-13, 1983).
2. D. R. Coltharp, "Blast Response Tests of Reinforced Concrete Box Structures," Symposium proceedings of the interaction of non-nuclear munitions with structures, U.S. Air Force Academy, Colorado (May 10-13, 1983).
3. H. Pahl and M. Kropatscheck, "Explosive Tests on Reinforced Concrete Elements Performed on Test Site 91 of the Federal Armed Forces At Meppen," *Infrastrukturstab der Bundeswehr, WWTB-80-17* (Summer 1980).
4. G. Loos and H. Pahl, "Explosive Tests on Under-Reinforced Model Structures in Incirlik (Republic of Turkey) and Meppen (Federal Republic of Germany)," *Infrastrukturstab der Bundeswehr, TB-82-01* (Jan 1982).
5. AFM 88-22, Department of the Air Force Manual "Structures to Resist the Effects of Accidental Explosions", Department of the Army, the Navy, and the Air Force (June 1969).
6. DOE/TIC-11268, "A Manual for the Prediction of Blast and Fragment Loading on Structures", U. S. Department of Energy, Albuquerque Operations Office, Amarillo Area Office, Amarillo, Texas (Nov 1980).
7. D. J. Yoviash and C. E. Canada, "REICON/SPALL," (To be published).
8. C. A. Ross et al, "Concrete Breaching Analysis," AFATL-TR-81-105 (DEC 81).
9. J. S. Rinehart, STRESS TRANSIENTS IN SOLIDS, Hyperdynamics, P. O. Box 392, Sante Fe, New Mexico, 87501 (1975).
10. H. E. Read and C. J. Malden, "The Dynamic Behavior of Concrete," 3 SR-707, Systems Science and Software, P. O. Box 1620, La Jolla, California (Aug 1971).

11. J. H. Oscarson and K. F. Graff, "SPALL Fracture and Dynamic Response of Materials, " Report #BAT-197A-4-3, (March 1968).
12. C. A. Kot, "Effects of Air Blasts on Power Plant Structures and Components, "NUREG/CR-0442, ANL-CT-78-41, Argonne National Laboratory (Oct 1978).

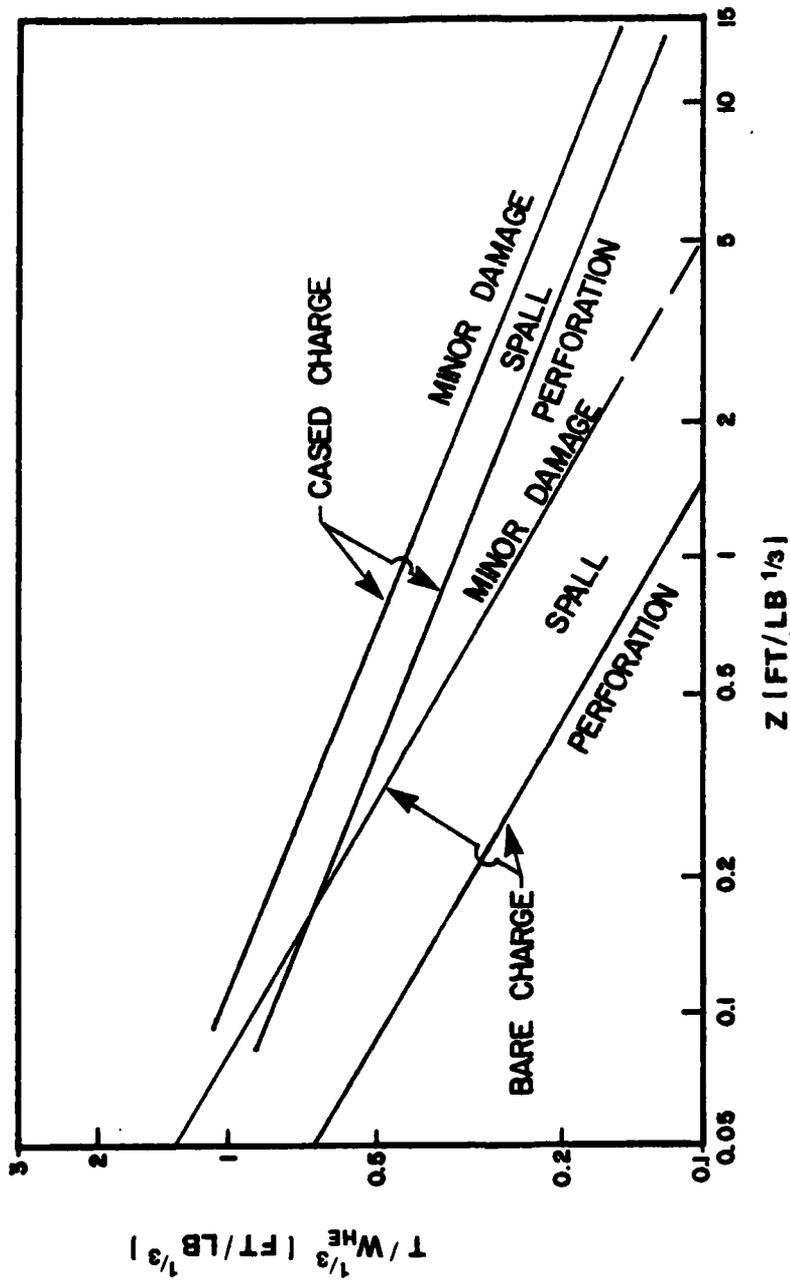
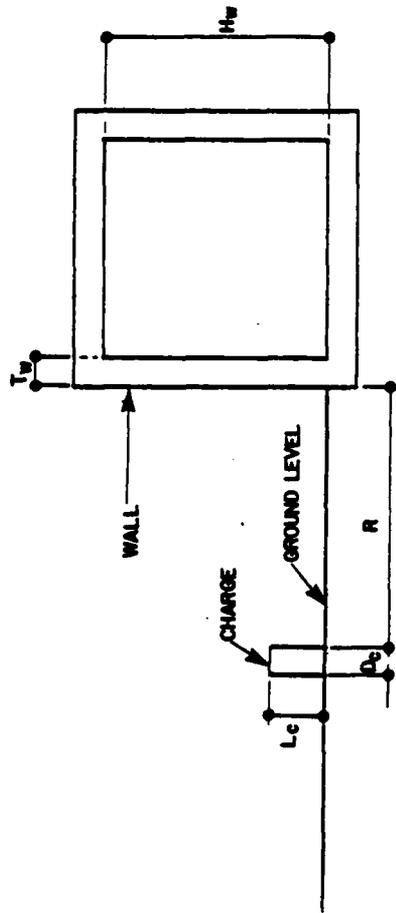
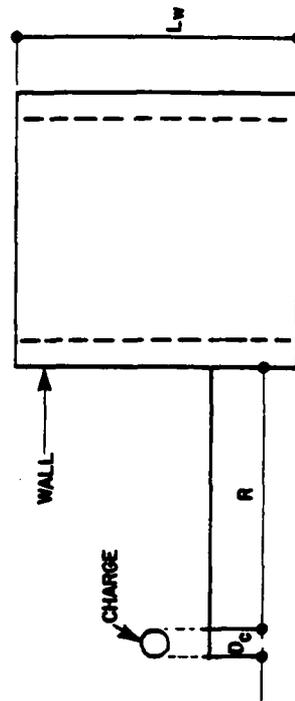


Fig 1. Comparison of Effects from Cased Charges and Bare Charges

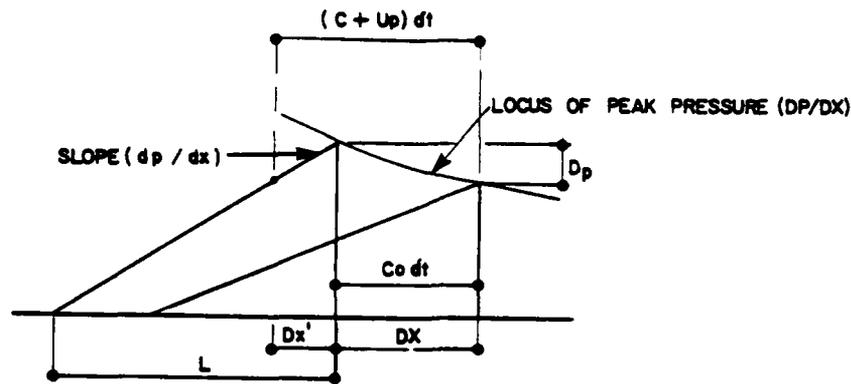


a. Elevation View



b. Plan View

Fig 2. Test Geometry



$$DX/Co = (Dx' + DX) / (Up + Co)$$

$$Dx'/DX = ((Up + C) / Co) - 1$$

$$Dp = (dp/dx) Dx'$$

$$DP/DX = (DP/Dx') (Dx'/DX)$$

$$= (dp/dx) \left(\frac{Up + C}{Co} - 1 \right)$$

Fig 4. Evaluation of Hydrodynamic Attenuation

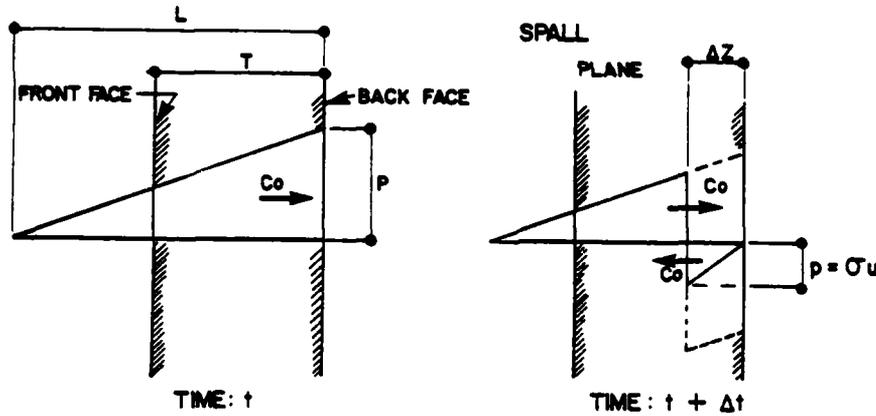


Fig 5. Evaluation of Concrete Spalling

TABLE I

CHARGE	DECAY CONST	SPALL		BREACH	
		R (ft)	i_T^* (psi-s)	R (ft)	i_T^* (psi-s)
BARE	0	2.4	5.21	2.8	3.76
	1	2.9	3.76	2.9	3.76
	5	3.4	2.86	3.0	3.76
CASED	0	5.2	2.60	4.5	2.69
	1	8.3	1.37	4.4	2.70
	5	21.0	0.36	4.4	2.70

* i_T : Impulse Applied to Front Face of wall

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

FINITE ELEMENT PRELIMINARIES IN EMP ENVIRONMENTS

Prepared by: Jack E. Chatelain
Academic Rank: Professor
Department and University: Physics
Utah State University
Research Location: Air Force Weapons Laboratory
Aircraft and Missile Division
Technology Branch
USAF Research: Dr. William E. Page
Date: 19 August 1983
Contract No: F49620-82-C-0035

FINITE ELEMENT PRELIMINARIES IN EMP ENVIRONMENTS

by

Jack E. Chatelain
Steven A. Roberts

ABSTRACT

The use of the finite element method is investigated to evaluate what functional forms would satisfy boundary conditions and asymptotic wave forms. It is shown that vector spherical harmonics used in conjunction with Bessel functions or negative powers of r could be useful. The possibility of multipole moments both in the source and wavefield as solution to the problem is indicated. Suggestions for further research are offered.

"This document was prepared under the sponsorship of the Air Force. Neither the US Government nor any person acting on behalf of the US Government assumes any liability resulting from the use of the information contained in this document or warrants that such use will be free from privately owned rights."

ACKNOWLEDGEMENT

The authors would like to thank the United States 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 spend a very worthwhile and interesting summer at the Air Force Weapons Laboratory, Kirtland Air Force Base, New Mexico. They would like to acknowledge the Laboratory, in particular the AFWL/NTAT Branch, for its hospitality and excellent working conditions.

Finally, we would like to express our appreciation for the leadership, encouragement and guidance offered by Dr. William E. Page. Our branch chief, Major James A. Kee, also left nothing to be desired in his administration of our stay.

I. INTRODUCTION: The calculation of the electromagnetic pulse (EMP) environment (electromagnetic field intensity) has been an equally difficult and important problem since the threat was first recognized. The most complete treatments of the problem have been based on finite difference analysis. The non-linearity introduced by the field dependent conductivity is the culprit. The finite difference method, while it does give answers, also has its shortcomings and perhaps a compromise between an analytical approach and finite differencing might incorporate the best of the two methods.

The finite element method seems to occupy a "niche" between the finite difference method and the analytical method. This is especially so when one tries to anticipate a "natural" functional dependence, i.e. an analytical solution to the homogenous equation. This solution should satisfy the boundary conditions, if possible. In three dimensional time dependent situations the usual procedure is to use finite elements on the space part (with optimum analytical functions as coefficients) and finite difference the time dependent equations.

We will direct our efforts towards the solution of the problem of a grounded "hemispherical boss" on an infinite grounded plane with current distributions outside of the "boss." We will indicate how Maxwell's equations (in particular the wave equations for \vec{E}) can lead to an approximate solution.

II. OBJECTIVES: Our objectives are as follows:

- a. To find optimal analytical functions to use in the finite element approximation.
- b. To investigate (if the functions in objective "a" are truly optimal) how one might proceed to a series solution.

III. STATEMENT OF THE PROBLEM: The wave equation for the field strength \vec{E} is:

$$\nabla^2 \vec{E} - \nabla(\nabla \cdot \vec{E}) = \frac{4\pi H}{c^2} \frac{\partial \vec{J}}{\partial t} - \frac{\mu \epsilon}{c^2} \frac{\partial^2 \vec{E}}{\partial t^2} \quad (1)$$

Where \vec{J} is the current density. The free field will be described by (1) with $\vec{J} = 0$, and the radiation field with both $\vec{J} = \nabla \cdot \vec{E} = 0$. We will apply equation (1) to the problem of an infinite conducting plane with a perfectly conducting hemispherical "boss" in the plane. The current distribution \vec{J} is given by tables (or other) and will be assumed to have azimuthal symmetry. With respect to the center of the conducting hemisphere. The media above the plane and sphere (the atmosphere) has an \vec{E} dependent ionization level and is also the source of a current (the Compton current). So in general we put:

$$\vec{J} = \vec{j} + \sigma(\vec{E}) \vec{E} \quad (2)$$

The problem is to find an \vec{E} which satisfies the boundary conditions for a given \vec{j} and $\sigma(\vec{E})$.

IV. BOUNDARY CONDITIONS: The boundary conditions are such that:

a. $\vec{E} \times \hat{n} = 0$ where \hat{n} is the unit normal to the plane and sphere.

b. $\vec{E} \rightarrow f_r \left(\frac{\hat{r}}{r^2} \right) + f_\theta \left(\frac{\hat{\theta}}{r} \right)$ as $r \rightarrow \infty$ (3)

where \hat{r} is the usual "position vector" of spherical polar coordinate system. We will use spherical polar coordinates with the origin at the center of the hemisphere.

A vector spherical harmonics description of a wave field emanating from spherical distributions of charge and currents has been used by Blaff and

Weisskopf(1) and others in conjunction with the description of gamma radiation coming from excited nuclei. The various multipole modes of charge motion in the nucleus are the source of the multipole radiation field described by vector spherical harmonics. We will assume then that a multipole description of source currents and charge in our problem is a viable approach.

V. VECTOR SPHERICAL HARMONICS: The vector spherical harmonics that we will use are described by E.L. Hill(2) and a few of their properties which we have used explicitly are as follow:

$$\vec{V}_{\ell 0} = \hat{\theta}_1 \frac{1}{[(\ell+1)(2\ell+1)]^{1/2}} \frac{\partial Y_\ell^0}{\partial \theta} - \hat{r}_1 \left(\frac{\ell+1}{2\ell+1}\right)^{1/2} Y_\ell^0 \quad (4)$$

$$\vec{W}_{\ell 0} = \hat{\theta}_1 \frac{1}{[\ell(2\ell+1)]^{1/2}} \frac{\partial Y_\ell^0}{\partial \theta} + \hat{r}_1 \left(\frac{\ell}{2\ell+1}\right)^{1/2} Y_\ell^0 \quad (5)$$

$$\nabla^2 [F(r) \vec{V}_{\ell 0}] = L_{\ell+1} [F(r)] \vec{V}_{\ell 0} \quad (6)$$

$$\nabla^2 [F(r) \vec{W}_{\ell 0}] = L_{\ell-1} (F) \vec{W}_{\ell 0} \quad (7)$$

$$\nabla \cdot [F(r) \vec{V}_{\ell 0}] = -\left(\frac{\ell+1}{2\ell+1}\right)^{1/2} \left[\frac{dF}{dr} + \frac{(\ell+1)}{r} F \right] Y_\ell^0 \quad (8)$$

$$\nabla \cdot [F(r) \vec{W}_{\ell 0}] = \left(\frac{\ell}{2\ell+1}\right)^{1/2} \left[\frac{dF}{dr} - \frac{(\ell-1)}{r} F \right] Y_\ell^0 \quad (9)$$

$$\nabla \times [F(r) \vec{V}_{\ell 0}] = i \left(\frac{\ell}{2\ell+1}\right)^{1/2} \left[\frac{dF}{dr} + \frac{(\ell+2)}{r} F \right] \chi_{\ell 0} \quad (10)$$

$$\nabla \times [F(r) \vec{W}_{\ell 0}] = i \left(\frac{\ell+1}{2\ell+1}\right)^{1/2} \left[\frac{dF}{dr} - \frac{(\ell-1)}{r} F \right] \chi_{\ell 0} \quad (11)$$

$$\nabla \times [F(r) \vec{\chi}_{\ell 0}] = i \left(\frac{\ell}{2\ell+1}\right)^{1/2} \left[\frac{dF}{dr} - \frac{\ell}{r} F \right] \vec{V}_{\ell 0} + i \left(\frac{\ell+1}{2\ell+1}\right)^{1/2} \left[\frac{dF}{dr} + \frac{(\ell+1)}{r} F \right] \vec{W}_{\ell 0} \quad (12)$$

$$L_\ell = \frac{d^2}{dr^2} + \frac{2}{r} \frac{d}{dr} - \frac{\ell(\ell+1)}{r^2} \quad (13)$$

Where Y_ℓ^m is the scalar harmonic; $\hat{r}_1, \hat{\theta}_1$ are unit vectors in the direction of r and θ respectively. We have assumed azimuthal symmetry and put $m=0, \vec{\varphi}=0$ in Hill's results. We will assume then that E can be represented by a linear combination of the vector spherical harmonics.

VI. SPHERICAL BESSEL FUNCTIONS: The choice of $F(r)$, depends primarily upon satisfying the boundary conditions. Presumably one could accomplish this with a variety of choices, i.e. $F(r) \sim r^{-\ell}$, $F(r) \sim$ spherical Bessel function etc, however, the choice can also affect computational difficulties. The operator (13) is strongly suggestive of differential form defining spherical bessel functions, i.e.

$$\frac{d^2 j_\ell}{dr^2} + \frac{2}{r} \frac{dj_\ell}{dr} - \frac{\ell(\ell+1)}{r^2} j_\ell = -k^2 j_\ell \quad (14)$$

The spherical bessel functions have an asymptotic limit,

$$j_\ell(kr) \rightarrow \frac{1}{kr} \sin(kr - \frac{\ell\pi}{2}) \quad (15)$$

and are also regular at the origin. One can also determine k_0 so that $j_\ell(k_0) = 0$ to satisfy the boundary condition on the sphere. We have, therefore, decided on the spherical bessel function to describe the radial dependence.

VII. THE ELECTRIC FIELD ASSUMPTION: Good approximations to the space dependent functions to be incorporated in our finite element solution would be those that satisfy the boundary conditions. We therefore put

$$\vec{E} = \sum_{\ell} \frac{A_{\ell e}(t)}{r^{2\ell+1}} [j_{\ell-1}(k_\ell r) + j_{\ell+1}(k_\ell r)] P_\ell \hat{r}_i + \sum_{\ell} A_{\ell o}(t) j_\ell(k_\ell r) P'_\ell \hat{\theta}_i \quad (16)$$

where \hat{r}_i and $\hat{\theta}_i$ are linear combinations of \vec{V}_{e0} and \vec{W}_{e0} . The combination,

$$\frac{j_{\ell-1}(k_\ell r) + j_{\ell+1}(k_\ell r)}{2\ell+1} = \frac{j_\ell(k_\ell r)}{kr} \quad (17)$$

is picked to give us a coulomb dependence in the \hat{r} term. The limit in (15) gives the radiation limit for large r . If we pick odd ℓ the radial \vec{E} vanishes at $\theta = \frac{\pi}{2}$ and the $\hat{\theta}$ field does not vanish at $\theta = \frac{\pi}{2}$. Finally if $k_\ell a$ is the zero of $j_\ell(k_\ell a)$ we will have the $\hat{\theta}$ component of \vec{E} vanish on the sphere for ℓ odd. The coulomb component involves $j_\ell(k_\ell r)$ where $k_\ell r$ and $k_\ell a$ are distinct.

If, when (16) is substituted into (1), the result would separate, for each order of ℓ , into ordinary differential equations in time (t) then the expansion would be a possible solution. Our preliminary calculations of $\nabla^2 \vec{E} +$

$\nabla(\nabla \cdot \vec{E})$ indicated that, success depends, as one might presume on the behavior of \vec{E} . The orthogonality of vector harmonics, the scalar harmonics, the Bessel functions (both of orders and arguments) tend to lend hope to the series approach. Perhaps a direct calculation of $\nabla \times (\nabla \times \vec{E})$ using equations (10), (11) and (12) would be a more direct and simpler procedure.

VIII. FINITE ELEMENT METHOD: The "finite element" method starts with the so-called "weak form" of the differential equation which is in the form of an integral. The finite element series which makes the integral vanish is, hopefully, an approximate solution. The differential equation is the Euler-Lagrange equation of the integral. The usual procedure is to multiply the differential equation by a parameter and integrate by parts to form squared terms. In the Galerkin form the parameter is chosen to be the dependent variable. A weak form for our problem is

$$\int (\nabla \times \vec{E})^2 dV - 4\epsilon \int \left(\frac{\partial \vec{E}}{\partial t} \right)^2 dV + 4 \int \vec{E} \cdot \frac{\partial \vec{J}}{\partial t} dV + 4\epsilon \int \frac{\partial \vec{E}}{\partial t} \cdot \vec{E} dV - \int [\vec{E} \times (\nabla \times \vec{E})] \cdot d\vec{s} = 0 \quad (18)$$

which is the time derivative of the Poynting Theorem. One could substitute (16) into (18) (which might prove to be formidable) with each order of ℓ being an element to approximate a solution.

IX. RECOMMENDATIONS: Our objectives in this project was to find a good approximate function to describe the space dependence for the EMP problem. A really "good" function might be used as the basis of a series solution to the problem.

The next step in the calculation would be to calculate $\nabla \times (\nabla \times \vec{E})$ with (16). This would reduce the left side of (1) to a linear combination of Bessel functions and vector harmonics. The right side \vec{E} dependence would have time derivatives of \vec{E} . We have not yet used all of the orthogonality in the

Bessel functions, so at this point we cannot state that our amplitudes $A_{r,l}$ etc. will separate in a simple way, but they might. If they do we will have an ordinary differential equation in time which will be numerically integrated.

A conductivity dependence of say $\sigma(\theta) = q + \sin\theta$ looks hopeful but higher powers of $\sin\theta$ will very probably lead to a mixing of different orders.

The finite element method has the potential of a more subtle determination of the amplitudes with, hopefully, less orders. We would substitute (16) into (18) to get the ordinary differential equations (in time) for the amplitudes.

REFERENCES

1. J.M. Blatt and V. Weisskopf, "Theoretical Nuclear Physics," New York: Wiley (1952).
2. E.H. Hill, "Theory of Vector Spherical Harmonics," AM. J. Phys. 22, 211 (1954).

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

MULTIPLE COCKPIT COMBAT MISSION TRAINER NETWORK

Prepared by: John A. Cicero

Academic Rank: Instructor

Department and University: Department of Electrical Engineering
Illinois Institute of Technology

Research Location: AIR FORCE HUMAN RESOURCES LABORATORY
OPERATIONS TRAINING DIVISION
WILLIAMS AIR FORCE BASE, ARIZONA

USAF Research: Lt Col Peter A. Cook

Date: August 10, 1983

Contract No: F49620-82-C-0035

MULTIPLE COCKPIT COMBAT MISSION TRAINER NETWORK

by

John A. Cicero

ABSTRACT

The feasibility of a multiple cockpit combat mission trainer (CMT) is investigated. It is shown that a cable network can be used to connect together many CMTs over a large geographical area. The cable channel is divided into a critical portion and a noncritical portion. The critical portion uses a time division multiple access (TDMA) technique to broadcast CMT position and attitude to all other CMTs in the network. The noncritical data portion uses a random access protocol to broadcast weather, threat information, etc., to all CMTs in the network. An analysis technique is presented which easily evaluates the random access protocol. Suggestions for further research in this area are proposed.

Acknowledgement

I would like to thank the Air Force Human Resources Laboratory, the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for providing me with the opportunity to spend an interesting and enjoyable summer at Williams Air Force Base, AZ. Also, I would like to thank Mr Warren Richeson and Ms Marilyn Vickery for coordinating my appointment at Williams.

I would especially like to thank Lt Col Peter A. Cook and Capt Caroline Hanson for providing me with an interesting assignment and excellent support facilities. I would also like to acknowledge the valuable discussions with German exchange scientist Mr Uwe List.

I. INTRODUCTION

The Air Force has determined that large-scale combat simulations are necessary to maintain force readiness for future armed conflict [1]. The increasingly complex nature of warfare demands that difficult combat missions be practiced regularly in a simulated hostile environment. It is prohibitively expensive to stage large field exercises frequently enough to provide the required experience for all players. Also, field exercises cannot provide actual threats such as surface-to-air missiles and antiaircraft artillery.

A combat mission trainer (CMT) network will link together a large number of simulators to provide the ability to practice major campaigns without loss of life or equipment, at a greatly reduced cost. The network will enable pilots to interact with all other aircraft in a strike force, and to compete against enemy formations. This concept requires a network that interconnects many CMTs over a large geographical area. Individual CMT position and attitude can be transmitted to all other CMTs in the network. Also, the latest digital terrain data from the Defense Mapping Agency, threat information from various intelligence sources, and satellite weather data can be transmitted across the channel.

A central computer can be incorporated in the network to provide firing status, and kill/damage determination to all the players. Also, the central computer can be used to synchronize the individual CMTs and to determine any network faults. Once such a network is established, encrypted data transmission can be used to regularly conduct highly classified combat scenarios without fear of compromise.

II. OBJECTIVES

The main objective of this project was to investigate the feasibility of a multiple cockpit combat mission trainer (CMT) network which could simulate air-to-air and air-to-ground combat scenarios. Specific objectives included: 1) determining which communication channel best satisfies the CMT network concept, 2) determining the maximum physical separation between CMTs in a network, 3) determining how the available channel could be used most efficiently, 4) finding and evaluating a random access protocol that provides good delay-throughput charac-

teristics for a channel with long delays, and 5) determining the effects of security on message length and channel efficiency.

III. Physical Communication Link and the CMT Network

A transmission medium which provides reasonably good security, large signal-to-noise ratios and large bandwidth must be used for the CMT network. Some possible mediums are coaxial cable, fiber optics, twisted pair, microwave and satellite links. Twisted pair cable is susceptible to interference at high frequencies and does not provide the signal-to-noise ratios necessary for high reliability, large bandwidth communications. With microwave and satellite links transmitted signals can easily be intercepted by an unauthorized receiver. Also, it will be shown that the transmission delay times for a satellite link are not acceptable in the CMT network.

Coaxial cable and fiber optics are two good mediums for the CMT network. Coaxial cable will be used as an example in this paper because it provides good bandwidth, constant received signal strength, a high signal-to-noise ratio, good isolation, and reliable components and connection techniques. Also, it is commercially installed by many CATV (Community Antenna Television) companies at competitive prices [2].

A typical CATV modem provides for point-to-point transmissions at rates of 6 Mbps using a variety of modulation techniques. CATV supports signal-to-noise ratios of approximately 40 dB which results in uncorrected bit-error rates of better than 1 in 10^9 bits.

The CMTs can be arranged in any geographic distribution. (Fig. 1 shows an example of 18 CMTs arranged in three squadrons of six with a stand alone central computer.) Therefore, the network will be designed so that all CMTs can communicate across a single link. This allows any possible geographic design to utilize the minimum number of links. (i.e., If two squadrons are transmitting data in the same geographic direction, they can share a common communication link instead of requiring two separate links.)

The maximum response time, T_d , from a pilot's action in one CMT to that action being displayed in another CMT must be determined by experiment. It is assumed to be 100 mSec for this paper. The time required for sampling the pilot's action, T_{pa} , for processing the data

at the receiving CMT, Trdp, and for producing an image at each frame of the CIG (computer image generator), Tc₁, Tc₂, and Tc₃ must be measured in the actual system. Then, the maximum data transmission time from one CMT to another, Tdt, can be calculated:

$$Tdt = Td - Tpa - Trdp - Tc_1 - Tc_2 - Tc_3 \quad (1)$$

For example, if the following times are assumed: Tpa = Trdp = Tc₁ = Tc₂ = Tc₃ = 1/60 of a second, then Tdt is calculated as 1/60 of a second. Tdt will be considered one epoch. (See Fig. 2.)

IV. Possibility of a Satellite Channel

A satellite network would be desirable for long distance communications or for communications over rough terrain. The propagation delay, Tpd, must be calculated to see if the total transmission delay is acceptable for the CMT network.

A geosynchronous satellite must have a period equal to that of the Earth's rotations, a sidereal day of 23h 56min 4.09sec (mean value). A geostationary satellite is geosynchronous with an equatorial (in the plane of the equator), near-circular and direct (same direction of rotation as Earth) orbit. The period of an elliptical orbit is given by [3]

$$\text{period} = 2\pi(A^3/u)^{1/2} \quad (2)$$

where A = semimajor axis of the ellipse and u = gravitational constant. For earth u = 3.99 x 10⁵ km³/S². Solving for A in equation (2),

$$A = ((\text{period}/2\pi)^2 u)^{1/3} \quad (3)$$

Since the period is approximately 86,164 sec, A = 42.2x10³ km. Note that A (the semimajor axis of the ellipse) is the distance from the center of the earth to the satellite. Assuming that the radius of the earth, r, is approximately 6.4 x10³ km, the altitude of the satellite, h, (which is the distance between the surface of the earth and the satellite) can be computed as

$$h = A - r \quad (4)$$
$$h = 42.2 \times 10^3 \text{ km} - 6.4 \times 10^3 \text{ km} = 35,800 \text{ km}$$

Finally, the single hop satellite delay (i.e., the time required to transmit a packet of information from one earth station to another via a satellite link) can be calculated as

$$\text{Tp}_d \text{ (Single hop delay)} = 2h/c \quad (5)$$

Where c is the propagation velocity in free space (3×10^8 m/s). Substituting c and h in equation (5) yields $\text{Tp}_d = 240$ milliseconds. This 240 millisecond propagation delay would replace the propagation delay, Tp_d , in Fig. 2, and would create a total delay from pilot action to action displayed of approximately 1/3 of a second. A delay of this magnitude is totally unacceptable in the CMT network.

V. Need for Synchronization

Synchronization will be extremely important for two reasons. First, the proposed communication link for this network is a single (expensive) channel. To allow many CMTs to efficiently access a single channel, the channel must be divided up into different time slots. Each time slot occurs during a predetermined interval and each CMT transmits data only during its uniquely assigned slot. This partition scheme which is called time division multiple access (TDMA) guarantees the minimum access delay for all the CMTs in the network.

During a CMT's uniquely defined time slot, the CMT broadcasts its critical data (CMT position, attitude, etc.) to every other CMT in the network. If two or more CMTs attempt to transmit during the same time slot a collision (mutually destructive interference) will occur and the packets will have to be retransmitted. Collisions can be avoided if the network is well synchronized and the CMTs only transmit packets during their assigned time slot.

The second reason for synchronization is that a CIG requires one epoch to process its data, and new data must wait until the current frame is completed before it can be processed. The processing latency can introduce an additional epoch of delay in the network response

time. In a well synchronized network, data will always be presented to the CIG just before the start of its current processing frame, thus eliminating any additional delays. Note that the pilot's stick position does not have to be sampled more often than once every T_{c_1} because the first frame of the CIG accepts new information only once every T_{c_1} .

VI. Calculation of the Maximum Physical Separation between CMTs

Before the maximum physical separation between CMTs can be computed, several quantities must be determined. First a packet, length L , which is used to transmit critical data must be determined. (This packet length is independent of the packet length that will be used for noncritical data.) Next, the bit or data rate, R , must be determined. For this example, the bit rate is 6 Mbps and the packet length is 1000 bits. Then, the time required to transmit one packet, T_{pack} , is defined as

$$T_{pack} = L/R \quad (6)$$

For this example, $T_{pack} = .167\text{msec}/\text{packet}$. The time required for N CMTs to transmit the packet each is $N \times T_{pack}$ (or 3msec for $N=18$).

If the total length of the data transmission slot, T_{dt} , is known then the transmission propagation delay, T_{pd} , can be calculated. The data transmission slot is equal to the time required to transmit N packets plus the maximum propagation delay from the transmitter to the receiver. Therefore,

$$T_{pd} = T_{dt} - N \times T_{pack} \quad (7)$$

For this example $T_{pd} = 13.66\text{msec}$, and the maximum physical separation between CMTs is

$$D_{max} = (T_{pd} - D_{max}/D_{repeat} \times T_{repeat}) \times c_M \quad (8)$$

Where D_{repeat} is the average distance between repeaters, T_{repeat} is the average propagation delay of a repeater, and c_M is the propagation velocity of the physical medium.

Solving for D_{max} in equation (8) yields

$$D_{max} = T_{pd} / (1/c_M + T_{repeat}/D_{repeat}) \quad (9)$$

If a repeater spacing of 1 km and an average repeater delay of 1 usec are assumed, and a propagation velocity, c_M , of $1/2 c$ (where c is the velocity in free space) is used, then the maximum separation between any two CMTs, $D_{max} = 1781\text{km}$ or 1107 miles (for the 18 CMT network).

VII. Synchronization Techniques

A synchronization technique [4] is described that is particularly applicable to the CMT network. Synchronization consists of a coarse and a fine search. The coarse search is applied only during the initialization of the network while the fine search (or fine tune) is continuously applied during normal data communications. (Note that during coarse search no other data communication is permitted.)

Each one of the N CMTs is assigned a unique number from 1 to N . (This same assignment is used for the unique TDMA time slots needed for the critical data communications). The central computer generates windows during the synchronization period, where a window (or a mirror), L_w , is defined as a period of time (measured in bits) in which everything that is received by the central computer is returned to the CMTs. The length of the window is always smaller than the length of the packet, (i.e., $L_w < L$). The bits at beginning and end of the packet, L , which do not pass through the window, L_w , (and are not returned to the CMTs) are defined as gaps, L_g . The relationship between the window, the packet length and the gaps is defined as

$$L_w = L - 2L_g \quad (10)$$

If the packet is defined as 1000 bits, and the window is defined as 950 bits, then two gaps (one at the beginning and one at the end of the packet) must be 25 bits each. During coarse search only one window is provided per epoch, but during normal operation N windows are provided per epoch. The N windows provide guard bands between packets and are used for the fine search.

Coarse Synchronization begins when the central computer notifies the network that it is going to provide a window corresponding to time slot number one. During the coarse synchronization period the central computer provides one window per epoch. This window is provided for X epochs, where X is the number of epochs needed to guarantee coarse synchronization. The first CMT will transmit a burst of Lw bits, then be idle for 2Lg bits. This pattern of Lw bits transmitting and 2Lg bits idle is repeated continuously for one epoch. Because the central computer provides only one Lw bit window per epoch, only Lw bits maximum can be returned to the transmitting CMT. Actually, when all Lw bits are returned to the transmitting CMT, the first phase of coarse synchronization for CMT number one is complete. Normally, what is returned during the first Lw bit window is a gap (of up to 2Lg bits maximum) somewhere in the window.

In the next few epochs, Y, the CMT must advance or retard its Lw bit bursts so that the gap is completely eliminated. At this point (which is the end of the first phase of coarse synchronization), all but one of the Lw bit bursts must be eliminated so that the unique time slot reserved by the central computer for CMT number one can be identified by CMT number one. The unwanted bursts can be eliminated by trial and error without knowing the actual distance between the CMT and the central computer. A binary elimination scheme provides an effective elimination technique. Note that an epoch can be divided into $Tdt \times R/L$ L bit slots. Therefore, the number of epochs required to eliminate the unwanted bursts, W, using binary elimination is

$$W = \log_2(Tdt \times R/L) , \text{ and} \quad (11)$$

$$X = Y+W \quad (12)$$

In other words, a CMT should be able to coarsely locate its unique time slot in X epochs. For this example, Y is assumed to be 3 epochs and X is calculated as 10 epochs. Simulations showed that coarse synchronization is always completed in 10 epochs or less for the network in this example.

After X epochs, the central computer must notify the second CMT to

prepare for coarse synchronization and it must move its window to the second time slot. This entire procedure is repeated for each of the N CMTs. The time required to coarsely synchronize the entire network is $(N) \times (X) \times (Tdt)$. For this example, the network can be coarsely synchronized in 3 seconds.

Fine Synchronization begins after coarse synchronization is completed. During normal operation, the central computer provides N contiguous windows per epoch. Each window is L_w bits long with $2L_g$ bits between windows. Fine tuning is performed to compensate for any drift that may occur between the clock of an individual CMT and the clock of a central computer. The fine search can be accomplished by designating the first and last L_f bits of the L_w bit packet as a pseudo noise (PN) sequence. This PN sequence will have a unique combination of 1's and 0's. When a CMT transmits a packet to the central computer, the packet is returned by the central computer to the transmitting CMT (in addition to being broadcast to all other CMTs). The transmitting CMT can compare the received PN sequence to the transmitted PN sequence, and thus determine the accuracy of the synchronization. If the sequences are not exactly the same (due to truncation by the window) the transmitting CMT can advance or retard the start of its transmission during the next epoch to compensate for any error.

The PN sequence is used only for fine tuning and does not contain any useful information. If some of the PN sequence does not pass through the window, the data portion of the packet will not be adversely affected. The PN sequence should be constructed so that it can easily be distinguished from the data portion of the packet. Also, it should immediately indicate the amount of correction needed to maintain synchronization. If it is assumed that a gap is interpreted as logic 0's the following might be a useful PN sequence:

```
GAP 1001000100001000001000000 DATA - - - - -  
- - - - - DATA 0000001000001000010001001 GAP
```

With this PN sequence there is a block of six logic 0's separating the data portion of the packet from the closest logic 1 in the PN sequence. This block of logic 0's is helpful in locating the start of the data

portion of the packet. Also, the PN sequence gives a quick indication of the amount of information that passed through the window. For example, if 1000010 - - - is the first piece of the PN sequence that passed through the window, the receiver can immediately determine that it must retard its timing by seven bits. Fine tuning continuously resynchronizes the network without disrupting the critical data communications link.

VIII. Noncritical Data Communications

The data communications considered in the previous sections are critical communications in the sense that every pilot's action must be displayed at every other CMT with the minimum delay. If there is excessive delay between CMTs, the network could not be used for air-to-air or air-to-ground scenarios. However, other network communications are possible which could tolerate some delay. For example, weather conditions could tolerate delays up to several hundred milliseconds without any noticeable degradation in network performance.

It is seen from equation (7) that there are T_{pd} seconds out of the T_{dt} second data transmission slot in which no critical communications can be performed. During this T_{pd} second interval (which is actually the transmission delay of the critical data communications) noncritical data communications can be performed. The noncritical data is not sent to a CIG and is not concerned with the CIG processing time.

Random access schemes [5] provide acceptable performance if some delays in noncritical message transmission can be tolerated. A random access or demand access protocol will be introduced which utilizes the channel efficiently for bursty (i.e., irregular) traffic in a network with large transmission delays. The noncritical data communications together with the critical data communications should utilize one hundred percent of the data transmission slot in an efficient network.

This random access scheme requires a CMT to make a reservation request before it transmits its noncritical data packet. The reservation request slot, L_r , should be as small as possible to provide the most efficient noncritical communications. L_r represents the minimum number of bits necessary to accurately discern one reservation

request from another. Given a physical communication link with a raw bit rate, R, a bandwidth, BW, and a carrier-to-noise ratio of C/N, Lr can be calculated according to [6] and [7] as

$$L_r = (2B_r + D_c + J_l + Q_t) \times R \quad (13)$$

Where $B_r = 1/R$, the basic resolution or bit time of the network, D_c is the synchronizing clock drift, approximately 6 nsec, J_l is the logic jitter in the network, approximately 5nsec, $Q_t = 2/(BW \times (C/N)^{1/2})$.

For example, if R is 6Mbps, BW is 6MHz, and C/N is 40db, then $Q_t = 3.3$ nsec and $L_r = 2.1$ bits. A reservation slot of 5 bits will be used in this example to provide an added margin of safety. Fig. 2 provides an illustration of a noncritical transmission slot. The number of bits contained in this Tpd interval is

$$L_{pd} = T_{pd} \times R \quad (14)$$

The L_{pd} bit slot is divided into N L_r bit intervals and one L_d bit interval where

$$L_d = L_{pd} - N \times L_r \quad (15)$$

L_d is the noncritical data packet length. For this example, calculations yield $L_{pd} = 82,000$ bits, $N \times L_r = 90$ bits and $L_d = 81,910$ bits.

In this random access scheme N CMTs compete for a single L_d bit slot every epoch. Since each CMT will not have noncritical data to transmit every epoch, this random access scheme is employed. The TDMA scheme used for critical data results in excessive average delays when the input traffic to the network is bursty.

Each CMT is initially assigned a number from 1 to N. During the first epoch any CMT that has information to transmit makes a reservation request during its assigned reservation time slot. For example, if CMTs five and seven have information to transmit during the first epoch, they will each transmit a L_r bit reservation request during the fifth and

seventh reservation time slots, respectively. Note that the reservation requests consist only of an Lr bit carrier burst (i.e., no modulated information is sent during the reservation slot). At the beginning of the second epoch, all CMTs examine the requests made during the first epoch. The reservation request closest to slot one wins the right to transmit its data packet during the second epoch. All the losers from the first reservation epoch together with all the new messages generated during the first epoch, make new reservation requests during the second epoch's reservation slot.

In a priority reservation scheme, the reservation requests in subsequent epochs follow the same priority assignments as the first epoch. The CMT with the lowest fixed assignment will always win the right to transmit its data first.

In a random reservation scheme, the priority of every CMT in the second epoch will vary from the priority in the first epoch according to the following algorithm: $\text{new assignment} = (\text{old assignment})_{\text{MOD } N} + 1$. Similarly, the priority assignment of subsequent epochs is modified according to this algorithm. It should be noted that if a packet collision is detected (possibly due to an error in synchronization) the CMTs will revert back to their original assignments of a number from 1 to N.

With either scheme the reservation process continues in every epoch. Requests are always made in the previous epochs for the right to transmit data during the present epochs. CMTs continuously make reservation requests until they win the right to transmit their packets.

Assume that every CMT has packet generations that follow a Bernoulli distribution, and that packets cannot be generated by a busy CMT (i.e., CMTs contain a single packet buffer). A CMT generates a new packet with probability, p, and remains idle with probability, 1-p. Also, assume that a packet can be generated by the CMT only if the CMT's packet buffer is empty at the beginning of the epoch. This transmission scheme can be modeled as a Markov process. Fig. 3 shows an example of the possible states in which a four CMT network can be found during one embedded Markovian interval. The designation, u_i indicates that there are i CMTs at the beginning of an embedded Markovian interval waiting to transmit a packet to the network during the interval. Since time is

slotted (i.e., the network is synchronized) the system can be described by the following Markov chain

$$\vec{u}(t+1) = A \vec{u}(t) \quad (16)$$

where t represents discrete time, $\vec{u}(t) = [u_0(t), u_1(t), \dots, u_N(t)]^T$ is the state probability vector, the sum of all u_i equals one, and A is the state transition matrix.

The structure of Fig. 3, shows that u_1 cannot be decreased by more than one packet per epoch (since there is a maximum of one packet transmission per epoch) and that the network cannot move from state u_i to state u_4 ($i > 0$) because busy CMTs cannot generate new packets. Noting that transmissions and arrivals are a function of the state of the network at the beginning of the embedded Markovian interval, a general state transition matrix, A , can be obtained

$$A^T = [a_{ij}] \quad (17)$$

where a_{ij} represents the probability of moving from state i to state j during one embedded Markovian interval. A general formula for a_{ij} for an N CMT network is

$$a_{ij} = \begin{cases} 0, & \text{for } k > N-1 \text{ or } i > j+1, \\ \binom{N-i}{k} p^k (1-p)^{N-1-k}, & \text{otherwise,} \\ \text{where, } k = j \text{ for } i \leq 1, \text{ and } k = j-i+1 \text{ for } i > 1. \end{cases} \quad (18)$$

The steady state solution of equation (16), denoted as u^* , can be obtained by substituting (18) into (17) and (17) into (16) then solving (16) iteratively until $\vec{u}(t+1) = \vec{u}(t)$. The average steady state throughput, S , is defined as the average number of packets transmitted during one Markovian interval. S is computed as

$$S = \sum_{i=1}^N u_i^* \quad (19)$$

The average number of busy CMTs in the network, G , is

$$G = \sum_{i=1}^N i u_i^* \quad (20)$$

And the average delay time for packets to reach their destination is

$$\text{Delay} = (G/S+1)Tdt \text{ sec} \quad (21)$$

Solving equation (16) becomes increasingly difficult for large CMT networks. For a 100 CMT network, a 100 by 100 array is required. To eliminate this problem, an averaging technique was investigated. It was shown in [8] that a stochastic system can be represented by a deterministic averaged equation such as

$$y(t+1) = y(t) + (Parr(t) - Ptr(t))/N \quad (22)$$

where $(N)(y(t))$ represents the average number of busy CMTs in the network at time t , N is the total number of CMTs in the network, $Parr(t)$ is the average number of new packets generated during one epoch, and $Ptr(t)$ is the average steady state throughput. It can be shown mathematically that $Parr(t)$ is equal to $(1-y(t))(N)(p)$, where p is the Bernoulli probability of arrival used in (18). It was determined experimentally that $Ptr(t)$ is equal to $(1-\exp(-y(t)))$. Then equation (22) can be rewritten as

$$y(t+1) = y(t) + ((1-y(t))(N)(p) - (1-\exp(-y(t))))/N \quad (23)$$

From the definition of G and S it can be seen that $G = (N)(y)$ and $S = Ptr$. Simulations were run to verify the predicted results and average errors of less than four percent were obtained.

IX. Data Security in the CMT Network

The effect of a security code [9] on packet length and transmission delay will be presented. A data encryption algorithm can utilize pseudo-random number generators, which generate uniformly distributed random numbers over a period of time that is considerably shorter than the period of the random number generator. The random number generator is called "pseudo-random" because the next random number is generated by

applying an algorithm to the present random number. In other words, the sequence of numbers generated by the pseudo-random number generator isn't really random, it just appears random over a period of time that is shorter than the period of the generator. When the transmitting CMT and the receiving CMT use the same algorithm in their pseudo-random generators, the encrypted message can easily be decoded. (The encrypted message generated by the transmitting CMT, does not appear to be random at the receiving CMT since both CMTs follow the same deterministic algorithm.) However, if an unauthorized receiver does not know the encryption algorithm the received information will appear to be random.

A pseudo-random number generator requires an initial "seed" (or number) from which it can generate a random sequence of numbers. The transmitting CMT and receiving CMT must use the same initial seed (and the same algorithm) to properly exchange information. A problem with this scheme is that the seed (or key) must be secretly transmitted to all CMTs in the network before communication can begin. The advantage of this scheme is that with sophisticated encryption algorithms an unauthorized receiver could know the algorithm but still not be able to decode the received sequence because the seed is not known.

This random number is added (modulo two) to the original data at the transmitting CMT and subtracted (modulo two) from the received message at the receiving CMT. This encryption scheme does not increase the length of the original data, (see Fig. 4) and the amount of delay added to the network is negligible.

The feasibility of a multiple CMT network over a large geographical area hinges on the question "how secure is a secure network?" Certainly the sophisticated encryption codes available today could not be decoded in real time without the afore knowledge of the encryption algorithm and the initial seed. However, given enough time (i.e., possibly days, weeks, months, etc.), most (if not all) encryption schemes could be cracked. This means that the data transmitted during CMT combat scenarios could be received and saved by the enemy and eventually decoded (at which time the entire scenario could be re-enacted).

This problem of unauthorized receivers essentially eliminates microwave, satellite and telephone links because there is easy access to these channels. The cable network could be secured by enclosing it in

some type of protected conduit which must be capable of detecting a break in the line.

X. NETWORK COST

The approximate network cost can be computed for a cable network (excluding CMTs and computers). Assume that the network has a single link connecting two nodes (each node may contain several CMTs). If a maximum distance of 1100 miles (calculated earlier) is used and connection costs of \$200 per repeater and \$5 per foot of installed unsecured cable are assumed then the cost of the network is: 1100 repeaters (approximately one per mile) @ \$200 each plus 1100 miles of cable @ \$5 per foot yields a total cost of \$29,260,000.

Certainly, existing networks could be used for a fraction of the cost, but it would be difficult to guarantee secure communications. It should be noted that a cable of this type typically has 300MHz of useable bandwidth. The 6 MHz bandwidth used by the CMT network occupies only 1/50 of the total available bandwidth. Therefore, many channels of audio and video could be added to this cable to make it more cost effective. For example, the images seen by all the pilots could be sent to a central location to be evaluated by the instructors.

XI. RECOMMENDATIONS

This paper presented a two-fold result. First, the feasibility of a CMT network was described. Second, a random access protocol for large delay networks was presented and evaluated. The protocol was evaluated using a traditional Markov chain and a simple approximation technique.

Future work could include the actual implementation of this network. Several CMTs could be connected at a single site with delays introduced to emulate an actual network. Response times could be measured and the random access protocols could be tested.

From a theoretical point of view, the evaluation technique presented in this paper should be extended to a CMT network with multiple packet buffers. The approximation technique would produce M equations instead of one equation, where M is the number of packets per CMT. Stability, delay, and throughput could then be evaluated for any possible CMT network with any size packet buffers.

BIBLIOGRAPHY

- [1] P.A. Cook, "Aerial Combat Simulation in the U.S. Air Force," Astronautics and Aeronautics, Sept., 1982, pp. 60-65.
- [2] G.T. Hopkins, "Local Computer Networking on CATV Coaxial Cable," Data Communications and Computer Networks, North-Holland Publishing Co., 1981, pp. 275-279.
- [3] V.K. Bhargava et al., Digital Communications by Satellite, John Wiley and Sons, Inc., N.Y., 1981.
- [4] C. Carter, "Survey of Synchronization Techniques for a TDMA Satellite-Switched System," IEEE TRANS. COMM., vol. COM-28, No 8, August 1980, pp. 1291-1301.
- [5] F.A. TOBAGI, "Multiaccess Protocols in Packet Communication Systems," IEEE TRANS. COMM., vol. COM-28, No 4, April 1980, pp. 468-488.
- [6] W.L. Pritchard, "Satellite Communications - An Overview of the Problems and Programs," Satellite Communications, Edited by H.L. Van Trees, IEEE Press, 1979, pp. 2-15.
- [7] R.A. Rapuano, N. Shimasaki, "Synchronization of Earth Stations to Satellite-Switched Sequences," Communications Satellite Technology, Progress in Astronautics and Aeronautics, vol. 33, 1974, pp. 411-429.
- [8] S.M. Meerkov, "Decentralized Control by Rational Controllers," Journal of Optimization Theory and Applications, vol. 32, Dec. 1980, pp. 499-520.
- [9] D.W. Davis, "Data Security in Computer Networks," Data Communications and Computer Networks, edited by S. Ramami, North-Holland Pub., N.Y., 1981, pp. 45-56.

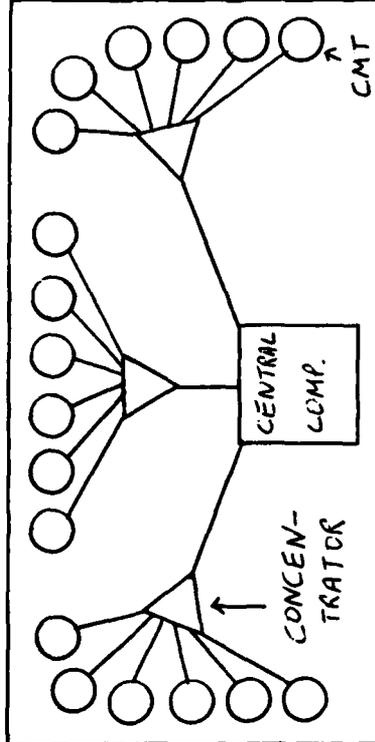


Fig. 1. An 18 CMT Network

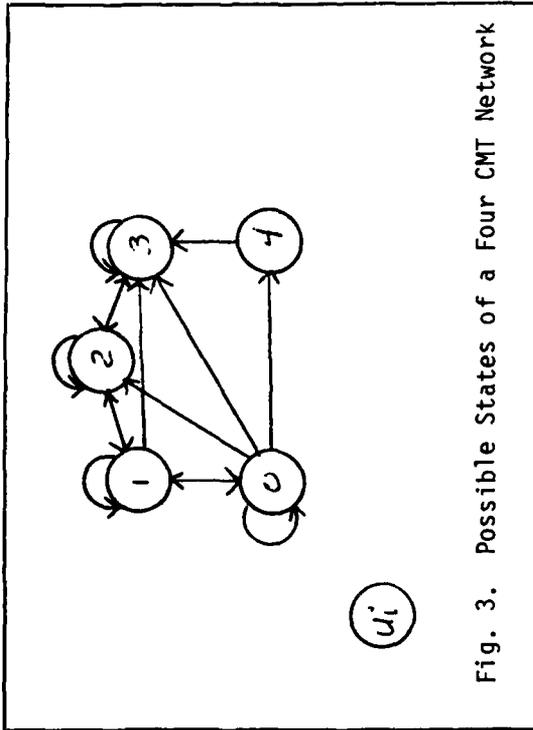


Fig. 3. Possible States of a Four CMT Network

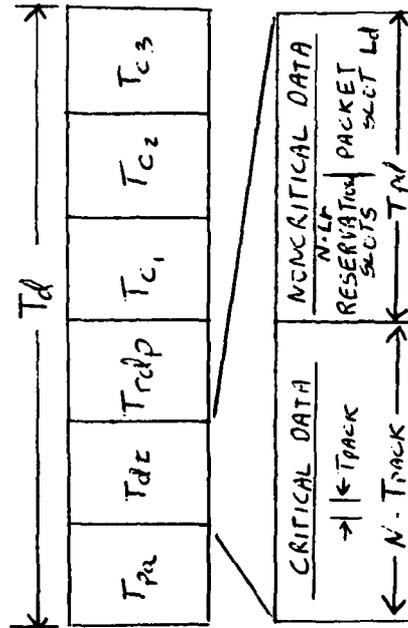


Fig. 2. System Response Times

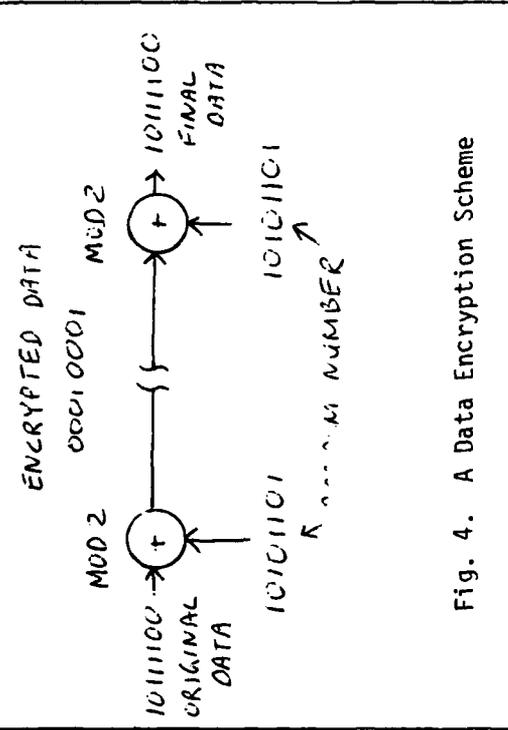


Fig. 4. A Data Encryption Scheme

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 USE OF A UNIQUE HEAT TRANSFER PROBE TO MEASURE
SPACECRAFT ROCKET PLUME CONTAMINATION

Prepared by: Dr. Frank G. Collins

Academic Rank: Professor

Department and University: Department of Mechanical and Aerospace Engineering, The University of Tennessee Space Institute

Research Location: Arnold Engineering Development Center, Calspan Field Services, Inc. Von Karman Facility, Advanced Test Diagnostics Branch

USAF Research: Allan Bailey

Date: September 1, 1983

Contract No: F49620-82-C-0035

THE USE OF A UNIQUE HEAT TRANSFER
PROBE TO MEASURE SPACECRAFT ROCKET
PLUME CONTAMINATION

by

Frank G. Collins

ABSTRACT

Hot wire probes have failed to find the same usefulness in rarefied flows that they have found in continuum flows because the convective cooling is masked by the conductive loss to the wire supports and, at lower pressures, by the radiative loss. In the present work, a free-molecule probe was constructed which had head thermistor heaters placed at its ends to keep the wire at a uniform temperature, thus eliminating the support conduction loss. The radiation loss was comparable to the convection loss at high Mach numbers for $n U < 10^{24}/m^2 s$ (n = number density, U = velocity). Under those conditions the probe was shielded, with the inner radiation shield maintained at the probe temperature.

The probe was designed to measure the boundary layers inside a Mach 4 conical nozzle and the nozzle external backflow region. $p_o = 10$ Torr, $T_o = 500^\circ K$ and the gas will be CO_2 . These tests will be performed in the near future.

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 for spending a very interesting and challenging summer at the Arnold Engineering Development Center, Arnold AFB, TN. He would like to thank the laboratory, and in particular the Advanced Test Diagnostics Branch of the Von Karmen Facility, for its hospitality and stimulating working conditions.

Finally, he would like to thank Allan B. Bailey for suggesting this area of research and for his collaboration and he would like to acknowledge the assistance of Dr. Leroy Brewer, Don Williams, Charles Nelson, Carl Kidd, Linwood Price, Bryan Seiber and Marshall Kingery.

I. INTRODUCTION

The potential contamination of spacecraft surfaces and components by the exhaust products from spacecraft thrusters is of considerable concern to the spacecraft design community and, in particular, to the U. S. Air Force. Contamination and subsequent degradation of sensitive spacecraft optical and radiative surfaces and components can compromise the mission objectives and lead to a reduced spacecraft life. Evidence of spacecraft contamination from backflow from spacecraft thruster plumes was provided by measurements made during flight STS-4 of the Shuttle. Large peaks in water vapor were measured with a mass spectrometer that was positioned in the open cargo bay (pointed outward) each time the PRCS thrusters were fired.¹ These thrusters fire perpendicularly to the Shuttle axis.

The concern for thruster contamination also exists because the more sophisticated spacecraft systems are designed to perform multiple, noninterfering missions for longer periods of time in space. However, the thrusters can leave contamination clouds of great size which remain in space for the spacecraft to pass through on successive orbits. This problem is demonstrated by the recent IUS 2nd stage burn at 23,000 miles altitude. At the end of the burn the contamination cloud was several hundreds of miles long. These examples clearly indicate that contamination from thruster rocket firings does occur. However, its fundamental causes must be understood before it can be effectively controlled.

Some ground-based experiments of thruster contamination have been performed in cryogenically pumped vacuum chambers. The mass flux pattern of the exhaust from several rocket motors was measured with quartz crystal microbalances (CCM). Significant mass fluxes were found in the region beyond the theoretical Prandtl-Meyer expansion angle², even at angles up to 140° from the nozzle centerline³. It has been found that the back flow region is mostly populated by molecules which come directly from the nozzle lip region and by the main flow molecules from very near the lip⁴.

Bird^{5,6} has computed the lip region using the direct simulation Monte Carlo method. For two-dimensional Prandtl-Meyer flows, the

non-equilibrium effects were found to be confined to the gas which entered the expansion within about twenty mean free path lengths from the corner. For underexpanded axisymmetric flows, the sonic line in the boundary layer was predicted to interact the nozzle lip and the plume backflow region was primarily determined by the flow in the immediate vicinity of the nozzle lip. Experimental verification of these predictions has not been made.

A computer program, CONTAM, has been developed to compute the backflow from thruster firings⁷. However, serious problems exist with this program when applied to low thrust motors which operate in rarefied conditions. In this regime the code predictions do not compare with the observed physics of the flowfield. For example, the program assumes that condensation occurs in a plane at some axial distance from the nozzle exit plane but recent low density tests at AEDC indicate that condensation occurs on a conical surface which emanates from the nozzle exit⁸. On the other hand, the direct simulation Monte Carlo technique has the opportunity for incorporating in a direct way the best understanding of the physics of molecular phenomena involved in these low density thruster contamination problems. Such a Monte Carlo code already exists but a very carefully controlled experiment using a small model nozzle must be performed for validation of the predictions of this code. This is one of the goals of the experiments that are being performed at AEDC, under the sponsorship of the Rocket Propulsion Laboratory.

The Monte Carlo code requires information upstream of the nozzle lip as starting conditions⁹. This is a region of intense gradients, and, as discussed previously, is also thought to be the region from which the backflow molecules come. Nonintrusive measurement techniques cannot be used to make measurements inside the nozzle but probes can. In the present study this region was examined with a specially designed hot wire probe and an impact probe.

The hot wire probe had very good spatial resolution and was capable of measuring the large gradients near the nozzle wall in the lip region. Very special design techniques had to be used to eliminate the

support conduction losses and to correct for the radiation losses. The combination of the two probe measurements led to values of the speed ratio, static temperature, pressure, density, velocity and enthalpy. By special design the probe was also capable of making measurements in the backflow region where the density was too low for some of the optical techniques to be successful without excessively long integration times.

The opportunity existed to compare the probe measurements with the measurements of the other techniques in the isentropic core of the nozzle expansion. U V Raman was used to obtain an absolute measurement of the number density and the velocity was obtained with a pulsed electron beam. An electron beam was used to obtain number density and rotational temperature and the mass flux was measured with a quartz crystal microbalance.

II. OBJECTIVES

A thin heated wire in a rarefied flow can be used to measure the heat transfer coefficient h , for the convective heat transfer to the probe wire and the adiabatic wall temperature of the wire, T_{aw} . It is assumed that the wire is in free molecule flow and that the flow approaching the wire is uniform. h and T_{aw} cannot be directly measured because of conduction losses to the wire supports and radiation losses from the probe. In addition, h and T_{aw} are generally not independent. However, a unique probe was designed and constructed which had the capability for measuring these quantities accurately. These two quantities, along with an independent measurement, can be used to obtain complete information about the flow field. Previous investigators had either coupled the hot wire measurements with a) impact pressure measurements¹⁰ or electron beam density measurements.¹¹

The objectives of the work were:

- (1) To determine the applicability of using a free-molecular hot wire probe to measure the flow field within and downstream of a conical supersonic nozzle, which is exhausting into a low vacuum.

- (2) To examine all previous uses of free-molecule hot wire probes and determine which design and operation would lead to the most accurate measurement of the flow field properties, using such a probe.
- (3) To design and have a free-molecule hot wire probe constructed, including necessary control electronics.
- (4) To develop all operational procedures for the probe and measure all calibration constants that are required for its operation.
- (5) To use the probe, coupled with an independent measurement, to carefully measure the flow field inside the exit of the conical nozzle, especially in the high gradient region in the boundary layer and near the nozzle lip.
- (6) To use the probe, coupled with an independent measurement, to measure the flow field external to the exit of the conical nozzle, especially in the region of backflow outside the shock wave system.

Due to a slip in the scheduled test of the rocket nozzle at AEDC not all of these objectives were accomplished during the ten week period when the investigator was at AEDC.

III. PREVIOUS USE OF FREE MOLECULE HEAT TRANSFER PROBES

The cross-stream cylinder was one of the first shapes to be examined theoretically in free-molecule flow¹⁹ and then these results were almost immediately used to obtain flow-field properties from a heated cross-stream cylinder. Sherman¹² measured T_{aw} for a very long unheated wire that was passed through a plane shock wave and obtained the shock wave profile and thickness. Wong¹³ developed a more traditional heated wire that was used to make measurements of h and T_{aw} in $M = 4$ and 6 flows. Coupled with an assumption of isentropic flow he was able to use the measurements to obtain the Mach number (speed ratio) and stagnation temperature. Theoretical expressions were used to make corrections for conduction losses to the supports and radiation losses were found to be unimportant [$T_{aw} \sim 345^{\circ}\text{K}$, $T_w - T_{\infty} < 100^{\circ}\text{K}$]. Laurman

and Ipsen¹⁴ used this hot wire probe to make flow field measurements around two simple two-dimensional shapes. They use an 8 inch long wire to get a very large length-to-diameter ratio to minimize the conduction error and found that large overheat ratios were required to accurately measure h and T_{aw} .

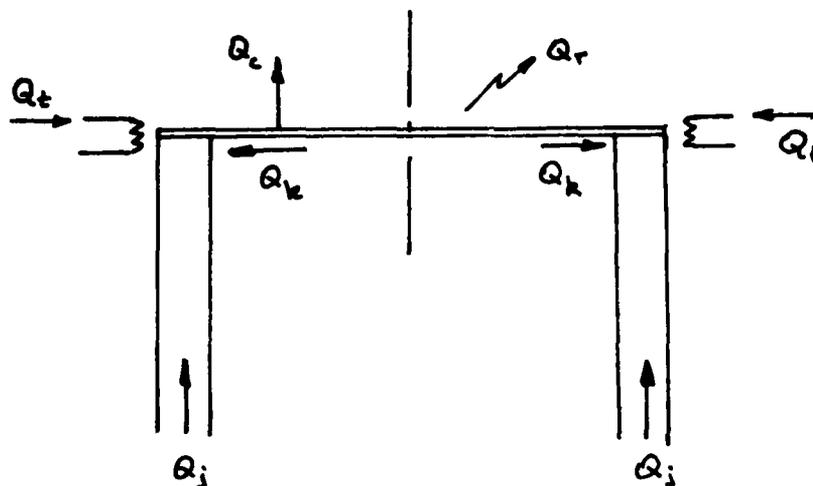
The conduction loss down the support prongs was the major problem with the free-molecule heat transfer probe and prevented its adoption as a standard measurement tool. Brown and Thomas¹⁵ pointed out that more heat was conducted through the supports when the probe was in a gas than when it was in a vacuum (while keeping the average probe temperature the same). By ignoring radiation Rajasooria and Brundin¹⁶ developed a technique for directly determining the conduction losses by calibration in a free stream and in a vacuum. The procedure eliminated the need for directly determining the probe dimensions or average material thermal conductivity. Lord¹⁷ greatly extended the technique for determining the conduction end loss correction, again ignoring radiation, and extended the corrections to fluctuating flows¹⁸. Muthoo and Brundin¹⁰ used Lord's correction technique and coupled the hot-wire measurements with those of the impact probe to obtain measurements of the speed ratio, pressure, temperature, stagnation temperature, density, velocity and enthalpy behind a sphere at a Mach number of 5.6. They resistively heated the support posts to minimize the conduction loss.

A major advance was made by Gottesdiener¹¹ who placed small bead thermistors at the ends of the heated wire and used the thermistors as nonlinear heat sources in a temperature control system that essentially allowed the entire length of wire to be at a constant temperature. This eliminated the conduction loss and allowed the radiation corrections to be measured under vacuum conditions and to be applied to each measurement. He used the probe in a free-jet flow where $M = 2$, $p = 2 \times 10^{-5}$ Torr. Coupled with electron beam density measurements, he was able to determine the flow speed ratio at the probe location. His design has been basically copied for the probe that was constructed at AEDC. However, radiation presented a much greater problem than was confronted by any previous investigators and the probe was shielded for measurements in

the lowest density flow, for the AEDC experiment.

IV. ANALYSIS OF HEAT TRANSFER TO A HEATED WIRE

Consider the steady state heat transfer from a finite length wire mounted upon supports which have heat sources attached to their ends.



In this figure Q_c = rate of convective heat transfer, Q_k = rate of conduction heat transfer to supports, Q_r = rate of radiative heat transfer to surroundings, Q_j is the rate of Joulean heating and Q_t is the rate at which heat is added by the end heaters (thermistors). The following assumption will be made:

$$Q_k = Q_t \quad (1)$$

Therefore, conduction can be neglected and an energy balance gives

$$Q_j = Q_c + Q_r \quad (2)$$

Introducing the heat transfer coefficient leads to the equation

$$h = \frac{Q_j - Q_r}{A_s (T_w - T_{aw})} \quad (\text{watts/m}^2\text{oK}) \quad (3)$$

where T_w is the wire temperature and A_s is the probe surface area. The wire temperature is related to its resistance by the relation

$$R_w = R_o [1 + \alpha_r (T_w - T_o)] \quad (4)$$

where R_0 is the probe resistance at a reference temperature T_0 and α_r is the temperature coefficient of resistivity. α_r is a function of the condition of the wire and must be determined by calibration for each probe.

The method for measuring h and T_{aw} will be described in Section VII.

V. FREE-MOLECULE HEAT TRANSFER TO AN INFINITELY LONG CROSS-STREAM CYLINDER

The free-molecule flow over an infinitely long circular cylinder at right angles to the flow field can be determined by assuming that the cylinder wall temperature is constant over the entire surface, that the thermal accommodation coefficient is constant over the surface and that each molecule carries $j/2$ kT units of internal energy, where

$$j = \frac{5 - 3\gamma}{\gamma - 1} \quad (5)$$

The thermal accommodation coefficient α , is usually assumed to be the same for all modes of energy transfer since experimental evidence is lacking for information about the accommodation for internal energy. The flow is assumed to be uniform. Then, the Stanton number and recovery factor are given by¹⁹

$$St = \frac{h}{\rho C_p U} = \alpha \frac{\gamma + 1}{\gamma} \frac{1}{12\pi^{3/2}} \frac{g(s)}{s} \quad (6)$$

$$r = \frac{T_{aw} - T}{T_0 - T} = \frac{6\gamma}{\gamma + 1} \frac{f(s) - g(s)}{s^2 g(s)} \quad (7)$$

$$s^2 = \frac{U^2}{2RT} \quad (8)$$

The functions of $f(s)$ and $g(s)$ are given in ref. 20. Also,

$$\frac{T_{aw}}{T} = \frac{f(s)}{g(s)} \quad (9)$$

Muthoo and Brundin¹⁰ combined these equations with the equation describing the measurement of an impact probe

$$p_i = pK(s) \quad (10)$$

to get the result

$$\frac{h\sqrt{2\pi T_{aw}}}{3 p_i \sqrt{R}} = \alpha \left(\frac{\gamma + 1}{\gamma - 1} \right) \frac{\sqrt{f(s) g(s)}}{18\pi K(s)} \quad (11)$$

For $s > \sqrt{\gamma/2}$

$$K(s) = \left(\frac{\gamma + 1}{\gamma} s^2 \right)^{\frac{\gamma}{\gamma-1}} \left[\frac{\gamma + 1}{4s^2 - \gamma + 1} \right]^{\frac{1}{\gamma-1}} \quad (12)$$

Thus, measurements of h , T_{aw} and p_i yield the speed ratio, if α is known.

In contrast, Gottesdiener¹¹ combined the hot-wire measurements with an independent measurement of the density, using the relation

$$F = 12\pi^{3/2} \frac{\gamma}{\gamma + 1} \left\{ \frac{f(s)}{[g(s)]^3} \right\}^{1/2} = \alpha \sqrt{2R} C_p \frac{\rho \sqrt{T_{aw}}}{h} \quad (13)$$

The velocity was then determined from a plot of

$$\frac{U}{\sqrt{2R T_{aw}}} = s \sqrt{\frac{g(s)}{f(s)}} \quad (14)$$

versus F .

Once the speed ratio is known plus either p_i or ρ , the remaining flow properties follow from equations (9), (10), (8) and assumptions of a perfect gas, namely,

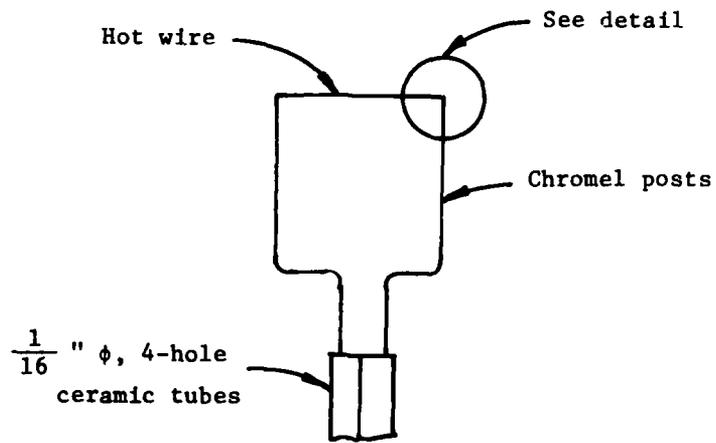
$$p = \rho R T \quad (15)$$

and

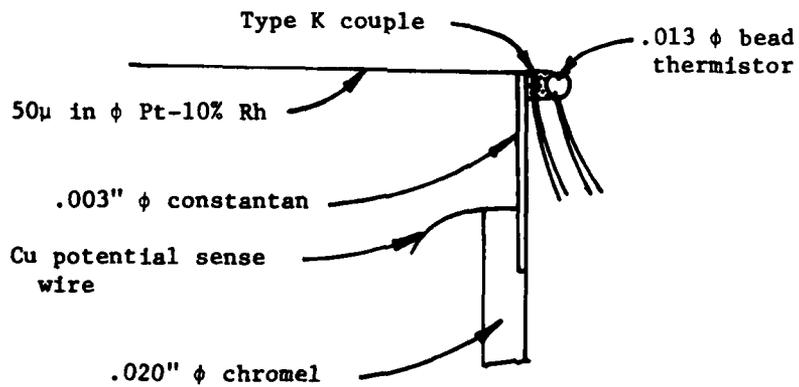
$$h = \frac{\gamma}{\gamma - 1} R \quad (16)$$

VI. DESCRIPTION OF FREE-MOLECULE PROBE

The probe is shown in the Figure. The heated element consisted of a 50×10^{-6} inch diameter Pt-10% Rh wire that was approximately 1/4 inch long. This wire was manufactured by the Wollaston process and had a silver sheath that was removed by a 50% nitric acid solution. The wire was attached to 0.003 inch diameter constantan posts which were in turn mounted to 0.020 inch diameter chromel supports. Normally the wires are attached to nickel supports but constantan has one-fourth the thermal conductivity of nickel and its use minimized the conduction loss from the ends of the wire. The constantan was silver soldered to the chromel with a eutectic solder and the Pt-10% Rh wire was attached to the constantan posts with EPO-TFK H20S, which is a high temperature (300°K to 400°K) silver-filled epoxy, manufactured by Epoxy Technology, Inc.



a) Hot wire probe



b) Detail

Construction of hot wire probe

Following Gottesdiener¹¹, the conduction losses were eliminated by heating the ends of the wire. Thermistors were found to be useful as the heating elements because of their rapid variation of resistance with temperature. The thermistors used were VECO 0.013 inch diameter bead thermistors with a resistance of 1000 ohms and dissipation constant of 0.10 $\text{mw}/^{\circ}\text{C}$. The end temperatures were sensed with small copper-constantan thermocouples. The thermocouples and thermistors were attached to the posts with the epoxy.

The chromel supports and all other wires went to three small multi-holed ceramic tubes which were mounted inside a 1/4 inch diameter stainless steel tube. This tube was supported on a mount vacuum chamber that could be rotated and translated with respect to the plane of symmetry of the nozzle. The short length of the wire (1/4 inch) allowed the probe to be positioned to within several thousands of an inch of the nozzle wall. This was more than adequate to resolve the properties through the nozzle wall boundary layer and its small size was more than adequate to resolve the high flow gradients in the nozzle lip region.

At very low values of the product nU (number density times velocity) the radiation heat transfer is much larger than the convection. When the probe was used in these circumstances it was surrounded by a double cylindrical radiation shield. The first shield consisted of two layers of copper wire (30 mesh, 37% open area) which were laid at 45° to one another. This shield contained a nichrome wire which heated it to the wire temperature. If the shape factor of this shield, from the probe, had been 1.0 then heating the shield to T_w would have eliminated the radiation error. However, the probe would no longer have been in free molecule flow. Therefore, the wire mesh was used; the molecules which struck the mesh were scattered in many directions, with little probability of returning to the wire, and those that proceeded through the mesh were pumped by the second cylinder, which was cooled with liquid nitrogen (CO_2 is effectively pumped at 77°K). This shielded arrangement allowed the probe to be used in regimes never reached by previous investigators.

The wire is in free-molecule flow ($Kn > 10$) if the number density is less than $8 \times 10^{16}/\text{cm}^3$. The highest density in the flow examined was $5 \times 10^{15}/\text{cm}^3$. Other properties of Pt-10% Rh are: $\rho_o = 18.4 \times 10^{-6} \text{ ohm cm}(0^\circ\text{C})$, $\alpha_r = 0.017/^\circ\text{K}$, $\epsilon_w \leq 0.1$.²¹

The heat transfer also depends on the thermal accommodation coefficient of the wire. The accommodation of both the translational and internal motion must be considered. Very little evidence exists about the accommodation of internal energy but that which does exist indicates that the accommodation coefficients for CO_2 for translational and rotational energy are both equal to 1.0 while the accommodation coefficient for vibrational energy is small, so that the overall accommodation coefficient is 0.81.²² The possibility exists for nonequilibrium to occur in the exhaust of the nozzle, especially in the boundary layer and the lip and backflow regions. Since the lowest vibrational temperature of CO_2 is 961°K and the stagnation temperature is only 500°K , there will be very little excitation of any of the vibrational modes and only rotational nonequilibrium needs to be considered. Assuming equal values of the translational and rotational accommodation coefficients, Sherman¹² derived the result

$$\frac{T_{aw}}{T_t} = \frac{f(s)}{g(s)} + \frac{1}{3} \left(\frac{T_r}{T_t} - 1 \right) \quad (17)$$

which gives the same result as equation (9) when the translational static temperature, T_t , equals the rotational temperature, T_r . Using this result, he calculated the difference in T_{aw} downstream of a shock wave for a diatomic gas, assuming that T_r remained at the upstream temperature, and compared it to the result that would be obtained for equilibrium flow. Even at an upstream Mach number of 4, the difference in wall temperature for the two cases was only about 2%. Thus, even in such a severe non-equilibrium flow as a normal shock wave, the effect on the wire temperature can be neglected as long as the accommodation coefficients for translational and rotational energy are equal.

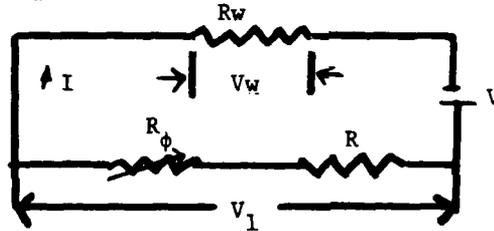
A greater error than that caused by the nonequilibrium flow is caused by flow gradients along the wire length. Traditionally free-molecule wire probes have been long (up to 8 inches) to minimize the

conduction loss. The present probe is considerably shorter because the conduction loss was reduced with the end heaters but it was felt that any further reduction in length would compromise its performance.

VI. PROBE OPERATION

The hot wire probe basically measures the convective heat transfer coefficient and the adiabatic wall temperature. These measurements must be coupled with the measurement of another property to yield other flow properties. It can either be coupled with measurements of the impact pressure or density, the latter from an electron beam. The first method is the only method applicable inside the nozzle while the latter is most useful in the nozzle backflow region.

The procedure for using the hot-wire probe to measure h and T_{aw} will now be briefly described. For purposes of illustration consider the following wire circuit (R_w = wire resistance).



The following procedures are applied at each probe location and for each flow condition. First, choose a value of T_w ($> T_{aw}$), then R_w is given by equation (4). Vary R_ϕ and the thermistor heating power until the wire is uniformly at T_w . Accurately measure V_1 and R_ϕ . The following equations then apply.

$$I = \frac{V_1}{R + R_\phi} \quad (18)$$

$$R_w = \frac{V_w}{I} = \frac{V_w}{V_1} (R + R_\phi) \quad (19)$$

$$Q_j(T_w) = I^2 R_w = \frac{V_w V_1}{(R + R_\phi)} \quad (20)$$

The amount of the total heating power that is due to radiation is determined by keeping the probe at T_w and the same location, turning off the flow, and reducing the pressure to 10^{-6} Torr or so. Then the above

measurement is repeated, leading to

$$Q_r(T_w) = \left[\frac{V_w V_1}{(R + R_\phi)} \right]_{\text{vacuum}} \quad (21)$$

This procedure is repeated, for a given flow condition and probe location, for at least four more values of T_w .

h and T_{aw} are determined from a plot of R_w versus $(Q_j - Q_r)$. The intersection gives R_{aw} , and then T_{aw} is determined from equation (4). From equations (3) and (4) the slope of the curve is

$$\frac{dR_w}{d(Q_j - Q_r)} = \frac{\alpha_r R_o}{h A_s} \quad (22)$$

Therefore, with accurate measurements of α_r , R_o and A_s , the heat transfer coefficient can be determined from the slope of the curve. The least accurate of these quantities is A_s , the surface area of the probe. The end correction method of Lord¹⁷ eliminates the need to measure A_s but that method cannot be extended to circumstances where radiation must be included.

Before the flow properties can be determined, the thermal accommodation coefficient of the wire must be measured. It can be determined from static gas tests at pressures such that $Kn > 10$. For the present probe, if the gas temperature is ambient, then the pressure must be less than 2.5 Torr. This is in the range of the Baratron, which yields accurate measurements of pressure. The above tests must be repeated at various pressures to give a measurement of the heat transfer coefficient. The $s = 0$ limit of equation (6) gives

$$h = \frac{Q_j - Q_r}{A_s (T_w - T)} = \alpha \left(\frac{\gamma + 1}{\gamma - 1} \right) \frac{p}{\sqrt{8\pi}} \sqrt{\frac{R}{T}} \quad (23)$$

This equation can then be used to measure the accommodation coefficient.

At each probe location either the impact pressure or the density was measured. If p_i was measured then equation (11) was used to determine the speed ratio and if ρ was measured then equation (13) was used. Once the speed ratio is known, the temperature can be determined from equation (9). The velocity can be determined from the definition of the speed ratio or equation (14). Perfect gas expressions then yield the density or pressure, respectively, and the enthalpy.

Radiation is a considerably greater problem for the present tests than those of previous investigators because of the high stagnation temperature and corresponding adiabatic wall temperature (500°K and 583°K , respectively) and because the walls of the chamber were cooled with liquid nitrogen. A simple analysis would assume that the probe was surrounded by walls at a temperature T_w and that all walls were gray. The rate of radiative heat transfer would then be given by

$$Q_r = \epsilon_w A_s \sigma (T_w^4 - T_{\infty}^4) \quad (24)$$

where ϵ_w is the emissivity of the probe (less than 0.1 for the present probe).

A detailed analysis indicates the importance of radiation for the present tests. The ratio Q_r/Q_c increases as T_w departs from T_{∞} , so measurements in the backflow ($T_{\infty} = 77^{\circ}\text{K}$) are more affected by radiation than those in the nozzle ($T_{\infty} = 470^{\circ}\text{K}$). This ratio is proportional to $1/nU$, which means that Q_r/Q_c increases in the boundary layer where the velocity is low and in the backflow region where the number density is low. Also, for a given value of nU , the ratio decreases as $(T_w - T_{aw})$ increases so radiation is less of a correction at higher overheat ratios. This discussion indicates that radiation cannot be ignored anywhere in the flow examined and the radiation correction had to be made. In the backflow region the radiation dominated the convection and the probe had to be shielded, as described in the previous section.

To get a feel for the magnitude of the heat transfer, the following formulas apply for the following conditions: gas - CO_2 , $U = 815 \text{ m/s}$, $n = 5.85 \times 10^{21}/\text{m}^3$, $s = 3.35$, $\alpha = 0.9$, $\epsilon_w = 0.1$.

$$\begin{aligned} Q_c &= 1.47 \times 10^{-5} [T_w - T_{aw}] \quad \text{watts} \\ Q_r &= 1.44 \times 10^{-16} [T_w^4 - T_{\infty}^4] \quad \text{watts} \end{aligned} \quad (25)$$

VII. RESULTS

It was determined that a free-molecule heat transfer probe could be useful for measuring the flow field inside and outside an underexpanded rarefied nozzle flow field. It is the only measurement tool that can make measurements inside the nozzle. At high stagnation temperatures it

must be surrounded by a radiation shield when used external to the nozzle.

Due to an unavoidable change in testing schedule the probe was not constructed but was being constructed when the research period ended and the experiments are now being planned for early winter. The investigator will be able to participate in these tests through a cooperative agreement between AEDC and the University of Tennessee Space Institute.

VIII. RECOMMENDATIONS

The results obtained indicate that the free-molecule hot wire probe offers great promise for measuring rarefied flow field properties. It should be especially useful for measuring the flow inside a nozzle, where optical techniques cannot be applied. Therefore, it is strongly recommended that the probe designed be used in the planned nozzle experiments. In addition, the following things need to be explored further: determine the effect of non-equilibrium flow on the heat transfer to a wire, caused by a change in the accommodation coefficient and determine if there might be some way to incorporate the radiation correction in such a way that the probe surface area does not have to be measured (similar to the technique of Lord¹⁷). Also, it is recommended that future probes be made of similar materials to eliminate thermocouple junction emf's (the first probe was constructed from available materials).

Investigators who make fundamental turbulence measurements in continuum flow have recently learned that the non-uniform wire temperature, caused by the conduction to the supports, causes errors in the results. It is strongly recommended that this probe be used under these conditions, to see if this error is eliminated. Also, the larger thrust spacecraft nozzles have turbulent wall boundary layers and this probe could be used to measure their flowfields.

REFERENCES

1. Miller, E.R., editor, "STS-2,-3,-4 Induced Environment Contamination Monitor (IECM) Summary Report," NASA TM-82824, February 1983.
2. Chirivella, J.E., "Hydrazine Engine Plume Contamination Mapping," AFRPL-TR-75-16, October 1975.
3. Chirivella, J.E., "Mass Flux Measurements and Correlations in the Back Flow Region of a Nozzle Plume," AIAA Paper No. 73-731, presented at AIAA 8th Thermophysics Conference, Palm Springs, CA, July 16-18, 1973.
4. Chirivella, J.E., E.C. Baroth and C.S. Guernsey, "Nozzle Lip Flow and Self-Scattering Molecular Collisions as Contributors to Plume Backflow," JANNAF 13th Plume Technology Meeting, April 27-29, 1982.
5. Bird, G.A., "Breakdown of Continuum Flow in Freejets and Rocket Plumes," 12th Int. Symp. on Rarefied Gas Dynamics, Vol. 74, Prog. in Astron. and Aeron., AIAA, pp. 682-694, 1980.
6. Bird, G.A., "The Nozzle Lip Problem," Proc. Ninth Int. Symp. on Rarefied Gas Dynamics, DFVLR Press, Porz-Wahn, West Germany, Vol. I, ref. A. 22, 1974.
7. Trinks, H. and R.J. Hoffman, "Experimental Investigations of Bipropellant Exhaust Plume Flowfield, Heating and Contamination and Comparison with the CONTAM Computer Model Predictions," AIAA 83-1447, June, 1983.
8. Bailey, A.B., private communication, June 1983.
9. Brock, F., private communication, April 12, 1983.
10. Muthoo, S.K. and C.L. Brundin, "Near Wake Flow Field Measurements Behind Spheres in Low Reynolds Number Hypersonic Flow," 9th Int. Sym. on Rarefied Gas Dynamics, DFVLR Press, Porz-Wahn, West Germany, ref. B.10, 1974.
11. Gottesdiener, L., "Hot Wire Anemometry in Rarefied Gas Flow," J. Phys. E. 13, 908-13 (1980)

12. Sherman, F.S., "A Low-Density Wind-Tunnel Study of Shock-Wave Structure and Relaxation Phenomena in Gases," NACA TN 3298, July 1955.
13. Wong, H., "Design and Development of a Free-Molecule Heat-Transfer Probe," Univ. of Calif., Berkeley, Report No. HE-150-143, Oct. 15, 1956.
14. Laurmann, J.A. and I.C. Ipsen, "Use of a Free Molecule Probe in High Speed Rarefied Gas Flow Studies," WADC Tech. Rept. 57-440, Oct. 1957; J.A. Laurmann, "The Free Molecule Probe and Its Use for the Study of Leading Edge Flows," Phys. Fluids 1, 469-477 (1958).
15. Brown, R.E. and L. Thomas, "The Assignment of the Heat Flow From Electrically Powered Filaments to Gas Conduction for Accurate Measurement of Thermal Accomodation," 7th Int. Rarefied Gas Dynamics Symposium, Vol. I, Editrice Technico Scientifica, Pisa, Italy, pp. 347-355.
16. Rajasooria, G.P.D. and C.L. Brundin, "Use of Hot Wires in Low-Density Flows," AIAA J. 9, 979-81 (1971).
17. Lord, R.G., "Hot Wire Probe End-loss Corrections in Low Density Flows," J. Phys. E 7, 56-60 (1974).
18. Lord, R.G., "The Dynamic Behavior of Hot-Wire Anemometers with Conduction End Losses," J. Phys. E 14, 573-578 (1981).
19. Stalder, J., G. Goodwin and M.O. Creager, "Heat Transfer to Bodies in a High-Speed Rarefied-Gas Stream," NACA Rept. 1093, 1952.
20. Stalder, J.R., G. Goodwin and M.O. Creager, "A Comparison of Theory and Experiment for High-Speed Free-Molecule Flow," NACA Rept. 1032, 1951.
21. Seban, R.A., WADD-TR-60-370, Part II, August 1962.
22. Devienne, F.M., "Low Density Heat Transfer," Adv. Heat Transfer 2, Academic Press, 1965, pp. 272-356.

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 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

Prepared by: Richard V. Conte, Ph.D.
Academic Rank: Assistant Professor
Department and University: Mechanical Engineering Department
Manhattan College
Research Location: Aero Propulsion Laboratory
Turbine Engine Division
Components Branch
USAF Research: Dr Kervyn Mach
Dr Richard Rivir
Date: 22 July 1983
Contract No.: F49620-82-C-0035

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

By

Richard V. Conte, Ph.D.

ABSTRACT

A program was undertaken to enable the utilization of advanced computer codes to investigate the heat transfer characteristics of arbitrary high-temperature turbine blades. The success of such a procedure begins with the modeling of the region in space occupied by the turbine blade together with the specification of appropriate boundary conditions. A computer code was written that would accept an arbitrary set of turbine blade coordinates and display them on a plotter. The blade data is then manipulated to achieve the desired orientation and appropriate mathematical functions are fitted between the blade coordinates. A cubic-spline, circle and straight line were used as the interpolating functions. When all the data points are satisfactory fitted, a two-dimensional mesh is fitted to the interpolated blade coordinates. The grid generated is customized to the particular test conditions being investigated.

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 the opportunity for him to spend an enriching summer at the Aero Propulsion Laboratory, Wright-Patterson Air Force Base, Dayton, Ohio. He would like to acknowledge in particular the Components Branch of the Turbine Engine Division for providing the equipment and atmosphere conducive to research.

Finally, he would like to thank Dr Richard Rivir for suggesting this area of research and Dr Kervyn Mach for his invaluable assistance in mastering a new computer system. He would also like to acknowledge Mr Charles MacArthur for his many helpful discussions, Mr John Straus for his help in inputting the computer programs, and Miss Ann Gerstner for typing this report.

I. INTRODUCTION:

The recent trend towards higher turbine inlet temperatures with the resulting higher power output continues unabated. Unfortunately, the thermal design of modern high-pressure turbines clearly represents one of the more difficult engineering tasks in the design of the aircraft gas turbine. The aerodynamic and thermal procedures currently available to the turbine designer are insufficient in that they do not allow a priori design that achieves design goals without expensive experimental development iterations. Hence, a more refined procedure to determine heat transfer characteristics on turbine components will be needed for the operating environment to be encountered in the years ahead.

Stepka¹ has reviewed the state-of-the-art and concluded that the uncertainty of predicting the heat transfer rates from hot gas streams to turbine blades was about 35 percent. Stepka, along with many others, were of the opinion that improved evaluation of the heat transfer rates is a high-priority task in the high-temperature turbine technology improvement program.

Research in heat transfer is invariably linked to research in fluid dynamics. With an improved understanding of the fluid dynamical phenomena in turbine cascades, a better understanding of the heat transfer characteristic should follow. Factors affecting the heat transfer have been discussed by Graham². The most dominant ones are acknowledged to include:

Lack of precision in the prediction of the inviscid flowfield around the airfoil, particularly in the forward, high accelerated stagnation region.

Uncertainties regarding the surface location at which transition is initiated as well as the surface extent of the transition zone.

Uncertainties regarding the influence of free-stream turbulence on local heat transfer rates in the laminar region as well as on initiation and extent of the transition region.

Limited understanding of the role of airfoil surface curvature on turbulence production/dissipation and boundary layer stability.

The complex environment described above suggests the need for an improved design approach based on numerical methods and enlightened by improved turbulence models to handle the influences described previously. Conceptually, a computer code marches down the blade surface from the leading edge toward the trailing edge in small increments. Along its way, the fluid dynamical characteristics and heat transfer rates are computed based on prior value of velocity and temperature. The effectiveness of the various computer programs currently available has been examined by Daniels and Browne³ and found lacking. A same conclusion was reached in a study by Gaugler⁴. The importance of continued heat transfer research for turbine blade applications has clearly been demonstrated.

II. OBJECTIVES:

The main objective of this project is to investigate the phenomena of turbulent heat transfer on turbine blades. It was our intent to utilize advanced computer codes which have not been previously used for this purpose. It was not our intent to actually run the codes but to construct a series of computer programs that will facilitate creation of the input data to the codes. The specific objectives of the computer programs were:

- (1) To accept arbitrary turbine blade data and display the results on a graphics plotter;
- (2) To manipulate the raw data to achieve any desired

two-dimensional orientation; (3) To mathematically fit an appropriate function between the turbine blade coordinates; (4) To generate a grid of mesh points to describe the region surrounding the turbine blade.

The unique feature of the above project is that it will interactively work on arbitrary turbine blade data.

III. MODELING THE TURBINE BLADE DATA:

To achieve the first objective, a series of data files were created which contained turbine blade data from a variety of sources. The data consisted of the x and y coordinates of the turbine blade sequentially numbered with the restriction of an odd number of data points. To accommodate the odd requirement, the first and last blade coordinates may be identical. If available, the file may also contain the leading and trailing edge radius centers. Program GRD will ask the name of the data file, input the coordinates and display the points on the plotter. Figure 1 is an example of this plotted raw data.

To achieve the second objective, GRD provides a series of manipulating options including y-axis mirroring, x-axis mirroring, translation in y-direction, translation in x-direction, rotation in the x-y plane, and aligning a two-blade coordinate with the x-axis. Using these routines the raw blade data can be manipulated into any desired orientation. Figure 2 shows the raw data of Figure 1 manipulated into a completely different orientation. The data points inside the blade represent the leading and trailing edge radius centers, and the starred coordinate is the first coordinate. The grid generating routine to follow requires the first

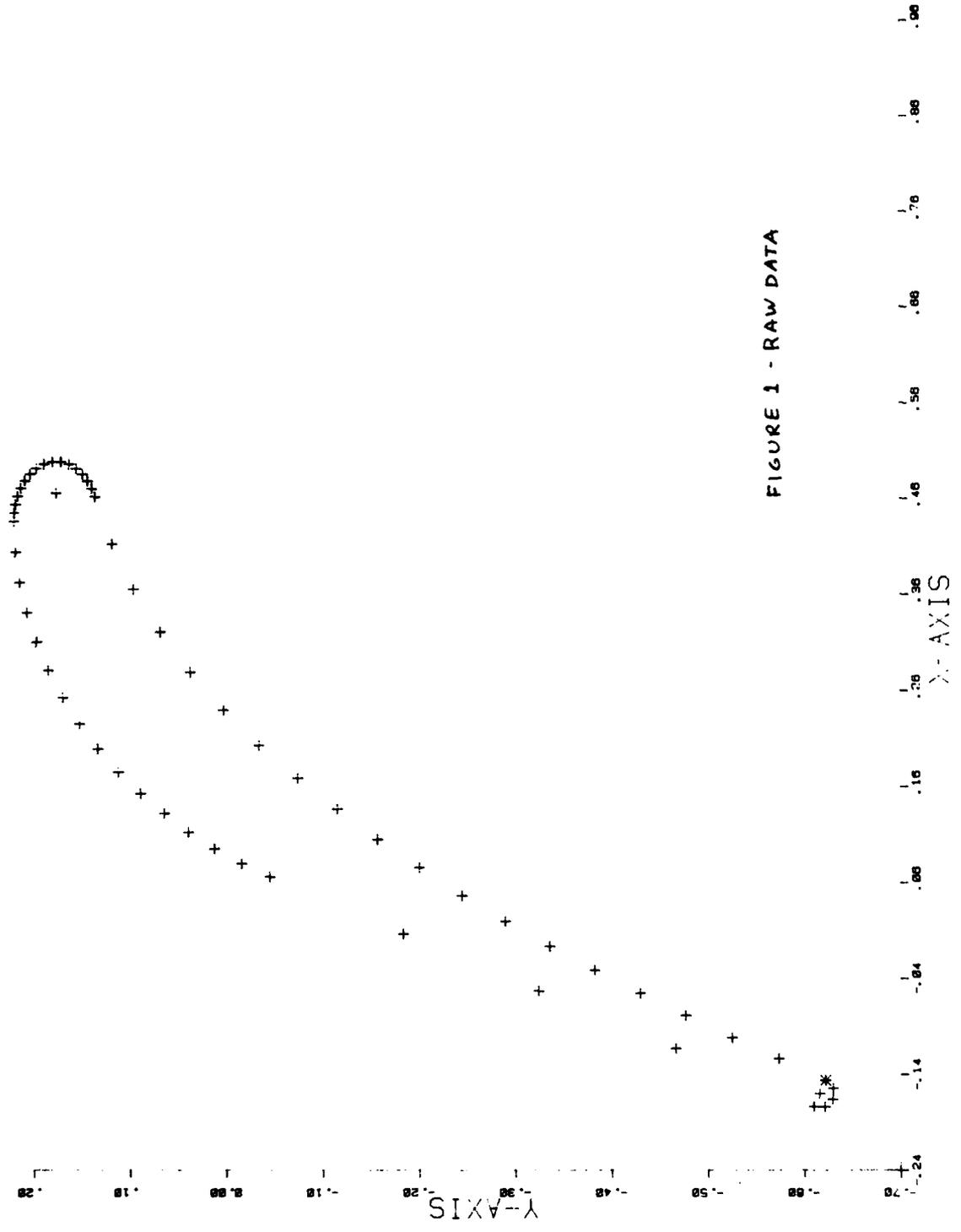


FIGURE 1 - RAW DATA

Y-AXIS
 .50
 .40
 .30
 .20
 .10
 .00
 -.10
 -.20
 -.30
 -.40
 -.50

18-8

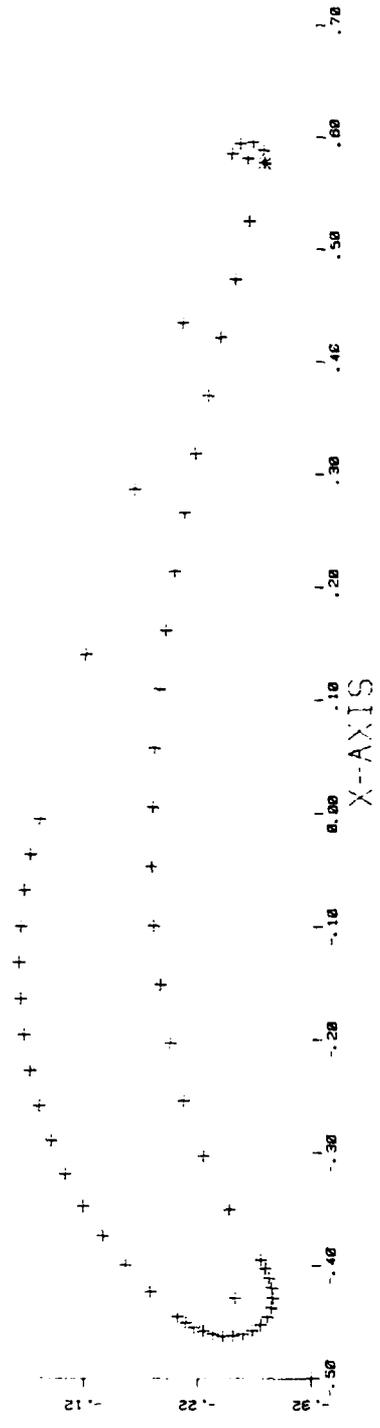


FIGURE 2 MANIPULATED DATA

coordinate to coincide with the stagnation point of the blade. To facilitate this requirement program NSNP will shift the coordinates to make the first coordinate the stagnation coordinate.

To achieve the third objective, GRD fits appropriate mathematical functions between the turbine blades. A cubic-spline was selected as the primary function since it is very successful in fitting the smooth data encountered in turbine blade profiles. The theory of cubic-spline functions is described in Hornbeck⁶. This procedure is quite sensitive to the relative orientation of the data. In particular, problems are encountered when the program tries to interpolate data "around a corner". To remedy this problem alternative interpolating functions were included. When asked, the program will fit either a straight line or a circle between specific blade intervals (a blade interval consists of three consecutive blade coordinates). The remaining blade intervals will continue to be fitted with the cubic-spline. The circle was quite useful in fitting data at the leading and trailing edges as these regions are circular by nature. This part of the program is highly interactive, with the user selecting which intervals are to be fitted with other than cubic-spline functions. This procedure is continued until all the coordinate points have been satisfactorily fitted. It should be mentioned that one input to the program is the number of interpolated (points between the original blade coordinates) data points. This value will vary depending on the spacing of the raw data. Figure 3 shows the blade of Figure 2 with the interpolated functions plotted over the blade coordinates. To arrive at this figure, eight intervals were defined to be non-cubic-spline. This is not always the case. Figure 4 shows an example where all the blade coordinates were fitted with the cubic-spline function.

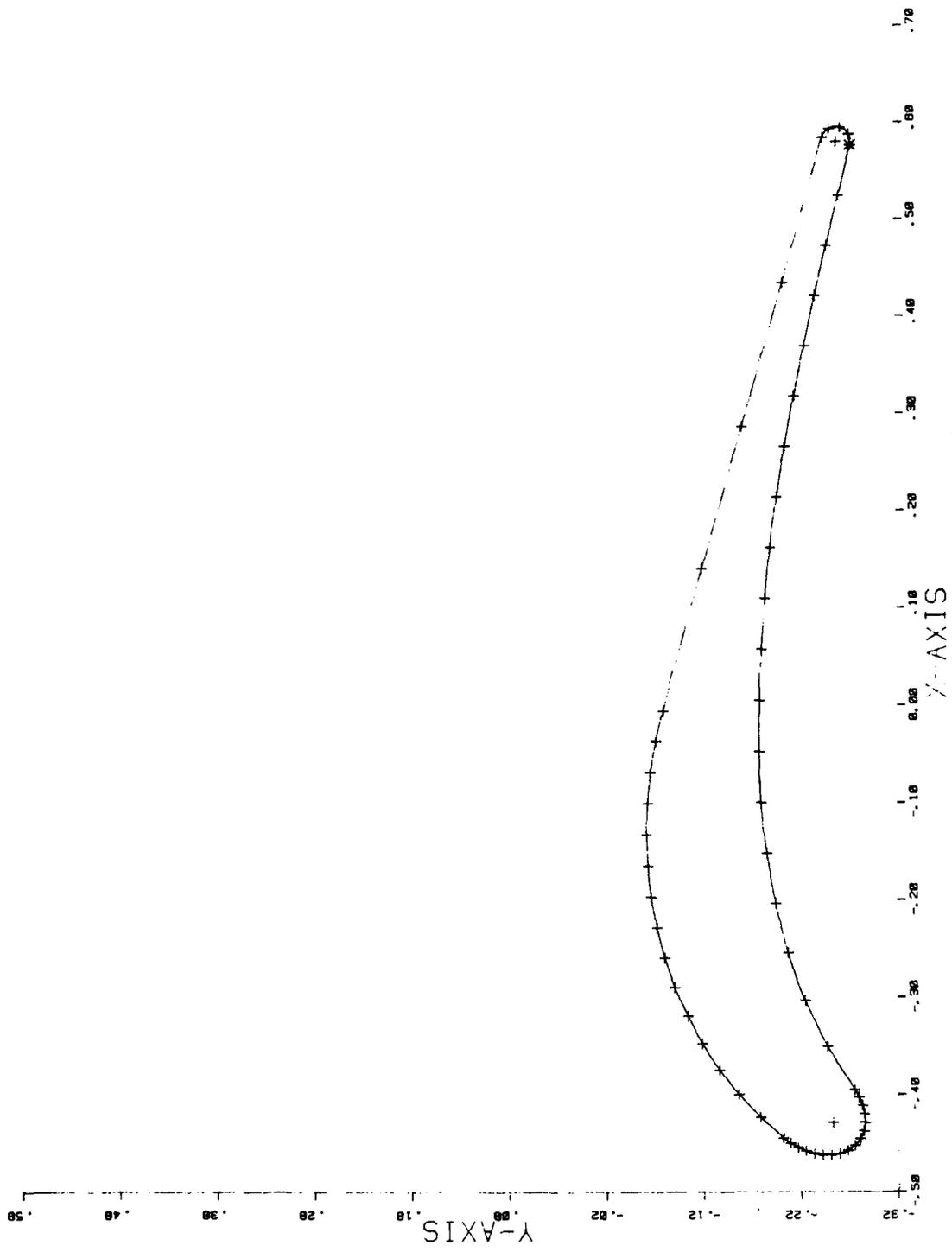


FIGURE 3 - FITTED BLADE

11-81
Y-AXIS
17.20
15.20
13.20
11.20
9.20
7.20
5.20
3.20
1.20
-0.80

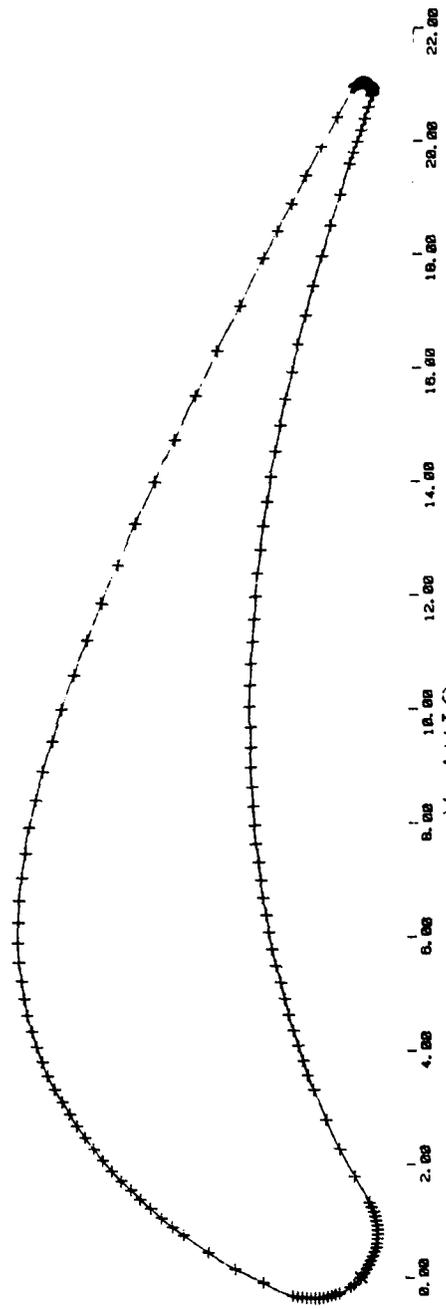


FIGURE 4 - FITTED BLADE

To achieve the fourth objective, program MESH was written. The interpolated coordinates of GRD are saved in a file to make them available to MESH. The program begins by computing a series of rays leaving the interpolated coordinates normal to the blade surface. Three distinct regions are defined: the boundary layer, the intermediate grid region, and the course grid region. The boundary layer will contain the highest concentration of points since it is the region where the velocity and temperature gradients are most severe. The boundary layer thickness is approximated at the mid-chord value for a given value of Reynold's number. Since the stagnation point contains the region of greatest velocity and temperature gradients, a higher concentration of grid points is allowed in this region. Following the recommendation of Bradshaw⁷, the grid points are spaced closer to the surface with ever increasing intervals as they move away from the surface. Each interval is a constant K greater than the previous interval. If $DEL1$ represents the intergal spacing between the first two grid points, then $K(DEL1)$ will represent the spacing between the next set of grid points. Once the fine grid is defined, a course grid is defined by first finding an ellipse which encloses the blade. The intersection of this ellipse and the rays define the beginning of the course grid. Succeeding ellipses are so defined as to degenerate into circles before terminating at approximately one chord length on all sides of the blade. Again, the same interval spacing technique is employed as is used with the boundary layer grid. With the fine grid and the course grid defined, the intermediate grid is defined in a manner analogous to the fine grid: the interval spacing technique is applied from the end of the fine grid to the beginning of the course grid. The final computed grid values are stored in a file for future use.

IV. RECOMMENDATION:

As mentioned previously, the broad area of which this project is part has great practical value. With a successfully generated grid, the next step would be to model the initial conditions of the flowfield. All these results, together with information regarding turbulence, would then be inputted in the advanced computer code and the numerical results can be compared with the experimental data. This will ultimately lead to the new design procedures mentioned in the beginning of this report.

My main recommendation is for follow-on research to complete the work started by this project. In particular, the grid generated must be supplemented with initial conditions and a description of the turbulence flowfield. These can be utilized by advanced computer codes to assess the validity of this procedure.

REFERENCES

1. Stepka, F. S., "Uncertainties in Predicting Turbine Blade Metal Temperatures", ASME Paper, 80-HT-25, 1980.
2. Graham, R. W., "Fundamental Mechanisms That Influence the Estimate of Heat Transfer to Gas turbine Blades", ASME Paper, 79-HT-43, 1979.
3. Daniels, L. D., and Browne, W. B., "Calculation of Heat Transfer Rates to Gas Turbine Blades," Int. J. Heat Mass Transfer, Vol. 24, No. 5, pp. 87-879, 1981.
4. Gaugler, R. E., "Some Modification to, and Operational Experiences with, the Two Dimensional, Finite Difference, Boundary-Layer Code, STAN 5," ASME Paper, 81-GT-89, 1981.
5. Han, L. S., Cox, W. R. and Chait, A., "Investigation of the Boundary Layer Behavior on Turbine Airfoils," AFAPL-TR-79-2011, 1971.
6. Hornbeck, R. W., Numerical Methods, (Quantum Publishing, New York, 1975), p. 47-50.
7. Bradshaw, P., Cebeci, T. and Whitelaw, J., Engineering Calculation Methods for Turbulent Flow, (Academic Press, New York, 1981), P. 171-173.
8. Han, L. S., Chait, A., Boyce, W. F., Rapp, J. R., "Heat Transfer on Three Turbine Airfoils", AFWAL-TR-82-2124, 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 and Infrared Spectroscopy of Extrinsic P-Type Silicon

Prepared by: Dr. Billy C. Covington

Academic Rank: Assistant Professor

Department and University: Department of Physics
Sam Houston State University

Research Location: Air Force Wright Aeronautical Laboratories,
Materials Laboratory, Electromagnetic Materials Division,
Laser & Optical Materials Branch

USAF Research Colleague: Dr. P. M. Hemenger

Date: July 29, 1983

Contract No: F49620-82-C-0035

Raman and Infrared Spectroscopy of Extrinsic P-Type Silicon

by

B. C. Covington

ABSTRACT

We present the results of a Raman spectroscopy study made at room and liquid helium temperatures of silicon samples conventionally doped with one of the Group IIIA elements boron, aluminum, gallium, indium, or thallium.

We attempt to observe a previously reported Raman electronic scattering line in boron-doped silicon. The failure to observe this line at approximately 23.4 meV is attributed to the type of laser used and perhaps to the quality of the sample. We report several temperature dependent lines of unknown origin.

A detailed study as a function of annealing temperature is made of neutron transmutation-doped silicon (gallium). Comparisons are made to a previous study of neutron transmutation-doped pure silicon and silicon (boron).

Suggestions for additional research are made.

Acknowledgement

The author would like to express his appreciation to 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 do research at the Air Force Wright Aeronautical Laboratories, Materials Laboratory, Wright-Patterson AFB, Ohio. He wishes to thank the Materials Laboratory staff for their support and in particular to thank Drs. Roger Becker and Dietrich Langer for their assistance with the Raman study. He especially wishes to thank Dr. Patrick Hemenger for his collaboration and helpful assistance.

I. INTRODUCTION

The development and spectral characterization of high quality extrinsic silicon material for fabrication of infrared detectors for use in missile, satellite, and aircraft imaging systems is of continuing importance to the Air Force. This report provides additional data on the spectral properties of silicon doped with one of the Group IIIA elements boron, aluminum, gallium, indium, or thallium. The research effort consists of two parts.

Part (1) Raman Spectroscopy

Raman spectroscopy has been used extensively to study the phonon spectra of pure silicon and silicon doped with p and n-type impurities¹⁻⁶. Comparatively few researchers⁷⁻¹¹ have reported studies of electronic Raman scattering by p and n-type impurities in silicon. Of particular interest, is the report by Wright and Mooradian⁷ of an electronic Raman scattering line in boron-doped silicon located at 23.4 meV from the 1.06 μ line of a Nd:YAG laser. This electron scattering line is attributed to a transition from the boron $P_{3/2}$ ground state of Γ_8^+ symmetry to a higher excited two fold state. One such possible transition suggested by Klein¹² is $1SP_{3/2}(\Gamma_8^+) \rightarrow 1SP_{1/2}(\Gamma_7^+)$. The presence of this line in the spectrum of boron-doped silicon and the failure by Wright and Mooradian¹³ and other researchers^{10,11} to observe a similar line in gallium and aluminum-doped silicon is a theoretical and experimental problem which deserves additional investigation.

Due to the extreme similarity of the Group IIIA acceptor absorption spectra, it would seem reasonable to expect a transition similar to the electronic boron transition to occur in aluminum, gallium, indium, and thallium-doped silicon. To the best of our knowledge, attempts to observe this transition have only been made in aluminum and gallium-doped

silicon. In addition, these previous experimental efforts were not made within the past few years and therefore did not have the benefits of higher quality silicon material and improved experimental apparatus. With this in mind, we have attempted to observe the boron-type electronic Raman transition in boron, aluminum, gallium, indium, and thallium-doped silicon.

Part (2) NTD Silicon (gallium)

Most silicon contains boron in concentrations sufficient to degrade detector efficiency and to require colder operating temperatures. In order to compensate (render infrared inactive) the unwanted boron, neutron transmutation-doping (NTD), a process suggested by Lark-Horovitz¹⁴ and discussed in detail by Tanenbaum and Mills¹⁵, is used. The process employs the (N, γ) reaction to convert ^{30}Si to ^{31}P , which then compensates the boron.

The major disadvantage of this process is that it produces undesirable radiation induced defects in the silicon crystal. These defects produce unwanted absorption centers as well as rendering the ^{31}P inactive. To cure the radiation damage and to activate the ^{31}P the crystal must be annealed at high temperature (900°C) for 15 to 30 minutes.

Covington¹⁶ and Watson and Covington¹⁷ have done studies of NTD pure silicon and NTD silicon conventionally doped with boron. We have extended the study to include NTD silicon conventionally doped with gallium.

II. OBJECTIVES

The specific objectives of this research project were:

- (1) Attempt to observe the previously reported electronic Raman scattering line at 23.4 meV in boron-doped silicon.

- (2) Attempt to observe the boron-type electronic Raman scattering line in aluminum, gallium, indium, and thallium doped silicon.
- (3) Obtain high resolution infrared absorption spectra as a function of annealing temperature for neutron transmutation-doped silicon (gallium).
- (4) Make comparisons between the NTD silicon (gallium) spectra and the previously obtained NTD pure silicon and NTD silicon (boron) data.

III. EXPERIMENTAL TECHNIQUES

Part (1) Raman Study

Samples measuring 5mm x 10mm by 1mm thick were cut from silicon boules doped with one of the Group IIIA elements. The impurity concentrations were: Si(B) - 1.6×10^{17} Bcm⁻³, Si(Al) - 5×10^{16} Al cm⁻³, Si(Ga) - 6.5×10^{16} Ga cm⁻³, Si(In) - 5×10^{16} In cm⁻³, and Si(Tl) - 5×10^{16} Tl cm⁻³.

For all scans, the samples were mounted so that the front surface of the sample was at 90° with respect to the spectrometer entrance slits. The laser light was incident to the sample surface at an angle of approximately 45°. This procedure produced a backscattering configuration for collecting the Raman signal. A Spex model 1401, 0.85 meter double monochromator was used for all measurements. The laser was a Spectra Physics krypton laser operating at a wavelength of 6471 Å and at one watt of power. The detector was an RCA photomultiplier cooled with chilled water.

For the low temperature scans, the samples were mounted to the cold finger of an Adonian dewar. The cold finger was at 4.2K but due to the laser light the sample temperature was significantly higher. We had no thermometer for direct measurement of the sample temperature.

Part (2) NTD Silicon (gallium)

All infrared absorption spectra were collected with a Digilab FTS Model 20 spectrophotometer operating at a resolution of 1 cm^{-1} . The spectral range was $4000\text{-}400 \text{ cm}^{-1}$.

The two samples used were cut from a boule of gallium-doped silicon grown by Virginia Semiconductor. The samples were irradiated at the Texas A&M Research Reactor to phosphorous concentrations of $1 \times 10^{14} \text{ P cm}^{-3}$ and $1 \times 10^{15} \text{ P cm}^{-3}$, respectively. A third sample cut from the same boule but not irradiated was used as a reference.

The annealing procedure consisted of 100°C annealing stages beginning at 200°C and ending with 900°C . The samples were annealed for 30 minutes at each stage. The annealing apparatus consisted of a Lindburg furnace, silicon annealing tube, silicon boat, and argon gas for purging.

IV. RESULTS AND CONCLUSIONS

Part (1) Raman Study

An attempt was made to observe the previously reported⁷ electron Raman scattering line in boron-doped silicon. A spectrum of Si taken at liquid helium temperature is shown in Figure 1. The strong line at approximately 103 cm^{-1} has not been identified. The same spectrum with an expanded scale to enlarge the region greater than 100 cm^{-1} is shown in Figure 2. As can be seen in these two spectra, there is no evidence of the 23.4 mev (187.2 cm^{-1}) electronic boron line. We think the absence of the electronic boron line can be attributed to either the quality of the sample or to the wavelength of the Krypton laser. With

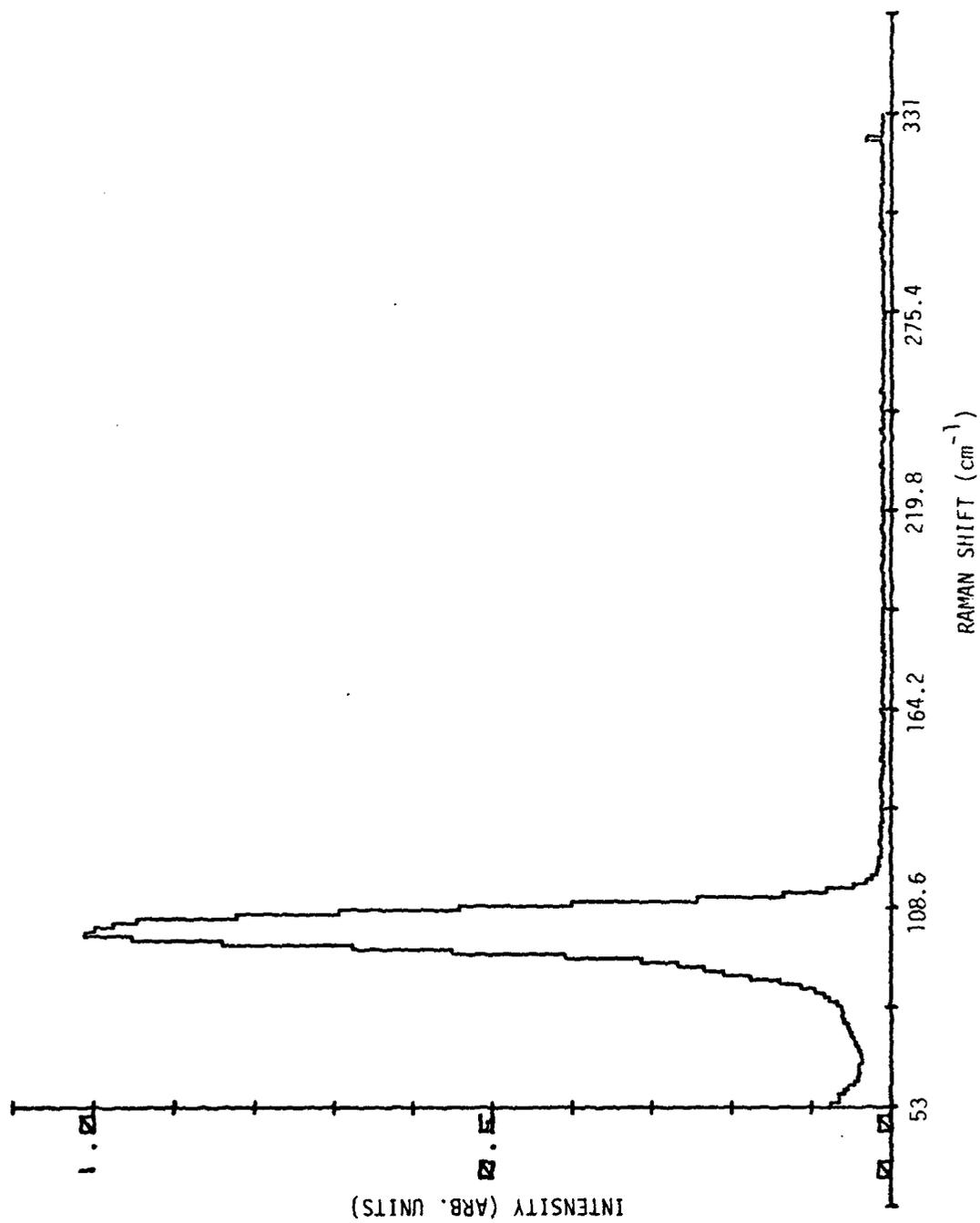
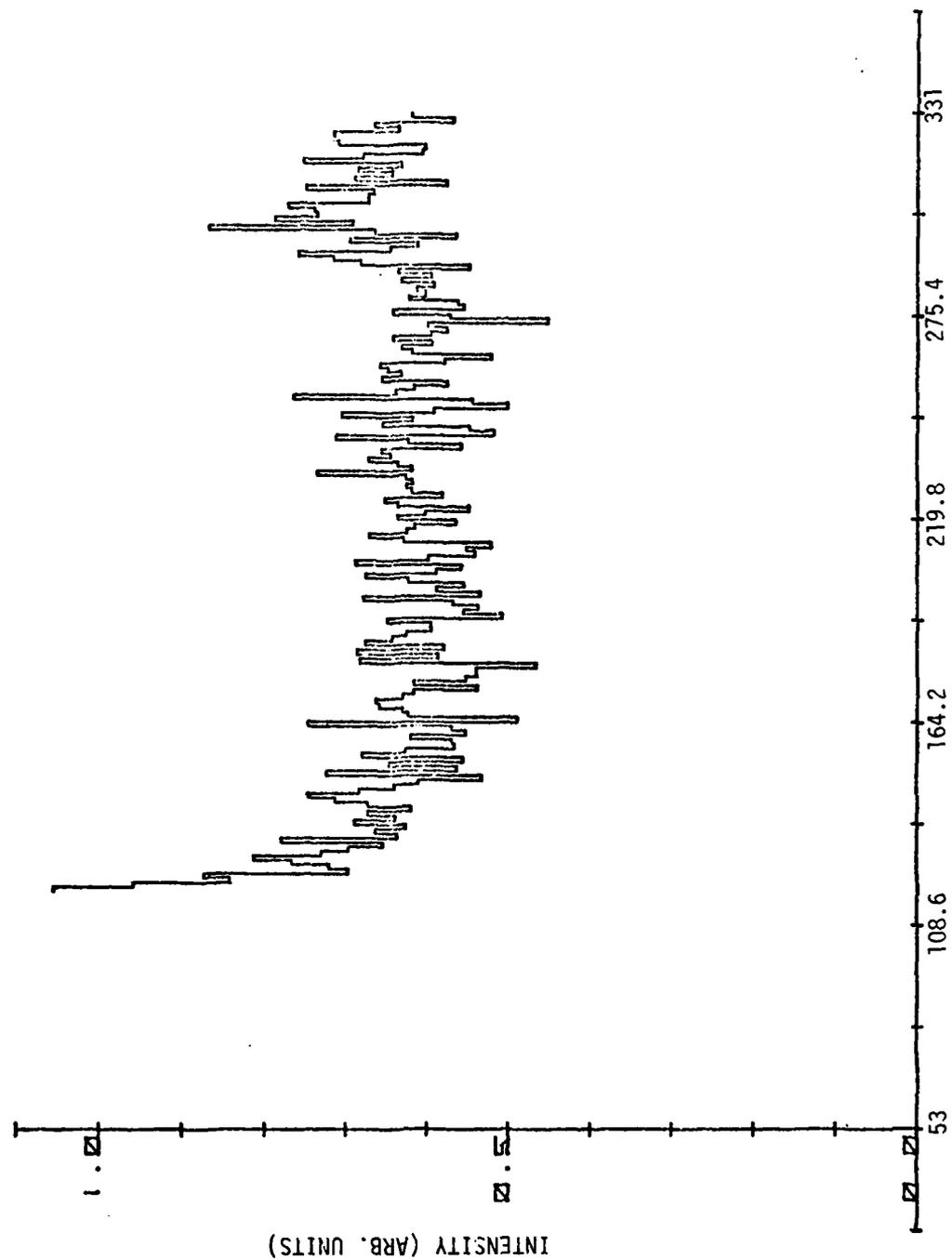


Figure 1. Silicon with $1.6 \times 10^{17} \text{ B cm}^{-3}$ at 4.2K.



RAMAN SHIFT (cm⁻¹)

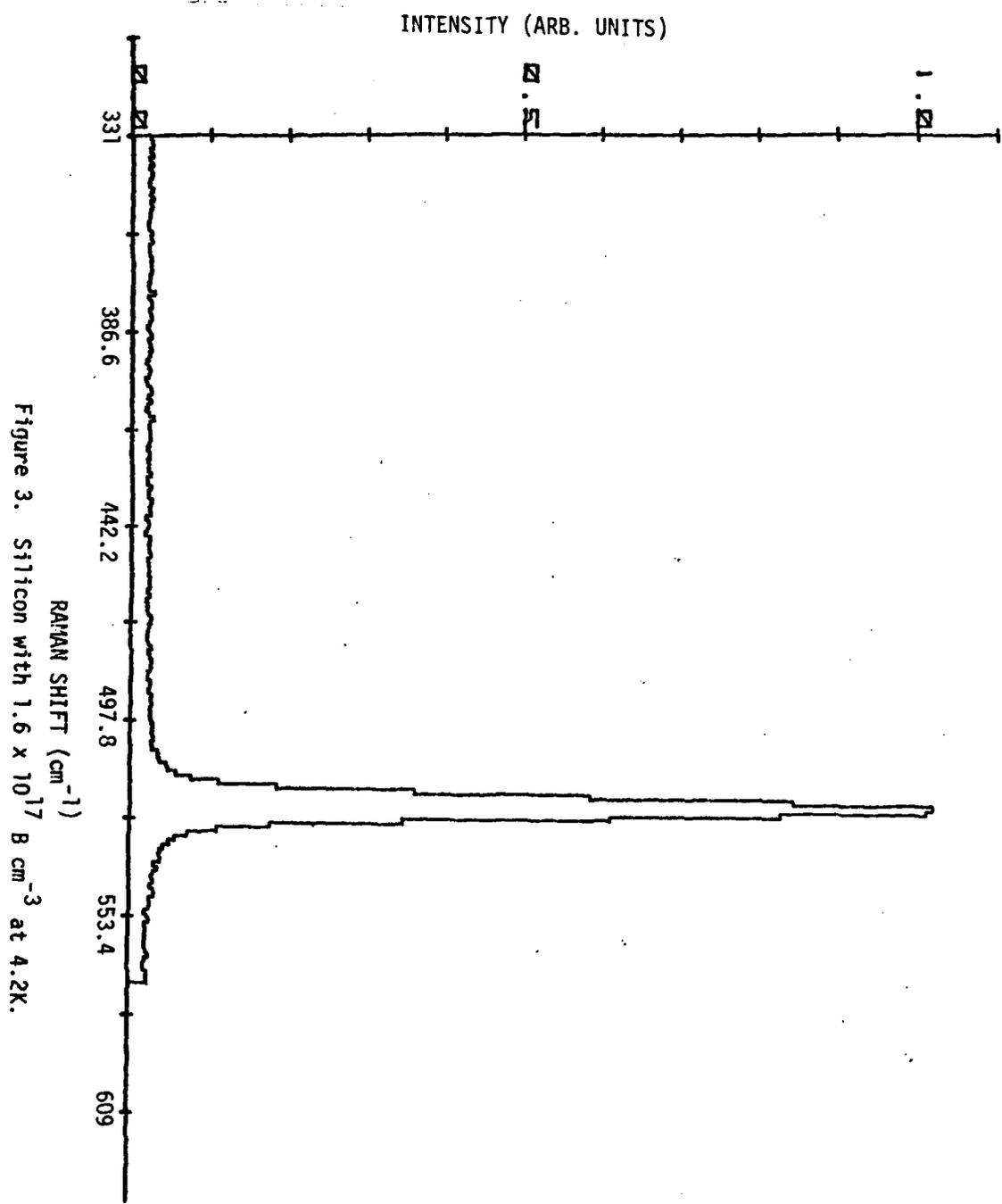
Figure 2. Silicon with 1.6×10^{17} B cm⁻³ at 4.2K.

regard to the quality of the sample, it was discovered after using this sample that the boron had been introduced by ion implantation which was followed by a float zone treatment to distribute the boron. Therefore, it is possible that the boron is not uniformly distributed throughout the sample. With respect to the Krypton laser, it should be pointed out that previous researchers observed the boron line using Nd:YAG lasers operating at 1.06μ . We are not aware of any researcher who has observed the boron line using a Krypton laser.

The zone centered optical phonon line located at approximately 520 cm^{-1} is shown in Figure 3. This spectrum was obtained at liquid helium temperature and is indicative of the sensitivity and line position accuracy of the Raman system we used.

The Raman spectrum for Si(In) obtained at 4.2K is shown in Figure 4. The lines at approximately 111 and 134 cm^{-1} have not been identified. The line at 304 cm^{-1} is a previously reported^{1,2} lattice phonon line. It is evident from the spectrum, that no electronic line exists in the spectral region around 187 cm^{-1} . Depending on which Group IIIA element is the dopant, it is possible that the boron-type electronic Raman line would be shifted to higher or lower frequency. At this time, we think any shift would be observable in the spectral region covered by Figure 4.

The spectrum for Si(B) at 77k is shown in Figure 5. The step shoulder is due to the 6471 \AA krypton line. The features of interest in this spectrum are three distinct lines located at 77 , 138 , 233 cm^{-1} , and the broad feature centered at approximately 190 cm^{-1} . This feature was also observed in several other Si spectra obtained at 77K. Even though the feature at 190 cm^{-1} is close to the location of the boron electronic line (187.2 cm^{-1}), the relatively high sample temperature would seem to rule out the possibility of this feature being produced by an electronic transition.



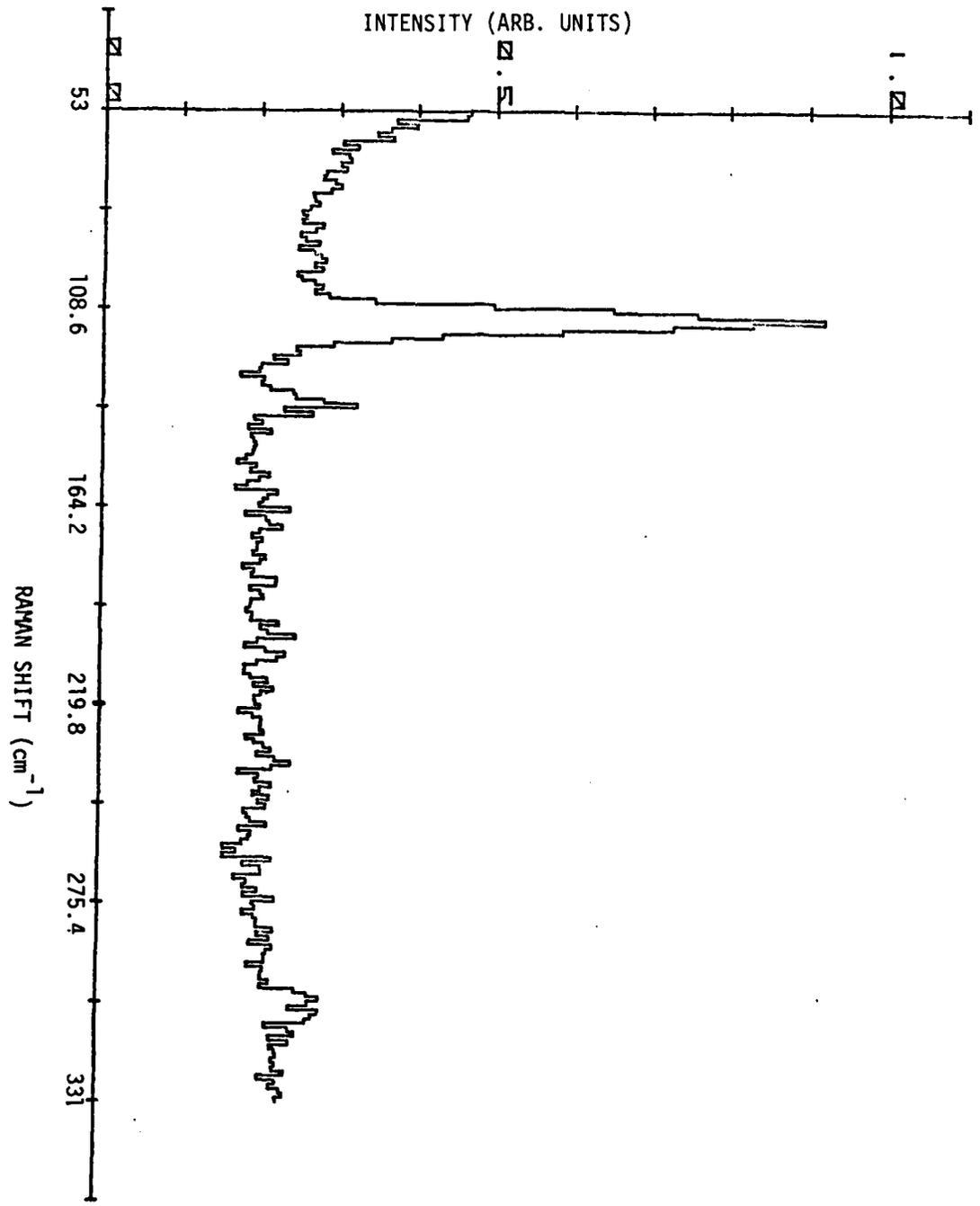


Figure 4. Silicon with 5×10^{16} In cm^{-3} at 4.2K.

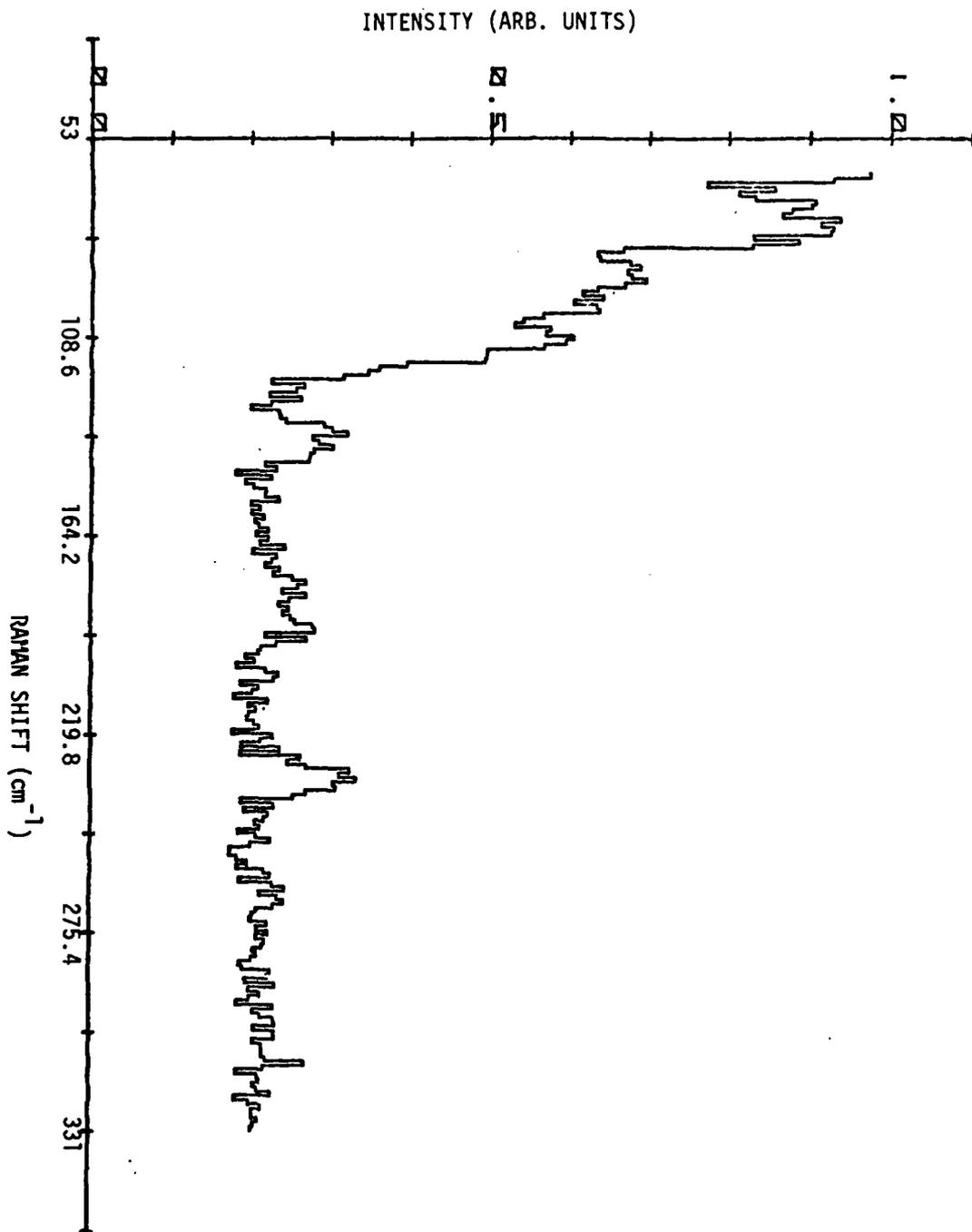


Figure 5. Silicon with 1.6×10^{17} B cm⁻³ at 77K.

The spectral region from 153 to 253 cm^{-1} for the same Si(B) sample at a temperature higher than 77K is shown in Figure 6. The intensity of the line at 233 cm^{-1} has greatly increased with the increased temperature of the sample. At room temperature, the intensity of this line has significantly decreased as can be seen in Figure 7. Previous researchers, such as Uchinokura et. al.¹ and Temple et. al.² have reported a phonon line at approximately 230 cm^{-1} , but they do not report the type of temperature dependence that we have observed. We have also observed a similar line at 233 cm^{-1} in several additional samples doped with other Group IIIA elements. We feel that a lattice line does exist in this region, but at present we cannot explain the temperature behavior that we have observed for this line. (See Addendum)

In order to avoid a lengthy presentation of spectra, we have summarized all lines that we have observed in silicon doped with Group IIIA elements into Table I.

Part (2) NTD Silicon (gallium)

The infrared absorption study of NTD silicon (gallium) is yielding results that for the most part are in good agreement with the previous studies by Covington¹⁶ and Watson and Covington¹⁷. The divancies in the spectral region around 3000 cm^{-1} have been observed prior to any annealing. After the 300°C, anneal essentially all of the divancies lines have been removed from the spectrum. After the 500°C anneal, gallium lines were detected in the sample containing 1×10^{14} P cm^{-3} . At present, spectra are being collected for the samples after annealing at 600°C. A more detailed final report will be submitted in the near future by the principal investigator John Kainer.

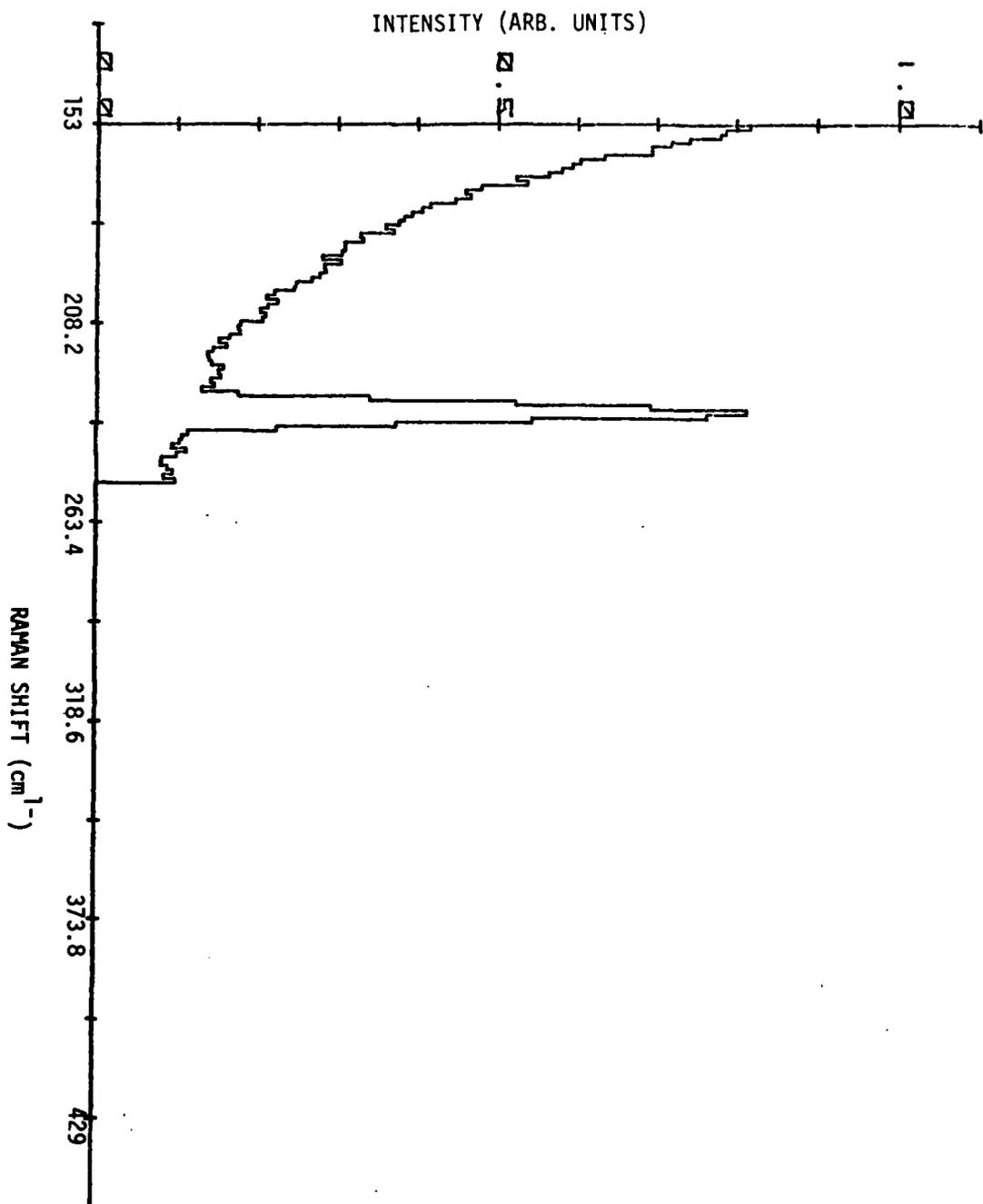


Figure 6. Silicon with 1.6×10^{17} B cm⁻³ at $T > 77$ K.

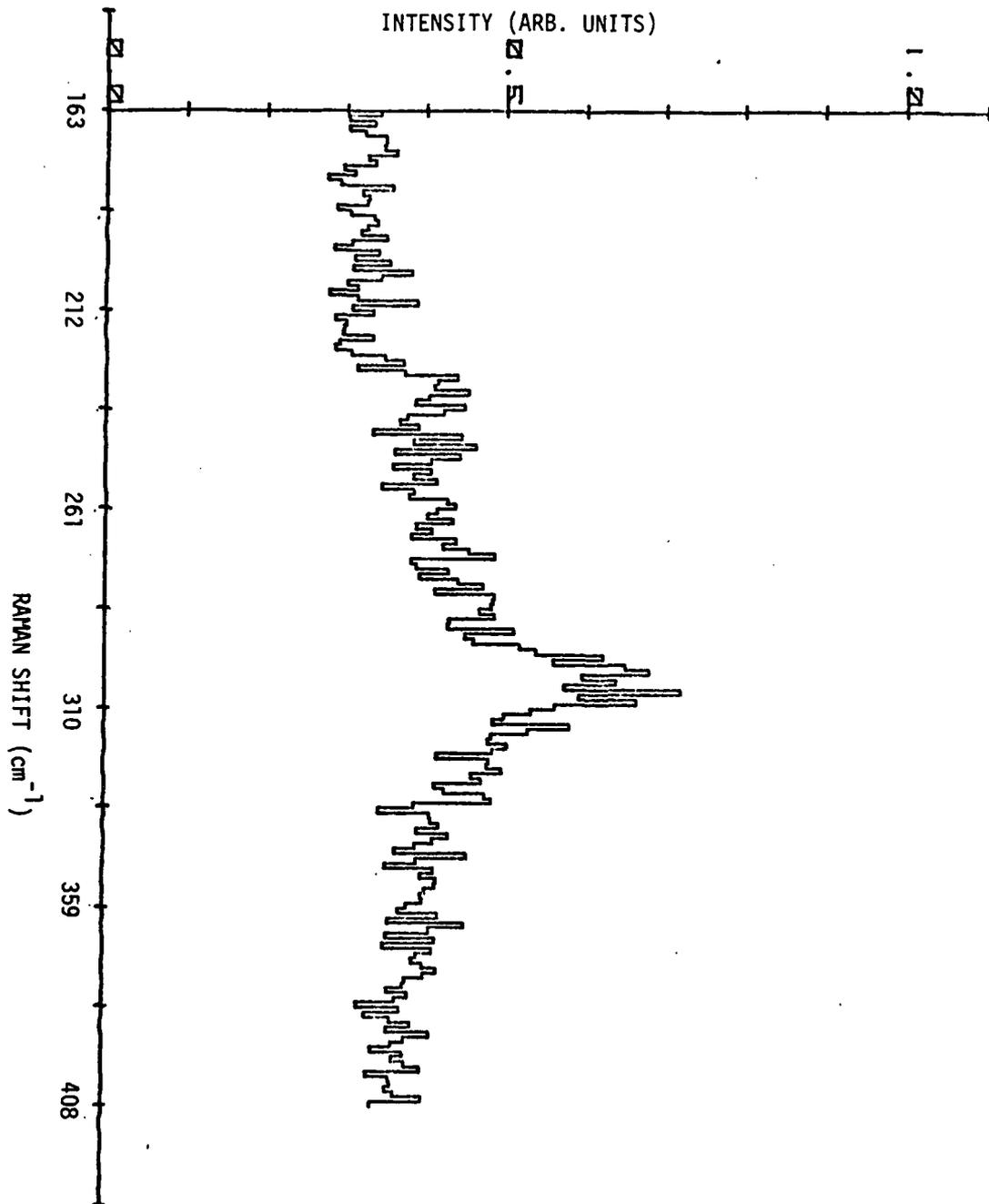


Figure 7. Silicon with 1.6×10^{17} B cm⁻³ at $T \approx 300$ K.

Table I.
Raman Shifted Lines (cm^{-1}) in Extrinsic P-Type Silicon

B	A1	Ga	In	Ti	Pure Si
<u>300K</u>	<u>300K</u>	<u>300K</u>	<u>300K</u>	<u>300K</u>	<u>300K</u>
77K	73	233	232	98	92
77	93	295	301	230	186
103	303	520	520	301	240
137	521			521	303
190					520
233					
520					
<u>4K</u>			<u>4K</u>		
102			96		
233			111		
304			134		
520			304		
			522		

V. RECOMMENDATIONS

Part (1) Raman Study

We feel two research approaches should be taken with respect to silicon doped with Group IIIA elements.

First, the room and liquid helium temperature spectral characterization that has been initiated should be completed using the krypton laser operating at $6471 \overset{0}{\text{Å}}$. This study should include the effects of different types of surface preparation upon the Raman spectrum. In addition, several samples should be investigated using an Argon laser operating at $4880 \overset{0}{\text{Å}}$. The comparison of the Krypton and Argon Raman spectra should determine which of these lasers are best suited for studying silicon.

Secondly, a Raman study of silicon doped with Group IIIA elements should be initiated using a Nd:YAG laser operating at 1.06μ . Previously reported studies^{7-9,11}, indicate that this particular wavelength is advantageous for observation of electronic Raman scattering. We feel that this particular study has the greatest probability for observing the boron-type electronic transition in silicon doped with gallium, aluminum, indium, or thallium.

Part (2) NTD Silicon (gallium)

We feel that the investigation of NTD silicon (gallium) should be expanded to include samples containing varying concentrations of gallium and phosphorus. A study of this type would provide additional information about the interactions taking place between the gallium and phosphorus atoms as well as the interactions between the impurities, the radiation induced defects, and the lattice. This type of information is essential to the continuing effort to improve detector grade silicon.

Addendum

We experimentally measured the position of the krypton laser lines that occur in the spectral region of interest to this study. With respect to 6471 Å line, we found three lines located at 93, 233, and 313 cm^{-1} . This compares favorably with the positions of 92.6, 232.8, and 308.5 cm^{-1} reported by Weber¹⁸.

We feel that it is very probable that the 93 and 233 cm^{-1} lines we have observed in the silicon spectra are due to these krypton lines. At this point, we still do not understand the temperature dependence observed for the 233 cm^{-1} line.

REFERENCES

1. K. Uchinokura, T. Sekine, and E. Matsuura, *J. Phys. Chem. Solids*, 35, 171 (1974).
2. Paul A. Temple and C. E. Hathaway, *Phys. Rev. B*, 7 (8), 3685 (1973).
3. H. Wendel, *Solid State Communications*, 31, 423 (1979).
4. K. Arya, M. A. Kanehisa, M. Jouanne, K. P. Jain, and M. Balkanski, *J. Phys C: Solid State Physics*, 12, 3843 (1979).
5. Raphael Tsu and Jesus Ganzalea Hernandez, *Appl. Phys. Lett.*, 41 (11), 1016 (1982).
6. M. Chandrasekhar, U. Rossler, and M. Cardona, *Phys. Rev. B*, 22 (2), 761 (1980).
7. G. B. Wright and A. Mooradian, *Phys. Rev. Lett.*, 18 (15), 608 (1967).
8. G. B. Wright and A. Mooradian, *Proceedings of the 9th Int. Conf. on the Physics of Semiconductors, Moscow, 1968*, pg. 1067.
9. J. M. Cherlow, R. L. Aggarwal, and B. Lax. *Phys. Rev. B*, 7 (10), 4547 (1973).
10. H. R. Chandrasekhar, A. K. Ramdas, and Sergio Rodriguez, *Phys. Rev. B*, 12 (12), 5780 (1975).
11. Kanti Jain, Shui Lai, and Miles V. Klein, *Phys. Rev. B*, 13 (12) 5448 (1976).
12. M. V. Klein, *Light Scattering in Solids*, edited by M. Cardana (Springer - Berlag, HeideTber, 1975).
13. G.B. Wright and A. Mooradian - private communication.
14. K. Lark-Horovitz, *Semiconducting Materials*, Edited by H. K. Henisch (Butter worths Scientific Publications, Ltd., London. 1951) pg. 47.
15. M. Tanenbaum and J. Mills, *J. Electrochem. Soc.*, 108, 171 (1961).
16. Billy C. Covington, Final Report, 1980 USAF - SCEEE Summer Faculty Research Program.
17. Jeffrey B. Watson and B. C. Covington, *Proceedings of the Fourth International Neutron Transmutation Doping conference, June, 1982* (Plenum Press).
18. *CRC Handbook of Laser Science and Technology, Vol. II - Gas Lasers*, Ed. Marvin J. Weber, CRC Press Incorp., Boca Raton, FL, 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

HUMAN FACTORS COMPARISON OF TOUCH SCREEN AND VOICE COMMAND DATA

ENTRY ON A COMMAND, CONTROL, AND COMMUNICATIONS SYSTEM

Prepared by: Peter M. Crane, Ph.D.
Academic Rank: Assistant Professor
Department and University: Department of Psychology
University of Pittsburgh at Johnstown
Research Location: Air Force Aerospace Medical Research Laboratory
Human Engineering Division
Technology Development Branch
USAF Research: Capt. David Leupp
Date: September 9, 1983
Contract No.: F49620-82-C-0035

HUMAN FACTORS COMPARISON OF TOUCH SCREEN AND

VOICE COMMAND DATA ENTRY ON A COMMAND,

CONTROL, AND COMMUNICATIONS SYSTEM

by

Peter M. Crane

ABSTRACT

Two data entry devices, touch sensitive screens and voice command, were evaluated on a command, communications, and control (C³) simulator. The simulator incorporated color coded, interactive graphics and required subjects to acknowledge incoming messages, interpret them, and send out summary reports. Data were collected on task performance from eight subjects first using touch screen input and then voice command input. Subject preferences were also collected. Results show that both systems worked well. The only consistent problem observed was that one word was frequently rejected by the voice recognition system. Subjects expressed a slight preference for voice input but both systems were acceptable. It is recommended that both input systems be incorporated into the C³ simulator after the vocabulary set and screen positions are optimized.

Acknowledgement

I would like to thank the Air Force Systems Command, Air Force Office of Scientific Research, and the Southeastern Center for Electrical Engineering for supporting this work. I would also like to thank the Air Force Aerospace Medical Research Laboratory, Wright-Patterson AFB, Ohio and particularly, the Human Engineering Division for providing me with a rewarding and productive summer. I express particular thanks to Capt. David Leupp, Capt. Richard Poturalski, Mr. Walter Summers, and Ms. Suzanne Kelly for their help.

I. INTRODUCTION

In order to effectively direct and deploy forces, commanders must have access to a wide variety of information. The function of a command, control and communications (C³) system is to provide this information quickly and to provide it in a way that is useful to people making decisions. A complete C³ system includes: the sensing devices, the hardware and software necessary to transmit and present the information, the system operator, and the decision makers. The operator's primary task in this high-information environment is to receive information, evaluate it, and to choose an appropriate response. A secondary task is to operate the C³ system per se quickly, accurately, and with a minimum of training. To assess how emerging technology may aid C³ systems, the Aerospace Medical Research Laboratory has developed a Simulator for C³ Operator Performance Evaluation (SIMCOPE) in which the operator is to simulate the duties of a warning officer at an air defense facility; the officer's task is to monitor incoming data for missile launches and to send out reports on these events. The new technologies incorporated into this simulator include interactive graphics, color coded displays, voice command data entry, and touch sensitive screens.

Advanced technology simulators such as SIMCOPE allow at least four levels of human performance problems to be studied under controlled conditions. The most complex problems involve the quality of the operator's decisions (e.g., hit and false alarm rates) when all relevant information has been gathered correctly. Variables such as ambiguous or conflicting data would be appropriate for studying these issues. A second level of problems involves incomplete or incorrect collection of data which the operator fails to detect. A third level concerns errors in system operation which are detected and require backtracking and re-entry of data. The fourth level of problems which can be studied using SIMCOPE involves system configurations which discourage errors but are slow or tedious to use.

Mountford, North, Metz, and Warner¹ describe an example of how the third and fourth level of problems can be studied. Mountford et al developed several dialogue alternatives for a voice data entry task. The dialogues encouraged either rapid entry by using brief prompts and little feedback or error-free data entry by using more explicit prompts and feedback messages. It was expected that the briefer format would be completed more quickly but prone to errors. It was found, however, that with little training, error

rates were not significantly different and that subjects strongly preferred the briefer dialogue format.

The present experiment investigated problems at this level using the SIMCOPE task comparing touch screen and voice command data entry. Both touch-screen and voice-command input modes are expected to encourage rapid, error-free, and congenial operation, but like Mountford et al's dialogue alternatives, they should be evaluated.

II. BACKGROUND

Dunsmore² reviews a number of studies which compared newly developed input systems to the traditional QWERTY keyboard. In general, data entry was faster and less prone to errors when a limited menu was used with a positioning device such as a lightpen, joystick, or mouse than if data had to be typed from a keyboard. English, Englebart, and Berman³ found the lightpen superior to a joystick in that the lightpen utilized a more practiced response of pointing. One problem, though, was that extended sessions with a lightpen produced arm fatigue and discomfort.

Pfauth and Priest⁴ review the advantages and disadvantages of touch screen devices over other positioning systems. They report that touch screens provide "direct visual to tactile control, ...fast data entry and control, ...minimal training, ...(and) high operator acceptance," (p. 502). Reported disadvantages were primarily technical and included glare, parallax error, and increased programmer time. When these disadvantages can be overcome, touch screens used with menu-selection dialogues provide a powerful combination particularly for the naive user. Another positive aspect of touch sensitive screens is that since the output display and the input device are the same, considerable space can be saved in the workstation. Stammers and Bird⁵ used touch screens for menu interactive data transfer and modification with experienced air traffic controllers. All subjects learned to use the system without difficulty and were in favor of placing such devices in operational use. Although little work has been done with touch screens and interactive graphics, Beringer and Maxwell⁶ expect improved performance over other input devices.

More extensive literature is available for voice data entry. Dunsmore² and Cochran, Riley and Stewart⁷ point out that voice command systems are expensive, have limited vocabularies, require extensive training and are

sensitive to ambient noise. However, where feasible, voice entry allows hands-off operation, simultaneous use of eyes, hands, and voice, and simple data entry. Funk and McDowell⁸ recommend that voice control be used only where discrete commands are sufficient and some aid to short term memory is available. Mountford and North⁹ found that keyboard data entry significantly degraded performance on a simultaneous tracking task but that voice entry did not. Finally, Gould, Conti and Hovanyczyk¹⁰ report very high user acceptance for a simulated listening typewriter in which a person would dictate text and have it appear on a screen.

III. OBJECTIVES

The present study compared operator performance on a simplified version of the SIMCOPE task using two data entry techniques -- touch screen and voice command. In relation to the SIMCOPE task, both touch screens and voice command data entry systems offer similar advantages. Since the operator's task involves input of short strings of data and menu manipulation rather than continuous tracking or text entry, both systems seem appropriate. Touch screens will place more load on short-term memory while the limitations of voice recognition technology may prove frustrating. The present study investigated how the operator can best interact with this system; the objective was to evaluate how well operators can use touch-screen or voice-command input in terms of data entry rate, response time, and user preference.

IV. EXPERIMENTAL PLAN

a. Equipment. The SIMCOPE console at AMRL consists of two, color video monitors with touch sensitive overlays (TSD Display Products, Model TF). A Threshold 600 voice recognition unit (Threshold Technologies, Inc., Deltran, NJ) was used for voice data entry with a headset mounted, close-talking microphone (Shure Model SM10A). A PDP 11/34 minicomputer controlled the simulator and collected data.

b. Subjects. The subjects were paid volunteers, five males and three females between 22 and 29 years old. All subjects had participated in other research at AFAMRL but none reported any previous experience with touch sensitive displays or computer voice recognition.

c. Procedure. Subjects were tested individually in four separate sessions. All sessions lasted from one hour to 90 minutes and were conducted

on different days. The first session involved orientation to the experiment and training on the simulator using touch input. The second session consisted of two 25-30 minute periods operating the console using touch input only separated by a five minute break. After the second period, subjects filled out a brief questionnaire (described below). In the third session, subjects trained the voice recognition unit to respond to their voices by speaking each command ten times into the microphone. The training data was saved on tape and reloaded at the start of subsequent sessions. The fourth session consisted of two 25 minute periods operating the console using voice data input. After the second period, subjects filled out a brief questionnaire.

d. Subject's Task. The following instructions were read to the subjects.

The basis of this simulation is two hypothetical countries, US and THEM. You are the Warning Officer on duty at the Command Warning Center facility in US. Your job is to monitor the data coming in from two sensor systems. The Advanced Detection System (ADS) and the Barrier Surveillance System (BSS). The ADS system can provide information about possible missile launches quickly but with very little detail. The BSS system provides much more information, but it takes longer to get that information. Every time information comes in from one of these systems, your job is to evaluate it and then to send a report to Command Defense Center concerning the event. Since a missile might be on the way, it is important that the report be completed quickly and accurately.

Look at the screen in front of you. Across the bottom of the screen are a number of areas containing messages such as DETAIL MAP and OUTPUT FORMAT. Activating one of these by touching the screen will bring that page of information into the area on the upper left part of the screen. Try it with the DETAIL MAP. Look at the map. First you will see US on the left and THEM on the right. Within US you will see nine cities plus several other potential targets. Within THEM there are five geographic regions. There are known launch sites marked with white triangles and suspected launch sites marked with open, green triangles. The two large circles on THEM indicates the coverage areas of the two ADS systems: North and South. The smaller circles along the coast of US are the coverage areas of the BSS stations.

An EVENT is the detection of something of interest by the ADS or BSS systems; i.e., a potential launch. When an event occurs, you will hear a beeping sound and see a flashing message on your screens telling

you what type of event is occurring.

Subjects were further instructed that their job was to fill out three different reports on incoming events; these reports were called ADS 1, ADS 2, and BSS. All reports consisted of several fields to be completed by entering information from a menu or from the image of a keypad displayed on the screen. All fields could be completed from information available to the subject by looking at different pages of data on the screens; no inferences or decisions were required. In voice data entry sessions, only responses to these forms (plus the acknowledge response to a new event) were input by voice command. Other interaction with the simulator such as changing pages or assigning events required touch input.

e. Experimental Design. Data were collected on three dependent measures. The questionnaire asked subjects to report any physical discomforts experienced at the end of a session including eyestrain, back, shoulder, or arm pain, and stress. After the voice data entry session subjects also rated the ease of use of the two data entry systems and their preference.

The remaining dependent measures concerned performance on the SIMCOPE task. Incoming events generated both flashing message on the screen and an audible prompt; these events had to be acknowledged before any other task could be performed. Latency to acknowledge was therefore recorded as a baseline measure of response time for touch and voice input. Forty-eight such latencies for both touch and voice were recorded for each subject. Overall task efficiency was measured by recording time to fill out reports. To assure that fill out times reflected only differences between voice and touch data entry, and not differences among events, only two sets of events were used. Each set included eight events and required three reports per event (ADS 1, ADS 2, and BSS). Subjects responded to both sets using touch input and later using voice input. Data were analyzed using four way within subjects analyses of variance (ANOVA): 2 input modes x 3 report types x 2 sets x 8 events; events were treated as a random factor.

V. RESULTS

a. Questionnaire. After both data collection sessions, subjects were asked to describe any symptoms of back, shoulder, or arm discomfort, stress or eyestrain. A few subjects mentioned eyestrain and arm discomfort but most reported no discomfort. These data were not analyzed further.

At the end of the voice data collection session, subjects expressed their opinions about the input devices on a set of five-point attitude scales. Subjects answered the questions, "Did you find the touch screen/voice recognition system easy to use?" on a scale of 1 = very difficult, 3 = neutral, and 5 = very easy. Mean rating for touch screens was 3.6 (between Neutral and Easy) while mean rating for voice recognition was 4.1 (just above Easy); $t(7) = 1.08$, ns. When asked, "Which system was harder to use, i.e. slower, made more mistakes, or didn't do what you intended?", two subjects rated voice "somewhat harder to use," four rated touch "somewhat harder," one rated touch "much harder" and one was neutral, $M = 3.5$ (between "neutral" and "touch somewhat harder"). Finally, subjects were asked, "Which system do you prefer using?" Three subjects expressed "strong preference for voice," two "moderate preference for voice," two were neutral and only one expressed a "moderate preference for touch."

b. Latency to acknowledge and report fillout times. The distributions of these variables were highly skewed. For latency to acknowledge, skewness = 8.04 and kurtosis = 100.21; for fillout time, skewness = 3.39 and kurtosis = 20.92. A reciprocal transform was therefore employed to reduce the skew: reciprocal acknowledge time ($\times 100$) skewness = 0.67, kurtosis = 0.50; reciprocal fillout time ($\times 1000$) skewness = 0.69, kurtosis = 0.27. Four way, within-subject ANOVAs were performed on each variable with events treated as a random factor. Using the procedure described by Winer¹⁴, interactions not significant at $p = .75$ were pooled.

For latency to acknowledge there were three significant main effects and three significant two-way interactions. There was a significant difference among Subjects, $F(7, 49) = 4.24$, $p < .01$, and between Input modes, $F(1, 63) = 332.5$, $p < .01$, with the mean for touch ($M = 54.49$) being faster than the mean for voice ($M = 32.41$). There was also a significant difference among events, $F(7, 49) = 2.62$, $p < .05$. The significant interactions were: Subject by Set, $F(7, 49) = 2.77$, $p < .05$; Input mode by Set, $F(7, 49) = 8.64$, $p < .01$; and Event by Message, $F(14, 98) = 2.84$, $p < .01$.

For fill-out time, there were three significant main effects, five significant two way interactions, and one three-way interaction. These effects are listed in Table 1. It should be noted that input mode did not produce a significant main effect, however, the Subject by Input mode interaction was significant, $F(7, 49) = 5.61$, $p < .01$.

VI. DISCUSSION

The results of the ANOVAs on latency to acknowledge and fill-out time seem to be contradictory. Although not statistically significant, voice input was slightly faster than touch for report fill-out ($M_{\text{voice}} = 25.35$, $M_{\text{touch}} = 22.51$, $F(1, 7) = 5.15$, $p = .057$), but significantly slower for latency to acknowledge. An analysis of the significant interactions plus a discussion of problems with the voice recognition system will clarify the results. Since the presence of significant interactions indicates that main effects and lower order interactions are confounded, only the unconfounded, significant interactions will be discussed.

There were three significant interactions for latency to acknowledge: Event by Message type, Subject by Set, and Input mode by Set. The Event by Message type interaction is of minor importance and was not analyzed in detail. Inspection, however, shows that the slowest acknowledge time was for the first message (ADS 1) for event 1. Acknowledge times were uniformly faster after that.

For the Subject by Set interaction, simple main effects test for set for each subject showed that one subject significantly improved performance from set 1 to set 2, $F(1, 7) = 9.21$, $p < .05$. The difference in latency to acknowledge between sets was not significant for all other subjects. These two interactions suggest that practice effects for latency to acknowledge were limited to the first message for all but one subject. The Input mode by Set interaction, however, shows that for touch input performance improved from set 1 ($M = 55.6$) to set 2 ($M = 59.33$) while for voice input, latency to acknowledge slowed from set 1 ($M = 33.8$) to set 2 ($M = 30.9$).

These effects can best be explained by looking at how the two input systems operate. Both the touch screen and voice recognition systems are serial input devices. Like a keyboard, these devices convert the operator's action into ASCII code which is then sent to the host computer. Two types of errors can occur with these systems. In a misrecognition, a code different from the operator's intent is sent to the host while in a rejection, no output is sent because the device could not interpret the operator's input. Misrecognitions could produce strange results but were rare in this experiment; rejections, however, occurred frequently. A subject using touch input only had to press the screen lightly. When the system recognized an input, it immediately responded with both audible and visual prompts. A rejection was

immediately obvious because the expected prompts would not occur. At this point, the subject had to lift the finger and then touch or tap the screen again. A subject using voice input spoke into the microphone and then waited for feedback. When a rejection occurred, the voice recognition did nothing for approximately one second and then signaled the rejection with an audible prompt. The subject had to then repeat the word until it was recognized.

When responding to an incoming message, the subject spoke the word, "acknowledge". The choice of this word was unfortunate but enlightening. While input errors were not recorded, it was the experimenter's observation that the spoken word, "acknowledge" was rejected more frequently than any other input. Subjects often had to repeat the word several times before it was recognized. Since Input mode did not affect report fill-out time, it was the rejection of "acknowledge" that accounts for the slow latency to acknowledge for voice input. Also, the Input mode by Set interaction shows that practice did not improve performance for voice input. Due to the large number of rejections, no conclusion can be drawn regarding which Input mode might be faster to acknowledge a message. This portion of the experiment should be repeated using pretested vocabulary items (see Section VII, Recommendations).

There are four unconfounded interactions for fill-out time. The Set by Event by Message and Subject by Message interactions are of minor importance. Both interactions reflect differences in the time taken to fill out ADS 1, ADS 2, and BSS reports. These interactions do not reflect variations due to Input mode and were not analyzed further.

The other unconfounded, significant interactions were Subject by Set and Subject by Input mode. Simple main effects tests on each subject's response times indicate that one subject showed significant improvement in fill-out time from set 1 to set 2; the other seven subjects showed no change. The most interesting effect was the Subject by Input mode interaction. Overall, voice input ($\bar{M} = 25.3$) was slightly faster than touch ($\bar{M} = 22.5$) but the difference was not significant, $F(1, 7) = 5.15$, $p = .057$. For five of the eight subjects there was no difference between the two modes; two subjects performed slightly better with touch but the difference was not significant. Only one subject showed significantly better performance for voice, $F(1, 7) = 9.21$, $p < .05$. (It should be noted that the three interactions which were significant for only one subject involved three different subjects.)

The objective of the present study was to compare touchscreen and voice

command input modes on a C³ task. It was expected that both devices would show similar advantages and no hypothesis was made regarding which would be superior. The results showed that voice input was subject to recognition errors for some commands but otherwise neither input mode provided superior performance. The SIMCOPE task is well suited for both input systems. The task requires rapid manipulation of menus and inputting short strings of data; it is not expected that C³ operators would be experienced typists. For this type of task and subject population, voice command and touch screen input should provide significant advantages over more traditional data entry devices.

VII. RECOMMENDATIONS

The results of the present research should be considered when designing future SIMCOPE studies. Three items which should be investigated are: the use of touch and voice by SIMCOPE operators, b) changing the vocabulary set to maximize voice recognition reliability, c) investigating the position of the touch screens to minimize fatigue.

a. Voice and Touch Input. Although the majority of subjects in this experiment preferred voice over touch input the difference is neither strong nor unanimous. One subject said that voice input required her to stay in one position so she could see both screens and that this static work posture was annoying. Subjects found voice entry occasionally frustrating when a word would be rejected several times. I see no reason, therefore, why one system should be preferred. Future SIMCOPE researchers may want to allow overcapping voice and touch input particularly for longer sessions in which a single input mode could produce boredom or fatigue.

b. Vocabulary Set. The voice recognition system can produce two types of errors, misrecognitions and rejections. When a rejection occurs the system was unable to match the utterance to a template and signals the operator with an audible prompt. On a misrecognition, the system matches an utterance to a similar but unintended response, e.g., "known" for "no." In the present study, this generally had no effect on the task but occasionally produced unexpected responses such as entering the "edit" mode (not used in this study) when the subject said "send". Although both types of errors occurred in this study, neither was recorded. French¹¹ presents a table listing correct recognitions, misrecognitions, and rejections for a 50 word vocabulary set. Future SIMCOPE researchers may wish to collect similar data for the SIMCOPE

vocabulary set and then perform a factor analysis or cluster analysis to determine the characteristics of commands which produce errors. Difficult items could then be changed or eliminated to improve task performance. As an example, the word "acknowledge" was troublesome for many subjects; a synonym or a unique pronunciation would have made their lives easier.

c. Touch Screen Position. As SIMCOPE is presently configured, the touch sensitive screens are placed vertically with the top edges at eye level approximately 16-18" from the operator's eyes. Although few subjects in the present study complained of hand or arm fatigue this has been mentioned in the literature by Pfauth and Priest⁴; Beringer and Maxwell⁶ cite arm fatigue as a particular problem with vertical screens. While design guidelines exist for placement of video display screens (e.g., Cakir, Hart, and Stewart¹²) no standards are available for touch sensitive screens. The variables of screen height relative to the eyes, declination angle, and distance from the operator should be manipulated to minimize fatigue in the hand, arm, shoulder, and neck. There may not be a single solution to this problem in that a screen at eye level and a high declination angle may be easy to view but produce arm and shoulder fatigue; a lower screen with decreased declination may produce less arm fatigue but cause neck and upper back discomfort. Unfortunately, task performance may not reflect operator discomfort (Beringer and Maxwell⁶). Future research, therefore, should employ more sensitive measures of stress and fatigue than performance rate and self-reports of discomfort. Physiological measures of muscle tension and detailed analysis of response times may provide the best indicators of how touch sensitive screens should be placed.

REFERENCES

1. Mountford, S. J., North, R. A., Metz, S. V. and Warner, N. Methodology for exploring voice-interactive avionics task: Optimizing interactive dialogues. Proceedings of the Human Factors Society - 26th Annual Meeting, 1982, pp. 207-211.
2. Dunsmore, H. E. Data entry. In: Kantowitz, B. H. & Sorkin, R. D., Human Factors: Understanding People-System Relationships. New York: John Wiley & Sons, 1983.
3. English, W. K., Englebart, D. C. & Berman, M. L. Display-selection techniques for text manipulation. IEEE Transactions on Human Factors in Electronics, March 1967, Vol. 8, pp. 5-15.
4. Pfauth, M. & Priest, J. Person-computer interface using touch-screen devices. Proceedings of the Human Factors Society - 25th Annual Meeting, 1981, pp. 500-504.
5. Stammers, R. R. & Bird, J. M. Controller evaluation of a touch input air traffic data system: An "indelicate" experiment. Human Factors, 1980, Vol. 22, pp. 581-589.
6. Beringer, D. B. & Maxwell, S. R. The use of touch-sensitive human-computer interfaces: Behavioral and design implications. Proceedings of the Human Factors Society - 26th Annual Meeting, 1982, p. 435.
7. Cochran, D. J., Riley, M. W. & Stewart, L. A. An evaluation of the strengths, weaknesses, and uses of voice input devices. Proceedings of the Human Factors Society - 24th Annual Meeting, 1980, pp. 190-194.
8. Funk, K. & McDowell, E. Voice input/output in perspective. Proceedings of the Human Factors Society - 26th Annual Meeting, 1982, pp. 281-222.
9. Mountford, S. J. & North, R. A. Voice entry for reducing pilot workload. Proceedings of the Human Factors Society - 24th Annual Meeting, 1980, pp. 185-189.
10. Gould, J. D., Conti, J. & Hovanczyk, T. Composing letters with a simulated listening typewriter. Proceedings of the Human Factors Society - 25th Annual Meeting, 1981, pp. 505-508.
11. French, B. A. Some effects of stress on users of a voice recognition system: A preliminary inquiry. Master's Thesis, Naval Post Graduate School, Monterey, CA, 1983.
12. Cakir, A., Hart, D. J. & Stewart, I. F. M. Visual Display Terminals. New York: Wiley-Interscience, 1979.
14. Winer, B. J. Statistical Principles in Experimental Design, 2nd Edition. New York: McGraw-Hill, 1971.

Table 1
Summary of Significant Effects for Report Fill-out Time

Source	Sum of Squares	df	Mean Square	F (or F')	
Subjects	10163.72	7	1451.96	13.57	p > .01
Error	5241.29	49	106.97		
Set	4456.89	1	4456.89	17.89	p > .01
Error (Quasi F Ratio)		10 (Est.)	249.73		
Message	7484.81	2	3742.41	14.07	p > .01
Error (Quasi F Ratio)		19 (Est.)	266.05		
Subject x Input Mode	2095.15	7	299.31	5.61	p > .01
Error	2614.76	49	53.36		
Subject x Set	876.24	7	125.18	3.04	p > .01
Error	2017.06	49	41.16		
Set x Event	1157.52	7	165.36	4.02	p > .01
Error	2017.06	49	41.16		
Subject x Message	2627.57	14	187.68	3.81	p > .01
Error	4826.19	98	49.25		
Event x Message	1786.77	14	127.62	2.59	p > .01
Error	4826.19	98	49.25		
Set x Event x Message	1634.81	14	116.77	3.31	p > .01
Error	3952.53	112	35.29		

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

PERFORMANCE ANALYSIS AND EVALUATION IN A LOCAL AREA NETWORK

Prepared by: Dr. Carolyn J. Crouch

Academic Rank: Associate Professor

Department and University: Department of Computer Science
The University of Alabama

Research Location: Computer Sciences Directorate
Armament Division
Air Force Systems Command

USAF Researcher: Dr. Robert N. Braswell

Date: July 22, 1983

Contract No.: F49620-82-C-0035

PERFORMANCE ANALYSIS AND EVALUATION

IN A LOCAL AREA NETWORK

by

Carolyn J. Crouch

ABSTRACT

Performance evaluation is of particular concern in a local area network composed of nonhomogeneous mainframe nodes for which only perfunctory performance evaluation statistics are known. This report identifies performance measures that should be applied in such circumstances, discusses modes of implementation, and investigates approaches to and implications of performance evaluation in local area networks. Recommendations are made for future research.

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 her with the opportunity to research an important and interesting topic in a "real world" environment.

Special thanks must be rendered to the Computer Sciences Directorate, Armament Division, at Eglin Air Force Base and to Dr. Bob Braswell, its director, for his suggestion of the research topic, his efforts in providing a congenial atmosphere, and his support.

The author also wishes to express her appreciation to the staff of the Eglin Technical Library, whose excellent aid in obtaining desired resources was essential to the success of this project.

I. INTRODUCTION

In order to measure system performance, tune a system to its particular environment, or make decisions about the possible reconfiguration of a system as its workload grows or changes, various data items must be collected. These data are the parameters of the measurement and evaluation algorithms which must be utilized in the complex timesharing environments of current systems in order to measure system performance effectively.

One of the basic principles of performance monitoring is that measurements and evaluations should be made throughout the life of the system, beginning with its initial design phases, and continuing with installation and throughout its daily processing activities. Appropriate measurement tools must be provided through proper system design to be most effective; i.e., to generate the data desired with the minimum cost in terms of system interference and/or degradation. Whereas system performance measurement techniques are incorporated as a design feature in many systems of recent vintage, older systems may possess only a subset of the desired features or lack such features altogether. Interconnecting such disparate systems via various topologies introduces an additional level of complexity; the measurement and evaluation of network performance now becomes increasingly difficult, due not only to the additional complexity introduced by networking but also to an inability to evaluate component node performance satisfactorily.

There are three common purposes for performance evaluation, namely:

1. Selection evaluation - deciding whether a particular system or system component, sold by a particular vendor, is appropriate.
2. Performance projection - estimating the performance of a system (or a hardware or software component of that system) that does not currently exist.
3. Performance monitoring - accumulating performance data on an existing system or component, insuring that a system is meeting its performance goals, aiding in estimating the impact of planned changes, and/or providing the data needed by management to make "strategic" decisions [1].

The emphasis of this discussion will be evaluation for the purpose of performance monitoring (item 3, above).

Consider now the two methods of performance evaluation as defined by Svobodova [2]: analytic and comparative. Whereas, as the name implies, the comparative method is used to evaluate the performance of one system relative to another, an analytic evaluation is made when the performance of a computer system is evaluated with respect to the various system parameters and system workload. Its purpose includes: (1) improving performance of an existing system (i.e., performance tuning) (2) maintaining performance of an operational system within specified limits (performance control), and (3) the design and implementation of new systems. For the purpose of this discussion, the emphasis is on performance tuning and control.

II. OBJECTIVES

The purposes of this paper are: (1) to analyze the performance evaluation techniques currently being used within a specific local area network (ECONFT); (2) to investigate alternate approaches to performance evaluation in terms of measures, tools, and methods; and (3) to recommend, on the basis of the aforementioned analysis and investigations, guidelines or approaches to performance evaluation which are appropriate for the current network topology and utilizable over time as the network structure and topology change.

To accomplish these objectives, the following tasks are undertaken:

1. An overview of the techniques which are used for performance monitoring, tuning, and control, as discussed within the literature, is presented (Section III).

2. Performance evaluation measures which have been found to be of value in similar systems, or which have been suggested and/or defined within the literature, are reported (Section IV). These measures are compared with those currently being used to evaluate specific ECONET components (see Section VII).

3. The contributions of software and hardware monitors, in terms of the types of data which can be measured, their advantages, and their disadvantages, are discussed (Section VI).

4. Since performance evaluation techniques appropriate to the individual components have been surveyed, the discussion is expanded to include performance evaluation in local area networks (LANs). Investigations of approaches which are currently being utilized are reported (Section VII).

5. Conclusions are drawn and recommendations made (in view of Sections III - VII, above) in terms of the current methodology of performance evaluation and its applicability to ECONET (Section VIII).

III. PERFORMANCE EVALUATION TECHNIQUES

"Performance of a system can be discussed only in the context of what the system is required to do." [2] Thus performance tuning and control are possible only in environments where the workload of the

system can be analyzed so that an accurate reflection of system activity can be obtained. Since workload fluctuates constantly, over both short intervals (in an interactive environment) and longer time periods (as the user population and the nature of the job mix change), continual monitoring of the system is required.

Lucas, in his discussion of performance monitoring, presents an evaluation of the techniques which have previously been used with some degree of success for both hardware reconfiguration and software modification [3]. A relevant portion of this analysis is reproduced in Table 1. These techniques can be generally characterized as system models (e.g., analytic models), workload models (benchmarks, synthetic programs), and a combination of the two (simulation).

Table 1

Performance Monitoring techniques

	Hardware Reconfiguration	Software Modification
1. [Analytic] Models	insufficient	not applicable
2. Benchmarks	insufficient	insufficient
3. Synthetic Programs	insufficient	insufficient
4. Simulation	satisfactory	satisfactory
5. Monitors (hardware and software)	satisfactory	satisfactory

A brief overview of these techniques follows.

1. Analytic Models

Analytic models are mathematical representations of computing systems (frequently based on queueing theory). Although such models may provide some assistance in hardware analysis, particularly for a certain component (such as a disk unit), their development and revision is "tedious and time-consuming," and they are frequently incapable of accurately representing a complex hardware-software environment [3]. An excellent overview of analytic models is given by Svobodova [2], who recommends a combination of analytical and empirical modeling techniques to produce the most useful results.

2. Benchmarks

Benchmarks are programs coded in a specific language with a variety of instructions; when executed on a particular system they give indications of both hardware and software performance. Their value in a

particular environment depends largely on how well they represent the job mix or the typical workload of the system. They are of limited use in performance monitoring except as before-and-after tests to monitor performance following system changes [3].

3. Synthetic Programs

Synthetic programs (synthetic benchmarks) can be written in both assembly and high-level languages and coded to cover a wide range of activities (I/O, file, operating system, computational). Such programs have many of the disadvantages of benchmarks; their value in performance monitoring lies largely in their use as before-and-after tests relative to system modification [3].

These techniques (1-3, above) are considered by Lucas to give some assistance in the area of performance evaluation, but each is in itself insufficient for that purpose.

4. Simulation

Simulation is a technique whereby a computerized model of the system being evaluated (along with its workload) is developed. This model is then run on an existing system to provide a reflection of the system being modeled and its behavior in certain circumstances. Simulators are of two basic types: event-driven (controlled by events which are made to occur according to probability distributions) and script-driven (controlled by empirically derived data, manipulated to reflect anticipated behavior) [1]. Construction of models via simulation requires considerable expertise on the part of the evaluator, can be very expensive to run, usually produces large amounts of data which must then be further analyzed either manually or by machine, and must be carefully validated prior to use [1]. The greatest drawback of simulation is its relatively high cost; the simulation model must usually be tailored to the specific environment.

The advantages of simulation are pointed out by Svobodova; he considers it "the most general, most flexible, and most powerful technique for studying and predicting system performance . . . a real asset in system design and development." [2] However, simulation is viewed by Lucas as helpful but "not sufficient" in monitoring system performance. He concludes that it is "unlikely that one would ever construct a simulation just for the purpose of monitoring a system and making changes, whereas the hardware and software monitoring techniques [see 5, below] may be used fairly easily." [3]

5. Performance Monitoring

Performance monitoring may be defined as a method of "collecting data on the performance of an existing system." [3] Monitoring is usually used to locate bottlenecks when either reconfiguring the existing hardware or attempting to improve the execution of software. It is also useful in gathering system usage profiles [3].

Techniques for performance monitoring are generally grouped into two classes, software and hardware. Rose [4] and Svobodova [2] include in addition a third type of monitor, the hybrid.

Two basic policies govern the way in which data is normally collected for performance monitoring purposes. An event trace monitor records information at the time of a particular event, such as the initiation of a job, task creation, (main) storage allocation, I/O activity, supervisor interrupt, or job completion [4]. Sampling monitors, on the other hand, initiate data collection activities based on the signal of the real-time clock over a (usually) constant sampling period. Information collected by a typical sampling monitor would include figures on CPU and channel utilization, the number of jobs in execution, etc.[4] The large volumes of data collected via monitoring must then be passed to data reduction routines for analysis and output in appropriate forms.

(1) Software Monitors

A software monitor is a special program which is incorporated into the software of the host system. As such it competes with user jobs for system resources and usually interrupts normal system processing so that the monitor can collect the required information. In its simplest form, a software monitor may be essentially a job accounting routine.

Software monitors can generally access and record information that is available to the application program and/or the operating system; the penalty paid is the overhead (artifact) generated by the introduction of additional software which competes for resources with other processes. For more information on software monitors, see Section VI.

(2) Hardware Monitors

Hardware measurement tools can be classified as internal or external. Internal hardware tools (such as the "Monitor Call(MC)" instruction on the IBM S/370, various timing meters, counting meters, and traps) must generally be incorporated in the system at design time. An external hardware monitor, on the other hand, is defined by Svobodova as "a free-standing device that senses electronic signals in the circuitry of the measured system and processes and records them externally to the measured system." [2] Such a device is a "non-interfering observer;" it requires no assistance from the host system [2].

A simple hardware monitor consists of sensors to monitor the state of the host, a logic plugboard for performing various Boolean functions on the monitored state bits, a set of counters that record event occurrences or the time durations of various activities, and a display or recording device. More sophisticated monitors may include data comparator features, sequencers, random access memory, and associative memory [2].

Some hardware monitors are programmable and include a mini- or microcomputer for more elaborate processing. For further information, see Section VI.

(3) Hybrid Monitors

Hybrid monitors combine the methods of both software and hardware monitors; "measurement information . . . is moved by the system software routines into special monitoring registers and is then measured by hardware monitors." [4] Such an approach may depend on the presence of specialized hardware for the monitoring registers, which is infeasible unless such hardware is included in the original design of the machine. As a result, such monitors are seldom used [4]. The presence of external rather than internal monitoring instrumentation makes this approach feasible, but it is necessarily restricted in scope by limits on the amount and complexity of information (per event) that it can process [2].

IV. BASIC PERFORMANCE MEASURES

The term "performance" in itself is a relative term, meaning in general "the manner in which or the efficiency with which a computer system meets its goals." Whereas some aspects of performance are not quantifiable (such as "ease of system use"), others, such as turnaround time, response time, and system reaction time (as defined in Table 2), are commonly recognized quantifiable measures of performance [1].

Performance evaluation starts with the selection of a proper set of parameters upon which the evaluation is to be based, i.e., the performance measures. Svobodova describes performance measures in terms of effectiveness (those indicating the capability to process a given workload and meet the time requirements of individual users) and efficiency (which is measured by internal delays and utilization of individual system components vs. demand) [2]. The former are the "prime performance measures" in his view; examples of both, including the most frequently used measures as well as some "interesting" ones used in one or very few studies, are seen in Table 2. For additional references to these measures, all of which can be used for tuning (and some for control), see [2].

Since, as Svobodova states, "Without understanding the workload, performance analysis and prediction is at best inaccurate [and] possibly completely invalid," examples of parameters utilized in workload characterization are reproduced in Table 3 [2]. A "unit of work" is herein defined in terms of a job (or command in an interactive environment), task (i.e., a representative operation carried out by a particular processor [CPU, channel]), instruction, or bit [2].

In an effort to measure the factors representing the quality or quantity of interactive service from the user's viewpoint, Abrams and Ireu identify and assess the implications of over thirty measures based

Table 2
Examples of Performance Measures

<u>Performance Measure</u>	<u>Description</u>
SYSTEM EFFECTIVENESS	
Throughput	Amount of useful work completed per unit of time with a given workload [c]
Relative throughput	Elapsed time required to process a given workload on system S1 / elapsed time required to process the same workload on system S2
Capability (capacity)	Maximum amount of useful work that can be performed per unit of time with a given workload (a measure of the maximum throughput) [c]
Turnaround time	Elapsed time between submitting a job [a] to a system and receiving the output [c]
Response time	Turnaround time of requests and transactions in an interactive or real time system [c]
Availability	Percentage of time a system is available to users
SYSTEM EFFICIENCY	
External delay factor	Job turnaround time/job processing time [c]
Elapsed time multi-programming factor (ETMF)	Turnaround time of a job under multiprogramming / turnaround time of this job when it is the only job in the system
Gain factor	Total system time [b] needed to execute a set of jobs under multiprogramming / total system time needed to execute the same set sequentially
CPU productivity	Percent of time a CPU is doing useful work (used a measure of throughput) [c]
Component overlap	Percent of time two or more system components operate simultaneously [c]
System utility	Weighted sum of utilization of system resources [c]
Overhead	Percent of CPU time required by the operating system
Internal delay factor	Processing time for a job under multiprogramming / processing time of this job when it is the only job in the system
Reaction time	Time between entering the last character on a terminal or receiving the input in the system and receiving the first CPU quantum [c]
Wait time for I/O	Elapsed time required to process an I/O task
Wait time for CPU	Elapsed time required to process a CPU task
Page fault frequency	Number of page faults per unit of time [c]

Table 3
Workload Characteristics

<u>Workload Parameters</u>	<u>Description</u>
Job CPU time	Total CPU time requested by a single job [a]
Job I/O requests	Total number of I/O operations requested by a single job
CPU service time	CPU time required to process a single CPU task
I/O service time	I/O time required to process a single I/O task
Interarrival time	Time between two successive requests for a system service
Priority	Priority assigned to a job by the user
Blocked time	Time a job is incapable of receiving CPU service
Memory requests	Amount of memory requested by a single job
Working set size	Number of pages of a single job that must be kept in main memory
Locality of reference	Times for which all memory references made by a single job remain within a single page or set of pages
User response time	Time needed by a user at an interactive terminal to generate a new request (think and type time)
User intensity	Processing time per request / user response time
Number of simultaneous users	Number of interactive users logged in concurrently
Number in the system	Number of jobs or tasks being serviced or waiting in queues for system resources
Instruction mix	Relative frequencies of different types of instructions the system must execute

[a] In an interactive system, each command issued from a user's terminal is assumed to represent a new job.

[b] System time is the time when at least one of the system's processors (CPU, channels) is busy.

[c] Can be used for control as well as tuning.

Source: [2]

on time, length and volume of transactions, rates of service, various ratios, etc.[5] In terms of user satisfaction (especially with reference to interactive users in a database community), measures of online application availability (the percentage of time that all operational components, including the database(s), are available), partial service interruption, and database currency may be of value [6].

Other indicators which may be used to evaluate performance are:

- CPU, I/O device, and channel mean service times [4]
- Mean and variance of response time [2]
- N percentile response time (i.e., the time limit that guarantees that the response times of N percent of all requests are shorter than this limit) [2]
- Utilization (in terms of the fraction of time the resource is in use, e.g., CPU utilization in terms of the time spent in the idle, supervisor, or program states) [2]
- Percentage interactive tasks (percentage of online tasks which are completed with one swap) [7]
- time mix (ratio of compute-bound task time to interactive task time) [7]
- translator utilization. [7]

Performance measures are usually expressed as mean values, which as Svobodova points out are in many cases clearly inadequate measures of system performance; e.g., even if the mean response time is short, a user in an interactive environment will be dissatisfied if the variance of the response time is large. Thus, if the exact distribution of the response time is unknown, variance as well as the mean response time should be considered as performance measures. Svobodova recommends N percentile response time (defined above). Moreover, (percentile) response time should be assessed separately for different classes of requests, since trivial ("edit") requests can be expected to take less time than involved ("compile") requests [2].

V. SOFTWARE AND HARDWARE MONITORS

Although some aspects of performance evaluation (e.g., CPU utilization) can be measured via either software or hardware instrumentation, in most cases the domains of these monitors are very different. Whereas software monitors usually record information accessible to the operating system or application programs, hardware

monitors measure the state of the computer (i.e., changes in the system status, usually reflecting component utilization) via electronic sensors, or probes.

Svobodova presents a cogent overview of both hardware and software monitors. This analysis reveals the following strong points and inadequacies [2].

Hardware monitors can record events (i.e., changes in system state) at high rates of speed, can detect and process several events occurring in parallel, exhibit high timer resolution, are machine independent, operate asynchronously to the measured system, can monitor several independent hardware units simultaneously, and require no system resources. Such monitors cannot, however, directly monitor information stored in memory (e.g., wait queues). They have limited decoding capabilities and can deal with software components only in an extremely limited way. A very significant disadvantage is that intimate knowledge of the hardware is required in order to develop a probe point library for a particular machine; such a library may not in fact be documented for older systems. The physical process of probe attachment is also difficult, and the resultant load may cause system failure [2].

Software monitors have an input rate (the maximum frequency at which events can be recognized and recorded) which is tied to instruction execution rate and may be degraded due to monitor overhead. They can detect and process events only sequentially; since timer resolution is tied to the host, it may be low in older systems. They are CPU and operating system dependent, may have problems if the timer becomes synchronized with the monitored events, cannot monitor I/O device and switch events, and consume system resources as reflected by the corresponding artifact. Such monitors can, however, access any main memory data. They have powerful decoding capabilities and can monitor software level components regardless of their locations in memory. Software instrumentation, however, also requires "intimate knowledge of software operations." [2]

Hardware monitors are superior to software monitors in measuring the duration and overlap of system activities. Software monitors, on the other hand, are "more powerful and flexible," but they may generate significant overhead [2].

The role of simple software monitors in producing useful performance data should not be ignored. Jain and Turner, for example, find that "image accounting" in the VAX environment produces valuable statistics on workload [8]; applying a statistical analysis package to data extracted from accounting records has likewise been found useful in characterizing workloads in a Univac shop [9].

VI. PERFORMANCE EVALUATION IN LOCAL AREA NETWORKS

The term "local area network" (LAN) is frequently applied to an interconnection of nonhomogeneous systems or devices which is limited geographically (usually to a few kilometers), has a specified range of data rates, and is controlled by a single organization or group. Performance evaluation methodology for such networks is currently in the process of development as LANs themselves rapidly proliferate.

Performance data in LANs can be collected in much the same way as for the component nodes, i.e., via hardware or software monitors. (Software monitors, supplied by particular vendors, are system and/or architecture dependent.) Terplan discusses the network monitor, a microcomputer-based monitor designed to collect performance data simultaneously over a large number of communications lines and terminals. Typical monitors provide status information at the network, application group, line, control unit, and device levels. Information available within levels includes response time, transaction rate, traffic counts, error rate, and component utilization. Terplan evaluates hardware and software monitors in terms of their capabilities to obtain performance data. He recommends combining hardware and software techniques in a hybrid approach [10].

Herskovitz, in evaluating the performance of a specific LAN, attempts to determine the performance characteristics of major subsystem components (e.g., the host, network adapters, and interfacing channels). The LAN in question uses NSC HYPERchannel processor adapters for communication support to transfer large volumes of data at high rates of speed between various hosts (e.g., VAX 11/780, CYBER /30 and 74, PDP 11). The adapters, positioned at nodal points in the network, intercept and forward communications and data transfers between hosts and serve as focal points in evaluating network performance.

Network performance characteristics chosen for measurement by Herskovitz include program-to-program network transfer time, data transfer rates, network software overhead, effective data trunk transfer rates, contention delays, network protocol timing measurements, and network bottlenecks. Hardware monitoring, transparent to the network, is the mode of instrumentation. For further details, see [11].

Multiple hardware and software tools are collected into a single system for the analysis and measurement of performance data in the complex LAN at Los Alamos National Laboratory. The idea here is to create a network performance analyst's workbench, supported by a single host (VAX 11/80) that is allowed to communicate with any other element in the network. A major objective is to isolate the tools and their management from evolutionary changes in the network composition and topology. One of these tools is a relational database management system which is employed to free personnel from data management details [12].

The National Bureau of Standards has defined nine general measures for evaluating network performance: availability, reliability, transfer

rate, accuracy, channel establishment-time, network delay, channel-
turnaround delay, transparency, and network security [13]. Several more
specific measures are suggested above.

VII. CONCLUSIONS AND RECOMMENDATIONS

ECONET, the local area network servicing the user community of
Eglin Air Force Base, connects a variety of nonhomogeneous systems; the
particular unique mainframes of current interest are the VAX 11/80, CDC
CYBER 176, and CDC 6600. Some of these nodes are scheduled to change
over a several year period as the older generation machines are replaced
by newer ones. The network services a large number of users (i.e.,
customers) in a heavily interactive environment; most processing
involves scientific or database management applications (disregarding
all classified operations). Performance evaluation of the ECONET CDC
nodes revolves around several basic statistics (i.e., number of jobs,
CPU time, I/O time, memory utilization, disc utilization, and connect
time) which are gathered via simple software monitoring of the
accounting records. The VAX system's software monitor provides much
more detailed information (with respect to memory management, I/O, file
management, job scheduling, process management, and processor
utilization). (For a more complete description of ECONET, including an
analysis of current performance evaluation techniques, suggested modes
of presentation for performance statistics, and an assessment of
specific hardware and software monitors, see [14].)

On the basis of the investigations detailed in Sections III - VII
of this paper, the following conclusions are drawn and recommendations
made:

1. Performance evaluation statistics for older ECONET components
are clearly inadequate. A basic statistic, namely the amount of CPU
time spent in the supervisor state (i.e., overhead), is unknown. There
is no information (other than indications through an EIMF approach) on
response time, a fundamental aspect of performance evaluation which is
even more important in a highly interactive environment where user
(i.e., customer) satisfaction is of considerable interest. (Some
statistics on turnaround time would also be helpful.) Resource request
queues should be monitored in order to detect bottlenecks. There is no
information on component overlap. A number of other performance
evaluation statistics would be helpful in controlling and tuning these
systems (see Section IV).

Considering the methods available for obtaining these data (via
modeling, or software or hardware monitoring), it is clear that unless
the expertise is made available for constructing the model(s), the only
viable option is monitoring. Both software and hardware monitors would
be applied under optimal circumstances, but due to the scarcity of
commercial products, the factors involved in generating such a monitor
in-house, and concern for monitor artifact, a hardware monitor would

provide the most readily available aid in producing meaningful performance statistics. (Depending upon when older-generation components are scheduled to be phased out, the expenditure of considerable effort to develop models or software dependent monitors is questionable.)

2. In order to obtain as accurate a picture as possible of a system as it operates, all types of statistics must be gathered. Neither hardware nor software monitors are sufficient in themselves for this purpose. To be truly effective, a variety of performance evaluation techniques must be applied -- analytic and empirical models, simulation models, and both hardware and software monitoring procedures. A "Performance Analyst's Workbench," a set of tools for use in evaluating the performance of each system, should be collected and maintained.

3. The interconnection of nonhomogeneous mainframes in a LAN greatly exacerbates the situation with respect to performance analysis and evaluation. Performance analysis in previously independent nodes, once not of critical interest, is now essential for network performance evaluation. To ameliorate this situation, a set of performance evaluation measures common to the disparate nodes should be identified. Methods for gathering the required metrics must be decided upon. A method for maintaining (and archiving) the resultant data, though the design of an appropriate database and the application of DBMS technology, should be constructed. When the performance of each node in a LAN can be measured and evaluated, it is possible to turn one's attention to network performance and know that obtaining a valid picture of network activity is now possible.

4. Performance evaluation in LANs is still in its infancy. Although some measures of network performance (and methods for obtaining them) have been identified (see Section VI), network architectures are so dissimilar that it is difficult to identify a common set of measures. Networking introduces an additional level of complexity that makes the approach taken at Los Alamos [12] even more appealing. It is through the application of this type of approach and under the direction of a properly qualified performance analyst who is responsible for dealing only with this complex aspect of the network that the best potential of local area networks can be achieved. As is indicated by these investigations, much research remains to be done in identifying appropriate performance measures and their modes of implementation in specific local area networks.

REFERENCES

1. Deitel, H. M., An Introduction to Operating Systems, Addison-Wesley, Reading, Mass., 1982.
2. Svobodova, Liba, Computer Performance Measurement and Evaluation Methods: Analysis and Applications, Elsevier, New York, 1976.
3. Lucas, Henry C., Jr., "Performance Evaluation and Monitoring," ACM Computing Surveys, Vol. 3, No. 3, pp. 79-91, Sept., 1971.
4. Rose, C. A., "A Measurement Procedure for Queueing Network Models of Computer Systems," ACM Computing Surveys, Vol. 10, No. 3, pp. 264-80, Sept. 1978.
5. Abrams, Marshall D., and Treu, Siegfried, "A Methodology for Interactive Computer Service Measurement," Comm. of the ACM, Vol. 20, No. 12, pp. 936-944, Dec. 1977.
6. "User-Oriented Performance Measures," Auerbach Reports, Auerbach Publishers, Inc., 1982.
7. Shemer, Jack E., and Robertson, John B., "Instrumentation of Time-Shared Systems," Computer, Vol. 5, No. 4, pp. 39-48, July/Aug., 1972.
8. Jain, Rajendra K., and Turner, Rollins, "Workload Characterization Using Image Accounting," Proceedings of the Computer Performance Evaluation Users Group 18th Meeting -- CPEUG82, pp. 111-20, Oct., 1982.
9. Bays, Walter N., and Voegeli, Dawn, "A Univac Workload Characterization," Proceedings of the Computer Performance Evaluation Users Group 18th Meeting -- CPEUG82, pp. 259-73, Oct., 1982.

10. Terplan, K., "Network Performance Reporting," Computer Network Performance Symposium: Performance Evaluation Review, Vol. 11, No. 1, pp. 156-68, Spring, 1982.
11. Herskovitz, Jonas, "Evaluating Local Network Performance," Proceedings of the Computer Performance Evaluation Users Group 18th Meeting -- CPEUG82, pp. 389-96, Spring, 1982.
12. Brice, Richard, and Alexander, William, "A Network Performance Analyst's Workbench," Computer Network Performance Symposium: Performance Evaluation Review, Vol. 11, No. 1, pp. 138-45, 1982.
13. Crouch, C. J., "ECONET: Performance Analysis and Evaluation in a Specific Local Area Network," unpublished report, Computer Sciences Directorate, Eglin Air Force Base, Florida, July 1983.
14. Grubbs, D. S., and Cotton, I. W., "Criteria for the Performance Evaluation of Data Communications for Computer Networks," NBSTN-882, Government Printing Office, 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

PERFORMANCE ANALYSIS AND EVALUATION IN A LOCAL AREA NETWORK

Prepared by: Dr. Carolyn J. Crouch
Academic Rank: Associate Professor
Department and University: Department of Computer Science
The University of Alabama
Research Location: Computer Sciences Directorate
Armament Division
Air Force Systems Command
USAF Researcher: Dr. Robert N. Braswell
Date: July 22, 1983
Contract No: F49620-82-C-0035

APPROVED:



Robert N. Braswell
Director of Computer Sciences
Eglin AFB, Florida
(904) 882-5315

JUL 29 1983
JM

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

IMPACT OF ADA ON USAF COMPUTATIONAL SUPPORT SERVICES

Prepared by: Dr. Donald B. Crouch
Academic Rank: Associate Professor
Department and University: Department of Computer Science
The University of Alabama
Research Location: Computer Sciences Directorate
Armament Division
Air Force Systems Command
USAF Researcher: Dr. Robert N. Braswell
Date: July 22, 1983
Contract No: F49620-82-C-0035

THE IMPACT OF ADA ON
USAF COMPUTATIONAL SUPPORT SERVICES

by

Donald B. Crouch

ABSTRACT

The feasibility and desirability of using Ada as the standard language for virtually all applications in the computational support organizations of the USAF are investigated. It is concluded that the power and flexibility of the language make it ideally suited for use in such organizations but that the programming environment of these organizations will have to be enhanced considerably before it becomes feasible to adopt it extensively. Ada's impact on the existing environment is explored in detail. Suggestions for further research in this area are made.

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 two of his students and himself with the opportunity to participate in the summer research program at the Computer Sciences Directorate at Eglin AFB, Florida. The personnel of this laboratory were very cooperative and extremely hospitable.

In particular, the author thanks Dr. Bob Braswell for his suggestion of this area of research and for the latitude provided to him in performing this work.

The author acknowledges the excellent support received from the librarians in the Technical Library at Eglin AFB and especially thanks Mary Watson, June Stercho, and Mary Murphy for the assistance provided during the literature review.

I. INTRODUCTION

The Air Force has embarked on a four-phase plan for the introduction of the language Ada as the standard programming language of the Department of Defense (DoD). The final phase of this strategy, which occurs in the 1990's, will result in regulations making the use of Ada mandatory for virtually all Air Force systems development [1]. The goal of this initiative is "to improve [software] productivity while achieving greater system reliability and adaptability." [2]

Ada is being adopted as the standard language for the Air Force in response to a growing concern over the high cost of software. The Air Force has estimated, for example, that by 1985 software expenditure will represent over 90% of total system costs [3]. The use of Ada in software development encourages modern programming practices based on the principles of software engineering. Adherence to these principles reduces the overall cost and complexity of developing and maintaining software systems by improving the clarity, reliability, efficiency, and maintainability of the software.

The adoption of Ada alone is not sufficient to enable the Air Force to reach its overall goal. The nature of the programming environment (i.e., tools, methodologies, management policies, and operating conditions) in which software systems are constructed generally influences the development of software to a far greater extent than the programming language per se. To reap the full benefits of Ada as the standard language, an integrated Ada programming environment is also required.

This project focuses on the problems which Air Force computational support organizations will face as a result of the mandated adoption of Ada and explores emerging research issues brought about by such language standardization.

II. OBJECTIVES OF THE RESEARCH

The overall objective of this effort is to determine the feasibility and desirability of using Ada as the standard programming

language in the computational support organizations of the Air Force.

Specific objectives are:

- To assess the impact that adoption of Ada will have on computational support organizations.
- To define environmental attributes and conditions which support successful incorporation of Ada into the organizations.
- To develop a strategy which organizations may follow in the introduction of Ada.
- To investigate broader areas of research which will emerge as a result of the DoD's standardization of Ada.

III. APPROACH

In order to meet the designated objectives, we first must identify the types of services provided by computational support organizations in general and, more importantly, ascertain the environmental conditions under which such services are performed. This information is then evaluated to reveal potential problems which may exist in current programming environments in relation to the major phases of the software life cycle. The sources of all such problems are then identified, and Ada is analyzed to determine whether it can lend support to solving these problems.

This report draws directly on discussions with technical personnel of the Computer Sciences Directorate of the Armament Division at Eglin AFB, a review of the Directorate's planning documents, and a survey of the literature.

IV. CURRENT PROGRAMMING ENVIRONMENTS

Air Force computational support organizations differ considerably in the extent of services they provide. The Directorate of Computer Sciences at Eglin AFB is representative of the most comprehensive service organization; the services that it provides "include both scientific and business computer systems operation and management, real-time range control and support, test data analysis and reduction, and systems engineering/consultation support of a wide variety of computer applications." [4] Software development for computational

scientific applications is done almost exclusively in FORTRAN; in some cases, assembly language subroutines and/or language extensions are utilized to implement certain aspects of a system. Regardless of the range of services provided by such organizations and the languages employed, we find that each organization exhibits similar environmental conditions. [4,5]

Due to the universality of these conditions and the general design objectives of Ada, the adoption of Ada as the standard language within the Air Force will prove to be beneficial to computational support organizations. To ascertain these benefits, we need to analyze in detail the environment in which such organizations operate in relation to the phases of the software life cycle support provided by the organizations.

The phases of software life cycle support include the following:

- Problem requirements analysis
- Software solution specifications
- Software design
- Coding
- Testing and acceptance
- Maintenance
 - Correcting latent program errors
 - Correcting deficiencies revealed through use
 - Responding to environmental changes
 - Optimizing for more efficient operation.

These phases of software support are generally carried out in programming environments operating under the following conditions:

- Shortage of software specialists
- Personnel migration
- Increased demand for software development
- Inadequate enforcement of documentation standards
- Inadequate software tools
- Rapid pace of technological advances.

Previous studies have concluded that these conditions lead to "software that is nonresponsive to user needs, unreliable, excessively expensive, untimely, inflexible, difficult to maintain, and not

reusable." [6] A review of the environment of the support organizations in relation to services provided indicates why these conclusions were reached.

The inability to recruit and retain well qualified software specialists is a major problem resulting from increased competition from the private sector for such individuals. In an attempt to alleviate the personnel shortage, many organizations are hiring engineers and mathematicians with only an introductory knowledge of a programming language (usually FORTRAN). An engineer or mathematician who "uses a computer" is not comparable to a computer engineer or computer scientist [7]. "Users" are generally not knowledgeable of the concept of software engineering design as it relates to all aspects of a computer system, the tools and methodologies supporting design and implementation, and algorithms, language constructs and man-machine processes necessary for effective and efficient coding.

To enable these junior level personnel to gain experience about the organization's support service operations, they are generally initially assigned to software maintenance tasks as opposed to software development. Yet, previous studies have found that over the lifetime of a software project, software maintenance exceeds in time and cost all other phases of life cycle support [8]. Thus, maintenance, one of the more critical areas of software support services, is relegated to the less experienced employees. The potential ramifications of such policies are immense. Consider the problems that confront the maintenance staff:

- The original software design/development team for a project has usually been dissolved and its members reassigned. Team members cannot be assigned maintenance tasks due to the demand for higher priority software development projects.
- Original requirements specifications and design specifications are not available or the documentation is so fragmented as to be of no use to the maintenance staff.
- Inadequate documentation exists for the code itself. Documentation should be an integral and required part of each phase of the life cycle, but in an environment of tight costs

and schedules, documentation is deferred, ad infinitum.

- Prior maintenance tasks have not been properly documented.
- Software packages are poorly structured and therefore not easily readable.
- The demand for modification to existing software is overwhelming due somewhat to the changes brought about by advances in technology.

These conditions when compounded create the most bottleneck in software support, that is, the inability of the maintenance staff to understand comprehensively the software system to which modifications are to be made [9]. In general, these conditions are a result of the lack of use of appropriate techniques for controlling and measuring the progress of software throughout its life cycle. Techniques are needed which lead to better control over the work of the design team, improved productivity of the software development team (via automation and/or reusable code), and enhanced quality and timeliness of the tasks performed by the maintenance team.

V. GENERAL SOLUTIONS TO THE ENVIRONMENTAL PROBLEMS

The maintenance problem, and thus the problems which create it, can be solved by improving the environment which pervades software development and support. This improvement could be brought about through a combination of technical and administrative changes. Specifically, more emphasis needs to be placed on applying engineering concepts to software development. Unfortunately, a gap exists between software engineering technology as it exists within certain academic institutions and as it is practiced by the (normal) professional programmer [10]. As long as software developers continue to use archaic tools and methodologies for solving complex problems, the software development and maintenance problems will continue to exist. The use of Ada for systems development is seen as a starting point in the solution of this problem, since Ada "encourages and enforces the basic principles of software engineering." [11]

Software engineering implies the "disciplined and skillful use of suitable software development tools and methods, as well as a sound

understanding of certain basic (programming) principles." [12] It encompasses the design, development, implementation, test, evaluation, and maintenance of software over its entire life cycle. The four fundamental goals of software engineering in relation to the characteristics programs should possess are modifiability, efficiency, reliability, and understandability. The software engineering principles that support these four goals are modularity, localization, abstraction, hiding, uniformity, completeness and confirmability.

The principles of software engineering aid in decomposing complex software systems into manageable entities. Several methodologies exist for designing software systems by applying these principles in a diversified manner [13]. Methodology alone, however, is not sufficient to solve the software problems. Software development tools are also needed. At present, software tools range in complexity from simple editors and debuggers to sophisticated programming environments [14]. Of all the available tools, however, the programming language utilized and the environment in which the software is developed are by far the most important.

We will investigate specifically the benefits to be derived from using Ada as one of the software tools in the programming environment of USAF support organizations and the characteristics of other software tools which enhance support services.

VI. THE ADA PROGRAMMING LANGUAGE

The design of Ada is strongly influenced by the goals and principles underlying software engineering. In fact, Ada encompasses all the principles previously delineated.

The language is designed with three overriding concerns: "program reliability and maintenance, programming as a human activity, and efficiency." [15] Emphasis is placed on program readability over ease of writing due to the established need for languages that promote software reliability and simplify maintenance. In general, Ada is

a modern, high-level, strongly typed language with many constructs similar to those in Pascal. Ada has special features for controlling visibility of identifiers and handling concurrent processes and errors efficiently. It also allows controlled access to the machine-level

architecture, provides for data abstraction, and has special constructs (generic units) for building libraries of routines for various data types....The language will be used for operating systems, compilers, numerical analyses, simulations, communications systems, real-time applications, and a variety of multiprocessing environments [16].

The overall goal of the Air Force's adoption of Ada as the standard language is to improve productivity while achieving greater system reliability and adaptability. The intent is to provide development and in-service support that is faster, less expensive and more predictable, and results in software that is more powerful, reliable and adaptable [2].

The power and flexibility of Ada make it ideally suited as the language to be used by USAF computational support organizations. The comprehensive set of applications that such organizations provide can be performed well in this language, including its real-time applications. Furthermore, Ada encourages utilization of software engineering principles, a very desirable goal of any computer organization.

The principal advantage to be gained from using Ada as the standard language is that it will eventually substantially reduce software support costs. In particular, Ada's package construct will expedite the software development process and simplify program maintenance, since it is the means by which true software components can be developed. Generics can also be used to reduce development time by enabling programmers to build reusable components which can be tailored to fit a variety of applications. Ada's strong typing features will facilitate testing and debugging and lead to quicker acceptance of software packages. Parallel tasking and intertask communication support real-time programming. Exception handling provides a uniform facility for defining, raising, and handling special conditions; this ability is essential for real-time systems.

Generally, Ada is a highly structured language; this feature permits one to comprehend more easily the scope and organization of programs written in it. Furthermore, Ada's syntax is designed to be very readable; emphasis is placed on the writing of visually appealing code.

The desirability of using Ada is sound. However, due to the complexity of the language, various administrative practices will have to be established and supporting tools acquired in order to manage this complexity and to make the language feasible for use in computational support organizations. Only then will Ada significantly reduce the cost and enhance the reliability of software support activities.

VII. ADA'S IMPACT ON THE ENVIRONMENT

In order for organizations to use Ada effectively, the environment in which software development occurs and support services exist must be enhanced considerably.

Personnel skills will have to be upgraded. Improvement is needed not only in the knowledge and use of Ada but also in the purpose and concepts of software engineering.

Programming standards and methodologies must be defined and enforced. An organizational-wide systematic approach to designing, coding, and documenting software is essential in the use of Ada. Even though organizational standards currently exist to help promote quality and reliability and enable programmers to migrate among projects, the complexity of enforcing such standards is overwhelming. Consequently, they often are ignored.

Since development of standards and methodologies alone is not sufficient, software tools must be defined, selected, and applied correctly and consistently in an Ada programming environment to ensure that standards are adhered to. Although many tools have been designed to facilitate software development and maintenance, most current tools focus on very narrow aspects of software development; in fact, a comprehensive, integrated set of tools has not been developed for any one particular programming language [17]. Thus, Ada is just the beginning; the real challenge is to develop software tools to support Ada software management tasks.

The essence of a software support environment is an integrated set of tools that continuously aid the user in all phases of the software support life cycle. Osterweil lists five characteristics that a true "environment" must exhibit: breadth of scope and applicability, user

friendliness, reusability of internal components, tight integration of capabilities, and use of a central repository [17]. Some Ada support tools will be provided by the Ada Programming Support Environment (APSE).[18] Minimally, APSE will consist of a database reflecting information about a project over its entire life cycle, tools such as editors, debuggers, etc. to provide the desired support, and user interfaces to provide access to the databases and the tools. Specifically, APSE facilities will consist of the following:

- A self-contained Ada oriented environment which conceals the host operating system from the user.
- Database facilities for the storage of information in various categories relating to the project's life cycle. The purpose of the database is to provide a central control point for project management.
- Software tools such as text editors, linkers, set-use static analyzers, control-flow static analyzers, terminal interface routines, and a file administrator.
- A front-end compiler which provides separate compilation of source modules into intermediate code. This feature facilitates transportability of software and retargeting of software to hardware other than the host.

Other tools that would be of particular value in an Ada programming environment include those which ensure adherence to established managerial and technical policies, as well as documentation assistance, style analyzers, and change propagation detectors [9]. Such tools would permit the programmer to work more effectively by aiding in the program structuring process, assisting in the creation, structuring and maintenance of documentation, checking programs for predetermined styles and guidelines, and analyzing and isolating effects of program changes.

The trend in software tools is toward the development of intelligent, knowledge-based workstations. Such workstations would provide its users the type of assistance that software specialists now provide. Initial work in this area has focused on the software acquisition manager's workstation, a microcomputer-based, interactive

system designed with an easily understood user interface [19]. Ultimately, this type of workstation will provide expert software acquisition knowledge covering the entire software life cycle. The development of similar tools as aids in program design and development would definitely increase programmer productivity and program reliability.

This general area of knowledge-based expert systems and the impact Ada will have on the development of such systems are explored in more detail later. First, let us consider a strategy which computational support organizations may follow in order to prepare now for the adoption of Ada as the standard language of the Air Force.

VIII. INTRODUCTION STRATEGY FOR ADA

Computational support organizations need to begin to prepare now for the integration of Ada and state-of-the-art software tools into their environment. The following systematic approach is recommended for doing so:

- Establish immediately a small Ada support group (at least two individuals) whose sole responsibility is to prepare for the move to Ada. As a first priority, members of the support group must become experts in software engineering concepts and the language Ada.
- Contract for DoD-approved Ada compiler(s) for the organization.
- Establish organization-wide design and programming methodologies.
- Define the minimal tools necessary for the Ada programming environment. Develop or contract development of such tools.
- Review current application areas and define reusable software components. Initiate development of generic packages.
- Establish training curricula in software engineering, Ada, support tools, methodologies and policies for retooling of existing personnel.
- Establish similar orientation programs for all new employees.
- Choose one or two new projects and designate their implementation in Ada. Close supervision must be given to

ensure that the capabilities of Ada are utilized and programming policies adhered to.

- Gradually retrain existing employees and phase in Ada as the standard language.
- As appropriate, recode existing applications in Ada. This involves more than a simple translation from one language to another. A complete redesign is necessary to take advantage of the features and capabilities of Ada.
- Continue to define and refine the Ada programming environment.

This strategy is a systematic approach to the introduction and use of Ada which emphasizes personnel skill improvement, modification of the existing programming environment to promote software quality rather than "defect identification," definition and creation of reusable software, and gradual integration of the language Ada into the system. If such a strategy is followed in a dedicated, conscientious fashion over the next seven years, then the organization will be well prepared to adopt Ada as the standard language in the early 1990's, the time designated for such standardization.

IX. ADA AND KNOWLEDGE-BASED SYSTEMS

As previously indicated, a surge of interest is now being expressed in the potential applications of expert systems technology. A rapid growth in the application of expert systems to a diversity of areas is anticipated during the next decade, especially in military usage. Since Ada has been selected as the standard language for DoD applications, we need to explore the role it will have in this evolving field. To do so, we need first to survey briefly the general area of knowledge-based expert systems.

Artificial intelligence is the "subfield of computer science concerned with the concepts and methods of symbolic inference by a computer and the symbolic representation of the knowledge to be used in making inferences." [20] Researchers in this area have discovered that the ability of a computer to perform well those tasks indicative of intelligence depends primarily on the quantity and quality of knowledge that the machine possesses about the task [21]. Thus, due to the

paramount importance of knowledge, research in artificial intelligence has shifted its focus during the last few years from inference-based models to knowledge-based models. The result has been a profusion of so-called "knowledge-based expert systems."

Expert systems are "problem-solving computer programs that can reach a level of performance comparable to that of a human expert in some specialized problem domain." [22] Such systems render to their users intelligent advice on the topics comprising the system's knowledge-base. Thus, expert systems are generally used in applications involving reasoning, as opposed to calculating. A major feature of such systems is that they are able to explain the lines of reasoning that lead to the decisions. This is a particularly important property since "the ultimate use of an expert system will depend upon its credibility to its users." [20]

Current applications of expert systems include medical diagnosis, engineering design, computer assisted instruction, structural analysis, mineral exploration, identification of organic compounds, and systems management [23]. The success of such systems has led many scientists to conclude that knowledge-based expert systems will be the application area for the 1990's.

This view is widely held by the Japanese, British and United States governments. In April, 1982, Japan embarked on an extensive research project with the aim of developing by 1990 a prototype fifth-generation computer system -- a knowledge-information processing system. This computer system will represent the culmination of four research areas: knowledge-based expert systems, very high level programming languages, decentralized computing and VLSI technology [24]. Japan plans to invest 45 million dollars in the project during the first three years; the total project cost may exceed several billion dollars [25].

Britain, on learning of the Japanese project, initiated its own research program in May, 1983, to develop a fifth-generation computer. The government, in cooperation with the British electronic industry, is investing 550 million dollars over a five year period in the project [26]. The project is directed toward four key areas, namely, software

engineering, VLSI, man-machine interfaces, and knowledge-based systems. The British objective closely parallels that of the Japanese.

The United States Defense Advanced Research Projects Agency "plans to do everything the Japanese have set out to accomplish and more." [27] It plans to push for advanced supercomputing and artificial intelligence technologies that may cost as much as one billion dollars.

Although knowledge-based techniques are still only in their infancy, the trend is obvious: the move is toward knowledge-based expert systems. These systems will be particularly beneficial in military-defense systems. Several pilot projects are currently on-going in this area.

The Navy has developed a medical diagnosis program to be used on board nuclear submarines where physicians may not be available and where radio silence must be maintained [28]. The Rand Corporation has been investigating the possibility of applying knowledge engineering techniques in the construction of an expert system for tactical targeting [29]. Researchers at the University of Illinois are focusing on in-flight fault diagnosis with the aid of an on-board intelligent computer system [30]. In the not-too-distant future, expert systems will be embedded in military systems as an integral part of their operation.

Initially, Ada was selected by DoD as the language to be used in the implementation of embedded systems. As we have seen, the flexibility and power of Ada are such that it can be used for most computational support activities. Since the trend is ultimately toward the incorporation of knowledge based systems into military systems and the development of knowledge based work stations for programming environments, the question naturally arises as to the feasibility of using Ada in the implementation of such systems.

At present, adequate software tools do not exist for building expert systems [22]. The design of these tools involves many considerations, including issues of generality, completeness, language features, database structure, and control methods. Suggested maxims for this design task are:

- Provide a high-level representation language.

- Provide built-in explanation and interaction facilities.
- Provide local operating system accessibility.
- Provide a basic data representation scheme that is as general as possible.
- Provide an accessible control mechanism if generality is more important than efficiency.
- Provide a constrained control mechanism if learning, self-modification, or sophisticated explanation is required.[31]

Functional and logic languages are generally used to program expert systems because the complexity of expert systems dictates "a rich, interactive editing and debugging environment." [21] In many cases, the rules and heuristics forming the knowledge base must be treated as data to be reasoned about and in other cases as code to be executed.

Research must be conducted in the area of programming languages to ascertain the applicability of Ada to the building of knowledge-based expert systems.

X. RECOMMENDATIONS

The power and flexibility of the Ada language make it ideally suited for use as the standard language of the USAF computational support organizations. These organizations need to prepare now for the implementation of Ada in order to minimize disruption of their services and to ensure and enhance their professional stature and leadership role in the computing environment. Proper usage of this language will lead to improved personnel skills and increased quality of software support services.

To integrate Ada properly into the organizations, additional research must be invested in the study and design of comprehensive software tools and techniques for future Ada programming environments. This is a major research area of particular importance now.

Future embedded military systems and software tools for the Ada environment will have to be implemented in the Ada language itself. Since the trend is toward the incorporation of expert systems into military systems and into the programming environment, research needs

to be directed toward the analysis of Ada as a language for implementing expert systems. The properties of a language for implementing such systems need to be established and the features of Ada evaluated in light of these properties. If Ada proves to be satisfactory for such applications, then the data structures, functions, procedures, and parsing techniques needed to implement expert systems in Ada need to be defined and implemented in Ada.

REFERENCES

1. David A. Hammond, "AF Ada Intro Plan is 4 Phased, Systematic," AFSC Computer Resource Newsletter, pp. 4-6, April, 1983.
2. Allan Kopp, "Strategy for a DoD Software Initiative," Paper presented at JOVIAL - Ada Users Group Meeting, Santa Clara, California, January, 1983.
3. Charles A. Muntz, "NSW Executive Enhancements II," Massachusetts Computer Associates, Inc., Report No. CADD-8301-3001, March, 1982.
4. "Five Year Plan FY 83-87," Computer Sciences Directorate, Armament Division, Eglin AFB, March, 1983.
5. M. Davis, S. Glaseman, and W. Stanley, "Acquisition and Support of Embedded Computer System Software," RAND Report R-2831, September, 1981.
6. D. A. Fisher, "A Common Programming Language for the DoD - Background and Technical Requirements," Report P-1191, Institute for Defense Analysis, June, 1976.
7. Herbert Freeman, ed., "Research Directions in Computer Engineering," Report of a Workshop, Washington, D. C., November, 1982.
8. Robert L. Glass, "Persistent Software Errors," IEEE Transactions on Software Engineering, Vol. 7, pp. 162-8, 1981.
9. J. S. Dean and B. P. McCune, "Advanced Tools for Software Maintenance," Advanced Information & Decision Systems, Report No. TR 3006-1, December, 1982.
10. Michael J. McGill, "Consideration in the Transfer of Software Engineering Technology," Proceedings of the 1980 National Computer Conference, AFIPS Press, 1980.
11. Grady Booch, Software Engineering with Ada, (Benjamin/Cummings, New York, 1983).
12. D. T. Ross, et al. "Software Engineering: Process, Principles and Goals," IEEE Computer, Vol. 8, pp.17-27, 1975.
13. G. D Bergland, "A Guided Tour of Program Design Methodologies," IEEE Computer, Vol. 14, pp. 13-37, 1981.
14. P. C. Houghton, Jr., "Software Development Tools: A Profile," IEEE Computer, Vol. 16, pp. 63-70, 1983.
15. Military Standard 1815A, "Ada Programming Language," Department of Defense, January, 1983.
16. "The Programming Language Ada," Auerbach Reports, Auerbach Publishers, Inc., 1982.
17. Leon Osterweil, "Software Environment Research: Directions for the Next Five Years," IEEE Computer, Vol. 14, pp. 35-43, 1981.
18. "Requirements for the Program Environment for the Common High Order Language," STONEMAN, Department of Defense, February, 1980.
19. Andrew B. Ferrentino, "Software Acquisition Manager's Knowledge Based Expert System," Software Architecture & Engineering, Inc., Report No. N14-82-C-0130, June, 1982.
20. Edward A. Feigenbaum and Pamela McCorduck, The Fifth Generation, (Addison-Wesley, Massachusetts, 1983).

21. Randall Davis and Douglas B. Lenat, Knowledge-Based Systems in Artificial Intelligence, (McGraw-Hill, New York, 1982).
22. Dana S. Nau, "Expert Computer Systems," IEEE Computer, Vol. 16, pp. 63-85 (1983).
23. Bruce G. Buchanan, "Partial Bibliography of Work on Expert Systems," Department of Computer Science, Stanford University, Report No. HPP-82-30, November, 1982.
24. Philip C. Treleaven and Isabel G. Lima, "Japan's Fifth-Generation Computer Systems," IEEE Computer, Vol. 15, pp. 79-88, 1982.
25. K. H. Kim, "A Look at Japan's Development of Software Engineering Technology," IEEE Computer, Vol. 16, pp. 26-37, 1983.
26. Kevin Smith, "UK Pursues Fifth-Generation Computer," Electronics, pp. 101-102, May 31, 1983.
27. "The Race to Build a Super Computer," Newsweek, pp. 58-64, July 4, 1983.
28. J. Henderson, et. al. "Adaptation of a Computer-Assisted Diagnosis Program for use by Hospital Corpsman Aboard Nuclear Submarines," Proceedings of a Symposium on Computer Applications in Medical Care, pp. 576-93, November, 1978.
29. M. Callero, D. Gorlin, F. Hayes-Roth, and L. Jamison, "Toward an Expert Aid for Tactical Air Targeting," Rand Report N-1645, January, 1981.
30. D. L. Waltz, et al. "An Expert Distributed Robotics System with Comprehension and Learning Abilities in the Aircraft Domain," Report T-116, Coordinated Science Laboratory, University of Illinois, July, 1982.
31. D. A. Waterman and F. Hayes-Roth, "An Investigation of Tools for Building Expert Systems," RAND Report No. R-2818, June, 1982.

1983 USAF - SCEEE SUMMER FACULTY RESEARCH PROJECT

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

A MOLECULAR ORBITAL STUDY OF $\text{OH}^- \cdot \text{H}_2\text{O}$, $\text{H}^+(\text{CH}_3\text{CN})_k (\text{H}_2\text{O})_m$,

AND $\text{H}^+(\text{HCN})_n$ CLUSTER IONS

Prepared by: Dr. Carol A. Deakyne

Academic Rank: Assistant Professor

Department and University: Department of Chemistry
College of the Holy Cross

Research Location: Air Force Geophysics Laboratory, Aeronomy Division,
Atmospheric Structure Branch

USAF Research Colleague: Dr. John F. Paulson

Date: October 11, 1983

Contract No.: F49620-82-C-0035

A MOLECULAR ORBITAL STUDY OF $\text{OH}^- \cdot \text{H}_2\text{O}$, $\text{H}^+(\text{CH}_3\text{CN})_k(\text{H}_2\text{O})_m$,

AND $\text{H}^+(\text{HCN})_n$ CLUSTER IONS

by

Carol A. Deakyne

ABSTRACT

The structures, energetics, and charge distributions of $\text{OH}^- \cdot \text{H}_2\text{O}$, $\text{H}^+(\text{H}_2\text{O})_2$, $\text{H}^+(\text{H}_2\text{O})(\text{CH}_3\text{CN})$, $\text{H}^+(\text{CH}_3\text{CN})_2$, and $\text{H}^+(\text{HCN})_n$, $n=1-4$, have been investigated ab initio at several basis set levels. Fully optimized geometries of symmetric and asymmetric $\text{OH}^- \cdot \text{H}_2\text{O}$ have been obtained utilizing the MP2/6-31+G** basis with the frozen core approximation. At this level of calculation, the asymmetric form is slightly more stable. Partial geometry optimizations have been initiated for the $\text{H}^+(\text{H}_2\text{O})_m(\text{CH}_3\text{CN})_k$ ions at the 3-21G level and compared with the 4-31G results. Both basis sets yield a symmetric hydrogen bond in $\text{H}^+(\text{H}_2\text{O})_2$ and asymmetric hydrogen bonds in $\text{H}^+(\text{CH}_3\text{CN})_2$ and $\text{H}^+(\text{H}_2\text{O})(\text{CH}_3\text{CN})$. For the latter cluster, the proton is positioned closer to the CH_3CN in accordance with the relative proton affinities of H_2O and CH_3CN . Completely and partially optimized structures of the $\text{H}^+(\text{HCN})_n$, $n=1-4$, clusters have been calculated using the 4-31G and 6-31G** basis sets with and without inclusion of electron correlation. The ions are linear, and for $n \geq 4$ there will be mixtures of linear ions of nearly equal stability present at equilibrium. Both of the dimers $[\text{HCNH} \dots \text{NCH}]^+$ and $[\text{HNCH} \dots \text{NCH}]^+$ may be in the system at the lower temperatures at which the trimers and tetramers are formed. The computed proton affinities and solvation energies, $\Delta E_{n-1,n}$, are in good agreement with experiment, with the exception of $\Delta E_{2,3}$. The theoretical data suggest that $\Delta H_{2,3}^\circ$ should be larger than the value reported. The structural changes, charge redistributions and proton migrations obtained for these complexes at each successive HCN is bound indicate that: 1) the central ion is HCNH^+ , 2) there is a cooperativity effect among the hydrogen bonds, and 3) the first solvent shell is not filled until two more HCN's are added to HCNH^+ . Suggestions are made for follow-on research in these areas.

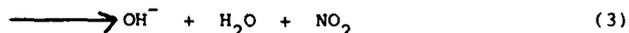
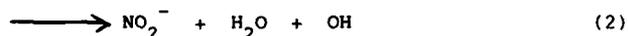
ACKNOWLEDGEMENT

The author would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research, the Southeastern Center for Electrical Engineering Education and the Atmospheric Structure Branch of the Air Force Geophysics Laboratory at Hanscom AFB, MA for enabling her to enjoy a very interesting and productive summer. In particular, the author would like to express her gratitude to Dr. John F. Paulson for suggesting this research project and for his friendly collaboration and helpful comments. She would also like to thank all the people working at the Atmospheric Structure Branch for their hospitality.

I. INTRODUCTION

Recently balloon-borne mass spectrometers have been utilized to determine the ionic composition of the earth's ionosphere and stratosphere.¹⁻⁷ These studies indicate that both the positive and negative ions are clustered species. Consequently, the properties of some of these cluster ions, such as their structures, energetics, and charge distributions have been under extensive experimental and theoretical investigation.⁸⁻¹¹ Information on these properties is needed for several reasons. First, data on the structures of these complexes will 1) lead to a more complete understanding of the forces between ions and neutral molecules, in particular the role of hydrogen bonding, which is believed to be significant,^{10,12,13} in the solvation of positive and negative ions and 2) provide information on the number of solvent molecules which can be accommodated by an ion and on its inner and outer solvation shells. This data is useful since the kinetic properties, electron affinities, mobilities, and products of the ionic recombination reactions of these clusters are believed to be dependent on their sizes.¹⁴ Second, thermodynamics data on clustering reactions is required in the theoretical development of aerosol formation¹⁰ and in the theoretical development of the chemical mechanisms involving negative ions which cause radio and radar blackout, i.e. photodetachment and photodissociation.¹⁴ Third, the calculated charge distributions will indicate whether the positive or negative charge is relatively evenly distributed or whether it is localized as is assumed in some electrostatic models used to compute bond energies.

The cluster ions investigated in this work, via ab initio molecular orbital theory are $\text{OH}^- \cdot \text{H}_2\text{O}$, $\text{H}^+(\text{H}_2\text{O})_n(\text{CH}_3\text{CN})_m$ and $\text{H}^+(\text{HCN})_n$. The structure and reactions of $\text{OH}^- \cdot \text{H}_2\text{O}$ have received considerable attention lately. Paulson and Dale¹⁵ have performed a double mass spectrometer and SIFT study of some gas phase reactions of $\text{OH}^- \cdot \text{H}_2\text{O}$ and NO_2 . Three of the reactions they observed are:

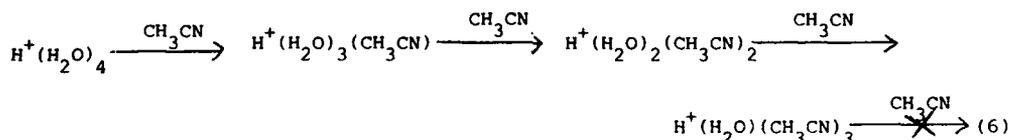
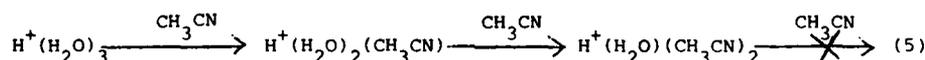
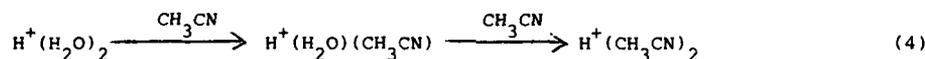


In order to probe the accuracy of ab initio methods in calculating energy changes for reactions involving negative ions, this set of reactions has been studied in

this laboratory.¹⁶ However, reliable thermodynamic data cannot be obtained unless the structure of $\text{OH}^- \cdot \text{H}_2\text{O}$ is known. Although several research groups¹⁷⁻¹⁹ have carried out ab initio calculations for $\text{OH}^- \cdot \text{H}_2\text{O}$, none of the calculations included diffuse functions which are known to be important for a proper description of anionic species.^{20,21}

Two types of positive ions found in the stratosphere are proton hydrates, $\text{H}^+(\text{H}_2\text{O})_n$, and $\text{H}^+(\text{H}_2\text{O})_m \text{X}_k$ cations, where X has a mass of 41 ± 1 amu and must have a proton affinity greater than 175 kcal/mole and an abundance greater than $7 \times 10^4 \text{ cm}^{-3}$ in order to enter into the ion chemistry.¹⁻⁴ The identity of X in the positive cluster ions is still uncertain. Two proposed candidates are MgOH ²² and CH_3CN , acetonitrile.² Smith et al.²³ have carried out studies using a selected ion flow tube (SIFT) apparatus which indicates that CH_3CN is indeed a viable candidate for X. In comparison, data on the height variations of the abundance ratios for $\text{H}^+(\text{H}_2\text{O})_m \text{X}_k$ and $\text{H}^+(\text{H}_2\text{O})_n$ cations argue against X being MgOH .^{24a}

The SIFT experiments²⁹ have shown that CH_3CN rapidly replaces H_2O in the cluster ions $\text{H}^+(\text{H}_2\text{O})_n$ according to the following reactions.



An interesting result is that in reaction sequences (5) and (6) the last H_2O molecule is not replaced.

These experiments demonstrate that the presence of CH_3CN in the stratosphere would lead to the rapid formation of $\text{H}^+(\text{H}_2\text{O})_m(\text{CH}_3\text{CN})_k$ ions and provide support for the hypothesis that X is CH_3CN . Further support for this hypothesis is that ions with masses equivalent to those of various $\text{H}^+(\text{H}_2\text{O})_m(\text{CH}_3\text{CN})_k$ ions are observed in the stratosphere^{2,3,24a} and that ions with these masses can be generated from H_2^+ ions via clustering and ligand exchange reactions in laboratory drift tube experiments.^{24b} Of course, these results do not prove that CH_3CN is X but they do strongly suggest that it is a likely candidate. Additional evidence would be provided by

finding an adequate source of CH_3CN in the stratosphere. Although the kinetic data indicate that the required concentration of CH_3CN is relatively small ($\approx 10^7$ molecules cm^{-3} at ~ 35 km) no sources of CH_3CN have been identified as yet. Feasible possibilities are reactions of N^+ and N_3^+ with hydrocarbons such as CH_4 , neutral-neutral reactions between nitrogen atoms and carbon-containing molecules,²³ and reactions involving HCN which is known to be present in the stratosphere.²⁵ In addition, CH_3CN may be transported into the stratosphere from below.^{24a}

Meot-ner²⁶ has carried out gas phase studies of HNCH^+ solvated with HCN and $\text{H}^+(\text{CH}_3\text{CN})$ solvated with CH_3CN and has obtained enthalpy and entropy changes for the clustering reactions. From a comparison of these data, Meot-ner has drawn conclusions about the structures of the clustered species and on the number of solvent molecules accommodated in the inner solvation shell of these ions. Although thus far there is no evidence that $\text{H}^+(\text{HCN})_n$ ions exist in the stratosphere (or in interstellar clouds where HCN and HNC have been found in significant quantities and HCNH^+ is believed to be their immediate precursor²⁷), more detailed information on the structures, energetics and charge distributions of these clustered species will aid in understanding and predicting these properties for observed clustered species.

Hirao et al.²⁸ have examined the structures and energetics of several $\text{H}^+(\text{HCN})_n$ cluster ions via ab initio techniques. However, their study is not as extensive as this one and they made some assumptions about the structures of these species which appear to be invalid.

In addition to the studies mentioned above, other theoretical investigations of cluster ions include those on $\text{NO}_3^-(\text{H}_2\text{O})_n$,^{29,30} $\text{NO}_2^-\cdot\text{H}_2\text{O}$,^{29,30} $\text{HCO}_3^-(\text{H}_2\text{O})_n$,³¹ $\text{H}^+(\text{H}_2\text{O})_n$,³² $\text{H}^+(\text{CO})_n$,³³ and $\text{F}^-(\text{CH}_3\text{CN})_n$.³⁴ All of these calculations have been found to be of considerable value in interpreting and correlating experimental results and in yielding insight into the structure, bonding, stability and charge distributions of clusters.

II. OBJECTIVES

The specific objectives of this project were to use ab initio molecular orbital theory to determine the following:

- 1) the optimum structures and calculated solvation enthalpies of $\text{H}^+(\text{HCN})_2$, $\text{H}^+(\text{HCN})_3$, and $\text{H}^+(\text{HCN})_4$;
- 2) the optimum structures and calculated solvation enthalpies of $\text{H}^+(\text{H}_2\text{O})(\text{CH}_3\text{CN})$, $\text{H}^+(\text{CH}_3\text{CN})_2(\text{H}_2\text{O})$, $\text{H}^+(\text{CH}_3\text{CN})(\text{H}_2\text{O})_2$, $\text{H}^+(\text{CH}_3\text{CN})_2$, and $\text{H}^+(\text{CH}_3\text{CN})(\text{H}_2\text{O})_3$;
- 3) the optimized geometry of $\text{OH}^-\cdot\text{H}_2\text{O}$;

- 4) the most economical basis set which yields accurate results.
- The results obtained will be analyzed to ascertain and elucidate:
- 1) the relative stability of straight-chain, branched, and cyclic structures;
 - 2) why the last H_2O molecule in the $\text{H}^+(\text{H}_2\text{O})(\text{CH}_3\text{CN})_n$, $n = 3,4$, cations is not readily replaced;
 - 3) the role of hydrogen bonding in the ion-molecule interactions and the optimum type of hydrogen bond, i.e., linear or bent and symmetric or asymmetric;
 - 4) the amount of charge transfer through the hydrogen bonds and the nature of the cooperative interaction between them;
 - 5) the number of solvent molecules which can be accommodated in the inner and outer shells of the cations;
 - 6) the structural reorganization around the proton produced by the addition of solvent molecules and the relationships between structure, charge transfer, and energetics;
 - 7) the importance of diffuse functions in determining the structures and relative stabilities of symmetric and asymmetric $\text{OH}^- \cdot \text{H}_2\text{O}$ species.

III. COMPUTATIONAL DETAILS

The calculations were carried out *ab initio* using the Gaussian 80 computer program³⁵ on a DEC VAX 11/780 computer. The closed shell ground state wavefunctions for the species investigated are described by single Slater determinants composed of doubly occupied orbitals. The molecular orbitals are expressed as linear combinations of atomic basis functions. Moller-Plesset perturbation theory³⁶ terminated at the second order (MP2) was applied to compute electron correlation effects. The frozen core approximation, where only valence-shell orbitals were correlated, was used.

Several different basis sets were utilized in the work. The 3-21G and 4-31G basis sets are s,p split valence basis sets. 6-31G** is similar but includes also polarization functions of d type on atoms other than hydrogen and of p type on the hydrogens. When diffuse sp functions on atoms other than hydrogen are included in the basis, the symbol + appears in its designation, i.e., 6-31+G**.

Standard scale factors and exponents were employed for all basis functions other than the diffuse functions. The scale factor for the diffuse functions was 1.0. The exponent for the diffuse functions added to oxygen was 0.068.²⁰

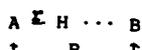
Geometry optimizations were done by the force relaxation method (HF)³⁷ or by the Fletcher-Powell method (MP2).³⁸ Reported bond lengths represent convergence to 0.001Å and bond angles to 0.1°. The total energies (E_T) obtained were then

used to compute stabilization energies ($\Delta E_{n-1,n}$) according to the following equation.

$$\Delta E_{n-1,n} = E_T(A^{\ddagger}B_n) - E_T(A^{\ddagger}B_{n-1}) - E_T(B) \quad (7)$$

No zero-point corrections were made for any of the calculated energies.

Full geometry calculations were carried out for $\text{OH}^- \cdot \text{H}_2\text{O}$, symmetric and asymmetric $\text{H}^+(\text{HCN})_2$, $\text{H}^+(\text{HCN})_3$, $\text{H}^+(\text{H}_2\text{O})(\text{CH}_3\text{CN})$, $\text{H}^+(\text{CH}_3\text{CN})$, and CH_3CN . Partial geometry optimizations were done for asymmetric $\text{H}^+(\text{HCN})_2$, $\text{H}^+(\text{HCN})_3$, and asymmetric $\text{H}^+(\text{HCN})_4$. In the latter optimizations, the calculated optimum geometries of the monomers, HCNH^+ and HCN , were retained in the complex ion and only the distances R and r (defined below) were optimized.



In the above diagram AH is the proton donor and B is the electron donor.

There are some systematic errors, compared to experimental values, in the geometrical parameters and some other molecular properties computed via the 3-21G and 4-31G basis sets.^{32a,39,40} Both basis sets generally underestimate bond lengths and overestimate bond angles, especially those angles involving atoms with lone pair electrons. The calculated dipole moments are often too large as are the calculated solvation enthalpies, since no corrections are made in the latter values for inaccuracies resulting from neglecting changes in zero-point energies, electron correlation, and a temperature term. (The computed enthalpies are for $T = 0^\circ\text{K}$.) Nevertheless, relative values of these properties along a series of molecules are well reproduced which warrants their use, particularly for qualitative comparisons in related systems. Including polarization functions and diffuse functions (for negative ions) in the basis set and calculating electron correlation effects improves the results. This has been done when it was feasible. Further improvement would be obtained by computing zero-point energies and carrying out a molecular dynamics study of these species. However, these calculations were not made because the required programs were not available.

IV. RESULTS AND ANALYSIS OF RESULTS

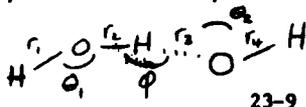
$\text{OH}^- \cdot \text{H}_2\text{O}$. One point of interest with respect to the geometry of $\text{OH}^- \cdot \text{H}_2\text{O}$ is whether the ion is symmetric or not. Earlier studies have shown that calculations carried out with a limited basis set^{17,18} and/or only partial geometry optimization⁴¹

are not sufficiently accurate to settle this point. Later work¹⁹ undertaken by performing a global geometry optimization with a correlated wavefunction indicates that this ion is symmetric. However, no diffuse functions were included in the basis of any of these prior calculations, and diffuse functions are known to be important for a proper description of anionic species.^{20,21} Consequently, optimized geometries and energies of structures with both symmetric and asymmetric hydrogen bonds were computed utilizing the MP2/4-31+G and MP2/6-31+G** basis sets. Calculations using the MP2/4-31+G basis were included to test whether this less accurate but more economical basis set would yield reliable results. The optimized structures and total energies of OH⁻·H₂O and OH⁻ and H₂O, for purposes of comparison, are listed in Table I.

TABLE I. Structure and Energies of OH⁻·H₂O, H₂O, and OH⁻.

Method	r ₁ ^a (Å)	r ₂ ^a (Å)	r ₃ ^a (Å)	r ₄ ^a (Å)	θ ₁ ^a (°)	θ ₂ ^a (°)	φ _{a (°)}	∠ _a (°)	Energy (hartrees)
OH ⁻ ·H ₂ O									
MP2/6-31G**(sym) ^b	0.965	1.223	1.223	0.964	99.1	99.1	180.0	110.0	-151.82012
MP2/4-31+G(sym) ^c	0.976	1.244	1.244	0.976	112.2	112.2	180.0	130.8	-151.53350
MP2/4-31+G(asy) ^c	0.971	1.048	1.556	0.980	108.1	121.5	174.4	136.9	-151.53589
MP2/6-31+G**(sym) ^c	0.966	1.222	1.222	0.966	105.2	105.2	180.0	110.9	-151.87531
MP2/6-31+G**(asy) ^c	0.964	1.085	1.420	0.967	102.4	108.0	177.4	111.7	-151.87574
MP2/6-31+G**(asy) ^{c,d}	0.964	1.073	1.452	0.968	103.1	113.3	179.4	0.0	-151.87328
MP2/6-31+G**(asy) ^{c,d}	0.964	1.083	1.426	0.968	102.8	108.3	178.2	60.0	-151.87474
MP2/6-31+G**(asy) ^{c,d}	0.964	1.087	1.415	0.968	102.5	108.1	177.7	90.0	-151.87557
MP2/6-31+G**(asy) ^{c,d}	0.964	1.080	1.430	0.967	102.9	111.8	177.1	180.0	-151.87526
H ₂ O									
MP2/4-31+G ^e	0.969	0.969			111.7				-76.04841
MP2/6-31+G** ^e	0.964	0.964			105.5				-76.23042
OH ⁻									
MP2/4-31+G ^e	0.988								-75.44139
MP2/6-31+G** ^e	0.971								-75.60197

^aSee Figure below for definitions. ^bReference 19. No stable asymmetric form could be located. ^cThis work. ^dThe dihedral angles were fixed at the values given in the table. All other parameters were optimized. ^eReference 16.



Examination of Table I shows that the MP2/6-31G** and MP2/6-31+G** optimized geometries are very similar with the exception of Θ_1 and Θ_2 . This demonstrates once again that including diffuse functions in the basis has very little effect on optimized geometries at this level of calculation. (This is not true when less accurate calculations are performed.)¹⁶ The MP2/6-31+G** computed values of r_1 , r_2 , Θ_1 , and Θ_2 , 0.966Å and 105.2°, respectively, in symmetric $\text{OH}^- \cdot \text{H}_2\text{O}$ are very close to the computed values of r , 0.064Å, and Θ , 105.5°, in H_2O . Since the theoretical Θ in H_2O is 1° larger than the experimental angle, it is likely that Θ_1 and Θ_2 are slightly too large as well. Nevertheless, 105.2° seems to be a more reasonable value for these angles than 99.1°.

Unlike the results of McMichael Rohlifing et al.¹⁹, a stable asymmetric form was located at the MP2/6-31+G** level. The following are some features of the asymmetric system. 1) The asymmetric form can still be considered as $\text{H}_2\text{O} + \text{OH}^-$ although the hydrogen in the hydrogen bond has moved 0.12Å along the bond axis. 2) r_1 is longer than r_4 as expected, since the O-H bond in H_2O is shorter than the O-H bond in OH^- .^{16, 42, 43} r_1 and r_4 are closer to the O-H distance in H_2O than in OH^- . 3) The O...O separation is greater in asymmetric than in symmetric $\text{OH}^- \cdot \text{H}_2\text{O}$. Other groups have observed this lengthening, also.^{19, 41} 4) The hydrogen bond is non-linear and Θ_2 is larger than Θ_1 , as a result of electrostatic effects.¹⁹ The difference in Θ_1 and Θ_2 is larger for these calculations than for other calculations.¹⁹ For this work Θ_2 is comparable to the HFH angle in $(\text{HF})_2$.⁴⁴ 6) The asymmetric form is non-planar as is the symmetric form.

In contrast to the MP2/6-31G** results,¹⁹ the MP2/6-31+G** results imply that $\text{OH}^- \cdot \text{H}_2\text{O}$ is asymmetric. However, McMichael Rohlifing et al.¹⁹ carried out their MP2 calculations using all MOs rather than the frozen core approximation utilized in this work. The near equivalence of the energies of the symmetric and asymmetric forms at the MP2/6-31+G** level suggests that inclusion of all MOs in the MP2 analysis could invert their relative energies, as could a CID calculation. Consequently, these computations are now underway in this laboratory.

The closeness of the total energies of the symmetric and asymmetric systems indicates a quite flat potential for transfer of the hydrogen within the hydrogen bond. An examination of the change in total energy as the dihedral angle of asymmetric $\text{OH}^- \cdot \text{H}_2\text{O}$ is varied shows that the potential surface for torsional motion is also quite flat. The theoretical barrier height is 0.067 eV. McMichael Rohlifing et al. obtained similar results.¹⁹

The MP2/4-31+G bond lengths and angles are too large compared to the MP2/6-31+G**

and MP2/6-31G** values, in accordance with trends observed for other molecules and ions.¹⁶ In addition, the energy difference between the symmetric and asymmetric forms is too large. However, the overall features of these systems are reproduced quite well at the MP2/4-31+G level.

The adiabatic electron affinity of $\text{OH}\cdot\text{H}_2\text{O}$ can be computed from the MP2/6-31+G** total energies of $\text{OH}^-\cdot\text{H}_2\text{O}$, OH , and H_2O by assuming that the bond dissociation energy (D) of $\text{OH}\cdot\text{H}_2\text{O}$ is negligible. The value obtained is 2.90 eV which is in good agreement with the experimental value of 2.91 eV.¹⁵ The latter is derived from E.A.(OH) and $D(\text{OH}^-\cdot\text{H}_2\text{O})$ assuming $D(\text{OH}\cdot\text{H}_2\text{O})$ is negligible. The calculated value of $D(\text{OH}^-\cdot\text{H}_2\text{O})$, 1.04 eV (corrected for zero-point energies) is also in good agreement with the experimental one, 1.08 eV.⁴⁵ MP2/6-31+G** atomic charges from population analysis⁴⁶ for symmetric and asymmetric $\text{OH}^-\cdot\text{H}_2\text{O}$ are given in the figure below.

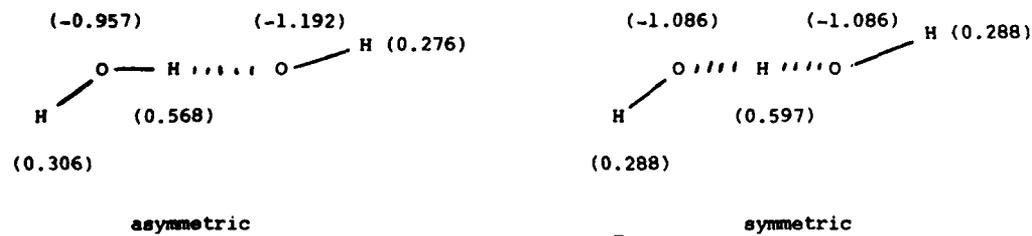


Fig. 1. Atomic Charges for $\text{OH}^-\cdot\text{H}_2\text{O}$.

The figure demonstrates that there is much less charge transferred from the OH^- to the H_2O upon formation of the hydrogen bond in the asymmetric form (0.084 e) than in the symmetric form (0.202 e). The electron density is transferred to the oxygen and terminal hydrogen of the water. However, most of the gain in electron density on these two atoms comes from the central hydrogen which has gained positive charge compared to its value in water (+0.370). Similar charge redistributions have been found in other studies of hydrogen bonded systems.⁴⁷

$\text{H}^+(\text{H}_2\text{O})_m(\text{CH}_3\text{CN})_k$ Ions. In order to study the charge redistributions and structure changes which occur upon substitution of $\text{H}^+(\text{H}_2\text{O})_n(\text{CH}_3\text{CN})_k$ ions, it is necessary to know the optimum geometries and atomic charge distributions of CH_3CN , $\text{H}^+(\text{CH}_3\text{CN})$, $\text{H}^+(\text{H}_2\text{O})_n$, and H_2O for reference. In an earlier report, the optimum geometries and total energies of some of these species were determined at the 4-31G basis set level.⁴⁸ This work has been repeated using the more economical 3-21G basis set (which was not available earlier), since the results obtained are comparable to the 4-31G results

yet considerably less CPU time is expended.⁴⁰ The latter will be especially significant when the larger cations are examined.

Wavefunctions for CH_3CN , H_2O and H_3O^+ have been calculated at the 3-21G level⁴⁹ but, to my knowledge, 3-21G wavefunctions have not been computed previously for the other systems. The 3-21G totally optimized structures of CH_3CN , H_2O , H_3O^+ , and $\text{H}^+(\text{CN}_3\text{CN})$ and preliminary results for $\text{H}^+(\text{H}_2\text{O})_2$, $\text{H}^+(\text{CN}_3\text{CN})_2$, and $\text{H}^+(\text{H}_2\text{O})(\text{CH}_3\text{CN})$ are presented in Figure 2. The 4-31G geometries (in parentheses, where available) are included also for reference. Table II tabulates the calculated total energy (E_T) of each of the compounds and selected proton affinities (P.A.).

TABLE II. Total Energies (a.u.) and Proton Affinities (kcal/mole)

System	E_T (3-21G)	E_T (4-31G)	P.A.(calc) 3-21G	P.A.(calc) 4-31G	P.A.(expt)
H_2O	-75.58596 ^a	-75.9086 ^c	191.6	183.2	170.3 ^e
H_3O^+	-75.89123 ^a	-76.2006 ^c			
$\text{H}^+(\text{H}_2\text{O})_2$	-151.55916 ^b	-152.1770 ^c			
CH_3CN	-131.19180 ^a	-131.7283 ^d	198.9	198.6	187.0 ^e
$\text{H}^+(\text{CH}_3\text{CN})$	-131.50875 ^b	-132.0445 ^d			
$\text{H}^+(\text{CH}_3\text{CN})_2$	-262.75653 ^b	—			
$\text{H}^+(\text{H}_2\text{O})(\text{CH}_3\text{CN})$	-207.15909 ^b	-208.0040 ^d			

^aReference 49. ^bThis work. ^cReference 32a. ^dReference 48. ^eReference 51.

In these computations, $\text{H}^+(\text{CH}_3\text{CN})$ was assumed to be linear and N-protonated in agreement with the 4-31G results.⁴⁸ The structure of each of the dimers was partially optimized according to the procedure described above in the Computational Details section. All hydrogen bonds were found to be linear. This procedure is reasonable as long as the hydrogen bond is not symmetric.^{47,48} It is clear from the results that this conformer of $\text{H}^+(\text{H}_2\text{O})_2$ has a symmetric hydrogen bond, and a global optimization of this conformer must be carried out now. In addition, several other conformations of $\text{H}^+(\text{H}_2\text{O})_2$, staggered $\text{H}^+(\text{CH}_3\text{CN})(\text{H}_2\text{O})$, and symmetric $\text{H}^+(\text{CH}_3\text{CN})_2$ must be considered. No other conformers of $\text{H}^+(\text{CH}_3\text{CN})_2$ or $\text{H}^+(\text{CH}_3\text{CN})(\text{H}_2\text{O})$ need to be calculated, however, since earlier results showed that the methyl hydrogen-methyl hydrogen and methyl hydrogen-water hydrogen non-bonded interactions are negligible and do not affect the total energy.⁴⁸

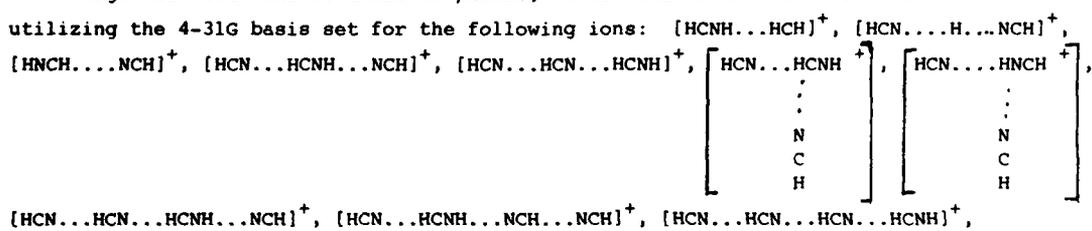
Examination of Table II and Figure 2 shows that the 4-31G relative energies along

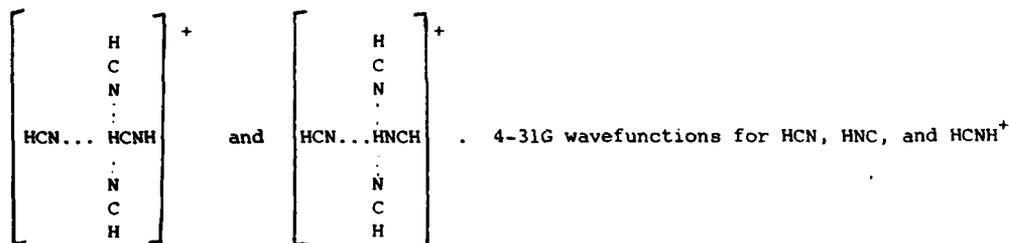
a series of molecules are reproduced quite well at the 3-21G level. Optimum bond lengths and bond angles are very similar as well for the two basis sets. One noticeable difference is the distance the hydrogen in the hydrogen bond is transferred along the hydrogen bond axis in $H^+(H_2O)(CH_3CN)$. The hydrogen has moved considerably further along the bond axis toward the oxygen in the 3-21G calculation than in the 4-31G calculation. This can be explained by comparing the relative proton affinities of CH_3CN and H_2O at the two basis set levels. The difference in the values of the 3-21G computed proton affinities is much smaller than that of the 4-31G values. Thus, at the 3-21G level, the H_2O competes more effectively for the hydrogen-bonded proton. Although the 4-31G result is closer to the experimental result, the 3-21G computations do yield the correct order of the proton affinities.

An analysis of the structural changes and charge redistributions which occur upon substitution of CH_3CN for H_2O in the water cluster ions and a rationalization for the number of water molecules that can be substituted in these complexes await further results.

$H^+(HCN)_n$. Recently Meot-ner²⁶ studied the solvation of $HCNH^+$ by HCN via a pulsed high pressure mass spectrometric technique. The general reaction is represented as $(HCN)_{n-1} H^+ + HCN \rightleftharpoons (HCN)_n H^+ (n=1-4)$. Temperature studies on the solvation reactions yielded values for $\Delta H_{n-1,n}^{\circ}$ and $\Delta S_{n-1,n}^{\circ}$. Meot-ner compared the $\Delta H_{n-1,n}^{\circ}$ values obtained for $HCNH^+$ with those found for H_3O^+ to gain some insight into the importance of hydrogen bonding in the ion-solvent interactions. From this comparison, he noted that the heat of solvation of $HCNH^+$ decreases more in going from the (1,2) to the (2,3) equilibrium. Meot-ner interpreted this larger decrease as evidence that $HCNH^+$, unlike H_3O^+ , has only one site for hydrogen bonding at the protonated center, i.e., NH^+ . This observation combined with the discontinuity found in the $\Delta H_{n-1,n}^{\circ}$ value going from the (2,3) to the (3,4) equilibrium led Meot-ner to predict that the $H^+(HCN)_n$ ions are linear, with each one-dimensional analogue of a solvent shell containing two HCN molecules. His overall conclusion was that hydrogen bonding is an important factor in stabilizing ion-solvent complexes.

In order to obtain additional information on the structures, solvent shells, and charge distributions of these complexes, wavefunctions have been calculated utilizing the 4-31G basis set for the following ions:





4-31G wavefunctions for HCN, HNC, and HCNH⁺

were computed previously.⁵²

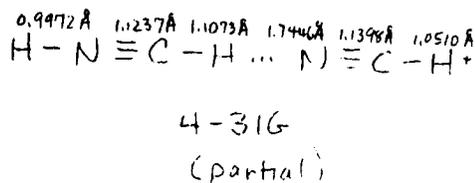
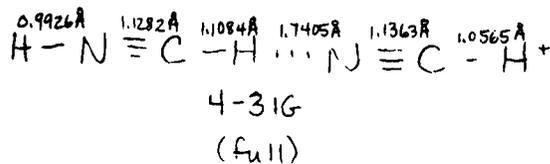
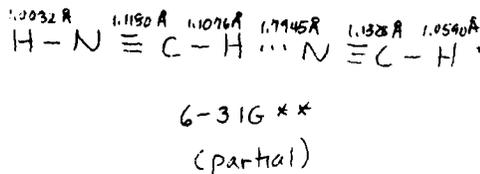
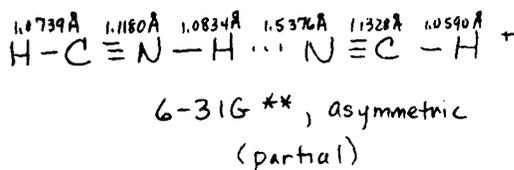
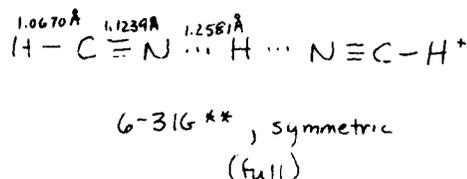
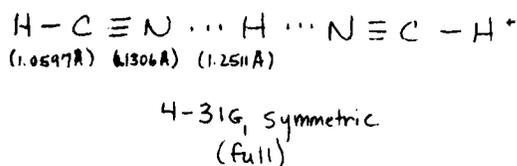
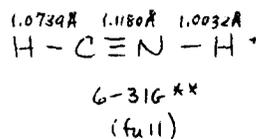
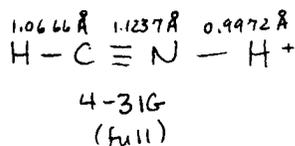
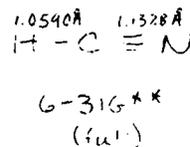
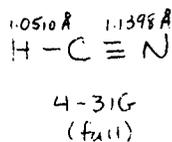
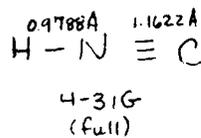
The effect of polarization functions on the geometries and relative energies of these complexes was investigated by calculating HCN, HCNH⁺, [HCN...H...NCH]⁺, [HCN...HNCH]⁺, and [HCN...HCNH]⁺ using the 6-31G** basis. Electron correlation effects on the relative energies were determined by performing MP2/4-31G//HF/4-31G and MP2/6-31G**//HF/6-31G** computations. (The notation MP2/4-31G//HF/4-31G designates that the total energy of the ion was calculated at the MP2/4-31G level using the optimized geometry obtained at the HF/4-31G level).

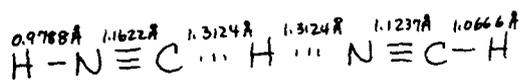
A partial geometry optimization was carried out for the asymmetric hydrogen-bonded systems; a complete geometry optimization was carried out for the symmetric hydrogen-bonded systems. (A global optimization of [HCN...HNCH]⁺, [HCN...HCNH]⁺, and [HCN...HNCH...NCH]⁺ showed that the bond lengths and total energies of the partially and completely optimized asymmetric ions are negligibly different.)

Hirao et al.²⁸ have also performed an ab initio study of H⁺(HCN)_n ions with the 4-31G basis set. However, their study is much less extensive than this one in terms of both the number of basis sets used and the number of complexes considered. For example, they did not consider any complexes involving [HNCH...NCH]⁺, which is a possible form of the dimer. The convergence levels for their optimizations must have been less stringent as well, since their fully optimized total energies are less stable than my partially optimized ones and there are differences of as much as 0.25^o in the structural parameters. In addition, in Hirao et al.'s.²⁸ work, the geometrical parameters were fully optimized only for the H⁺(HCN)_n ions where n ≤ 3. For the linear H⁺(HCN)_n ions where n=4,5 the geometries of HCN and H⁺(HCN)₃ were assumed to remain fixed at their calculated optimum values and only the distance between them was varied. This work shows that there are significant structural differences in the linear trimer and tetramer; thus, the above assumption is not reasonable. Furthermore, in order to be able to compare total and relative energies, it is necessary to use a consistent approach in computing them.

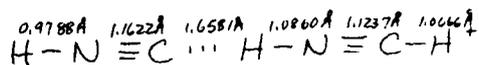
Figure 3 shows the HF/4-31G and HF/6-31G** totally and partially optimized structures of the various complexes. All of the hydrogen bonds were found to be linear indicating that they are formed via the directed lone pair of electrons on

the electron-donating nitrogen rather than through the π -electron system.⁵³
 Table III gives the total energies of the complexes together with calculated $\Delta E_{n-1,n}^{\circ}$
 and experimental $\Delta H_{n-1,n}^{\circ}$ data. Net atomic charges from population analysis⁴⁶ are
 presented in Figure 4.

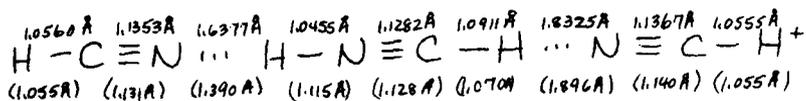




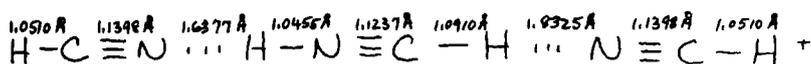
4-31G
(partial)



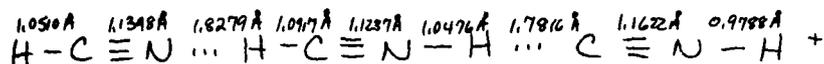
4-31G
(partial)



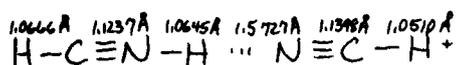
4-31G
(full)



4-31G
(partial)

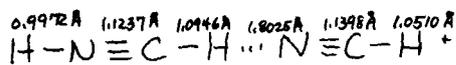


4-31G
(partial)



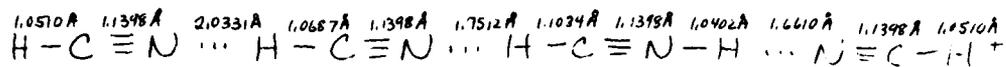
: 2.8963\text{\AA}
 * (2.972\text{\AA})
 N
 ||| 1.1398\text{\AA}
 C
 | 1.0510\text{\AA}
 H

4-31G
(partial)

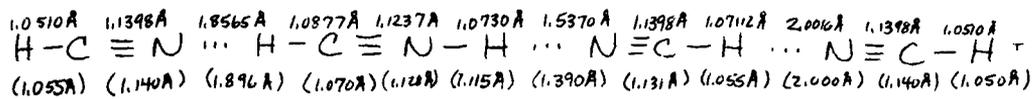


: 2.6542\text{\AA}
 N
 ||| 1.1398\text{\AA}
 C
 | 1.0510\text{\AA}
 H

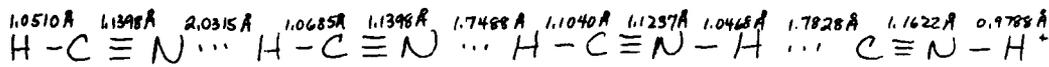
4-31G
(partial)



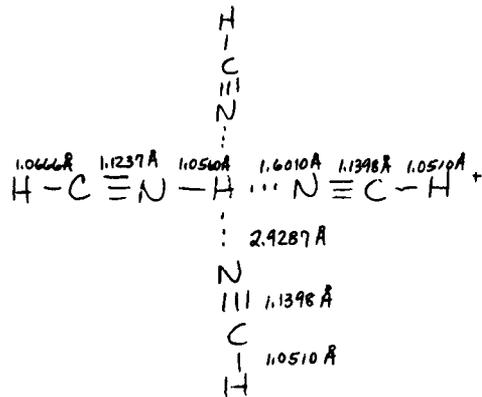
4-31G
(partial)



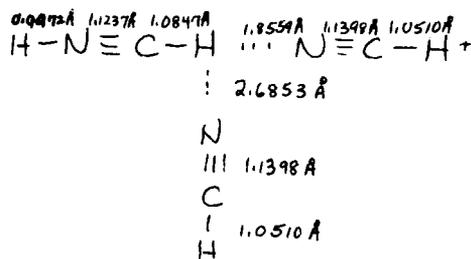
4-31G
(partial)



4-31G
(partial)



4-31G
(partial)



4-31G
(partial)

Fig. 3. The 4-31G optimized structures of HCN, HNC, and HCNH⁺ are from reference 52. Hirao et al's.²⁸ optimized parameters are given in parentheses. All other geometries are from this work. For each complex it is noted whether a partial or full optimization was carried out.

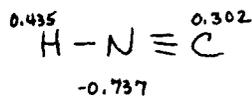
TABLE III. Total Energies (E_T) and Stabilization Energies ($\Delta E_{n-1,n}$ and $\Delta H_{n-1,n}$) of the $H^+(HCN)_n$ Ions.

Cluster	Method ^a	E_T (a.u.)	$-\Delta E_{n-1,n}$ (kcal/mole)	$-\Delta H_{n-1,n}^{\circ}$ (kcal/mole)
HCN	HF/4-31G(full)	-92.73193 ^b		
	HF/6-31G**(full)	-92.87714 ^c		
	MP2/4-31G(full)	-92.94190 ^c		
	MP2/6-31G**(full)	-93.16176 ^c		
HNC	HF/4-31G(full)	-92.71678 ^b		
HNCH ⁺	HF/4-31G(full)	-93.02235 ^b	182.2	174.5 ^d
	HF/6-31G**(full)	-93.16530 ^c	180.8	
	MP2/4-31G(full)	-93.22649 ^c	178.6	
	MP2/6-31G**(full)	-93.44515 ^c	177.8	
HCN...H...NCH ⁺ (symmetric)	HF/4-31G(full)	-185.79941 ^e	28.3	30.0 ^f
	HF/6-31G**(full)	-186.08313 ^c	25.5	
	MP2/4-31G(full)	-186.22089 ^c	32.9	
	CISD/4-31G(full)	-186.17081 ^e	30.3	
	MP2/6-31G**(full)	-186.65928 ^c	32.9	
HCNH...NCH ⁺ (asymmetric)	HF/4-31G(full)	-185.80012 ^e	28.8	
	HF/4-31G(full)	-185.80094 ^c	29.3	
	HF/4-31G	-185.80080 ^c	29.2	
	HF/6-31G**	-186.08436 ^c	26.3	
	MP2/4-31G	-186.22089 ^c	32.9	
	CISD/4-31G(full)	-186.17063 ^e	30.2	
	MP2/6-31G**	-186.65790 ^c	32.0	
HNCH...NCH ⁺	HF/4-31G(full)	-185.79367 ^e	24.7	
	HF/4-31G	-185.79353 ^c	24.6	
	HF/6-31G**	-186.07655 ^c	21.4	
	MP2/4-31G	-186.20736 ^c	24.5	
	MP2/6-31G**	-186.64366 ^c	23.1	
HNC...H...NCH ⁺	HF/4-31G	-185.78326 ^c	—	
HNC...HNCH ⁺	HF/4-31G	-185.78262 ^c	—	
HCN...HNCH...NCH ⁺	HF/4-31G(full)	-278.56202 ^e	18.8	13.8
	HF/4-31G(full)	-278.56512 ^c	20.2	
	HF/4-31G	-278.56493 ^c	20.1	
	MP2/4-31G	-279.19423 ^c	19.7	

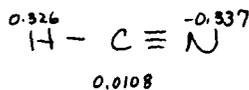
TABLE III. (cont'd)

Cluster	Method ^a	E_T (a.u.)	$-\Delta E_{n-1,n}$ (kcal/mole)	$-\Delta H_{n-1,n}^{\circ}$ (kcal/mole)
HCN...HCNH...CNH ⁺	HF/4-31G	-278.54789 ^c	14.0	
	MP2/4-31G	-279.16432 ^c	9.5	
HCNH...NCH ⁺	HF/4-31G	-278.54685 ^c	8.9	
NCH	HF/4-31G	-278.54584 ^e	8.7	
HNCH...NCH ⁺	HF/4-31G	-278.54390 ^c	11.6	
NCH				
HCN...HCN...HCNH...NCH ⁺	HF/4-31G	-371.31660 ^c	12.3	11.8
HCN...HCNH...NCH...NCH ⁺	HF/4-31G	-371.31813 ^c	13.2	
	HF/4-31G	-371.31699 ^e	14.5	
HCN...HCN...HCNH...CNH ⁺	HF/4-31G	-371.29906 ^c	12.1	
NCH				
HCNH...NCH ⁺	HF/4-31G	-371.29146 ^c	8.0	
NCH				
NCH				
HNCH...NCH ⁺	HF/4-31G	-371.29227 ^c	10.3	
NCH				

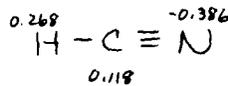
^aFull or partial geometry optimizations were performed using the HF/4-31G or HF/6-31G** basis set and then this structure was used to compute the energy at either the HF or MP2 level. CISD denotes a configuration interaction calculation where all single and double substitutions have been taken into account. ^bReference 52. ^cThis work. ^dReference 51. ^eReference 28. ^fReference 26.



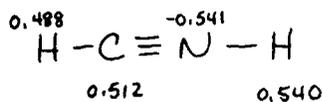
4-31G



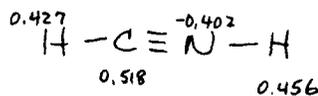
4-31G



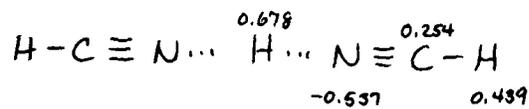
6-31G**



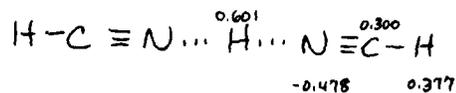
4-31G



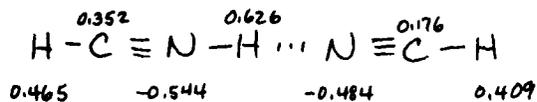
6-31G**



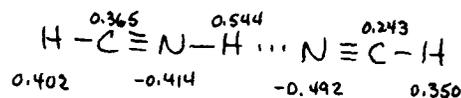
4-31G, symmetric



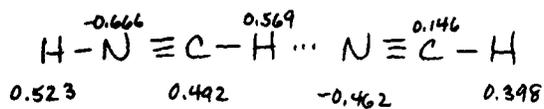
6-31G**, symmetric



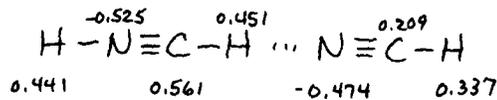
4-31G



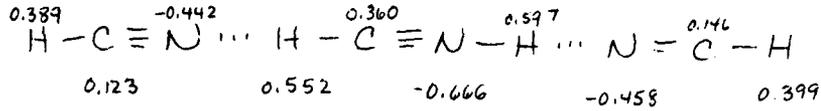
6-31G**



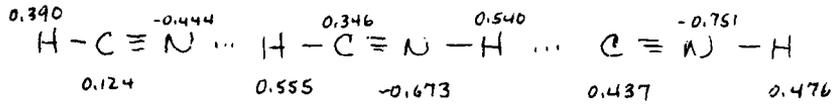
4-31G



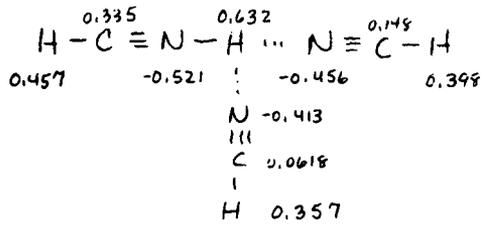
6-31G**



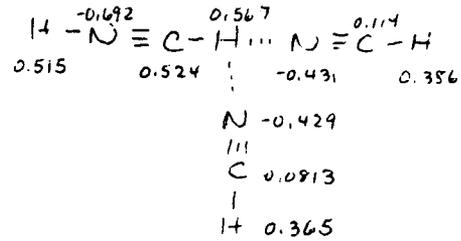
4-31G



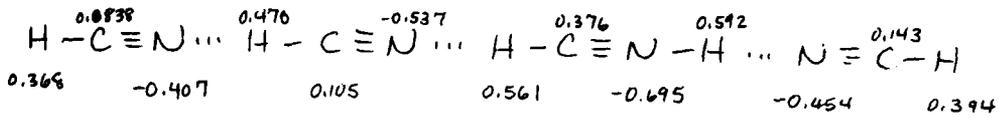
4-31G



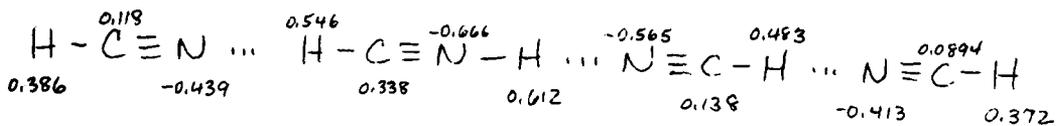
4-31G



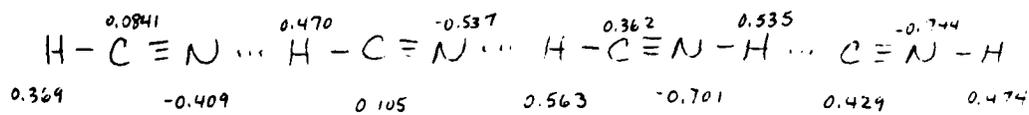
4-31G



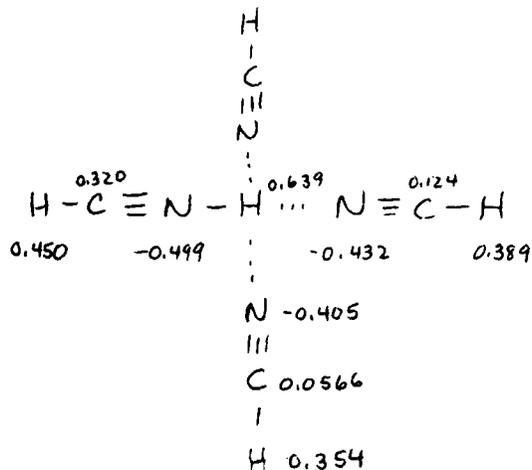
4-31G



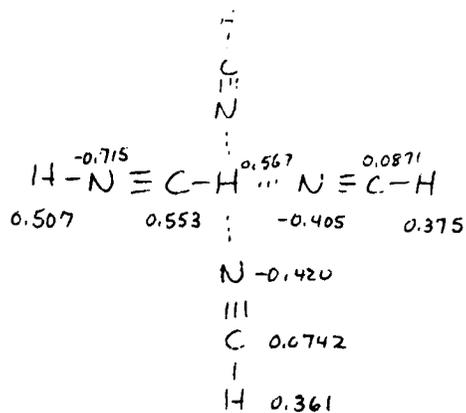
4-31G



4-31G



4-31G



4-31G

Fig. 4. Charge distributions for full optimized symmetric and partially optimized asymmetric hydrogen-bonded systems. Charge distributions for HCN, HCNH⁺, and symmetric H⁺(HCN)₂ are from reference 28. All other are from this work.

The HF/4-31G results for $[\text{HCNH}\dots\text{NCH}]^+$ and $[\text{HNCH}\dots\text{NCH}]^+$ (Table III) show an energy difference of 4.6 kcal for these ions. The $[\text{HNCH}\dots\text{NCH}]^+$ dimer is less stable since carbon is a poorer proton donor than nitrogen. Although the HF/4-31G $\Delta E_{1,2}$ for $[\text{HNCH}\dots\text{NCH}]^+$, 24.6 kcal, is considerably smaller than the observed $\Delta H_{1,2}^\circ$, the large magnitude of the solvation energy and the relatively small difference of 4.6 kcal in the $\Delta E_{1,2}$ values for these ions suggested that the $\text{H}^+(\text{HCN})_2$ dimer observed by Meot-ner²⁶ is a mixture of $[\text{HCNH}\dots\text{NCH}]^+$ and $[\text{HNCH}\dots\text{NCH}]^+$. To test this suggestion, it was decided to carry out more accurate calculations for these systems to obtain more reliable energies.

Adding polarization functions to the basis set (HF/6-31G**) lowers the computed $\Delta E_{1,2}$ values but changes the difference in them very little (by 0.3 kcal). When correlation effects are accounted for in the MP2/6-31G**//HF/6-31G** calculations, the magnitudes of the $\Delta E_{1,2}$'s increase again but, more importantly, $\Delta E_{1,2}$ for $[\text{HNCH}\dots\text{NCH}]^+$ is now ~ 9.0 kcal less stable than for $[\text{HCNH}\dots\text{NCH}]^+$. Whether the MP2 calculations reproduce the CISD potential surface was tested as follows. Hirao et al.²⁸ have reported CISD/4-31G total energies and $\Delta E_{1,2}$ values for symmetric and asymmetric $[\text{HCNH}\dots\text{NCH}]^+$. For comparison purposes, the MP2 calculations on the monomers and dimers were repeated at the MP2/4-31G//HF/4-31G level. (This was also necessary since it was not feasible to reoptimize the geometries of the trimers utilizing the HF/6-31G** basis set.) The change in the relative energies of the symmetric and asymmetric N...N dimer and the change in the magnitude of $\Delta E_{1,2}$ when correlation effects are included by using configuration interaction were reproduced by using the MP2 method. Furthermore, the solvation energies obtained from the MP2/6-31G**//HF/6-31G** and MP2/4-31G//HF/4-31G computations vary by only about one kcal and the trends in relative energies are the same. The reliability of the MP2 calculated energy differences is supported by this parallelism between them and with the CISD results.

The 9.0 kcal deviation in the solvation energies of the N...N and N...C dimers makes it unlikely that there is a significant amount of the N...C dimer present at the temperatures at which only dimers are formed. However, it does not rule out the presence of the N...C dimer at the lower temperatures at which the trimers and tetramers are formed. The linearity of the van't Hoff plot for trimer formation obtained by Meot-ner²⁶ implies that if there is a mixture of N...C and N...N dimers present, their relative concentrations are essentially constant over the temperature range considered.

There is good agreement between the MP2 $\Delta E_{0,1}$ and $\Delta E_{1,2}$ values and the experimental $\Delta H_{0,1}^\circ$ and $\Delta H_{1,2}^\circ$ values. The 3-4 kcal overestimation of the ΔE 's has been ascribed to the overestimation of the polarity of HCN at these basis set levels.²⁸

Taking the zero-point energies and the temperature effect into account should improve the correlation between ΔE and ΔH° . An approximate figure for the overvaluation of the MP2 trimer solvation energies is obtained by assuming that the temperature correction ($\sim RT$) and the zero-point energy correction are the same for each $\Delta E_{n-1,n}$ for $n > 1$.^{28,32}

The results at all basis set levels demonstrate that the potential for proton motion in the N...N dimer hydrogen bond is quite flat. In contrast, although the 4-31G computed figure of 6.85 kcal is most likely too high,¹³ the energy required to transfer the proton from the C to the N in the N...C dimer hydrogen bond is considerably larger. The potential for proton motion in the latter bond has a maximum value at the fully transferred position but more energy is needed to move the hydrogen from the carbon to the midway point than from the midway point to the nitrogen.

Of the four trimers that could be produced from the dimers, only $[\text{HCN}\dots\text{HCNH}\dots\text{NCH}]^+$ is likely to be formed. The total energies of the two T-shaped trimers and of $[\text{HCN}\dots\text{HCNH}\dots\text{CNH}]^+$ are too unstable compared to the total energy of $[\text{HCN}\dots\text{HCNH}\dots\text{NCH}]^+$ to be present in any significant concentration. Hirao et.al.²⁸ attribute the lower stability of $[\text{HCN}\dots\text{HNCH}]^+$ to the repulsive interaction between the lone pair on the

third HCN and the π -electron density on the two neighboring nitrogens of the HCN...HNCH group. A similar explanation applies to the other T-shaped trimer also. $[\text{HCN}\dots\text{HCNH}\dots\text{CNH}]^+$ is less stable due to the higher energy of CNH in relation to HCN.

At each basis set level considered, the $\Delta E_{n-1,n}$ values for the most stable complexes are within 3 kcal of the $\Delta H^\circ_{n-1,n}$ values, with the exception of $\Delta E_{2,3}$ and $\Delta H^\circ_{2,3}$ which vary by ~ 6 kcal. There does not appear to be any intrinsic reason for the energy of $[\text{HCN}\dots\text{HCNH}\dots\text{NCH}]^+$ to be overly stabilized compared to the other calculated energies. In fact on the basis of energy changes for $[\text{HCN}\dots\text{HNCH}]^+$ and $[\text{HCN}\dots\text{HCNH}]^+$ when electron correlation effects are included, one would expect the most reliable solvation energy, i.e., the MP2 computed $\Delta E_{2,3}$, to be even larger than it is. The difference in $\Delta E_{1,2}$ of the C...N dimer with respect to that of the N...N dimer is larger at the MP2 level than at the HF level, not because the C...N dimer is destabilized but because the N...N dimer is stabilized when electron correlation is incorporated. Combining these effects leads to a larger predicted MP2 solvation energy than the one obtained.

Although more accurate computations would be required for confirmation, the calculated $\Delta E_{2,3}$ values together with the estimated 3-4 kcal energy lowering due to zero-point energy and temperature effects suggest that $\Delta H^\circ_{2,3}$ is ~ 16 kcal rather than 13.8 kcal, as long as there is not a mixture of trimers present at equilibrium. The considerably

higher energy of the trimers other than $[\text{HCN}\dots\text{HCNH}\dots\text{NCH}]^+$ and the probability of rearrangement to the more stable complex should one or more of the other trimers be formed (as a result of the excess energy generated upon complex formation and the long-lived nature of the complexes⁵⁴), make it unlikely that such a mixture exists.

Two of the possible tetramers, $[\text{HCN}\dots\text{HCN}\dots\text{HCNH}\dots\text{NCH}]^+$ and $[\text{HCN}\dots\text{HCNH}\dots\text{NCH}\dots\text{NCH}]^+$, will be present at equilibrium. These two complexes deviate from each other in that the fourth HCN has been added to opposite ends of $[\text{HCN}\dots\text{HCNH}\dots\text{NCH}]^+$. Their total energies vary by 0.960 kcal indicating that there is a small variation in energy depending on which end of the complex is solvated. The more favorable situation is the one in which the N...N hydrogen bond is centrally located. As noted above, the calculated solvation energies are very close in magnitude to the observed solvation enthalpies. $[\text{HCN}\dots\text{HCN}\dots\text{HCNH}\dots\text{CNH}]^+$ and the +-shaped tetramers are less stable for the reasons cited above for the corresponding trimers.

For the linear pentamers, and for higher values of n, there will be a mixture of complexes with similar energies in the system at equilibrium, also. Again the complexes will differ in the position of the N...N hydrogen bond and the HNCH^+ group.

Several observations can be made regarding the data in Figure 3. 1) A partial optimization of the asymmetric hydrogen-bonded systems is all that is required, since there is very little difference in the fully and partially optimized structures. 2) There are only small differences in the 4-31G and 6-31G** structures, although the 6-31G** C-H and N-H bond lengths are consistently longer while the C≡N bond lengths are consistently shorter. The C...N distances are also longer at the 6-31G** level. 3) When the $\text{HCNH}\dots\text{NCH}$ dimer is formed, the N-H bond length increases compared to its length in HCNH^+ but the remaining bond lengths change very little. The same is true for the C-H (in the hydrogen bond) and other bond distances in the $\text{HNCH}\dots\text{NCH}^+$ dimer. Newton and Ehrenson³² observed a similar O-H bond lengthening in $\text{H}^+(\text{H}_2\text{O})_2$ compared to H_3O^+ . The longer C...N distance than N...N distance in these dimers is consistent with the weaker strength of the C...N hydrogen bond.⁴⁷ 4) Trimers can be formed by adding an HCN to either end or the middle of the two possible dimers. The two linear additions to the $\text{HCNH}\dots\text{NCH}^+$ dimer, $\text{HCNH}\dots\text{NCH}\dots\text{NCH}^+$ (1) and $\text{HCN}\dots\text{HCNH}\dots\text{NCH}^+$ (2), yield the same trimer, i.e. (2), since the hydrogen-bonded H migrates from the end HCN to the central HCN in trimer (1). This hydrogen migration demonstrates that the core of these complexes is HCNH^+ and not H^+ . Therefore, the complexes would be more correctly designated as $\text{HCNH}^+(\text{HCN})_m$. This is less true in Hirao et al.'s structures.²⁸ 5) Further evidence for the above is obtained from one of the trimers yielded from the linear addition of HCN to $\text{HNCH}\dots\text{NCH}^+$. $\text{HNCH}\dots\text{NCH}\dots\text{NCH}^+$ becomes $\text{HNC}\dots\text{HNCH}\dots\text{NCH}^+$, as again the hydrogen migrates to form the HCNH^+ core. The other linear addition to $\text{HNCH}\dots\text{NCH}^+$

produces trimer (2). 6) The hydrogen-bonded N-H and C-H distances are shorter and the N...N and C...N distances are longer in the linear trimers than they are in the dimers. These structural changes are also observed for the $H_3O^+(H_2O)_n$ ($n=2-5$) ions and demonstrates the cooperative interactions between hydrogen bonds.³² Shorter N-H or C-H and longer N...N or C...N bond lengths correspond to weaker hydrogen bonds.^{32,47} Therefore, both hydrogen bonds in the trimer are weaker than the analogous hydrogen bonds in the dimers. The formation of the first hydrogen bond to $HCNH^+$ reduces its capacity to act as a proton donor in the second hydrogen bond in the trimer.³² However, the second hydrogen bond is not as weak as it could be because a compromise is achieved whereby the first hydrogen bond is weakened relative to its strength in the dimer to allow the second hydrogen bond to strengthen somewhat. Thus, it is more favorable to form two moderately strong, partially protonated, hydrogen bonds rather than one strong, essentially fully protonated, and one weak, essentially neutral, hydrogen bond.³² 7) The distance of the third HCN from the linear dimer in the T-shaped trimers is $\sim 0.3\text{\AA}$ less for the C...N dimer than for the N...N dimer. The reason for this is that the C and N are farther apart in the C...N bond than the N's are in the N...N bond; consequently, the π -electron density on the C and N atoms does not block the approach of the third HCN molecule as much as it does on the neighboring N atoms. In the dimer portion of the T-shaped trimers, the hydrogen-bonded C-H and N-H bond lengths are shorter and the N...N and C...N bond lengths are longer than they are in the dimers. However, the changes are not as large as they are for the linear trimers. 8) The cooperativity effect in hydrogen bonding is exhibited in the tetramers in the following way. Bonding the fourth HCN to one of the end HCN's in the trimer both lengthens the C-H bond of the trimer HCN and makes its nitrogen atom a better electron donor. This in turn strengthens the hydrogen bond between it and the $HCNH^+$, which lengthens the C-H or N-H distance and shortens the C...N or N...N distance in this hydrogen bond. By way of compensation, the other hydrogen bond involving the $HCNH^+$ is weakened and the opposite changes occur for the hydrogen-bonded distances. 9) The C...N bond length for the last HCN added is essentially equivalent for all three linear tetramers studied and is $\sim 0.10\text{\AA}$ shorter than the C...N distance in neutral $HCN...HCN$, 3.2\AA .⁵⁵ The other C...N bonds in the complexes are considerably shorter than this. A similar correlation is seen between the neutral O...O bond length in $H_2O...HOH$ and the O...O bond length in the outer solvation shells of the $H_3O^+(H_2O)_n$ ions.^{32,55} 10) For the +-shaped tetramers, the addition of the fourth HCN causes the other HCN's to move away from each other, thereby weakening the other three hydrogen bonds. 11) Hirao et al.'s structures²⁸ have shorter N...N and C-H hydrogen-bonded distances and longer N...C and N-H hydrogen-bonded distances than those reported here. There is also very little change in their

optimized bond lengths when the above groups are involved in a second hydrogen bond. Their results are, therefore, less consistent with my conclusion that the central ion in these complexes is HCNH^+ rather than H^+ . However, my conclusion is substantiated by the following facts. Their results are also less consistent with Newton and Ehrenson's³² observations on the $\text{H}_3\text{O}^+(\text{H}_2\text{O})_n$ clusters, and the total energies calculated in this work are all more stable than those calculated by Hirao et.al.²⁸ Nevertheless, the relatively small differences in the two total energies is additional evidence for facile proton movement in most of these hydrogen bonds.

The HF/4-31G atomic charge distributions for the N...N and N...C dimers (Figure 4) demonstrate that electron density is transferred from the electron donor to the proton donor. In $[\text{HCNH} \dots \text{NCH}]^+$, $0.10e^-$ is transferred to the HCNH group; in $[\text{HNCH} \dots \text{NCH}]^+$ only $0.0820e^-$ is transferred to the HCNH group indicative of the weaker strength of the latter hydrogen bond. Most of the electron density lost from the electron donating HCN moiety comes from the hydrogen and carbon; the nitrogen actually gains electron density. The charge loss on the hydrogen of the HCN group makes this group a better proton donor in an additional hydrogen bond.⁵³ Most of the electron density gained by the HCNH is transferred to the heavy atom (N or C) and hydrogen not involved in the hydrogen bond. The hydrogen-bonded hydrogen actually loses electron density when the complex is formed. This charge redistribution makes the HCNH moiety a poorer proton donor in a subsequent hydrogen bond.⁵³ Similar charge gains and losses on the individual atoms are found in the higher order complexes as well.

In the linear trimer $[\text{HCN} \dots \text{HCNH} \dots \text{NCH}]^+$, the HCNH group now has an overall charge of +0.843, the HCN participating in the N...N hydrogen bond has an overall charge of +0.0870, and the remaining HCN has an overall charge of +0.0700. The HCNH has gained more electron density than it did in the dimers but it still retains 84.3% of the positive charge. Both HCN's donate less electron density to the HCNH moiety than they did in the dimers.

In the linear tetramers, $[\text{HCN} \dots \text{HCN} \dots \text{NCH} \dots \text{NCH}]^+$ and $[\text{HCN} \dots \text{HNCH} \dots \text{NCH} \dots \text{NCH}]^+$, the HCNH retains approximately 83% of the total positive charge, the fourth HCN donates $0.045 - 0.048e^-$, the other singly hydrogen-bonded HCN's donate slightly less electron density than they did in the trimer, and the doubly hydrogen-bonded HCN donates considerably less electron density than it did in the trimer. The latter change occurs because this HCN is acting now as both an electron donor and a proton donor.

Clearly there is no extensive delocalization of positive charge throughout the complexes and the positive charge gained by each HCN decreases as more HCN's are added. Similar charge redistributions were obtained by Desmeules and Allen⁴⁷ and Newton and Ehrenson.³²

One point of interest with respect to the properties of the $H^+(HCN)_n$ cations is whether the first solvent shell of $HCNH^+$ is completed with one or two additional HCN's. To answer this question, Meot-ner²⁶ compared the difference $\Delta\Delta H^\circ = \Delta H^\circ_{1,2} - \Delta H^\circ_{2,3}$ for $H^+(H_2O)_n$ (-12.1 kcal), $H^+(OHCH_3)_n$ (-11.8 kcal), $H^+(HCN)_n$ (-16.4 kcal), $H^+(CH_3CN)_n$ (-20.9 kcal), and $H^+(CH_3OCH_3)_n$ (~-21 kcal). From these data Meot-ner concludes that the first solvent shell of $HCNH^+$ is completed after adding one more HCN, since he interprets the larger drop in $\Delta\Delta H^\circ$ for $HCNH^+$ compared to H_3O^+ and $CH_3OH_2^+$ as indicative of the absence of an unoccupied hydrogen bonding site for the third HCN molecule. Hirao et al.²⁸ agree with this conclusion because they believe the core ion in these clusters is H^+ , and they find that for $n \geq 3$ the central proton is not the most reactive site for further clustering. However, on the basis of my results, I deduce that the first solvent shell is not filled until $HCNH^+$ has been solvated by two HCN's.

The evidence for my deduction is as follows. 1) The total energies, structural data, charge distributions, and proton migrations discussed previously all indicate that the central ion in these complexes is $HCNH^+$ not H^+ . 2) There is another possible interpretation of Meot-ner's data. The first and second solvent molecules added to H_3O^+ and $CH_3OH_2^+$ form the same type of hydrogen bond, i.e., O-H...O. In contrast, they form different types of hydrogen bonds with $HCNH^+$, i.e., N-H...N and C-H...N. The fact that the C-H...N hydrogen bond is inherently weaker than the N-H...N hydrogen bond accounts for the bigger drop in $\Delta\Delta H^\circ$ regardless of whether the first solvent shell is filled or not. Furthermore, the calculations reported in this work suggest that Meot-ner's value for this $\Delta\Delta H^\circ$ is too large. Since the $\Delta\Delta H^\circ$ value for $H^+(HCN)_2$ falls between the value for $H^+(H_2O)_2$ and $H^+(CH_3OH)_2$ (incomplete solvent shell) and for $H^+(CH_3CN)_2$ and $H^+(CH_3OCH_3)_2$ (complete solvent shell) and the C-H...N bond is inherently weaker, I conclude that the first solvent shell is not filled yet in $H^+(HCN)_2$.

Further support for this conclusion is obtained by comparing the structural changes, solvation energies, and charge redistributions found for $HCNH^+(HCN)_n$ with those found for the inner and outer solvent shells of $H_3O^+(H_2O)_n$. When the fourth HCN is added to the $HCNH^+(HCN)_2$ trimer, $\Delta E_{3,4}$ depends very little on which end of the complex is solvated. This is not true for the formation of the trimers. The hydrogen bonds involving the two CH groups in $[HCNH...NCH]^+$ are not of equivalent strength. The calculations show that when a third HCN is added to the above dimer to form $[HCNH...NCH...NCH]^+$, there is a proton migration to produce $[HCN...HNCH...NCH]^+$ so that the $HCNH$ group is solvated on both sides. This is consistent with the cooperativity effects cited previously between the hydrogen bonds in the $HCNH^+$ and H_3O^+ complexes. This cooperativity

yields two medium strong hydrogen bonds rather than one strong and one weak hydrogen bond demonstrating that the inner solvent shell is not filled in the $\text{HCNH}^+(\text{HCN})$ dimer.

For the two most stable linear tetramers, there are considerable differences in the bond lengths of the inner three HCN's but the bond lengths of the fourth HCN's are essentially equivalent. The N...N and N...C distances in the trimers and tetramers are within 0.1\AA of their distances in the dimers, with the exception of the C...N distance for the fourth HCN. This length is more than 0.20\AA longer than it is in the dimer. Analogous changes in the O...O bond lengths are observed when the inner ($0.1 - 0.18\text{\AA}$) and outer (0.34\AA) shell waters are added to H_3O^+ .³² Furthermore, the electron density donated to the HCNH moiety by the last HCN bound in the tetramers is similar to the electron density donated by the outer shell water in $\text{H}_3\text{O}^+(\text{H}_2\text{O})_4$.³² The inner shell waters and remaining HCN's also show a similar pattern of charge redistributions.

V. RECOMMENDATIONS

Specific recommendations for follow-on research include:

1) For $\text{OH}^-\cdot\text{H}_2\text{O}$, carrying out CID and MP2 calculations for the asymmetric and symmetric forms which include all MO's in the analysis of the electron correlation effects.

2) For $\text{H}^+(\text{H}_2\text{O})_m(\text{CH}_3\text{CN})_k$, extending the 3-21G investigation beyond the $\text{H}^+(\text{H}_2\text{O})_2$, $\text{H}^+(\text{CH}_3\text{CN})$, $\text{H}^+(\text{CH}_3\text{CN})_2$, and $\text{H}^+(\text{H}_2\text{O})(\text{CH}_3\text{CN})$ ions to higher order complexes in order to gain some insight into their chemical properties.

REFERENCES

1. Arnold, F., D. Krankowsky, and K.H. Marien, "First Mass Spectrometric Measurements of Positive Ions in the Stratosphere," Nature, 267, 30-31, 1977.
2. Arnold, F., H. Böhringer, and G. Henschen, "Composition Measurements of Stratospheric Positive Ions," Geophys. Res. Lett., 5, 653-656, 1978.
3. Arijs, E., J. Ingels, and D. Nevejans, "Mass Spectrometric Measurement of the Positive Ion Composition in the Stratosphere," Nature, 271, 642-644, 1978.
4. Arijs, E., D. Nevejans, and J. Ingels, "Unambiguous Mass Determination of Major Stratospheric Positive Ions," Nature, 288, 684-686, 1980.
5. Arnold, F. and G. Henschen, "First Mass Analysis of Stratospheric Negative Ions," Nature, 257, 521-522, 1978.
6. Arnold, F. and R. Fabian, "First Measurements of Gas Phase Sulfuric Acid in the Stratosphere," Nature, 283, 55-57, 1980.
7. Arijs, E., D. Nevejans, P. Frederick, and J. Ingels, "Negative Ion Composition Measurements in the Stratosphere," Geophys. Res. Lett., 8, 121-4, 1981.
8. Viggiano, A.A., R.A. Perry, D.L. Albritton, E.E. Ferguson, and F.C. Fehsenfeld, "The Role of H₂SO₄ in Stratospheric Negative Ion Chemistry," J. Geophys. Res., 85, 4551-4555, 1980.⁴
9. Yamdagni, R. and P. Kebarle, "Hydrogen-Bonding Energies to Negative Ions from Gas-Phase Measurements of Ionic Equilibria," J. Am. Chem. Soc., 93, 7139-7143, 1971.
10. Lee, N., R.G. Keesee, and A.W. Castleman, Jr., "The Properties of Clusters in the Gas Phase. IV. Complexes of H₂O and HNO_x Clustering on NO_x⁻," J. Chem. Phys., 72, 1089-1094, 1980.
11. Wu, R.L.C. and T.O. Tiernan, "Formation and Thermodynamic Properties of Hydrated NO₃⁻ Ions," 28th Annual Conference on Mass Spectrometry and Allied Topics, May 1980, New York, New York, pp. 145-146.
12. Kebarle, P., S.K. Searles, A. Zolla, J. Scarborough, and M. Arshadi, "The Solvation of the Hydrogen Ion by Water Molecules in the Gas Phase. Heats and Entropies of Solvation of Individual Reactions: H⁺(H₂O)_{n-1} + H₂O → H⁺(H₂O)_n + H⁺(H₂O)_n," J. Am. Chem. Soc., 89, 6393-6399, 1967.
13. Scheiner, S. and L.B. Harding, "Proton Transfers in Hydrogen-Bonded Systems. 2. Electron Correlation Effects in (N₂H₇)⁺," J. Am. Chem. Soc., 103, 2169-2173, 1981.
14. Keesee, R.G., N. Lee, and A.W. Castleman, Jr., "Atmospheric Negative Ion Hydration Derived from Laboratory Results and Comparison to Rocket-Borne Measurements in the Lower Ionosphere," J. Geophys. Res., 84, 3719-3722, 1979.
15. Paulson, J.F. and F. Dale, "Reactions of OH⁻·H₂O with NO₂," J. Chem. Phys., 77, 4006-4008, 1982.

16. Deakyne, C.A., "A Molecular Orbital Study of Atmospherically Important Species, Final Report, AFOSR-82-0198, 1983.
17. Depaz, M., S. Ehrenson, and L. Friedman, "Study of the H^+ and OH^- Hydrated Ions by the CNDO/2 Method," J. Chem. Phys., 52, 3362-3368, 1970.
18. Jorgensen, W.L. and M. Ibrahim, "Ab Initio Studies of $RO^- \dots HOR'$ Complexes. Solvent Effects on the Relative Acidities of Water and Methanol," J. Comput. Chem., 2, 7-11, 1981.
19. McMichael Rohlffing, C., L.C. Allen, C.M. Cook and H.B. Schlegel, "The Structure of $(H_3O_2)^-$," J. Chem. Phys., 78, 2498-2503, 1983.
20. Chandrasekhar, J., J. G. Andrade, and P.V.R. Schleyer, "Efficient and Accurate Calculation of Anion Proton Affinities," J. Am. Chem. Soc., 103, 5609-5612, 1981.
21. Dunning, T.H. and P.J. Hay, "Gaussian Basis Sets for Molecular Calculations," Methods of Electronic Structure Theory, Plenum Press, New York, pp. 1-27, 1977.
22. Ferguson, E., "Sodium Hydroxide Ions in the Stratosphere," Geophys. Res. Lett., 5, 1035-1038, 1978.
23. Smith, D., N.G. Adams, and E. Alge, "Ion-Ion Mutual Neutralization and Switching Reactions of Some Stratospheric Ions," Planet. Space Sci., 29, 449-454, 1981.
24. a) Arnold, F., G. Henschen, and E.E. Ferguson, "Mass Spectrometric Measurements of Fractional Ion Abundances in the Stratosphere - Positive Ions," Planet. Space Sci., 29, 185-193, 1981; b) Böhringer, H. and Arnold, F., "Acetonitrile in the Stratosphere, Implications from Laboratory Studies," Nature, 290, 321-322, 1981.
25. Paulson, J.F., personal communication.
26. Moet-Ner, M., "Solvation of the Proton by HCN and CH_3CN . Condensation of HCN with Ions in the Gas Phase," J. Am. Chem. Soc., 100, 4694-4699, 1978.
27. Watson, W.D., "Interstellar Molecule Reactions," Rev. Mod. Phys., 48, 513-552, 1976.
28. Hirao, K., S. Yamabo, and M. Sano, "Theoretical Study on the Stability and Structure of $H^+(HCN)_n$ and $M^+(CH_3CN)_n$ ($M^+ = H^+, Li^+,$ and Na^+) Clusters," J. Phys. Chem., 86, 2626-2632, 1982.
29. Lee, N., R.G. Keesee, and A.W. Castleman, Jr., "The Properties of Clusters in the Gas Phase. IV. Complexes of H_2O and HNO_x Clustering on NO_x ," J. Chem. Phys., 72, 1089-1094, 1980.
30. Howell, J.M., A.M. Sapse, E. Singman, and G. Snyder, "Ab Initio Self-Consistent Field Calculations of $NO_2^-(H_2O)_n$ and $NO_3^-(H_2O)_n$ Clusters," J. Phys. Chem., 86, 2345-2349, 1982.
31. Keesee, R.G., N. Lee, and A.W. Castleman, Jr., "Properties of Cluster Ions in the Gas Phase. 3. Hydration Complexes of CO_3^- and HCO_3^- ," J. Am. Chem. Soc., 101, 2599-2604, 1979.

32. a) Newton, M.D. and S. Ehrenson, "Ab Initio Studies on the Structures and Energies of Inner and Outer-Shell Hydrates of the Proton and Hydroxide Ion," J. Am. Chem. Soc., 93, 4971-4990, 1971; b) Newton, M.D., "Ab Initio Studies of the Hydrated H_3O^+ Ion. II. The Energetics of Proton Motion in Higher Hydrates (n=3-5)," J. Chem. Phys., 67, 5535-5546, 1977.
33. Yamabe, S. and K. Hirao, "A Theoretical Study of the Structure and Stability of $H^+(CO)_n$, $H^+(N_2)_n$ and $H^+(O_2)_n$ Clusters (n=1-6)," J. Am. Chem. Soc., 103, 2176-2179, 1981.
34. Yamabe, S. and K. Hirao, "Theoretical Study on the Structure and Stability of $X(CH_3CN)_n$ (X=F and Cl, n=1, 2, 3, 4) Clusters," Chem. Phys. Lett., 84, 598-603, 1981
35. Binkley, J.S., R.A. Whiteside, R. Krishnan, R. Seeger, D.J. DeFrees, H.B. Schlegel, S. Topiol, L.R. Kahn, and J.A. Pople, QCPE, 13, 406, 1981; K. Krogh-Jespersen, private communication.
36. a) Moller, C. and M.S. Plesset, "Approximation Treatment for Many-Electron Systems," Phys. Rev., 46, 618, 1934. b) Pople, J.A., J.S. Binkley, and R. Seeger, "Theoretical Models Incorporating Electron Correlation," Int. J. Quant. Chem. Symp., 10, 1-19, 1976.
37. a) Pulay, P., "Ab Initio Calculation of Force Constants and Equilibrium Geometries in Polyatomic Molecules. I. Theory," Molec. Phys., 17, 197-204, 1969. b) Schlegel, H.B., S. Wolfe, and F. Bernardi, "Ab Initio Computation of Force Constants. I. The Second and Third Period Hydrides," J. Chem. Phys., 63, 3632-3638, 1975.
38. Fletcher, R. and M.J. D. Powell, "A Rapidly Convergent Descent Method for Minimization," Comput. J., 6, 163-168, 1963.
39. Pople, J.A., "A Priori Geometry Predictions," Applications of Electronic Structure Theory, ed. H.F. Schaefer III, Plenum Press, New York, pp. 1-28, 1977.
40. Binkley, J.S., J.A. Pople, and W.J. Hehre "Self-Consistent Molecular Orbital Methods. 21. Small Split-Valence Basis Sets for First-Row Elements," J. Am. Chem. Soc., 102, 939-947, 1980.
41. Roos, B.O., W.P. Kraemer, and G.H.F. Diercksen, "SCF-CI Studies of the Equilibrium Structure and the Proton Transfer Barrier of $H_3O_2^-$," Theor. Chim. Acta, 42, 77-82, 1976.
42. Benedict, W.S., N. Gailar, and E.K. Plyler, "Rotation-Vibration Spectra of Deuterated Water Vapor," J. Chem. Phys., 24, 1139-1165, 1956.
43. Huber, K.P. and G. Herzberg, "Constants of Diatomic Molecules," Van Nostrand-Reinhold, New York, 1979.
44. DelBene, J. and J.A. Pople, "Theory of Molecular Interactions. II. Molecular Orbital Studies of HF Polymers Using a Minimal Slater-Type Basis," J. Chem. Phys., 55, 2296-2299, 1971.
45. Payzant, J.D., R. Yamdagni, and P. Kebarle, "Hydration of CN^- , NO_2^- , NO_3^- , and OH^- in the Gas Phase," Can J. Chem., 49, 3308-3314, 1971.

46. Mulliken, R.S., "Electronic Population Analysis on LCAO-MO Molecular Wave Functions," J. Chem. Phys., 58, 4733-4742, 1973.
47. Desmeules, P.J. and L.C. Allen, "Strong Positive Ion Hydrogen Bonds: The Binary Complexes Formed from NH_3 , OH_2 , FH , PH_3 , SH_2 , and ClH ," J. Chem. Phys., 72, 4731-4748, 1980.
48. Deakyne, C.A., "A Molecular Orbital Study of $\text{NO}_3^- \cdot \text{H}_2\text{O}$, $\text{OH}^- \cdot \text{HNO}_3$, and $\text{H}^+(\text{H}_2\text{O})_m$ ($\text{CH}_3\text{CN})_k$ Cluster Ions," AFOSR SFRP Research Associate/SCEEE Fellow research project report, summer 1981.
49. Whiteside, R.A., M.J. Frisch, J. S. Binkley, D.J. DeFrees, H.B. Schlegel, K. Raghavachari, and J.A. Pople, "Carnegie-Mellon Quantum Chemistry Archive," Department of Chemistry, Carnegie-Mellon University, Pittsburgh, PA. 15123.
50. Hopkinson, A.C., M.H. Lien, K. Yates, P.G. Mezey, and I.G. Csizmadia, "A Nonempirical Molecular Orbital Study on the Acidity of the Carbon-Hydrogen Bond," J. Chem. Phys., 67, 517-523, 1977.
51. Wolf, J. F., R.H. Staley, I. Koppel, M. Taagepera, R.T. McIver, Jr., J. L. Beauchamp, and R.W. Taft, "Gas Phase Basicities and Relative Proton Affinities of Compounds between Water and Ammonia from Pulsed Ion Cyclotron Resonance Thermal Equilibria Measurements," J. Am. Chem. Soc., 99, 5417-5429, 1977.
52. Ditchfield, R., W.J. Hehre, and J.A. Pople, "Self-Consistent Molecular-Orbital Methods. IX. An Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules," J. Chem. Phys., 54, 724-728, 1971.
53. DelBene, J.E. and F.T. Marchese, "Molecular Orbital Theory of the Hydrogen Bond. V. Hydrogen Bonding Through the Lone Pair and the Pi System in HF-HCN ," J. Chem. Phys., 58, 926-929, 1973.
54. Meot-ner, M., personal communication.
55. Kollman, P.A., "Hydrogen Bonding and Donor-Acceptor Interactions," in Modern Theoretical Chemistry, Vol. 4: Applications of Electronic Structure Theory, ed. by H. Schaefer, Chapter 3, Plenum Press, New York, 1977.

1983 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

FINAL REPORT

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

DESIGNS FOR EVALUATING THE VALIDITY AND ACCURACY OF

PERFORMANCE RATINGS

Prepared by: Terry L. Dickinson
Academic Rank: Professor
Department and University: Department of Psychology
Old Dominion University
Research Location: Human Resources Laboratory, Manpower and Personnel
Division, Force Utilization Branch, Productivity and
Measurement Function at Brooks Air Force Base, San
Antonio, Texas
USAF Research Contact: Dr. Jerry W. Hedge
Date: 19 September 1983
Contract No: F49620-82-C-0035

MODELS FOR EVALUATING THE VALIDITY AND
ACCURACY OF PERFORMANCE RATINGS

by

Terry L. Dickinson

ABSTRACT

Ratings are an important source of information about job performance. For many jobs, objective measures of performance are not available or are impractical to obtain so that ratings are the sole source of information. Unfortunately, performance ratings are a distorted source of information.

The multitrait-multimethod and person perception designs have been used to investigate the distortions in ratings. The purpose of using these research designs is to isolate the factors that distort the ratings and to use this knowledge to improve the quality of performance ratings. The goal of the present research was to develop a design that combined both the multitrait-multimethod and person perception designs.

Each design was discussed, and examples were presented to illustrate that design. The combination design was used to isolate the influence of rater, ratee and context factors on the quality of performance ratings. Future research was recommended to understand and utilize these factors to improve performance ratings.

ACKNOWLEDGMENTS

The work on any project must be accomplished with the aid of others. I wish to acknowledge the help provided by my colleagues at the Human Resources Laboratory. They made this project a pleasurable experience. Dr. Jerry Hedge provided professional support and stimulated my thoughts on performance measurement through our discussions. Lt. Colonel Roger Ballentine created formats for discussion with himself and others. Dr. R. Bruce Gould encouraged my efforts and provided administrative help to the project. To the many friends that I made at the Laboratory, thanks for your warmth, interest and pleasant company. Finally, I was privileged to work with two colleagues on a related project in performance measurement. They gave special stimulation to the present project. James B. Flynn helped to organize my thoughts and shared his own thoughts on several important issues. Catherine E. Hassett was always willing to listen to ideas and work to improve them. Her continual support contributed greatly to the success of the project.

I. INTRODUCTION

Performance ratings are an important method for measuring and defining human attributes. They have been used in research and applied contexts to describe a diversity of human attributes such as group leadership skills, problem solving ability and interpersonal skills. In some contexts, performance ratings serve as substitutes for more objective but expensive methods such as work sample testing, while in other contexts, ratings are the only practical measures of attributes.

Despite the utility of performance ratings, they must be interpreted with caution. Since they require human judgments, performance ratings are fallible measures. Several distortions in ratings have been identified that illustrate their fallibility including leniency, halo, and similarity errors (Landy & Farr, 1980). Such distortions limit inferences about human attributes and the amounts of those attributes possessed by the people who are rated.

The limitations to inferences have been addressed with research into the validity and accuracy of ratings (DeCotiis & Petit, 1978; Saal, Downey, & Lahey, 1980). The validity of ratings is investigated with a multitrait-multimethod design (Boruch, Larkin, Wolins, & MacKinney, 1970). The purpose in using such a design is to evaluate performance ratings against criteria that are logical requirements for measures of human attributes. In particular, variance components and intraclass correlations are computed to evaluate the individual differences in performance accounted for by the ratings. The accuracy of ratings is investigated with a person perception design (Cronbach, 1955). The

purpose in using such a design is to compare performance ratings against target ratings that have been specified by the investigator for the research context. In this design, accuracy statistics are computed to describe several discrepancies between the performance and target ratings.

Research on the validity of ratings was stimulated by Lawler's (1967) application of a multitrait-multimethod design. He emphasized that several sources are available for obtaining ratings (e.g., supervisors, peers and the self) and that these sources may differ in their ratings of performance. Lawler encouraged the application of a multitrait-multimethod approach to compare ratings from several sources. Subsequent research has used a multitrait-multimethod design to investigate formats for obtaining ratings (Burnaska & Hollmann, 1974), the nature of human attributes (Borman & Dunnette, 1975), and rater training (Borman, 1978).

Research on the accuracy of performance ratings has focused on the effects of rater training (Bernardin & Pence, 1980; Borman, 1977; 1979a; Hedge, 1982; McIntyre, Smith & Hassett, 1983). Borman (1977; 1979a) introduced the person perception design to assess training to avoid leniency and halo errors. He found that an admonishment to avoid these distortions was successful, but accuracy was not improved. Apparently, raters learned to avoid certain distortions but not how to rate accurately. Other studies have addressed the relationship between the accuracy of ratings and rater attributes such as personality, interests and observational skills (Borman, 1979b; Murphy, Garcia, Kerkar, Martin, & Balzar, 1982).

Although several factors have been investigated as determinants of the validity and accuracy of ratings (cf. DeCotiis & Petit, 1978; Landy & Farr, 1980), no comparison has been made of their influence on both validity and accuracy. The research has compared ratings against criteria for describing individual differences in performance or against target ratings that specify appropriate performance. The factors should be included in a research design that assesses their joint influence on the validity and on the accuracy of ratings. Such a design would employ a multifactor approach to investigate the limits that the factors place on inferences about human attributes.

II. OBJECTIVES OF THE RESEARCH EFFORT

The goal of this research project is to develop a design to guide investigations on both the validity and accuracy of ratings. The design will combine the multitrait-multimethod and person perception designs, and it will utilize the procedures of analysis of variance. Prior to the presentation of the combination design, the multitrait-multimethod and person perception designs will be described to provide a background. Examples will be discussed that illustrate all the designs.

III. VALIDITY OF PERFORMANCE RATINGS

Performance ratings measure attributes that are assumed to account for differences between individuals in their performance. While the attributes are identified and operationally defined through established procedures such as job analysis and criterion development (McCormick, 1976; Smith, 1976), the

assumption should be questioned in most ratings contexts. Job analysis and criterion development are inexact scientific procedures. Some attributes may be poorly defined, irrelevant, or redundant with other attributes, and the performance ratings for these attributes will reflect no meaningful differences in individual performance.

Multitrait-multimethod validation is a research strategy for assessing the individual differences accounted for by performance ratings (Kavanagh, MacKinney & Wolins, 1971). In this strategy, a rating measure is defined as a trait-method unit. A trait is conceived as a human attribute that is conceptually distinct from other attributes accounting for performance. Some examples of attributes include ability to facilitate group discussions, define acceptable work procedures, and provide negative feedback to others. A method is a procedure for operationally defining traits. Some methods include forced choice scales, checklist scales, and example-anchored scales. In sum, a rating measure taps a particular trait with a particular methodology.

The trait-method combinations in a research study are determined by the rating context. This context is dictated by the interests of the researcher and the nature of performance. For example, a researcher may use job analysis to define the traits that determine performance of jet mechanics for a commercial airline company, decide to measure that performance with two formats for ratings, and obtain ratings of the mechanics by their immediate supervisor. Thus, the researcher "designs" the multitrait-multimethod investigation.

Basic Design. Analysis of variance has been used to analyze the ratings from a multitrait-multimethod investigation (Boruch, Larkin, Wolins & MacKinney,

1970). The basic design includes the three factors of ratees, traits, and methods. As shown in Table 1, the variation in ratings is partitioned into seven sources. The researcher is not concerned with all of the sources of variation in the analyses. The fixed effects of Methods, Traits, and Methods x Traits are usually based on scales of convenience to the investigator and provide no information about validity. For example, two methods may differ because one method employs five-point scales while the other employs nine-point scales, and two traits may differ because one is more socially desirable. In contrast, the random effects of Ratees, Ratees x Traits, Ratees x Methods, and Error provide information about the validity of the measures. These sources allow inferences about the individual differences among ratees.

Insert Table 1 about here

The Ratees source of variation indicates the ability of the measures to order the ratees. This ordering can be due to either traits or methods, or both. Of course, the more the measures agree in their ordering of ratees, the more the measures describe individual differences between the ratees. The Ratee source of variation is said to reflect the convergent validity of the measures.

The Ratees x Traits interaction indicates differential ordering of the ratees by the traits. Since the traits should reflect different aspects of performance, the interaction is desirable. In fact, the stronger the interaction, the greater the number of distinct discriminations between the ratees with the traits. The Ratee x Traits source of variation reflects the discriminant validity of the measures.

The Rates x Methods interaction indicates differential ordering of the rates with the methods. This differential ordering is undesirable. The methods for rating should not influence the ordering of rates. Only the traits should determine the ordering of rates. The Rates x Methods source of variation reflects the method bias of the measures.

The Error source indicates residual variation due to sampling and measurement errors. The size of this effect relative to the remaining sources of variation suggests the extent of differences between the rates that cannot be accounted for by the traits and the methods.

The Error mean square may be used to compute F-ratios to establish statistical significance for the remaining sources. However, the F-ratios are based on large degrees of freedom, and the critical values to establish significance are frequently exceeded. A more appropriate strategy for assessing the relative variation in ratings explained by the sources is to compare variance components. These components provide a comparison of the relative sizes of convergent validity, discriminant validity, method bias, and error, while controlling for degrees of freedom. For a single research study, the variance components may be compared directly. However, since the variance component due to Error would differ from study to study, comparisons of the variance components from several studies is not appropriate. Rather, ratios of the variance components are formed to generate intraclass correlation coefficients. These ratios are expressed as a source's component divided by the sum of all variance components. Each ratio reflects the proportion of variance accounted for by that source relative to the variation accounted for by all sources.

Computations. The computations associated with a multitrait-multimethod design may be accomplished in several ways. First, the computations may be conducted directly on the ratings that are obtained in the investigation. These computations use the sums of squares formulas that are traditionally employed in the analysis of variance (e.g., Kirk, 1968; pp. 239-240).

Another computational strategy involves the use of the variance-covariance matrix among the measures (Stanley, 1961). This matrix can be used to compute the sum of squares for the various random effects of interest in the multitrait-multimethod investigation. This computational strategy has the advantage of displaying the contributions of each of the measures to the ordering of the rates. It is directly related to the use of the correlational matrix among the measures in a multitrait-multimethod investigation (Campbell & Fiske, 1959; Kavanagh, MacKinney, & Wolins, 1971).

Example. An issue of research in performance measurement is the choice of a method for obtaining performance ratings (Schwab, Heneman, & Decotiis, 1975). All methods are not equally desirable. Methods should be compared in terms of the individual differences that each summarizes in the ratings. One method is preferred to others if the ratings that are obtained with that method display more discriminant validity and less bias in ordering the rates.

As an illustration, suppose an investigator needs rating scales for his research on the performance of test administrators. The investigator has defined the traits and collected the items for constructing the rating scales (e.g., Dickinson & Zellinger, 1980). However, the investigator still needs to specify his method for obtaining the ratings. He has narrowed his decision to

be a choice between the two methods of example-anchored scales (Taylor, 1968) and checklist scales (Landy & Trumbo, 1980). To aid his decision making, the investigator has collected data in a multitrait-multimethod design. The data are displayed in Table 2. The analysis of the data that were used by the investigator in making his choice is presented below.

Insert Table 2 about here

The data were collected from a group of raters who viewed videotapes of ten test administrators who were played by actors according to ten scripts of performance. The tests that were given by the administrators were the same in each videotape. However, the performance of the administrators on the dimensions varied across the videotapes. The group of raters viewed each tape, discussed the performance of that administrator, and rated performance on each of three traits using the example-anchored and checklist methods for rating.

The investigator employed a traditional formulation in conducting his analysis of the variation in the ratings. A summary of the analysis is shown in Table 3. The variance components and intraclass correlation coefficients indicate that the measures can be used to order ratees with substantial validity and with little bias due to the method for rating. Convergent validity and discriminant validity account approximately for two-thirds of the variation that determines the ordering of ratee performance. The example-anchored and checklist methods for rating have little influence on the ordering of ratees. Both are equally desirable, and the investigator can choose either of the two methods on the basis of the results. Additional research or practical considerations must guide his choice between the two methods.

Insert Table 3 about here

Beyond the Basic Design. The basic design for a multitrait-multimethod investigation can be expanded to research the factors that distort the validity of ratings. Several theoretical models are available to guide such research (DeCotiis & Petit, 1978; Landy & Farr, 1980; Wherry & Bartlett, 1982). The models describe factors ranging from the ability and motivation of raters to organizational policies concerning the use and purpose of the ratings.

To continue in illustration, suppose the investigator decomposes his decision from a choice between two methods to a choice between two methods that can be used to collect ratings for two purposes. The investigator has collected data from two groups with the basic multitrait-multimethod design. One group is told that the purpose for the ratings is to research the validity of the tests for selecting new employees, while the second group is told that the purpose is to motivate employees by rewarding or punishing them for their past performance. Finally, the investigator only collected five ratings from each group. The research group viewed videotapes one through five, while the motivation group viewed the remaining videotapes.

A four-factor design was used to analyze the ratings that were collected by the investigator (cf. Winer, 1971; pp. 539-546). The design has factors of Purposes, Ratees nested within Purposes, Traits, and Methods. The psychometric interpretations for the sources of variation are summarized in Table 4.

Insert Table 4 about here

The expanded multitrait-multimethod design includes both fixed and random effects. As with the basic design, only the random effects in the expanded design provide inferences about individual differences among ratees. For example, the Ratees within Purposes effect represents the ability of the measures to order the ratees with both purposes for the ratings. This effect is a pooling of the Ratee effects available from the two purposes for ratings. This pooled effect includes variation due to convergent validity and the interaction of convergent validity with purpose for the ratings. Unfortunately, the nesting of ratees prevents separating the variation due to convergent validity from its interaction. The decision by the investigator to design the research with ratees nested within purpose groups produces this confounding. Similarly, the variation due to discriminant validity and method bias cannot be separated from their interaction with purpose for the ratings.

A summary of the analysis is shown in Table 5. The expanded research design suggests that purpose for the ratings has little influence on the multitrait-multimethod properties of the ratings. Convergent and discriminant validity again account for substantial differences in the ratings of performance. Little method bias is present; both methods of rating are equally desirable. Purpose for the ratings only influences the raters use of the scales to describe amounts of the attributes. In particular, the ratees were rated higher on trait number two when the purpose for the ratings was research than when the purpose was for motivation. Since the trait reflects the "warmth"

versus "coldness" of the administrator, the investigator suspects that the raters valued this attribute highly in test administrators. He believes that the raters were emphasizing high standards of rapport with the examinees for the test administrators.

Insert Table 5 about here

IV. ACCURACY OF PERFORMANCE RATINGS

Accuracy statistics have been described as the most appropriate criteria for assessing the distortions in ratings (Borman, 1978). While other statistics are available, most lack a meaningful standard for defining distortions (Saal, et al., 1980). In contrast, the person perception design for investigating accuracy requires the development of a standard. The standard is usually a set of target ratings that specifies the performance scores of ratees on several attributes.

Target ratings can be developed from the judgments of experts or other decision-making groups or from objective measures. For example, psychologists have rated the performance of actors as displayed in videotapes. These expert ratings were averaged to define the target ratings (Borman, 1979b). Supervisory ratings of performance have also been used to define target ratings in assessing the accuracy of self ratings (Mabe & West, 1982). Finally, life history information and paper-and-pencil tests have been used as objective measures to develop target ratings for assessing the accuracy of ratings of interviewee performance (Cline & Richards, 1960).

Cronbach's Formulation. The overall accuracy of a rater is defined as the sum of squared discrepancies between the rater's performance ratings and the target ratings for the ratees. Cronbach (1955) argues convincingly that overall accuracy should be broken down into four statistics that are mathematically independent components of overall accuracy.

Elevation is the component of accuracy due to the mean of the performance ratings for the group of ratees and the set of attributes. The rater whose mean is close to that of the target ratings will tend to rate the performance of the ratees more accurately. Although Cronbach (1955) states that elevation describes the way a rater uses the rating scale, this statistic is useful for describing the accuracy of the rater in judging the overall performance of a group of ratees (Murphy, et al., 1982).

Differential elevation is the component of accuracy associated with the mean ratings that a rater gives the ratees the set of attributes. In some rating contexts, a mean rating for the set could indicate the overall job performance of a ratee. This component of accuracy reflects a rater's ability to order ratees in comparison to their overall differences as specified by the means of their target ratings. Murphy et al., (1982) suggest that this component of accuracy is important for administrative decisions. For example, a supervisor is often required to nominate subordinates for training programs or to choose one for promotion.

Stereotype accuracy reflects the accuracy of a rater in using the attributes to describe the group of ratees. The mean ratings on the attributes

given by the rater to the group are compared to the mean ratings given to the group by the expert source. This component of accuracy is important in making administrative decisions. For example, a supervisor may need to diagnose relative strengths and weaknesses of his group of subordinates to choose training programs or other developmental activities for them. These decisions require accurate summary evaluations of subordinates on the attributes of performance.

Finally, the most important component of accuracy is differential accuracy (Cronbach, 1955). The target ratings for each ratee are compared to the performance ratings given to them by the rater. Differential accuracy reflects the rater's ability to rate accurately the individual ratee. In an organizational setting, differential accuracy is important for research purposes and for developing employees. Most research projects utilize the performance ratings of individuals, necessitating that each ratee be described with little distortion in the ratings. Employee development requires accurate feedback about an individual's performance, so that changes that are undertaken for improvement are appropriate to the individual.

Computations. The computations for the accuracy statistics were presented by Cronbach (1955) in his seminal article. These statistics are oriented to the descriptions of the accuracy of each rater, and so, the underlying research design is not emphasized. Indeed, subsequent research studies have utilized the accuracy statistics as measures of the rater's "ability" to perceive others (e.g., Borman, 1979b; Cline & Richards, 1960; Crow & Hammond, 1957). As a consequence, little attention has been given to the basic design underlying person perception investigations, and its extension to other areas of research.

Basic Design. Analysis of variance can be used to summarize the ratings that are obtained in person perception investigations. The basic design includes the three factors of rating sources, ratees, and traits. Table 6 displays the seven sources of variation in the basic design, and it summarizes the psychometric interpretations of those sources.

Insert Table 6 about here

The sources for ratings are the rater and the experts who provided the target ratings. The variation in ratings accounted for by Rating Sources reflects elevation accuracy. The larger this source of variation, the larger the difference between the overall mean rating of the rater and that of the experts, and the more inaccurate is the rater.

The Ratees effect indicates the ability of the rating sources to describe differences between ratees over the attributes. This effect can be due to the rater, the expert source for the target ratings, or both. Since the investigator will typically select the ratees to differ from one another on the attributes, the Ratees effect should account for substantial variation in the ratings. However, the more the rater agrees with the target ratings, the greater the Ratees effect. The rater who is accurate in ordering the ratees compared to the expert source enhances the convergent validity of the ratings.

The fixed effect of Traits reflects the relative amounts of the performance attributes possessed by the group of ratees. The investigator designs this effect into the research with the choice of the rating context and the selection

of the ratees. The rating context usually includes attributes that differ in their social desirability, and consequently, some attributes will have greater value to the rater than others. Furthermore, the ratees who are selected by the investigator may be chosen to have unequal amounts of the attributes. If the expert source for ratings provides target ratings that confirm the investigator's intentions, then, the Traits effect is likely to account for variation in the ratings.

The Rating Sources x Ratees interaction reflects differential elevation accuracy, and it indicates differential ordering of the ratees by the rater compared to the expert source for ratings. This differential ordering is undesirable. An accurate rater should order the ratees in a manner similar to that ordering provided by the expert source. Since the expert source serves as the standard for defining the differences between ratees, the effect can also be considered a reflection of differential convergent validity. A rater may describe more or fewer differences between the ratees in assessing their performance on the set of attributes. The larger the interaction, the more inaccurate is the rater in ordering the ratees.

The Rating Sources x Traits interaction indicates the stereotype accuracy of the rater. An accurate rater should agree with the expert source in the relative amounts of the attributes reflected in the group of ratees. The larger this interaction, the more inaccurate the rater.

The Ratees x Traits interaction reflects the extent of individual differences on the attributes perceived by the rater and expert source. Since the researcher should select the ratees for the investigation, the differential

ordering of the ratees on the attributes can be determined by the researcher. Of course, this assumes that the target ratings are close to the intended performance scores for the ratees (Borman, 1979). For example, the researcher can construct videotapes of actors who play ratees. Then, the performance of ratees can be acted to represent scaled amounts of the attributes. If the investigator selects ratees who differ in their ordering on the traits, then, discriminant validity will explain variation in the ratings. Moreover, the more the rater's ratings match those of the expert source, the stronger will be the interaction, and the more accurate will be the rater.

The Rating Sources x Ratees x Traits interaction reflects the differential accuracy of the rater. This is the ability of the rater relative to the expert source to describe individual differences among the ratees. This interaction is undesirable. The rater who is accurate should agree with the expert source on the differences among the ratees. If the rater disagrees with the expert source, the rater will possess more or less discriminant validity in his or her ratings. Since the target ratings serve as a standard, this differential discriminant validity is undesirable.

Computations. The sums of squares that are obtained from the analysis of the variance in ratings are closely related to the accuracy statistics developed by Cronback (1955). The accuracy statistics are contrasts between effects in the analysis of variance design. Each accuracy statistic is a contrast of effects of the rater to those of the expert source for ratings. Of course, an effect is a linear combination of means, and such combinations are used to compute sums of squares in the design.

Combination Design. The person perception design for the investigation of accuracy can be combined with the multitrait-multimethod design. The combined design includes the four factors of rating sources, ratees, traits, and methods. In essence, the person perception design has been expanded to include more than one method for obtaining performance ratings, while the multitrait-multimethod design has been expanded to include more than one source for the ratings. As shown in Table 7, the combined design includes the sources and psychometric interpretations of each separate design as well as several other sources with their psychometric interpretations.

Insert Table 7 about here

The Rating Sources x Ratees x Methods interaction reflects the differential ordering of the ratees provided by the rater using the methods for rating compared to the ordering provided by the expert source using the methods. This differential ordering is undesirable. An accurate rater should order the ratees regardless of the method for rating in a manner similar to that ordering provided by the expert source. Of course, the expert source for ratings may order ratees differently depending on the method for rating. Since the expert source serves as the standard for defining the differences between ratees, this result can be considered a method bias in the target ratings. However, a logical property for a standard is not to contain method bias. The target ratings should serve to evaluate the rater's ability to describe ratees regardless of the method for rating. Hopefully, the investigator can design the research so that the target ratings are defined to have no method bias.

The Rating Sources x Traits x Method interaction indicates the accuracy of the rater in using the attributes to describe the group of ratees by the methods. If the investigator designs the research so that the target ratings contain no method bias, this interaction suggests that the rater uses the attributes to describe the performance of the group differently with each method for rating. This interaction is again undesirable. No component of a rater's accuracy should depend on the method for rating.

Finally, the Ratees x Traits X Method interaction reflects an influence of the method for the ordering of ratees on the attributes summed over the rater and the expert source. This interaction is also undesirable. The ordering of ratees on the traits should not depend on the method for obtaining ratings. If the investigator designs the research so that the expert source orders the ratees on the traits in the same manner using the methods, the interaction is determined by the rater's inability to use the methods similarly. This differential use of the methods reflects differential discriminant validity by the rater, and it indicates inaccuracy by the rater. The rater should order the ratees on the attributes regardless of the method that the rater uses for making ratings.

Example. Consider an extension of the issue of the choice of a method for obtaining performance ratings. While methods should be compared in terms of the individual differences that each summarizes in ratings, another aspect is the accuracy with which the rater can use the methods to summarize individual differences in the ratings. The multitrait-multimethod design assumes that a method is preferred if it influences the ordering of ratees less than other methods. The combination design extends the assumption to consider accuracy. A

method is preferred if the rater can use it to obtain greater agreement with the expert source for ratings.

To finish in illustration, suppose that our investigator developed the scripts and videotapes in a series of workshops with a group of experts. The experts were highly familiar with the performance of test administrators. Scripts were modified and actors changed their performance until the experts were in high agreement in their ordering of the ratees with both methods for rating. In sum, the investigator designed the target ratings to contain no method bias.

The investigator employed the combination design to evaluate the accuracy of several raters. The results of the analysis of the ratings that were obtained from one rater are shown in Table 8. The data in Table 2 were used for this analysis. Furthermore, assume that the investigator only collected five ratings from each rating source. The expert source and rater each viewed and rated the same videotapes.

Insert Table 8 about here

The results of the research indicate that the rater was fairly accurate. Elevation and differential accuracy accounted for little variation in the ratings; both were not statistically significant. The mean of the performance ratings given by the rater for the group of ratees on the set of attributes compared favorable to the mean provided by the expert source. Importantly, the rater agreed for the most part with the expert source on the differences among the ratees. The Rating Source x Ratees x Traits interaction was negligible in magnitude suggesting discriminations by the rater comparable to those by the expert source.

The results do suggest some inaccuracies by the rater. Differential elevation accuracy and stereotype accuracy were both statistically significant. For most ratees, the rater and expert source agreed on individual differences across the set of attributes using both methods. However, test administrator three was given a much greater mean rating by the expert source. This ratee was the only female actor to play a test administrator, and the investigator suspects that sex may explain the greater rating. Perhaps, the rater is prejudiced against female test administrators. The Rating Source x Traits interaction indicated that the rater did not perceive the relative importance of the traits similar to the expert source. In particular, trait number two was seen as significantly less important by the rater. This trait reflects the "coldness" versus "warmth" of the test administrator, and the investigator suspects that the rater is insensitive to that attribute of test administration.

The investigator was quite pleased that the method for rating had little influence on the ratings. There was no method bias in ordering the ratees shown by the rater or the expert source. The investigator was successful in eliminating method bias in the target ratings, and the rater was able to order accurately the group of ratees utilizing the set of attributes. At least for this rater, the investigator is confident that either method for rating performance can be used to obtain accurate ratings of performance. Nonetheless, the investigator does recognize that the ratings obtained with the example-anchored method cannot be compared in absolute size to those obtained with the checklist method. The Trait x Method interaction suggests considerable scale bias in measuring the attributes.

V. DISCUSSION

Several models of the rating process outline variables that influence the accuracy of ratings (DeCotiis & Petit, 1978; Kavanagh, Borman, Hedge & Gould, 1983; Landy & Farr, 1980). However, none emphasizes the influence of logical requirements for performance measures on accuracy. The research studies that support the models have evaluated accuracy statistics against rater attributes such as personality and training experience. These studies illustrate a myopic research strategy (Cronbach, 1955). They are not connected to meaningful theory about the logical requirements for performance measures.

The combination design can provide a broader strategy for accuracy research. It emphasizes the assessment of accuracy in the framework of logical requirements for performance measures. The investigator can determine conditions under which ratings are obtained including contextual factors, rates, traits, methods for rating, and sources for target ratings. These conditions allow the investigator to design the amounts of multitrait-multimethod properties into the target ratings. Such logical requirements provide a rich framework for interpreting the accuracy of performance ratings.

Target ratings should be designed to possess the multitrait-multimethod properties found in practice. For example, criterion research consistently shows that job performance is a multidimensional concept (Landy & Trumbo, 1980). There are many routes to success in most work contexts, and so, several attributes are necessary to describe performance. Consequently, the investigator must design the target ratings to possess discriminant validity.

Several points are important to consider in the design. The discriminant validity of the target ratings should be representative of the rating context so that accuracy findings generalize beyond the particular research setting.

Brunswik's (1956) view of representative design underscores this point. Unfortunately, representative designs are apt to be expensive. Most accuracy studies have used videotapes of four to eight ratees who are each rated on six to twelve dimensions (e.g., Borman, 1977; 1979a; Hedge, 1982; McIntyre, et al., 1983; Murphey, et al., 1982). Such small combinations of ratees and dimensions restrict the amount of discriminant validity that can be designed into the target ratings, and subsequently the generality of the research findings.

The combination design can be expanded to consider the broad scope of research on performance ratings. Multiple raters can be included in the Rating Sources to consider rater characteristics such as sex, race, ability and motivation. Effects coding of the raters against the expert source will provide the statistics for each rater that are contained in the combination design (Kerlinger & Pedhazur, 1973). Ratee characteristics can also be studied in the combination design. Videotapes of actors can be constructed whose target ratings are identical but who differ in characteristics such as age, sex and race. Furthermore, manipulations of ratee and rater characteristics in the same design address important legal questions about equal employment opportunity and the quality of performance ratings (Cascio & Bernardin, 1981). Finally, contextual factors can be evaluated for their impact on the accuracy of ratings. Factors that can be studied include the intended use of the ratings (McIntyre, et al., 1983), the content of the attributes (Kavanagh, 1971), and the feedback given to raters on their accuracy (Ilgen, Fisher, & Taylor, 1979).

VI. RECOMMENDATIONS

No research study has implemented the combination design to investigate both the validity and accuracy of performance ratings. The design provides a rich framework for understanding the distortions in performance ratings, and

thereby, it can identify factors to control or remove to improve ratings. To date most research on the accuracy of ratings has focused on the training of raters to become more accurate in their ratings. Several programs have been used to train raters to make more accurate ratings. This line of research should continue; however, it must be expanded to address the influence of logical requirements for performance measures on accuracy training. For example, a study could consider the impact of accuracy training on performance measures that differ in their amounts of discriminant validity. The combination design developed in this report provides the research strategy for future studies.

REFERENCES

- Borman, W. D. Consistency of rating accuracy and rating errors in the judgment of human performance. Organizational Behavior and Human Performance, 1977, 20, 258-272.
- Borman, W. Exploring the upper limits of reliability and validity in job performance ratings. Journal of Applied Psychology, 1978, 63, 135-144.
- Borman, W. C. Format and training effects on rating accuracy, and rater errors. Journal of Applied Psychology, 1979, 64, 410-421. (a)
- Borman, W. C. Individual difference correlates of accuracy in evaluating others' performance effectiveness. Applied Psychological Measurement, 1979, 3, 103-115. (b)
- Borman, W. C., & Dunnette, M. D. Behavior-based versus trait-oriented performance ratings: An empirical study. Journal of Applied Psychology, 1975, 60, 561-565.
- Boruch, R. F., Larkin, J. D., Wolins, L., & MacKinney, A. C. Alternative methods of analysis: Multitrait-multimethod data. Educational and Psychological Measurement, 1970, 30, 833-853.
- Brunswik, E. Perception and the representative design of psychological experiments. (2nd ed.) Berkeley: University of Colorado Press, 1956.
- Burnaska, R. F., & Hollmann, T. D. An empirical comparison of the relative effects of rater response biases on three rating scale formats. Journal of Applied Psychology, 1974, 59, 307-312.
- Campbell, D. T., & Fiske, D. W. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 1959, 56, 81-105.
- Cascio, W. F., & Bernardin, H. J. Implications of performance appraisal litigation for performance appraisal decisions. Personnel Psychology, 1981, 34, 211-226.
- Cline, V. B., & Richards, J. M., Jr. Accuracy of interpersonal perception -- A general trait? Journal of Abnormal and Social Psychology, 1960, 60, 1-7.
- Cronbach, L. J. Processes affecting scores on understanding of others and assuming "similarity." Psychological Bulletin, 1955, 52, 177-193.
- Cronbach, L. J., & Gleser, G. C. Assessing similarity between profiles. Psychological Bulletin, 1953, 50, 456-473.
- Crow, W. J., & Hammond, K. R. The generality of accuracy and response sets in interpersonal perception. Journal of Abnormal and Social Psychology, 1957, 54, 384-390.

- DeCotiis, T. A., & Petit, A. The performance appraisal process: A model and some testable propositions. Academy of Management Review, 1978, 3, 635-646.
- Dickinson, T. L., & Zellinger, P. M. A comparison of the behaviorally anchored rating and mixed standard scale formats. Journal of Applied Psychology, 1980, 65, 147-154.
- Hedge, J. W. Improving the accuracy of performance evaluations: A comparison of the methods of performance appraiser training. Unpublished doctoral dissertation, Old Dominion University, 1982.
- Ilgen, D. R., Fisher, C. D., & Taylor, S. M. Consequences of individual feedback on behavior in organizations. Journal of Applied Psychology, 1979, 64, 359-371.
- Kavanagh, M. J. The content issue in performance appraisal: A review. Personnel Psychology, 1971, 24, 653-668.
- Kavanagh, M. J., Borman, W. C., Hedge, J. W., & Gould, R. B. A model of performance measurement quality for validation research in the military. AFHRL Technical Paper, Brooks Air Force Base, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory, 1983.
- Kavanagh, M. J., MacKinney, A. C., & Wolins, L. Issues in managerial performance: Multitrait-multimethod analyses of ratings. Psychological Bulletin, 1971, 75, 34-49.
- Kerlinger, F. N., & Pedhazur, E. J. Multiple regression in behavioral research. New York: Holt, Rinehart and Winston, 1973.
- Kirk, R. E. Experimental design: Procedures for the behavioral sciences. Belmont, CA: Wadsworth, 1968.
- Landy, F., & Farr, J. Performance rating. Psychological Bulletin, 1980, 87, 72-107.
- Lawler, E. E., III. The multi-trait-multi-rater approach to measuring managerial job performance. Journal of Applied Psychology, 1967, 51, 369-381.
- Mabe, P. A., & West, S. G. Validity of self-evaluation of ability: A review and meta-analysis. Journal of Applied Psychology, 1982, 67, 280-296.
- McCormick, E. J. Job and task analysis: In M. D. Dunnette (Ed.), Handbook of industrial and organizational psychology. Chicago: Rand McNally, 1976.
- McIntyre, R. M., Smith, D. E., & Hasset, C. e. Accuracy of performance ratings as affected by rater training and perceived purpose of rating. Journal of Applied Psychology, 1984, in press.
- Murphy, K. R., Garcia, M., Kerkar, S., Martin, C., & Balzer, W. K. Relationship between observational accuracy and accuracy in evaluating performance. Journal of Applied Psychology, 1982, 67, 320-325.

- Saal, F. E., Downey, R. G., & Lahey, M. A. Rating the ratings: Assessing the psychometric quality of rating data. Psychological Bulletin, 1980, 88, 413-428.
- Schwab, D. P., Heneman, H. G., & DeCotiis, T. A. Behaviorally anchored rating scales: A review of the literature. Personnel Psychology, 1975, 28, 549-562.
- Smith, P. C. Behaviors, results, and organizational effectiveness; The problem of criteria. In M. D. Dunnette (Ed.), Handbook of industrial and organizational psychology. Chicago: Rand McNally, 1976.
- Stanley, J. C. Analysis of unreplicated three-way classifications with applications to rater bias and trait independence. Psychometrika, 1961, 26, 205-219.
- Taylor, J. B. Rating scales as measures of clinical judgment: A method for increasing scale reliability and sensitivity. Educational and Psychological Measurement, 1968, 28, 747-766.
- Wherry, R. J., & Bartlett, C. J. The control of bias in ratings: A theory of ratings. Personnel Psychology, 1982, 35, 521-551.
- Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill, 1971.

Table 1

Summary Table for the Psychometric Interpretations of the
Basic Multitrait-Multimethod Design

Sources	Psychometric Intrepretation
Traits (T)	Trait Bias
Methods (M)	Scale Bias
T x M	Trait by Scale Bias
Rateses (R)	Convergent Validity
R x T	Discriminant Validity
R x M	Method Bias
Error	Sampling and Measurement Errors

Table 2

Example Data for Basic Multitrait-Multimethod Design

Test Administrators	Methods					
	1			2		
	Traits			Traits		
	1	2	3	1	2	3
1	4	7	2	3	6	3
2	3	5	1	3	5	4
3	7	9	6	6	8	6
4	6	6	2	4	5	3
5	5	5	1	4	4	4
6	8	2	5	5	5	7
7	4	1	1	3	4	5
8	6	3	4	6	2	2
9	7	5	2	8	6	4
10	7	1	1	4	2	2

Table 3

Summary Table for the Analysis of the Data for the
Basic Multitrait-Multimethod Design

Sources	df	MS	F-Ratio	VC	IC
Traits (T)	2	18.87	4.43*	-	-
Methods (M)	1	.82	.55	-	-
T x M	2	8.47	7.84**	-	-
Rates (R)	9	9.57	8.86**	1.42	.24
R x T	18	4.26	3.94**	1.59	.38
R x M	9	4.26	1.37	.13	.12
Error	18	1.08		1.08	.26

*p<.05 **p<.01

Table 4

Summary Table for Psychometric Interpretations of the
One Factor Design Beyond the Multitrait-Multimethod Design

Sources	Psychometric Interpretations
Purposes (P)	Research Conditions
Rates (R)/P	Convergent Validity Within Research Conditions
Traits (T)	Trait Bias
T x P	Trait Bias by Research Conditions
T x R/P	Discriminant Validity Within Research Conditions
Methods (M)	Scale Bias
M x P	Scale Bias by Purpose
M x R/P	Method Bias Within Research Conditions
T x M	Trait by Scale Bias
T x M x P	Trait by Scale Bias by Research Conditions
Error	Measurement and Sampling Error

Table 5

Summary Table for Analysis of Data for One Factor Design
Beyond the Multitrait-Multimethod Design

Sources	df	MS	F-Ratio	VC	IC
Purposes (P)	1	3.75	.36	-	-
Rates (R)/P	8	10.30	11.32**	1.56	.49
Traits (T)	2	18.87	10.14**	-	-
T x P	2	23.40	12.58**	-	-
T x R/P	16	1.86	2.04	.48	.15
Methods (M)	1	.82	.54	-	-
M x P	1	1.35	.90	-	-
M x R/P	8	1.50	1.65	.20	.06
T x M	2	8.47	9.31**	-	-
T x M x P	2	2.40	2.64	-	-
Error	16	.91		.91	.29

**p<.01

Table 6

Summary Table for the Psychometric Interpretations of the
Basic Accuracy Design

Sources	Psychometric Interpretation
Rating Sources (S)	Elevation Accuracy
Ratees (R)	Convergent Validity
Traits (T)	Trait Bias
S x R	Differential Elevation Accuracy (Differential Convergent Validity by Rating Sources)
S x T	Stereotype Accuracy
R x T	Discriminant Validity
S x R x T	Differential Accuracy (Differential Discriminant Validity by Rating Sources)

Table 7

Summary Table of Psychometric Interpretations of the
Combination Design

Sources	Psychometric Interpretations
Rating Sources (S)	Elevation Accuracy
Rates (R)	Convergent Validity
Traits (T)	Trait Bias
Methods (M)	Scale Bias
S x R	Differential Elevation Accuracy (Differential Convergent Validity by Rating Sources)
S x T	Stereotype Accuracy
S x M	Differential Scale Bias by Rating Sources
R x T	Discriminant Validity
R x M	Method Bias
T x M	Trait by Scale Bias
S x R x T	Differential Accuracy (Differential Discriminant Validity by Rating Sources)
S x R x M	Differential Elevation Accuracy by Methods (Differential Method Bias by Rating Sources)
S x T x M	Differential Stereotype Accuracy by Methods
R x T x M	Differential Discriminant Validity by Methods
Error	Measurement and Sampling Error

Table 8

Summary Table for the Analysis of the Data for the
Combination Design

Sources	df	MS	F-Ratio	VC	IC
Rating Sources (S)	1	3.75	.40	-	-
Ratees (R)	4	11.31	15.08**	1.19	.20
Traits (T)	2	18.87	8.50*	-	-
Methods (M)	1	.82	.39	-	-
S x R	4	9.29	12.39**	1.94	.33
S x T	2	23.40	15.60**	-	-
S x M	1	1.35	1.52	-	-
R x T	8	2.22	2.96	.55	.09
R x M	4	2.11	2.81	.34	.06
T x M	2	8.47	7.92*	-	-
S x R x T	8	1.50	2.00	.62	.10
S x R x M	4	.89	1.19	.15	.03
S x T x M	2	2.40	3.20	-	-
R x T x M	8	1.07	1.43	.34	.06
Error	8	.75		.75	.13

*p < .05 **p < .01

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

Searching for Precursors to Laser-Induced Damage

Prepared Jointly by: Dr Michael F Becker
Academic Rank: Associate Professor
Department and University: Department of Electrical Engineering
University of Texas at Austin

and

Dr Fred E Domann
Academic Rank: Associate Professor
Department and University: Department of Physics
University of Wisconsin - Platteville

Research Location: Optical Coating Damage Laboratory
Air Force Weapons Laboratory
Kirtland AFB, New Mexico

USAF Research: Dr Alan Stewart

Date: 1 August 1983

Contract No.: F49620-82-C-0035

Searching for Precursors to Laser-Induced Damage

by

Michael F. Becker

and

Fred E. Domann

Abstract

Several experimental techniques are investigated to try to find a precursor to permanent damage caused by laser radiation. These techniques include charge emission, neutral particle emission, surface photoconductivity, and surface potential measurements. The techniques are applied to a wide variety of optical materials: diamond-turned copper, silicon, and various dielectric films on fused silica substrates. Charge emission was found to be a useful technique, and its extension to multi-pulse experiments is recommended. Neutral particle emission was too insensitive to be useful. Photoconductivity data were gathered, but an ultrahigh vacuum system in conjunction with photon cleaning is probably necessary to gather meaningful data. Suggestions for further investigations are offered, including a surface potential probe.

Acknowledgements

The authors hereby acknowledge the Southeastern Center for Electrical Engineering Education; the Air Force Office of Scientific Research; the Air Force Weapons Laboratory, Kirtland Air Force Base, New Mexico; and Dr Arthur Guenther for providing the opportunity of making this an interesting and productive summer.

We would also like to thank Dr Alan Stewart for his helpful advice and generous contribution of his time and effort. We thank the rest of the staff of the Optical Coating Damage Laboratory for their assistance whenever it was needed.

Finally, one of us, F.E.D., would like to thank his wife, Lou Anna, and daughter, Kate, for relieving him of his usual summer duties.

MICHAEL F BECKER
Associate Professor
Department of Electrical
Engineering
University of Texas at Austin

FRED E DOMANN
Associate Professor
Department of Physics
University of
Wisconsin - Platteville

I. Introduction.

In the years that laser-induced damage has been studied, two important, fundamental questions still remain unanswered.¹ (1) What microscopic mechanism or mechanisms lead to laser damage? (2) Is there a reliable (non-destructive) precursor to laser damage? These two questions are clearly related, and an understanding of one will aid in understanding the other. In this summer research project, we will concentrate on the second question and seek precursor information in laser damage using a variety of new techniques. The study is to be carried out using a wide variety of samples, and the results will be correlated with other commonly used diagnostics for laser damage.

II. Objectives.

The problem of finding a precursor to laser damage will be addressed using the following experimental techniques: (1) charge emission, (2) neutral particle emission (detected by rapid pressure fluctuations), (3) surface photoconductivity, and (4) surface potential changes. The apparatus for each of these diagnostics is to be constructed, tested, and applied to as wide a range

of samples as possible. Ideally, samples should include metals, semiconductors, insulators, and dielectric thin films. A description of the apparatus and the experimental results for each of the four diagnostics will be found in each of the four subsequent sections of this report.

III. Charge Emission.

The emission of negative charge into vacuum during laser irradiation was the first diagnostic to be studied in this project. However, we will first describe the characteristics of the laser system which are common to all of the experiments which were conducted.

The laser was a Quantel Nd:YAG laser oscillator amplifier system. The 1.06 micron pulses were 5.07 ± 1.0 ns long and were produced by passive Q-switching. The pulse profiles were measured with a fast vacuum photodiode and were always smooth Gaussians. After amplification, the pulses passed through an attenuator consisting of a rotating half-wave plate and a polarizer. At the sample location, the laser beam was brought to a focus by a 2 m focal length lens. The focal plane of the lens was scanned with a narrow (<50 micron) slit in order to

determine the beam spot size on the target. By this method, the beam radius at the focus was found to be 180 microns. In addition, the beam profiles were reasonably Gaussian. The laser energy in the focal plane was measured by an energy meter which sampled part of the incident beam. This meter was calibrated against a second meter placed behind the focal plane. Using all the preceding information, the energy, the energy density (fluence), and the intensity (power/area) of the laser pulse incident on the sample could be determined for every laser shot.

A variety of samples were used in these experiments: metals, semiconductors, and dielectrics. Diamond-turned copper mirrors were tested as being representative of the best metal optics. Single-crystal silicon wafers polished on both sides represented the semiconducting materials, and fused silica substrates coated with half-wave thick dielectric coatings represented the wide band-gap insulators. Two high quality coating materials were tested in these experiments: Ta_2O_5 and Al_2O_3 . These dielectric films were also tested for photoconductivity, and these experiments will be described in Section V.

During the course of the work looking for precursors to damage, the thresholds for single-pulse damage were measured for all of the optics which we tested. These results are shown in Table I.

Sample	Single-Pulse Damage Threshold	Extremes of Damage-No-Damage
Silicon Wafer	1.6 J/cm ²	1.1-2.18 J/cm ²
Copper, Diamond- Turned	2.1 J/cm ²	2.0-2.4 J/cm ²
Cr-26 Film	10.4 J/cm ²	9.8-12.0 J/cm ²
Ta ₂ O ₅ Film	~ 10 J/cm ²	9.4-10.8 J/cm ²
Al ₂ O ₃ Film	Uncertain	50-125 J/cm ²

Table I. Single-pulse damage threshold for all the samples which were tested (except for the photoconductivity samples, see Section V.)

The main emphasis of this section is on the relation between charged particle emission and laser-induced damage. The data were collected by placing the samples in a vacuum of about 10⁻⁶ Torr. A charge collecting wire was placed about 2 cm above and slightly to the side of the sample surface. The wire was charged to +1,500 volts (sufficient to collect all emitted negative charge), and the negative charge pulses were amplified by an Amptek A-

203 and A-206 charge sensitive amplifier chip set. The sensitivity was calibrated as 1.75×10^{13} volts/coulomb. In practice, the ultimate sensitivity was limited by induced noise to about 10^{-15} coulomb, and the dynamic range was slightly more than two orders of magnitude. The charge emission waveform was then acquired by the Tektronics transient recorder and computer. We should point out that judging whether a sample is damaged or not is a controversial subject. The criterion we used was that damage occurred if any visible change could be observed under an optical Normarski microscope at up to 400x.

In general, the results of the charge emission experiments fell into two categories:

(1) No charge emission is observed until the material surface is damaged, or

(2) The surface emits charge prior to damage but there is no obvious correlation between charge emission and subsequent damage.

Silicon and the Cr-26 dielectric film fell into category 1, while copper and the remaining dielectrics fell into

category 2. For category 1, the quantity of charge emitted was always so great as to saturate the charge collection electronics.

Some experiments were conducted in which more than one pulse was incident on the same site. In this case, all the materials in category 2 showed a "cleaning" effect where the charge emission eventually became undetectable after a sufficient number of pulses of the same energy were incident on the site. Presumably, this effect is due to laser cleaning of surface contaminants by evaporation and ionization. The quantity of charge emitted showed no correlation with laser energy in all cases except one. For copper, the charge emission prior to damage was noisy but shows a monotonic increase with increasing laser fluence. This result is plotted in Figure 1. Because of the scatter in the data, there is no way to discriminate between a power law dependence of emitted charge on fluence or an inverse exponential dependence. Theoretically, the former dependence is associated with multiphoton processes, and the latter is associated with thermal emission over an energy barrier.

The morphologies of the damaged samples were studied using Normarski microscopy. In general, there was more

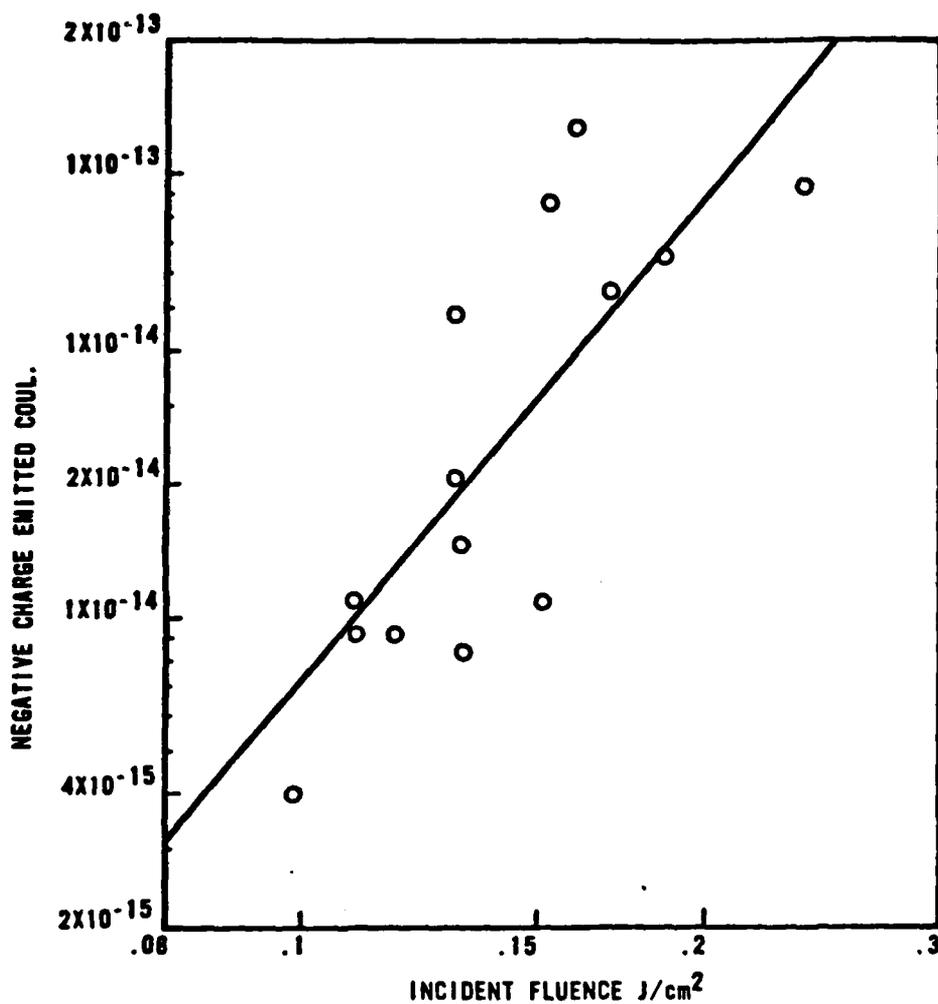


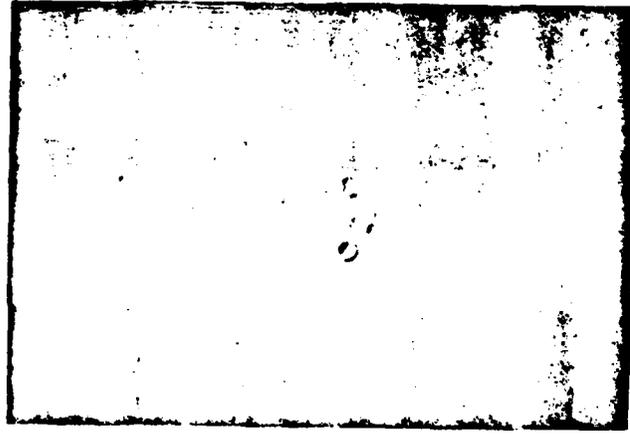
Figure 1. Negative charge emitted from a diamond-turned copper mirror versus laser fluence. Each point represents a single laser shot on a fresh site at an energy well below damage.

information to be gained from the more pure and ordered materials; i.e., more structure was visible on single-crystal silicon than on polycrystalline copper or on glassy dielectrics. Emphasis was placed on near damage threshold events only.

Typical results showed that the copper formed small melt pits on damage with no apparent correlation between the pits. At higher fluences, more uniform melting resulted in ripple and diffraction ring formation in addition to melt pits. Silicon showed a wealth of information and gave distinctly different results for single-pulse and multi-pulse damage. For single-pulse damage, shallow features are formed: grooves, ripples, and diffraction rings. But for multi-pulse damage, only correlated chains of pits are formed, and each chain always contains more than one pit. This damage morphology has been observed and models for it discussed for the case of picosecond laser pulses.² The high quality dielectric films damaged in what appear to be circular regions. These regions are of high contrast under the Normarski. A limited number of examples of the damage morphologies we have observed are shown in Figure 2.



a.



b.



c.

Figure 2. Typical single shot near threshold damage morphologies, Normarski micrographs.

a. Copper mirror, 524x, 5.60 J/cm²;

b. Silicon, 262x, 2.62 J/cm²;

c. Ta₂O₅ film on fused silica, photoconductivity sample, 262x, 21.6 J/cm².

In conjunction with the experiments on silicon, the nonlinear absorption of the 1-mm-thick wafers was measured at fluences up to the single-pulse damage threshold. The data are plotted in Figure 3 such that if two-photon absorption is present, a linear plot will result. This effect shows clearly in the figure. Unfortunately, the sample had very parallel faces, and 1.06 microns fell on an etalon resonance. If this were not the case the nonlinear absorption coefficient could be easily determined from the slope of the curve.

In summary, we find charge emission to be an indicator of surface cleanliness only for the first laser pulse. In some materials, it is a reliable indicator of the first damage event. Work remains on the significance of charge emission in the multi-pulse regime.

IV. Neutral Particle Emission.

Neutral particle emission data were collected from one sample each of copper and silicon. Neutral particles were detected with a fast ionization gauge Model FIG-1, manufactured by Beam Dynamics, Inc. The operating principles of the fast ionization gauge are identical with those of a standard Bayard-Alpert ionization gauge, but the dimensions

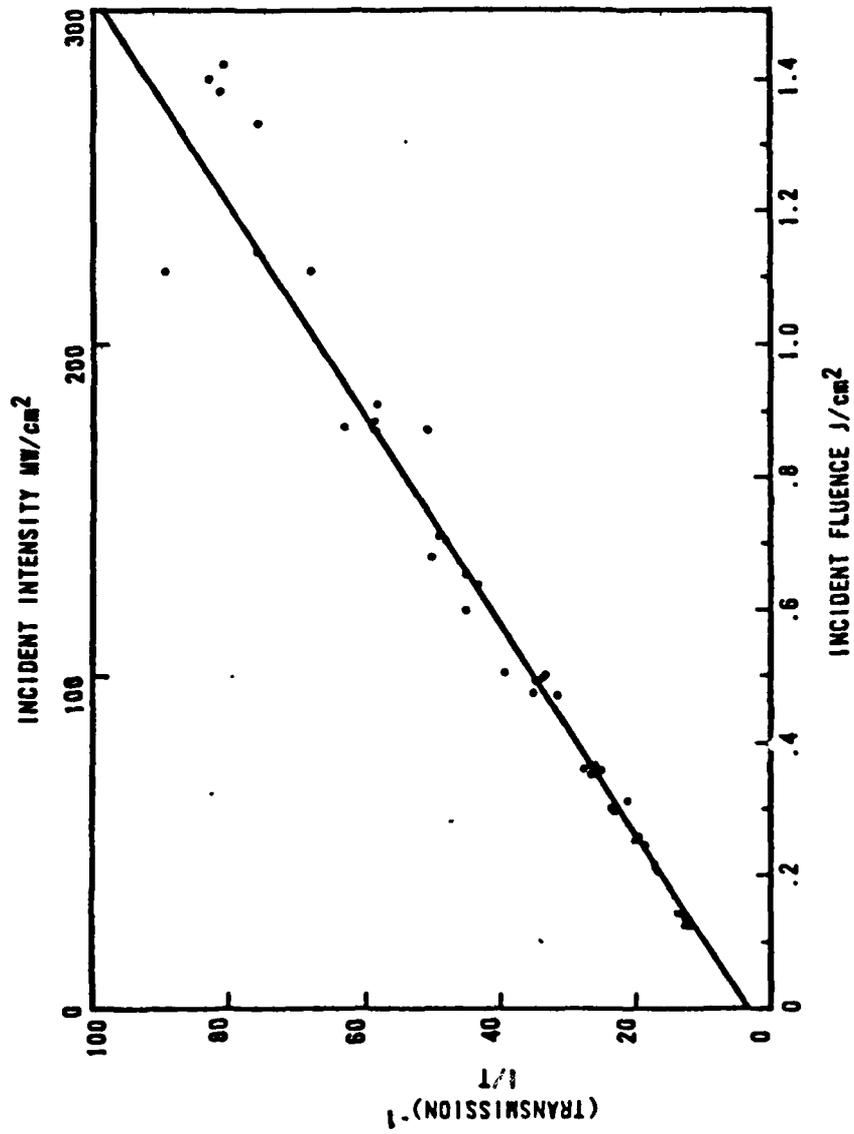


Figure 3. Inverse of sample transmission versus laser fluence for a 1-mm-thick silicon sample. Each point represents a single laser pulse, and the damage threshold is at the right edge of the figure. The laser pulse length was 5 nsec.

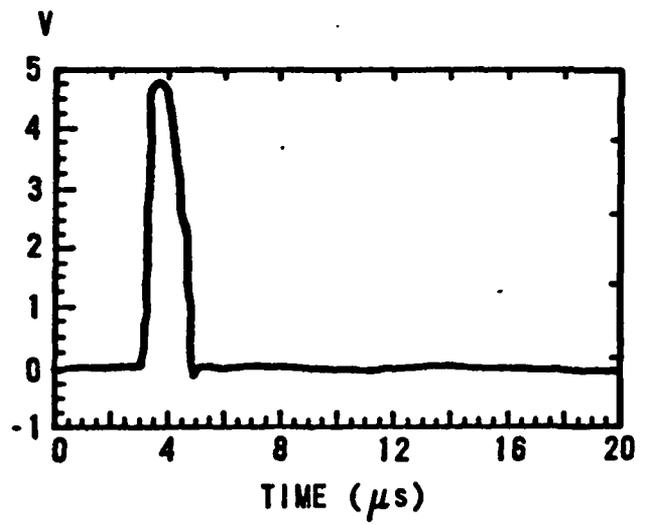
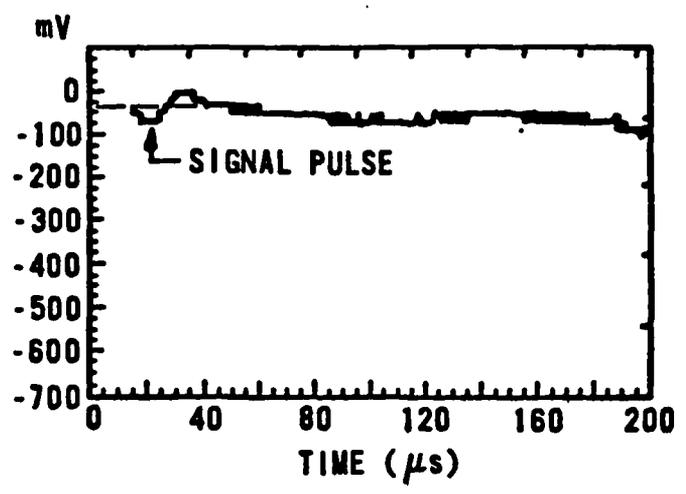
of the fast ionization gauge are significantly smaller, thus reducing the ion travel times from the grid region (where the ionization occurs) to the collector. The fast ionization gauge was controlled with a Model FC-1 Control Unit which regulated the emission current (continuously adjustable, $5\mu\text{A} - 3\text{mA}$) and provided a DC voltage output signal ($4\text{mV} - 10\text{V}$). The response time of the device was $3\mu\text{s}$, and the sensitivity was 1×10^5 volt Torr⁻¹ per mA of emission current. During these measurements, the emission current was nominally adjusted to 0.5 mA.

Figure 4 shows a typical output that was obtained when a copper sample was irradiated with a 3.1 J/cm^2 pulse, which is well above the single-shot damage threshold for copper at 2.1 J/cm^2 . The small neutral particle emission in Figure 4a can be compared to the saturated ion emission signal shown in Figure 4b.

Similar results were obtained for neutral particle emission from silicon.

The reason for the small signal was no doubt due to the small solid angle subtended by the fast ionization gauge, together with its small ionization volume (330 mm^3). This

(a) Neutral particle emission signal obtained with the FIG.



(b) Ion emission signal.

Figure 4. Neutral particle and ion emission from copper.

diagnostic was deemed too insensitive, based on the above results, and was not used in subsequent experiments.

V. Photoconductivity.

Photoconductivity data were collected for three samples of Al_2O_3 and one of Ta_2O_5 . Gold on chromium conducting strips 0.045 inches wide and $0.1\mu\text{m}$ thick were vapor-deposited on top of the dielectric coatings. The sample was mounted in a ceramic (MACOR by Corning) holder shown in Figure 5, and electrical connections to the conducting strips were made with gold-plated, electromechanical relay contacts. The conducting strips were deposited such that four different spacings between strips, called channels, resulted. These channel widths were 0.08, 1.21, 1.65, and 2.03 mm. Voltages were applied to the conducting strips such as to produce electric fields from about 5.0×10^4 V/m to 1.7×10^6 V/m. Charge flowing into the positive electrode was sensed with the charge amplifier described in Section III (Charge Emission). The output of the charge amplifier was sent to the transient digitizer.

Since the noise level of the charge amplifier output was approximately 0.020 mV and the digitizer saturated at approximately 2.2 V, the dynamic range of charge detecting

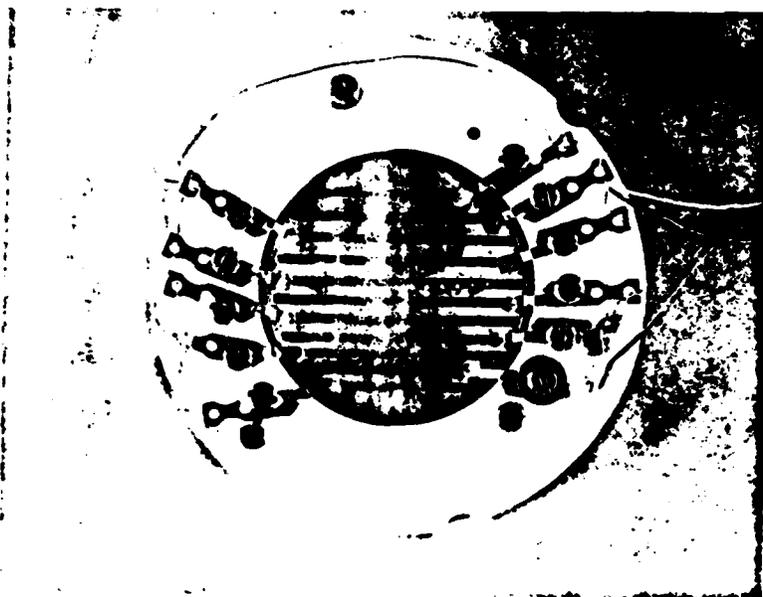


Figure 5. A photograph of a sample mounted in the ceramic sample holder. The small, evenly spaced dots across one of the conducting strips are laser-damage spots used as markers.

system covered about two orders of magnitude. It turns out that this range was not adequate to measure the amount of charge collected as a function of laser radiation for laser fluences adequate to cause damage. For laser fluences that resulted in an unsaturated charge collection signal, an attempt was made to find a functional relationship between the charge collected and laser fluence. The results for several samples (Al_2O_3 and Ta_2O_5) are shown in Table II. It was found that these data were uncorrelated on a linear, log-log, or exponential graph.

Laser Fluence (J/cm^2)	1.97	6.68	7.19	6.12	5.29	4.77	6.61	5.95	5.46	5.76	4.8
Charge Collected ($\times 10^{15}$ coulombs)	8.7	37.7	100	49.0	12.0	13.0	130	77.0	14.3	77.0	44.0

(a)

Laser Fluence (J/cm^2)	0.58	0.56	0.55	0.54	0.56	0.58	0.64	0.62
Charge Collected ($\times 10^{14}$ coulombs)	8.84	2.0	4.13	1.57	7.4	8.8	9.69	0.713

(b)

Table II. (a) Charge collected versus laser fluence
for Al_2O_3
(b) Charge collected versus laser fluence
for Ta_2O_5

The data in Table II show that for a given laser fluence, Ta_2O_5 conducts charge much more readily than Al_2O_3 , and that for either type of sample the charge conduction varied widely from site to site for a given radiation fluence.

In fact, for one Al_2O_3 sample (which was irradiated at 1.0 Atm), there was no apparent predamage charge conduction.

It was also found that for those samples of Al_2O_3 that were irradiated in a vacuum chamber ($P = 2 \times 10^{-6}$ Torr), predamage charge collection was readily apparent. This raises the question of whether the charge collected was actually due to the photoconductivity or charge emission. Since no provision was made in the experiment to distinguish between the two, one must conclude that either could have produced the observed signal.

To summarize the photoconductivity experiment then, several key items should be pointed out: (1) The dynamic range of the charge collecting instrumentation was too small to determine if a relationship between collected charge and laser fluence existed; (2) the data were very

site-specific, probably due to surface contaminants; and (3) some experimental provision should be made to determine whether the collected charge is due to photoconductivity or charge emission (see Recommendations).

VI. Electrostatic Probe.

It is possible that the accumulation of surface charge or surface dipoles or the change in surface potential might be a precursor to laser damage. An instrument to measure surface potential was researched, designed, and constructed during the course of this work. It utilizes a vibrating capacitor probe to measure surface potential. A recent method to do this is described in reference 3, and this type of device is reviewed in reference 4. Surface potential is sensitive to surface charge, surface dipoles, and adsorbed impurities. For this reason, it might be useful in the study of laser damage.

The probe utilizes an electromagnetic drive coil with a probe tip 1 mm in diameter. The tip may easily be vibrated from .1 to .25 mm excursion for voltages between 5 and 12 Vrms. A plan view of the probe is shown in Figure 6.

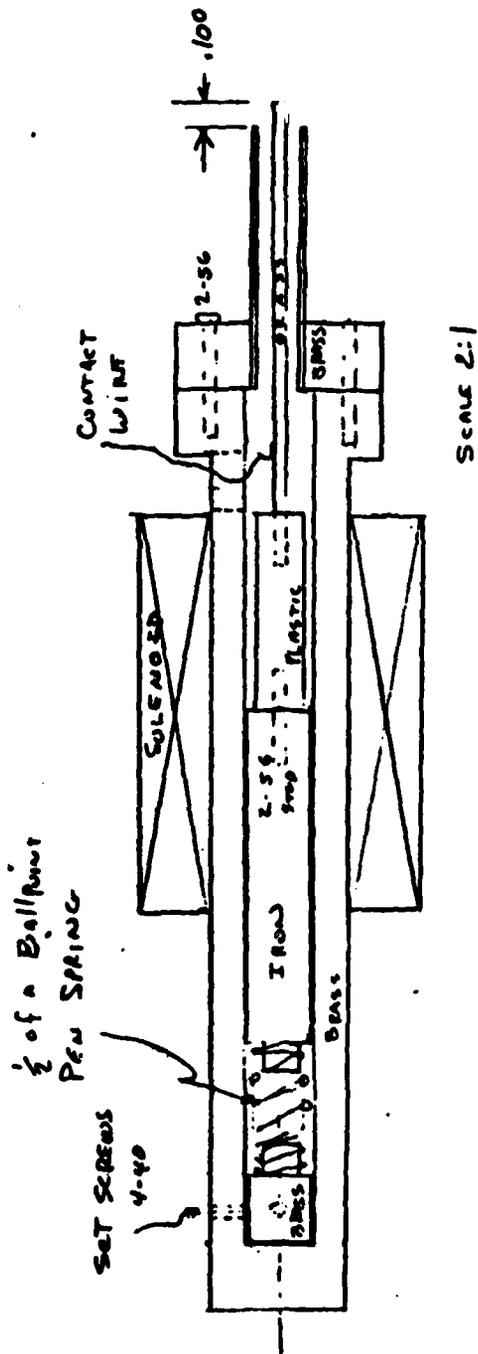


Figure 6. Plan view of the electrostatic vibrator probe for measuring surface potential.

At the end of the project, the probe had not been tested with a laser-damaged sample. The continuation of this investigation will be treated in Section VII.

VII. Recommendations.

Charge emission has been shown to be a useful diagnostic in laser damage studies. Some improvements can still be made, and we make the following recommendations:

a. It is recommended, of course, to raise the saturation level of the charge amplifier so that the dynamic range well exceeds the level at damage threshold.

b. It is recommended that multi-pulse experiments be conducted to separate cleaning effects from damage and pre-damage phenomena.

c. It is recommended that an energy analysis system be added to the system to facilitate photoelectron or photoion spectroscopy. Data obtained from such experiments may be very useful in determining the energetics of laser damage and lead to useful microscopic models.

d. It is recommended that a sensitive mass spectrometer be added to the vacuum chamber to help identify the liberated species. The mass spectrometer would also be useful to determine the cleanliness of the surface.

It is still uncertain whether surface conductivity will be a useful diagnostic for laser damage. Because conductivity is significantly effected by surface contaminants⁵ it is recommended that conductivity measurements similar to those described in Section V be carried out in a vacuum chamber capable of reaching pressures well below those achieved with the present system. To estimate a reasonable background pressure, we may use the fact that t_1 is approximately $10^{-6}/P$, where t_1 is the time in seconds for one monolayer of a contaminant to adsorb on an atomically clean surface and P is the background pressure in Torr. Thus, for example, if one assumes a perfectly clean surface in a background pressure of 10^{-9} Torr, one monolayer of contaminants will be adsorbed in approximately 10^3 seconds. At these pressures, a sample could be cleaned via photon desorption while in the vacuum chamber and afford the experimenter ample time to collect data.

Finally, we recommend that tests using the surface potential measuring device be conducted as this technique still looks promising.

References

1. The field is well documented by the proceedings of the 1st through 14th Annual Boulder Damage Symposia which are printed as NBS Special Publications.
2. See for example: M. F. Becker, R. M. Walser, Y-K. Jhee, and D. Y. Sheng, "Picosecond Laser Damage Mechanisms at Semiconductor Surfaces," SPIE Vol. 322, Picosecond Lasers and Applications, p. 93, Bellingham, Wash., SPIE (1982) and the references contained therein.
3. J. C. Campuzano and R. G. Greenler, "Instrument for Combining Reflection-Absorption Infrared Spectroscopy with Other Surface-Sensitive Techniques," Rev. Sci. Instrum. 52, May 1981, pp. 678-683.
4. N. A. Surplice and R. J. D'Arcy, "A Critique of the Kelvin Method of Measuring Work Functions," Journal of Physics E: Scientific Instruments 3, 1970, pp. 477-482.
5. K. L. Mittal, ed., Surface Contamination Vol. 2, New York, Plenum Press, 1979, p. 967.

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

HYPERBARIC OXYGENATION (HBO) ALTERATION OF METABOLISM

AND CARDIOVASCULAR FUNCTION

DURING AND FOLLOWING EXERCISE CONDITIONING

Prepared by: James W. Dooley
Academic Rank: Assistant Professor
Department and University: Division of Health, Physical Education and Recreation. Wright State University
Research Location: Hyperbaric Medicine Division, School of Aerospace Medicine, Brooks AFB, San Antonio, Tx
USAF Research: LTC Richard A. Henderson, M.D.
Date: 29 August 1983
Contract No: 49620-82-C-0035

HYPERBARIC OXYGENATION (HBO) ALTERATION OF METABOLISM
AND CARDIOVASCULAR FUNCTION
DURING AND FOLLOWING EXERCISE CONDITIONING

by

James W. Dooley

ABSTRACT

Qualification for compression chamber operations was secured by enrollment in and completion of the Compression Chamber Team Training (CCTT) course, B30ZY9300-007, PDS code WS3. Since time and equipment availability restrictions threatened the completion of the proposed study, the scope of the study was expanded, and a new proposal was developed for submission to the Advisory Committee on Human Experimentation. The newly proposed project will be attempted in the Summer of 1984 and entails measurement and evaluation of acute and chronic physiologic responses to eight weeks of aerobic exercise conditioning performed at 2 ATA in hyperbaric oxygenation (HBO) conditions. Physiologic and metabolic variables include heart rate, ECG, impedance derived stroke volume (ZCG) for cardiac output determination, blood pressure, $\dot{V}O_2$ max, serum lipids and lipoproteins, complete blood count (CBC) and percent body fat.

Title: Hyperbaric Oxygenation (HBO) Alteration of Metabolism and Cardiovascular Function During Exercise and Following Exercise Conditioning.

Project/Task/Work Unit: SUPTXXHM

Principal Investigator: James W. Dooley, Ph.D., SAM/HM (SFRP Fellow)/3281

Associate Investigators: Lt Col Richard A. Henderson, M.D., SAM/HM/3281

Chief, Clinical Investigations Branch

Loren Myhre, Ph.D., SAM/VNB/3814

James C. Miller, Ph.D., SAM/VNS/3464

Medical Consultant: Col Frederick S. Cramer, SAM/HM/3281

Statistician: Joseph R. Fischer, Jr., SAM/BRA/2818

Purpose/Objective of the Study:

The purpose of this prospective study is to determine the effects of hyperbaric oxygenation (HBO) on the metabolism and cardiovascular function of man during aerobic exercise conditioning. It is the primary objective of this study to obtain data that will expand our knowledge of the physiologic and metabolic responses to an HBO exercise conditioning regimen and to more clearly define the roles of oxygen transport and oxygen utilization in the limitation of human performance.

Background:

Hyperbaric oxygenation (HBO) involves the use of 100% oxygen (O_2) as a breathing medium while entirely enclosed in a hyperbaric chamber at a pressure greater than one atmosphere absolute (1 ATA). HBO has been proven effective in the treatment of various trauma and disease conditions. The risks of central nervous system (CNS) and pulmonary toxicity associated with HBO have been minimized through adherence to treatment schedules that have been developed from several decades of investigation into the effects of hyperoxic and/or hyperbaric conditions on human performance.

In therapeutic use, HBO enhances oxygenation of previously ischemic (hypoxic) tissue, thereby causing reduction or elimination of pathologic hypoxia and acceleration of cellular metabolism. This increased tissue oxygenation often occurs despite a reduction in blood flow, due to the tremendously increased O_2 availability per volume of blood. The increased blood O_2 concentration is primarily the result of an increase in dissolved O_2 during hyperbaric exposure. It is this characteristic of HBO that has interested research physiologists who have attempted to determine the effects of HBO on work performance.

Results of studies conducted on the effects of HBO on work performance have indicated a definite acute alteration of both maximal and submaximal performance parameters. The effects are directly related to the magnitude of the pressure and the duration of the HBO exposure. Physiologic variables studied include heart rate, O_2 consumption ($\dot{V}O_2$) and minute ventilation (\dot{V}_E).

During rest and/or submaximal exercise at pressures less than or equal to 2 ATA, HBO suppresses heart rate, maintains or increases $\dot{V}O_2$, and decreases \dot{V}_E . In comparison to exercise response in normal air conditions, maximal exercise under HBO conditions produces no change in maximal heart rate; a slight increase in $\dot{V}O_2$ max; and a decrease in \dot{V}_E max. Generally, HBO enhances work tolerance if excessive pressure or duration of exposure is avoided.

Prolonged O_2 exposure even at 1 ATA can result in pulmonary toxicity, and at 3 ATA, even mild exercise while breathing O_2 can elicit a convulsion (CNS toxicity). These toxic reactions vary with the individual, but guidelines have been established to minimize the risk of toxic responses. It is considered safe to utilize HBO during maximal exercise at pressures of 2 ATA (33 fsw) or less in the dry hyperbaric chamber. The wet (submerged in water) environment is more stressful, and vigorous exercise performed in water at pressures greater than 1.9 ATA (30 fsw) is contraindicated. Though O_2 is not recommended in extreme pressure conditions, O_2 tolerance at lesser pressures is generally improved when an intermittent O_2 -air breathing protocol is employed.

Although submaximal HBO exercise studies have identified acute physiologic responses, there is no documented evidence of chronic physiologic (or metabolic) responses to HBO exercise conditioning. Since maximal exercise at 2 ATA in dry HBO conditions is considered safe, a submaximal, intermittent O_2 -air exercise conditioning protocol at 2 ATA appears to be a safe, viable means of testing chronic physiologic and metabolic responses to repeated HBO exercise exposures.

Various non-invasive devices of established validity are available and can be utilized to assess physiologic and metabolic changes including those in heart rate, cardiac output, blood pressure, $\dot{V}O_2$ max, body composition, blood lipids and lipoproteins, and blood hemoglobin, hematocrit, and red and white cell counts. The use of impedance cardiography provides the means for determining acute and chronic responses of cardiovascular function to HBO exercise. Automated blood chemistry and complete blood count analyses, open-circuit spirometry assessments of $\dot{V}O_2$ max, and body volume determinations provide valid means for determining metabolic responses to HBO exercise conditioning.

Hazards: Possible hazards to human volunteers include:

1. Inherent cardiovascular risks related to exercise; heart rate and blood pressure will be elevated well above resting levels. Though not likely, the exercise may elicit abnormal heart rhythms or excessively elevated arterial blood pressure, which could result in heart attack or vascular lesion;

2. Possible toxic responses to oxygen (convulsion, pulmonary irritability) for HBO subjects. Convulsion (CNS toxicity) is not unlike the grand mal seizure of epilepsy. Injury may result from objects struck due to flailing during the convulsion. Pulmonary irritability with its associated labored breathing may develop from breathing 100% O_2 ;

3. Inability to escape in event of fire. Precautions have been taken to ensure that possibility of fire inside (or outside) the chamber is extremely remote;

4. Ear and sinus problems related to pressure changes;

5. Decompression sickness (DCS) and air embolism from hyperbaric chamber ascent following the pressurization period.

Precautions: Precautions to avoid the above possible hazards:

1. All subjects will be under 36 years of age who have passed a Class III physical examination with no contraindications to exercise as determined from an exercise stress ECG;

2. The short duration of the HBO sessions virtually eliminates the hazard of pulmonary toxicity. CNS toxicity risk will be minimized by keeping HBO conditions within the limits where there is no reported evidence of convulsion:

a. Exercise will be performed in an atmospheric pressure not to exceed 2 ATA (33 fsw). At 2 ATA there have been no reported convulsions in "dry" HBO conditions;

b. Total oxygen breathing time will be restricted to a maximum of 24 minutes (3 stages of 8 minutes each) for subjects in the HBO treatment protocols;

c. An intermittent work and breathing medium regimen will be followed. For HBO subjects, 6 to 8-minute O₂ breathing exercise bouts at a submaximal work load are followed by 2-minute air-breathing, work-relief (unloaded pedalling) intervals. So, not only is the O₂ removed periodically, but a simultaneous work-relief interval is provided. This intermittent O₂-air protocol extends O₂ tolerance in resting and exercising conditions;

d. Work loads, initially 49 to 98 watts (W) and increasing to a maximum of 147 W, are light to moderate work loads for healthy young subjects. Individualized work loads will be calculated not to exceed those that elicit greater than 55 % of the pre-determined maximum oxygen consumption (VO₂ max);

e. During all sessions, observers will monitor and assist the subject should any problem arise;

f. A flight surgeon (medical officer in charge) will be in the immediate chamber area during all sessions and will be responsible for any medical treatment should the need arise;

g. Subjects will be briefed on the usually occurring subjective feelings of cycle exercise and will then be instructed to report any "unusual" feelings at the time of occurrence. If these feelings continue following a relief interval the session will be terminated and the chamber will be ascended to ground level;

3. Fire hazard precautions include:

a. Removal of jewelry, watches, and other items that may cause sparking/electrical arcing;

b. Use of ventilated masks with breathing media intake and an overboard dump exhaust that vents CO₂ and expired O₂ to the outside air during hyperbaric dives, keeping the chamber air between 21 and 23 percent O₂.

c. A fire suppression system (water sprinklers) and a hand-held hose are located inside the chamber;

d. Protective flame-resistant dive suits will be worn by all persons in the chamber;

e. Plumbing of electrical equipment is such that possible spark-producing units are located outside the chamber;

4. As an extreme protection against development of decompression sickness, all persons will be advised not to fly for up to 24 hours following the hyperbaric dive, and vigorous exercise will be contraindicated for 12 hours following the dive (AFR 50-27, paragraphs 8b, c). As a precaution in the avoidance of air embolism, subjects will be instructed before and during each dive not to hold their breath or perform the valsalva maneuver during ascent of the chamber.

5. Middle ear and sinus pressure equalization problems will be minimized as follows:

a. Subjects and inside observers will not be allowed to dive if they have a severe cold or upper respiratory congestion;

b. Chamber descent will not exceed one foot per second, and the chamber operator will "level off", when directed by the inside observer, if the subject has difficulty with pressure equalization. Descent will be slow enough to allow equalization;

c. Subjects will be instructed in the usage of the valsalva maneuver during descent to keep sinuses and airways to the middle ears patent;

d. Chamber ascent will be performed at three feet per minute to minimize the remote possibility of air embolism;

e. Dry chamber dives will be restricted to 2 ATA (33 fsw) for a maximum time of 38 minutes. At this depth and time exposure, the risk of developing DCS upon ascent is virtually non-existent. Thirty-three feet is a no-decompression dive depth.

Emergency Procedures: Should precautions fail, the following emergency procedures will be followed.

1. Subjects complaining of any illness will be referred to the flight surgeon (medical officer in charge) for appropriate treatment or referral to the USAF Clinic Flight Surgeon.

2. Conditions requiring medical attention will be handled by the flight surgeon.

3. CNS toxicity, as evidenced by convulsion, obviates termination of O₂ breathing and the use of "loose restraint" of the subject. Once the seizure has subsided, the subject will be ascended to the surface. The flight surgeon will direct all therapeutic action.

4. Hyperbaric chamber technicians will follow published USAF safety guidelines (T.O. 43D8-3-4-61, paragraph 3-18; 21 September 1978 with Change 1, 12 August 1981) for emergency procedures for fire inside or outside the chamber. Emergency procedures are posted at the chamber operator location.

5. Appearance of symptoms of air embolism will call for initiation of the appropriate treatment dive table (Table 6A) under the direction of the flight surgeon (medical officer in charge).

Hazardous Duty Pay: Subjects will be selected from Brooks AFB personnel and will receive hazardous duty pay.

Normal Physiological State: The effects of 100% O₂ and exercise in a hyperbaric environment of 2 ATA are not related to a significant departure of subjects from the "normal" physiological state. However, a 24-hour period of non-flying should be observed following any hyperbaric exposure (AFR 50-27).

Equipment Malfunctions: Equipment malfunctions that would terminate the experiment are as follows:

1. Long-term loss of functional availability of the hyperbaric chamber due to loss of clean air and/or oxygen sources.
2. Structural damage to chamber, e.g., loss of chamber wall integrity.

Requirements for Human Volunteers:

1. Numbers Required: Eight (8) experimental subjects are required. Subjects will experience one of two exercise protocols. Four (4) control subjects will also be tested but will not participate in the exercise protocol.

2. Date and Total Number of Days Required: All subjects will be scheduled for an orientation session during which time the experimental protocol and subject responsibilities will be explained. Arranged appointment times for conditioning sessions will be suited as closely as possible to the optimal availability of the subject, between 0730 and 1630, Monday, Wednesday and Friday during a 8-week conditioning period from mid-June through August 1984.

3. Duties and Responsibilities of Human Volunteers: The duties and responsibilities are as follows:

a. Inform principal investigator and/or flight surgeon of any change in medical status, including the use of any prescription or non-prescription medication.

b. Keep scheduled appointments for the experimental sessions or reschedule in the case of unavailability for the previously scheduled appointment time;

c. Follow instructions regarding experimental protocol as given by the principal investigator.

4. Special Requirements of Human Volunteers:

- a. Subjects must have passed a Class III physical examination.
- b. Subjects must be males under 36 years of age.
- c. Subjects cannot be smokers.
- d. Subjects must not be under, or on medication for, treatment of a medical disorder.
- e. Subjects may be recreationally active but must be untrained (as determined by the investigators).

EXPERIMENTAL METHOD

Conditions of the experiment:

1. Drugs: 100% oxygen (O₂) will be administered intermittently as breathing medium to HBO subjects in the hyperbaric chamber at an ambient pressure of 2 ATA (33 fsw). The dosage will be 100% O₂ for six to eight minutes at 2 ATA for three intermittent exercise periods with two minutes of

unloaded exercise-relief on air between loaded ergometer exercise periods. Total maximum O₂ administration time is 26 to 32 minutes over a 32 to 38-minute total time period.

Grand mal seizure is the most apparent symptom of CNS toxicity. Dosage of O₂ is set to a level below that considered sufficient to evoke convulsion in exercising subjects in a dry hyperbaric environment. O₂ will be alternated with air to provide even greater protection against the toxic effects known to occur at greater depths and/or greater times of exposure.

2. Chamber Requirements: The chamber will be compressed to 2 ATA, a "depth" equivalent to 33 feet of sea water (fsw). The descent rate will not exceed one foot per second. Time at depth will be 32 to 38 minutes:

- a. 1 X 5-minute adjustment period (O₂) = 5 minutes (rest)
- b. 1 X 3-minute warm-up period (O₂) = 3 minutes
- c. 3 X 6 to 8-minute exercise period (O₂) = 18 to 24 minute (exercise)
- d. 3 X 2-minute recovery period (air) = 6 minutes (recovery)

Ascent and descent time will take an approximate total time of 11 to 12 minutes. The time of descent (1 to 2 minutes) will be dictated primarily by each subject's ability to equalize pressure in the middle ear and sinus cavities. Ascent will be at a rate of 3 feet per minute and will require approximately eight minutes.

To maintain temperature, an Environmental Control System (ECS) will "air-condition" the chamber, keeping the temperature at approximately 23°C during all excursions. Temperature will be set and monitored by the outside chamber operator.

All subjects and chamber technicians will wear vented noise suppression ear covers. Lighting will be provided from lamps outside the #2 chamber or from electrically insulated lamps inside chamber #1.

Humidity will be kept at an acceptable level by the ECS and will be recorded as percent relative humidity on a Bacharach SERDEX Hygrothermograph. Since the exercise is of such low intensity and short duration and breathing media are humidified by an in-line micronebulizer, no problem with dehydration is expected.

3. Objective Measurements:

a. **Electrocardiogram (ECG).** A lead II appearing ECG will be periodically recorded at rest, during exercise, and during recovery. The ECG monitor is provided as a component of the Minnesota Impedance Cardiograph used to simultaneously determine stroke volume. The ECG will provide heart rate (HR) data.

b. Impedance Cardiogram (ZCG). A tetra-polar thoracic impedance cardiogram (ZCG) will provide stroke volume data simultaneous to the HR recording for the ECG. This allows for calculation of cardiac output (CO). Interpretation of the ZCG also provides for the measurement of ventricular ejection time (VET), a measure of cardiac contractility.

c. Maximal Oxygen Uptake ($\dot{V}O_2$ max). This measurement of maximal aerobic power is provided by the open circuit spirometry determination of the minute volume of the subject's expired air (\dot{V}_E) and concentrations of oxygen (O_2) and carbon dioxide (CO_2) in the expired air at maximal exercise on the treadmill. $\dot{V}O_2$ max will be expressed as milliliters of oxygen per kilogram of body mass per minute ($ml\ O_2\ kg^{-1}\ min^{-1}$).

d. Serum lipids and lipoproteins. Cholesterol, triglycerides and high density lipoprotein (HDL) cholesterol will be determined from valid automated blood chemistry analyses of serum from blood samples drawn from fasting subjects (12 to 15 hours post-prandial). The ABA-100 will be utilized for determination of serum cholesterol and triglyceride levels. HDL cholesterol will be determined by the phosphotungstate procedure.

e. Complete blood count (CBC). Additional blood samples will be analyzed for red blood cell (RBC) concentration, hematocrit (Hct), hemoglobin (Hb), and white blood cell (WBC) concentration according to standard USAF Hospital laboratory procedures.

f. Blood Pressure (BP). Systolic blood pressure (SBP) and diastolic blood pressure (DBP) will be taken and recorded at rest and, periodically, during exercise, one minute prior to the recording of ECG and ZCG. To avoid the hazard of mercury spillage from a mercury sphygmomanometer in the hyperbaric chamber, an aneroid cuff and gauge will be used. The cuff will be wrapped and secured around the subject's arm throughout the exercise session. The head of an anaesthetist's stethoscope will also be secured to the subject's arm to maintain a constant auscultation monitoring site. DBP will be recorded as that pressure reading (mm Hg) corresponding to the fourth Korotkoff sound since the fifth sound is often difficult to distinguish during exercise. Pulse pressure (PP) and mean arterial pressure (MAP) will be calculated from systolic and diastolic measurements.

4. Research Design

a. Subjects: Twelve volunteer male subjects under 36 years of age will be randomly assigned to one of three groups of four subjects each. Subjects may be involved in recreational activities but will be untrained. Informed consent will be obtained from all subjects.

b. Physiologic and Metabolic Assessments: Control and exercise conditioning subjects will be tested for determination of $\dot{V}O_2$ max, body composition, blood lipids and lipoproteins, and complete blood count prior to and following the experimental period. Additionally, exercise conditioning subjects will be periodically tested for heart rate, blood pressure, stroke volume, and $\dot{V}O_2$ responses to submaximal work during the last two minutes of each of the three 6 to 8-minute exercise bouts.

c. Treatments

(1) Hyperbaric oxygenation (HBO). Following a 5-minute resting hyperbaric oxygenation adjustment period at 2 ATA (30), each HBO subject will pedal at a frequency of 50/minute on a cycle ergometer for three minutes at a work load setting of 300 kgm min⁻¹ (49w). Pedalling frequency will be maintained by use of a metronome. Following a 3-minute warm-up period, the work load will immediately be set to a calculated work load that represents 55% of his previously determined $\dot{V}O_2$ max. However, work loads will not exceed 900 kgm min⁻¹ (147w). Exercise bouts will last for six minutes during week one and will increase by one minute each week until week three when exercise bouts last for eight minutes. Exercise bouts will remain at eight minutes thereafter. The three successive exercise bouts will be interrupted by 2-minute zero-load pedalling recovery intervals during which time the subject's mask will be removed, and he will breathe chamber air.

(2) Normobaric air (NBA). NBA subjects will perform a 3-minute warm-up at 49w and will follow the same exercise protocol as the HBO subjects, except NBA subjects will not be dived and will breathe room air.

(3) Controls (c). Control subjects will not participate in the exercise conditioning portion of the study but will be tested along with HBO and NBA subjects prior to and following the exercise conditioning period. They will be instructed to maintain their routine activity and dietary patterns throughout the experimental period.

d. Statistical Analyses: Repeated measurements analyses of variance (ANOVA^S) will be used to test for acute and chronic HBO effects for each of the selected dependent variables. Tests of pairwise comparisons will be performed on the mean values of variables exhibiting significant variance. The .05 level will be accepted as statistically significant.

Support Personnel and Equipment:

a. Equipment:

(1). The number 1 or number 2 hyperbaric chamber in the Hyperbaric Medicine (HM) Division facility is available and will be utilized for the study. The chambers are plumbed for ECG, ZCG and defibrillator utilization and can easily accommodate an exercising subject on a cycle ergometer and an inside observer.

(2). Two mechanically-braked Monark cycle ergometers will be available. One each will be provided by HM and VNB of the USAF School of Aerospace Medicine.

(3). Two Marshall anaeroid sphygmomanometers will be utilized for blood pressure measurements.

(4). A Bird 500 cc in-line nebulizer will be used to humidify the O₂ supply for the HBO subjects.

(5) The O₂ breathing medium will be supplied by a liquid oxygen source and regulated (dialed to the "safety" setting) by a Diluter Demand Oxygen Pressure Breathing Regulator, Type A-14, manufactured by Firewel Co., Inc.

(6) An IFM/Minnesota Impedance Cardiograph (Instrumentation for Medicine, Greenwich, Connecticut), model 304B, will be provided by USAFSAM/VNS and utilized to obtain ZCG data for determination of left ventricular stroke volume. Eight silver-silver chloride spot electrodes will replace the conventional aluminized circumferential electrodes on mylar strips. The 304B provides ECG and heart sounds data as well as impedance wave form data for stroke volume determination.

(7) A Perkin-Elmer 1100 Medical Gas Analyzer (mass spectrometer) will be utilized to determine the percent O₂ and CO₂ in the HBO breathing medium.

(8) $\dot{V}O_2$ max will be determined from a stress ECG test performed by the medical personnel in the Clinical Sciences Division of the School of Aerospace Medicine according to a standard protocol.

(9) A Scott Fire-Fighter clean vinyl face mask will be worn by all subjects during exercise sessions. A tygon tube has been fitted to the exhaust port to provide an overboard dump of expired air for HBO subjects during 100% O₂ breathing periods. An intake hose receives O₂ from Diluter Demand Oxygen Pressure Breathing Regulator previously described.

(10) A Bacharach SERDEX Hygrothermograph will be utilized in the chamber to record temperature and relative humidity.

b. Personnel: The hyperbaric chamber team members have all completed the Compression Chamber Team Training Course (B30ZY9300-007) and/or hold the appropriate AFSC designation qualification.*

(1) Chamber Operator - has direct control of ascents and descents of chamber as required; keeps close visual (T.V. monitor) and intercom contact with persons inside the chamber; and selects appropriate breathing medium.

(2) Recorder - monitors time, coordinates the the O₂-air breathing schedule, and records the collected physiologic data.

(3) Crew Chief - prepares the chamber for each dive; assures equipment availability and proper functioning; checks and records percent O₂ in breathing media; and is available for troubleshooting during any emergency or malfunction.

(4) Inside Observer (Investigator) - keeps close personal contact and communication with the subject; assures subject compliance with experimental exercise protocol inside the chamber; takes and reports BP at the appropriate times as coordinated by the Chamber Operator.

(5) Aerospace Physiologist - overall officer in charge of the dive; may simultaneously act in any non-medical crewmember position outside the chamber; assures that established procedures (AFP 161-27 July 1983) are followed during compression operations.

(6) Medical Officer, Flight Surgeon - responsible for all therapeutic procedures in compression operations; functions as an advisor but may act in any crew position previously described.

*Compression Chamber Team Member Duties and Operating Procedures (with additionally specified duties for this project) AFP 161-27, July 1983, Chapter 14.

Pertinent References

SCOPE OF THE PROBLEM

Recent publications (66, 72, 82) have documented the effectiveness of hyperbaric oxygenation (HBO) therapy in the treatment of various trauma and disease conditions. Carbon monoxide poisoning, osteoradionecrosis, refractory osteomyelitis, burns, skin and bone grafts, gas gangrene, decompression sickness, and gas embolism are conditions most commonly treated in hospitals of this country (66, 72). Internationally, especially in the Soviet Union, these and other conditions including chronic ischemic heart disease, cerebral stroke, and bronchial asthma are routinely treated with HBO (82).

Although results have been gratifying, development of HBO therapeutic protocols has not been without controversy. Lack of reliable, objective evidence of HBO benefits in earlier years (72) and the old and recent documentation of the toxic effects of oxygen (O₂) in hyperbaric environments

(8-11, 13, 18, 21, 24, 40, 41, 56, 68, 70, 88, 95) have understandably required justification for the clinical use of HBO. But it is extremely important to understand that the very studies that have documented the toxic effects of O₂ in hyperbaric environments have been invaluable in providing the basis for the development of safe standards for O₂ use in diving, medical therapy, and future hyperbaric research (9,10). Safety and untoward effects of HBO exercise protocols will be addressed later in this proposal.

The benefits of hyperoxia, in both normobaric and hyperbaric conditions, with regard to exercise or work conditions have been addressed primarily in relation to cellular metabolism and cardiovascular function. Study of skeletal and cardiac muscle metabolism and the O₂ transport that supports it has provided researchers with an ever increasing knowledge of the mechanisms involved in the limitation of sustained physical effort.

While normobaric hyperoxia has resulted in increased performance times (2, 7, 14, 19, 25, 61, 84, 91, 92) there is a lack of agreement on the causative mechanisms. Several investigators (25, 29, 43, 59, 86) have reported increases in oxygen uptake ($\dot{V}O_2$) during O₂ breathing, concluding that O₂ transport via the central circulation was the limiting factor for maximal aerobic work. Others (2, 7, 14, 19, 61, 84, 91, 92) were either unwilling to attribute $\dot{V}O_2$ max increases to O₂ transport or felt that the limitation of maximal aerobic work was controlled at the tissue level. Two studies reported a decreased blood flow to working muscle either during (85) or immediately following (71) submaximal work while breathing hyperoxic gas

mixtures. Additionally, Welch, et al (85) reported no change in $\dot{V}O_2$ during hyperoxia and argued that a greater extraction of O_2 from the diminished blood flow was a reflex cellular adjustment to meet cellular demands.

Hyperoxia has been shown to reduce the energy cost of ventilation (14, 25, 44, 91, 92) and control the hydrogen ion concentration (2, 19, 61) during exercise. Increased muscle myoglobin storage of O_2 (17) may also be a factor in the improved work tolerance observed during hyperoxia.

Interpreting cardiovascular, ventilatory, and metabolic data from hyperbaric hyperoxia (hereafter referred to as HBO) studies is complicated by the direct effects produced by increased ambient pressure. Air is composed of nitrogen, oxygen, and carbon dioxide, all of which exert narcotic effects at increased atmospheric pressures (41, 46, 56). Disregarding narcosis, the increased density of air in hyperbaric conditions poses a major problem by increasing the ventilatory energy cost during exercise (17a, 27, 32, 33, 35, 65, 70, 73). When helium is substituted for nitrogen, there is a significant decrease in the density of the breathing medium and a corresponding decrease in energy cost of ventilation (16, 65, 73).

The decreased resting and submaximal exercise heart rate (bradycardia) observed while breathing air at depth (17a, 26, 27, 32, 33, 65, 73, 93) is also observed when helium is substituted for nitrogen (16, 27, 32, 73), indicating that heart rate suppression is not a response to nitrogen narcosis. In long "saturation dives", there are reports of an ultimate loss

of the phenomenon of bradycardia (16, 78, 93), a phenomenon in itself. Bradycardia and the eventual return to normal rate are without adequate explanation (16, 33, 78).

When O_2 is breathed in hyperbaric environments, resting and submaximal exercise heart rates (and in one case (6), maximal heart rate) continue to be suppressed, this suppression being somewhat greater than when air is breathed (6, 47, 79). There appears to be an O_2 -independent and an O_2 -dependent bradycardia effect.

Oxygen uptake ($\dot{V}O_2$) during maximal or submaximal work in HBO conditions has been reported to be greater than that observed in hyperbaric normoxic conditions (6, 31, 79). Others (47, 70) have reported no significant differences in $\dot{V}O_2$ during HBO exercise. One author (35) actually reported a decreased $\dot{V}O_2$ (submaximal) at 2 ATA. Whether or not $\dot{V}O_2$ was increased, the concurrent lowering of heart rate (relative bradycardia) has resulted in an increased O_2 pulse (ml O_2 /heart beat), an indication of increased O_2 transport to and/or uptake by metabolically active tissues despite a lowered heart rate (33).

Blood pressure and cardiac impedance measurements during HBO exercise have not been addressed. Behnke, et al (11) reported a rise in systolic and diastolic pressures in resting subjects after prolonged O_2 exposure at 4 ATA. A coincidental rise in resting heart indicated that the blood pressure and heart rate responses were O_2 toxicity responses. Transthoracic electrical impedance has been used to measure ventricular stroke volume in subjects during resting conditions on saturation dives (78).

The development of an electrical impedance cardiograph (53, 54) has permitted a passive, non-invasive determination of ventricular stroke volume from which cardiac output can be calculated. Recent advances in technology have enabled the development of automated impedance cardiography systems (34, 64, 75) that provide valuable cardiac function data. Impedance cardiography is not without limitations (48, 62, 74), however it has been established as a reliable method of determining intra-individual changes in stroke volume in healthy resting and exercising subjects (4, 12, 22, 28, 52, 58, 63, 77, 81). Impedance cardiography provides less reliable data when used with cardiac patients (36, 38). Comparisons with other methods of cardiac output determination (22, 28, 51, 62, 69) indicate that impedance cardiogram (ZCG) correlation values range from fair to good with regard to inter-subject variance, but are highly reliable with regard to intra-subject variance.

Although several authors (60, 89) found it necessary to use intermittent pauses during exercise to reduce the effects of upper limb ZCG movement artifacts, respiratory motion artifacts can be effectively screened, and ZCG data can be automatically collected and analyzed (23).

Several medical studies have attempted to evaluate the effects of hyperoxia and HBO on specific cardiac and cardiovascular function parameters in dogs (20, 37, 50, 59, 76, 83) and humans (1, 15, 49, 87). Hyperoxia decreased coronary flow and contractile force of the myocardium in open-chest, anaesthetized dogs (20). Heart rate and blood pressure were unaffected. HBO studies performed on resting dogs at atmospheric pressures ranging from 2 to 4 ATA (37, 50, 76, 83) indicated a decrease in heart rate, stroke volume (and

thus cardiac output), coronary flow, and myocardial contractility. Myocardial substrate utilization was not impaired (83), and heart metabolism at 3 ATA was the same as at 1 ATA (59). 100% O₂ at 2 ATA provided protection against ventricular fibrillation in dogs with induced occlusion of the main left coronary circumflex artery (76).

Hyperoxia decreased heart rate and stroke volume in eight men with irreversible chronic obstructive pulmonary disease (COPD) although no changes in pulmonary, bronchial arterial, or pulmonary wedge pressures were observed (1). In ten heart disease patients, 100% O₂ administration resulted in a significant fall in heart rate and myocardial blood flow with no changes in cardiac output, mean arterial blood pressure, or myocardial O₂ uptake (49). The maintenance of cardiac output despite a decreased heart rate indicates an increased stroke volume with an unaltered myocardial O₂ uptake.

HBO at 2 ATA produced decreases in cardiac output and stroke volume (but not heart rate) and increases in mean arterial blood pressure and systemic vascular resistance (15) in 10 male post-myocardial infarction patients. In ten healthy male subjects (87), cardiac output and heart rate decreased (but not stroke volume), and mean arterial blood pressure was unaffected after oxygen inhalation at 3.04 ATA. An informative report by Hitchcock, et al (42) suggested that despite reduced coronary flow, HBO therapy may prove beneficial to patients with cardiac disease due to increased oxygenation of cardiac tissue through "inter-coronary channelling".

UNTOWARD EFFECTS AND SAFETY

Bronchial irritability with decreased vital capacity and labored breathing (dyspnea) are symptoms of pulmonary O₂ toxicity (18, 56, 68). Pulmonary O₂ tolerance can be dramatically extended when the O₂-breathing periods are alternated with periodic air-breathing periods (40, 88). This intermittent O₂-air protocol is now employed by all USAF Hyperbaric Medicine facilities in all HBO treatment protocols.

Central nervous system O₂ toxicity is most dramatically manifested by convulsion not unlike that of epilepsy (8-10, 21, 56, 68). Convulsion represents a more serious and a more likely threat to submerged divers than to subjects in a "dry" hyperbaric chamber at the same barometric pressure (9, 10, 21, 56, 68). As cited by Dr A.R. Behnke (10), the U.S. Navy Diving Manual (1973) states "the employment of 100% oxygen for breathing during working dives is generally limited to shallow water operations because of the hazard of oxygen toxicity." Adding credibility to this concern is the recent report (13) of convulsion of a trained SCUBA diver during submaximal horizontal cycling after 72 minutes on an O₂-air breathing protocol (44 minutes on O₂) in 25 feet of sea water in a "wet-pot" facility. The subject had also experienced a convulsion earlier on a similar exercise protocol in 40 feet of sea water.

Tolerance to 100% oxygen in the "dry" hyperbaric chamber is greater than that in water. Several authors (6, 70, 79) have reported no untoward effects of subjects performing maximal exercise at 2 ATA (equivalent to 33 feet of sea water) while breathing oxygen. At 3 ATA (66 fsw), one of seven

subjects experienced convulsion while breathing oxygen during cycle exercise to exhaustion (47, 70). In all reviewed cases of reported convulsion (9, 10, 13, 21, 47, 56, 68, 87), the CNS O₂ toxicity symptoms disappeared upon removal of the O₂ source. No permanent damage from the convulsion was noted in any report.

The reported convulsions occurring during "dry" HBO exercise (47, 70) were among subjects continuously breathing oxygen at maximal exercise. In the literature reviewed, there was no report of intermittent O₂-air administration during maximal exercise or, perhaps more practical, during submaximal exercise conditioning. The use of O₂ during submaximal dry chamber exercise conditioning at atmospheric pressures less than or equal to 2 ATA (33 fsw) appears to be safe for healthy subjects, especially if an intermittent O₂-air protocol is employed. The usefulness of hyperoxia (HBO) was well-stated in the report of the 1976 Navy-Wide Workshop on High Pressure Biomedical Research (68):

"...Although the toxic potential of hyperoxia should be respected, the useful properties of this gas can be exploited safely and effectively" (p. 43).

With regard to exercise risk in general, there were no reports of heart attack or other untoward effects in any reviewed study. There were also no reports of fire, decompression sickness or air embolism in the review of related literature.

CONSENT FORM

TITLE: Hyperbaric Oxygenation (HBO) Alteration of Metabolism and Cardiovascular Function During and Following Exercise Conditioning.

1. I hereby volunteer to participate as a test subject in this experimental study. The purpose of the study is to determine the effects of HBO and submaximal ergometer exercise on cardiovascular function.
2. As a participant in this study, I understand that I will undergo the following procedures:
 - a. Cycle Ergometer Exercise As an HBO subject, while inside the compression chamber I will pedal a stationary cycle at a rate of 50/minute against submaximal work loads. I understand that, following a 5-minute warmup, I will pedal for six to eight minutes for each of three exercise bouts. A 2-minute zero-load pedalling recovery interval follows each exercise bout. I also understand that I am to maintain the proper pedal frequency as set by a metronome cadence. As an air-breathing subject I will perform the same exercise outside the chamber.
 - b. Breathing Gas As an exercise subject, I understand that I will breathe either air or 100% oxygen (O₂) during the exercise sessions. The O₂ will be delivered through a clear vinyl face mask that will be worn during the exercise bouts; the mask will be removed during the recovery intervals during which time normoxic air will be breathed.

c. Hyperbaric Exposure If selected as an HBO subject, I will be exposed to a hyperbaric pressure of 2 ATA, equivalent to a sea water depth of 33 feet. As a subject on the normobaric air protocol, I understand that I will not be exposed to hyperbaric conditions and that I will breathe normal air.

d. Physiologic Measurements I will be prepared and assessed for the following physiologic measurements:

(1) An electrocardiogram (ECG) will be administered for the purpose of recording rate and electrical activity of my heart.

(2) The functioning of my heart will be assessed from the administration of two tests. The impedance cardiogram (ZCG) and a conventional gas analysis (CO_2 re-breathing) method will be used to determine ejection volume and contractile force of my heart.

(3) My expired air will be collected and analyzed for determination of maximal aerobic power ($\dot{V}\text{O}_2\text{max}$) during a treadmill exercise test before and after the exercise conditioning period of the study. An ECG will be simultaneously recorded as I walk to exhaustion on the treadmill. I understand that this is a test of maximal performance. Similar tests will be performed to test aerobic power at submaximal work during the conditioning period.

(4) An anaeroid pressure cuff will be attached around my arm for the purpose of determining systolic and diastolic blood pressures during the exercise tests and conditioning sessions.

cavities, with possible bleeding, may result. I have been briefed in the use of various methods to help pressure equalization and to help prevent the occurrence of ear drum rupture or sinus bleeding. During the decompression (ascent) at the end of each chamber treatment, I have been instructed not to hold my breath. If I do hold my breath, I am aware that the air in my lungs will expand and may cause a rupture of my lungs (pneumothorax) or the formation of air bubbles in my blood stream (air embolism) that could interfere with circulation to my brain or other vital organs. I have also been briefed as to the possibility of convulsion resulting from oxygen toxicity on the hyperbaric treatment in which I will be breathing 100% oxygen. While extremely remote, there is also a risk of fire in the chamber. The chamber is equipped with a fire suppression (sprinkler) system as well as a hand-held hose for use in case of fire.

d. Physiologic Measurements:

- (1) ECG - the ECG is non-invasive and involves no unusual risks.
- (2) ZCG - the ZCG involves the application of a small magnitude (4mA, 100KZ) electrical current to my thorax. Electrical shock potential exists, however, I understand that electrical shock is not involved in the testing procedure, and technical precautions have been taken to prevent its occurrence.
- (3) $\dot{V}O_2$ max - during tests of maximal exercise the possibility exists that problems arising from elevated heart rate and blood pressure may develop. I understand that a physician and qualified technicians administer and closely monitor the test for $\dot{V}O_2$ max and that the test will be terminated if indications of

intolerance should arise. I also understand that this test is intended to determine my tolerance to the exercise that will be performed during the conditioning sessions.

(4) BP - the inflation cuff, air gauge, and stethoscope used for measurements of blood pressure involves no unusual risk.

(5) Blood analyses - I understand that the vena puncture technique of blood sampling carries with it some risk of infection and hemorrhage. I have been informed that sterile, disposable needles will be used by persons trained to draw blood samples and no adverse reactions are intended or expected.

4. I understand that I will receive no direct benefit from my participation in this study. Not participating in this study may be equally or more beneficial to me.

5. I understand I must inform the principal investigator of any changes to my medical status and include information regarding any medications I have taken and any medical or dental care/treatment received since my last use as a test subject.

6. The decision to participate in this program is completely voluntary on my part. No one has coerced or intimidated me into participating in this program. I am participating because of my interest and desire to be involved with this experimental study. Dr. _____ has adequately answered any and all questions I have about this study, my participation, and the procedures involved. I understand that Dr. _____ will be available to answer any questions I have about procedures throughout this study. I further understand that I may withdraw this consent at

any time and discontinue further participation in this study. I also understand that the investigator of this study may terminate my participation in this study at any time if he believes this to be in my best interest.

Volunteer's Signature Date

Printed Name, Grade, SSAN

Medical Consultant Date

Principal Investigator Date

Witness Date

Privacy Act of 1974 applies. DD Form 2005 Filed in Clinical/Medical Records.

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

NATURAL FREQUENCIES AND MODE SHAPES

OF UNIFORM BEAMS

Prepared by:	George R. Doyle, Jr.
Academic Rank:	Associate Professor
Department and University:	Mechanical Engineering, University of Dayton
Research Location:	Aero Propulsion Laboratory, Compressor Research Facility
USAF Research:	John H. Gutman
Date:	July 1, 1983
Contract No:	F49620-82-C-0035

NATURAL FREQUENCIES AND MODE SHAPES

OF UNIFORM BEAMS

by

George R. Doyle, Jr.

ABSTRACT

Most machinery contains rotating shafts that possess some unbalance. This unbalance results in a self-excited lateral vibration, which reaches a maximum value when the operating speed is near a lateral natural frequency. This report contains a summary of the derivation of natural frequency and mode shape equations for uniform beams. In addition to lateral motions, longitudinal and torsional vibrations were also included. Each of these three modes of vibration were established with various classical (fixed, free, and pinned) and nonclassical (inertial and spring load) boundary conditions. Second order effects of rotary inertia, shear deformation, and axial load were also considered. Finally, a rigid body pitch plane analysis of a shaft supported by N bearings was made.

ACKNOWLEDGEMENTS

The author would like to acknowledge the Air Force Systems Command, Air Force Office of Scientific Research, for their sponsorship of the Summer Faculty Research Program. He would also like to thank Mr John H. Gutman and Mr Eugene P. Hoffman of the APL at Wright-Patterson Air Force Base for their interest.

I. INTRODUCTION: The Aero Propulsion Laboratory's (APL) Compressor Research Facility (CRF) has been under investigation and development since the mid 1970's. Its task is to operate compressors up to speeds of 30,000 rpm for the purpose of qualifying new designs and also for research in various compressor operating conditions. Along with data acquisition, reduction and analyses, the drive train system (motor, shafts, gear boxes, and bearings) have been the major areas of development.

Since the drive train system will deliver 30,000 horsepower at 16,000 rpm, its components are quite massive, e.g., the bull gear and shaft in the low speed gearbox weigh over 24,000 pounds, and several shafts range from 10 to 15 inches in diameter. The larger component parts of the drive train naturally lead to greater unbalance, more difficult shaft alignment, and potentially more bearing problems. All of these difficulties can lead to large amplitude vibrations, and in fact, that has been the experience of the facility in recent months. To reduce the vibration, shafts have been balanced to a degree significantly better than the manufacturer's specifications, shaft alignment has been closely checked, and new bearings have been installed. All of these improvements should provide a substantial reduction in vibration levels when the system becomes operational in the fall of 1983.

The work accomplished during the 1983 USAF-SCEEE Summer Faculty Research Program was directly related to vibration problems associated with high speed shafts. If any rotating component possesses an unbalance (and they always do), the component will typically experience a self-excited lateral vibration synchronous with its rotating speed. Simple vibration theory predicts that the magnitude of the oscillation

is linearly proportional to the unbalance, and to a transfer function which depends on the damping ratio of the system, and more importantly, on the ratio of the driving frequency to the shaft's natural frequency. If the operating speed occurs near the shaft's lateral natural frequency, and damping is low (which would be typical of steel shafts), relatively large lateral oscillations will occur. It is, therefore, desirable to determine the natural frequencies of the system; and if they occur within the overall operating range, provide damping or do not operate in the restricted range near the natural frequency. A better approach, if possible, would be to redesign the system so that none of its natural frequencies occur within the operating range.

The determination of the natural frequencies of a system should be properly done by considering the system as a whole. However, to do so is, at best, difficult and costly. Therefore, a series of transcendental equations were derived to determine the natural frequencies of continuous systems. With each frequency equation, a mode shape expression was also derived. For completeness, the natural frequencies of the pitch plane motions of a rigid shaft supported by N bearings (spring and damper in parallel) were also derived. From these expressions, longitudinal, torsional, and lateral natural frequencies of uniform beams with various combinations of classical and nonclassical boundary conditions can be obtained. These values can be used for a first order approximation to some of the natural frequencies associated with the total system, or they can be used to partially verify a more complex model. The rigid body modes are also important and would be present in the results of any large scale computer program.

II. OBJECTIVES OF THE RESEARCH EFFORT: The initial goals of this research effort were to understand the causes of large vibratory motions of the drive train of the CRF. One of the primary causes was thought to be operating speeds in the range of shaft natural frequencies. This observation resulted in narrowing the goal to a comprehensive review and further development of natural frequencies and mode shapes for uniform beams in longitudinal, torsional, and lateral motions. Second order effects such as rotary inertia, shear deformation, and axial load were also investigated. Rigid body pitch plane motions of a beam on N suspension systems were included for completeness.

III. MATHEMATICAL APPROACH: The development of the partial differential equations of motion and their general solution is given in many references (1)-(7). The derivation involves consideration of a small element of the beam acted upon by internal forces and moments. Both classical (fixed, free, pinned) and nonclassical (inertial and spring load) boundary conditions are applied to each end.

Equations of Motion

o Longitudinal Vibrations

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} \quad (1)$$

where u = Longitudinal displacement

x = Longitudinal position

t = Time

c = Normal stress wave propagation speed = $\sqrt{E/\rho}$

E = Young's modulus

ρ = Mass density

o Torsional Vibrations

$$\frac{\partial^2 \theta}{\partial t^2} = \bar{c}^2 \frac{\partial^2 \theta}{\partial x^2} \quad (2)$$

where θ = Angular displacement

\bar{c} = Shear stress wave propagation speed = $\sqrt{G/\rho}$

G = Shear modulus

o Lateral Vibrations

$$\frac{\partial^4 y}{\partial x^4} + p^4 \frac{\partial^2 y}{\partial t^2} = 0 \quad (3)$$

where y = Lateral displacement

$p^4 = \bar{m}/EI$

\bar{m} = Mass per unit length

I = Cross-sectional area moment of inertia

o Lateral Vibrations with Rotary Inertia

$$\frac{\partial^4 y}{\partial x^4} + p^4 \frac{\partial^2 y}{\partial t^2} - \frac{J_0}{EI} \frac{\partial^4 y}{\partial x^2 \partial t^2} = 0 \quad (4)$$

where J_0 = Rotary inertia per unit length

o Lateral Vibrations with Shear Deformation

$$\frac{\partial^4 y}{\partial x^4} + p^4 \frac{\partial^2 y}{\partial t^2} - \frac{\bar{m}}{k'AG} \frac{\partial^4 y}{\partial x^2 \partial t^2} = 0 \quad (5)$$

where k' = Shape factor

A = Cross-Sectional area

- o Lateral Vibrations with Rotary Inertia and Shear Deformation

$$\frac{\partial^4 y}{\partial x^4} + P^4 \frac{\partial^2 y}{\partial t^2} - \left(\frac{J_0}{EI} + \frac{\bar{m}}{k'AG} \right) \frac{\partial^4 y}{\partial x^2 \partial t^2} + \frac{J_0 P^4}{k'AG} \frac{\partial^4 y}{\partial t^4} = 0 \quad (6)$$

- o Lateral Vibrations with Axial Load

$$\frac{\partial^4 y}{\partial x^4} + P^4 \frac{\partial^2 y}{\partial t^2} - \frac{T}{EI} \frac{\partial^2 y}{\partial x^2} = 0 \quad (7)$$

where T = Axial load in beam (positive if tension)

Boundary Conditions

- o Longitudinal Vibrations

<u>Boundary Conditions</u>	<u>Left End</u>	<u>Right End</u>
Fixed	$u = 0$	$u = 0$
Free	$\frac{\partial u}{\partial x} = 0$	$\frac{\partial u}{\partial x} = 0$
Mass Load	$M_1 \frac{\partial^2 u}{\partial t^2} = AE \frac{\partial u}{\partial x}$	$M_2 \frac{\partial^2 u}{\partial t^2} = -AE \frac{\partial u}{\partial x}$
Spring Load	$K_1 u = AE \frac{\partial u}{\partial x}$	$K_2 u = -AE \frac{\partial u}{\partial x}$

- o Torsional Vibrations

<u>Boundary Condition</u>	<u>Left End</u>	<u>Right End</u>
Fixed	$\theta = 0$	$\theta = 0$
Free	$\frac{\partial \theta}{\partial x} = 0$	$\frac{\partial \theta}{\partial x} = 0$
Inertial Load	$J_1 \frac{\partial^2 \theta}{\partial t^2} = GI_p \frac{\partial \theta}{\partial x}$	$J_2 \frac{\partial^2 \theta}{\partial t^2} = -GI_p \frac{\partial \theta}{\partial x}$
Spring Load	$K_{\theta 1} \theta = GI_p \frac{\partial \theta}{\partial x}$	$K_{\theta 2} = -GI_p \frac{\partial \theta}{\partial x}$

o Lateral Vibrations

<u>Boundary Condition</u>	<u>Left End</u>	<u>Right End</u>
Fixed	$X = X' = 0$	$X = X' = 0$
Free	$X'' = X''' = 0$	$X'' = X''' = 0$
Pinned	$X = X'' = 0$	$X = X'' = 0$
Translational Mass	$X'' = 0$ $EIX''' = \omega^2 M_1 X$	$X'' = 0$ $EIX''' = -\omega^2 M_2 X$
Rotary Mass	$X''' = 0$ $EIX'' = -\omega^2 J_1 X'$	$X''' = 0$ $EIX'' = \omega^2 J_2 X'$
Translational Spring	$X'' = 0$ $EIX''' = -K_1 X$	$X'' = 0$ $EIX''' = K_2 X$
Rotational Spring	$X'' = 0$ $EIX'' = K_{\theta 1} X'$	$X'' = 0$ $EIX'' = -K_{\theta 2} X'$

o Lateral Vibrations with Rotary Inertia and/or Shear Deformation or Axial Load

Boundary conditions were the classical boundary conditions used for the Lateral Vibration Problems.

The partial differential equations were solved using a separation of variables technique in which the total solution was assumed to be a product of a function of time and a function of position. The time function resulted in harmonic functions in all cases. The position solution, which leads to the mode shape, was given by a harmonic function for the longitudinal and torsional vibration problems. For the

lateral vibration problems, the position solution included both harmonic functions and hyperbolic functions. A complete derivation is given in reference (7).

IV. RESULTS: Typically, an application of the boundary conditions resulted in a transcendental equation, from which "n" eigenvalues could be determined. The "n" eigenvalues were then related to the "n" natural frequencies through the time domain function. Also, knowing the eigenvalues completely determined the mode shapes through the position function. With both the natural frequencies and mode shapes established, all potential free vibration problems can be solved with the addition of initial conditions. For the present application, however, the primary goal was to determine natural frequencies, for it is at the natural frequency that a component will respond with greatest amplitude, especially in steel shafts, where damping is very low.

The actual transcendental frequency equations and mode shape equations were, in many cases, very lengthy, and will not be given in this report, except for a few examples. The total set of equations can be found in reference (7).

o Longitudinal Vibrations

<u>Boundary Conditions</u>	<u>Natural Frequencies</u> rad/s (n=1,2,3...)	<u>Mode Shape</u> (n=1,2,3...)
Fixed-Mass Load	$\tan(\omega_n L/c)$ $= AE/M_2 \omega_n c$	$\sin(\omega_n x/c)$

where L = Length of beam

ω_n = Natural frequency

o Torsional Vibrations (form is identical to longitudinal vibrations)

<u>Boundary Conditions</u>	<u>Natural Frequencies</u> (rad/s (n=1,2,3...))	<u>Mode Shape</u> (n=1,2,3...)
Inertial Load	$\tan(\omega_n L/c)$	$\frac{J_1 \omega_n c}{G I_p} \sin(\omega_n x/c)$
- Spring Load	$= \frac{G I_p c (K_{\theta 2} - J_1 \omega_n^2)}{(G^2 I_p^2 + J_1 K_{\theta 2} c^2) \omega_n}$	$-\cos(\omega_n x/c)$

o Lateral Vibrations

<u>Boundary Conditions</u>	<u>Natural Frequencies*</u> (n=1,2,3...)	<u>Mode Shape</u> (n=1,2,3...)
Pinned - (Fixed/Rotational spring)	$\frac{2EI(\beta_n L)}{L}$ $= K_{\theta 2} \left[\frac{1}{\tan(\beta_n L)} - \frac{1}{\tanh(\beta_n L)} \right]$	$\sin(\beta_n x) + \gamma_n [\sinh(\beta_n x)]$ where: $\gamma_n = - \frac{\sin(\beta_n L)}{[\sinh(\beta_n L)]}$

$$* \omega_n = \frac{(\beta_n L)^2}{L^2} \left[\frac{EI}{m} \right]^{1/2}$$

o Lateral Vibrations with Rotary Inertia

<u>Boundary Conditions</u>	<u>Frequency Equation*</u> (n=1,2,3...)	<u>Mode Shapes</u> (n=1,2,3...)
Pinned-Pinned	$\nu_n L = n\pi$	$\sin(\nu_n x)$

$$* \omega_r = \omega_n \left[1 + \frac{I}{A} \nu_n^2 \right]^{-1/2}$$

where ω_n = Natural frequency without rotary inertia effect

ω_r = Natural frequency with rotary inertia effect

o Lateral vibrations with shear deformation (form is identical to lateral vibrations with rotary inertia)

<u>Boundary Conditions</u>	<u>Frequency Equation*</u> ($n=1,2,3\dots$)	<u>Mode Shape</u> ($n=1,2,3\dots$)
Pinned-Free	$\frac{(\lambda_n L)}{L} \tan(\nu_n L)$ $= \frac{(\nu_n L)}{L} \tanh(\lambda_n L)$	$\sin(\nu_n X) + \delta_n \sinh(\lambda_n X)$ where: $\delta_n = \frac{L \nu_n^2 \sin(\nu_n L)}{[\lambda_n^2 \sinh(\lambda_n L)]}$

* ν_n and λ_n are the roots of the characteristic equation of the position function. Their expressions contain the natural frequency and other system properties.

o Lateral vibrations with rotary inertia and shear deformation (form is identical to lateral vibrations with rotary inertia).

o Lateral vibrations with axial load (form is identical to lateral vibrations with rotary inertia).

The frequency and mode shape equations for the lateral vibration with rotary inertia, shear deformation, and axial load appear to have two sets of eigenvalues, ν_n and λ_n . Actually, each variable is a rather complicated algebraic expression containing the natural frequency and the system parameters. Except in the pinned-pinned case, a computer program would be necessary to solve for the natural frequencies.

V. SUMMARY: The results of this research program have been to establish a series of tables that itemize frequency and mode shape equations for a large variety of boundary conditions associated with uniform

beams. Some of these results can be found in other references (1)-(6). As a result of this particular study, a report (7) has been written which details the derivation of the equations and gives the above-mentioned tables. These tables can be used to find natural frequencies of isolated components and to compare simple exact solutions to more complicated computer modeling.

VI. RECOMMENDATIONS: Most of the frequency and mode shape equations that were derived during this program can be solved by trial and error manually, or on one of the more sophisticated modern calculators. But those equations involving lateral deflection with rotary inertia or shear deformation or axial load would require a small computer program to be written. For those more complex expressions and even for the simpler ones, a computer program that accepted system characteristics and boundary conditions as input, and gave natural frequencies and corresponding mode shapes as output, would be useful for individuals dealing with that type of problem.

REFERENCES

1. Thomson, W. T., Theory of Vibrations with Applications, Second Edition, Prentice-Hall, Inc., 1981.
2. Tse, F. S., Morse, I. E., Hinkle, R. T., Mechanical Vibrations Theory and Applications, Second Edition, Allyn and Bacon, Inc., 1978.
3. Timoshenko. S., Young, D. H., Weaver, W., Jr., Vibration Problems in Engineering, Fourth Edition, John Wiley and Sons, 1974.
4. Meirovitch, L., Analytical Methods in Vibrations, MacMillan Co., 1967.
5. Harris, C. M., Crede, C. E., Shock and Vibration Handbook, Second Edition, McGraw-Hill Co., 1976.
6. Gorman, D. J., Free Vibration Analysis of Bearings and Shafts, John Wiley and Sons, Inc., 1975.
7. Doyle, G. R., Jr., "Natural Frequencies and Mode Shapes of Uniform Beams," University of Dayton, July 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

TIME-DEPENDENT CALCULATIONS OF SWIRLING

NOZZLE FLOW

Prepared by: Dr. J. Craig Dutton

Academic Rank: Assistant Professor

Department and University: Mechanical Engineering Department
Texas A&M University

Research Location: Aero Propulsion Laboratory
Ramjet Engine Division
Ramjet Technology Branch

USAF Research Colleague: Dr. Roger R. Craig

Date: August 12, 1983

Contract No: F49620-82-C-0035

TIME-DEPENDENT CALCULATIONS OF SWIRLING

NOZZLE FLOW

by

J. Craig Dutton

ABSTRACT

The effect of swirl on the mass flow, thrust, and flowfield characteristics of supersonic, converging-diverging nozzles has been investigated. A widely used and well tested time-dependent code (VNAP) for predicting two-dimensional, compressible, internal flows has been modified to include the swirl velocity component. The required modifications involve the integration of an additional momentum equation, the inclusion of additional terms in the original equation set, and re-formulation of the boundary conditions to reflect the addition of the tangential velocity component. The resulting modified code (SNAP) has been used to predict the flowfield and performance characteristics of a converging-diverging nozzle at various levels of swirl. The results indicate that for swirl numbers of approximately 0.4 or less, reductions in the nozzle mass flowrate and thrust on the order of 10 percent or less can be expected. The most pronounced effect of swirl on the nozzle flowfield is to cause a large increase in the axial velocity near the centerline as compared to the no-swirl case.

ACKNOWLEDGMENTS

I would like 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 providing me with the opportunity and financial assistance to participate in the Summer Faculty Research Program at the Aero Propulsion Laboratory, Wright-Patterson AFB, Ohio. I am particularly indebted to Dr. Frank D. Stull, Chief, and all of the members of the Ramjet Technology Branch for their help and hospitality during the course of my summer research effort. I would like to extend my special thanks to Dr. Roger R. Craig for suggesting this area of research and for providing valuable technical assistance and direction. Finally, I would like to thank Mr. Ken G. Schwartzkopf for his help in generating the computer plots used in this study.

I. INTRODUCTION

During the last several years the use of swirl to improve the performance characteristics of integral rocket/ramjet combustors has been investigated [1]. These studies have demonstrated that high combustion efficiencies, accompanied by significantly lower pressure losses, are possible for swirled combustors relative to typical flameholder configurations. In addition, shorter combustor lengths are possible with swirl. However, the combustor is not the only component of the ramjet propulsion system. Immediately downstream from the combustor is the nozzle whose purpose is to efficiently convert the random thermal energy of the combustion gases into the directed kinetic energy of the high speed exhaust jet. An obvious question which arises concerns the effect of the swirl generated in the combustor inlet on the performance of the nozzle. It is well known that the qualitative effect of swirl is to reduce the mass flow and thrust of nozzles at given stagnation conditions. However, the magnitude of these reductions and the effects of swirl on the nozzle flowfield have not been thoroughly investigated, particularly for configurations typical in ramjet applications. These configurations consist of relatively short, highly curved nozzles with the swirl generated by fixed vane inserts. As a result, the general purpose of this research investigation has been to study the effect of swirl on ramjet propulsion nozzles.

II. OBJECTIVES

In light of the preceding discussion, the specific objectives of this research were:

- (1) To conduct a literature review of previous work concerned with swirling, supersonic, converging-diverging, nozzle flows;
- (2) To become familiar with the "VNAP" computer code written by M. C. Cline at Los Alamos National Laboratory for predicting two-dimensional, time-dependent, compressible, internal flows;
- (3) To modify the two-dimensional VNAP code as necessary to account for swirl;
- (4) To use the modified code to predict the discharge coefficient, thrust efficiency, and flowfield characteristics of convergent-divergent

ramjet nozzles with swirl.

As will be discussed in the following sections, each of these topics has been addressed during the course of this investigation.

III. PREVIOUS WORK

The topic of swirling nozzle flow has been fairly extensively studied over the last 35 years with most of the work being done in the 1960's and early 70's. The impetus for the work in this era was provided by spin-stabilized rockets in an effort to determine the effects of spin on nozzle performance. The majority of the publications on this subject [2-21] have been quasi-one-dimensional theoretical investigations in which the effects of swirl have been included in an otherwise conventional one-dimensional nozzle analysis. Both analytical and numerical approaches have been developed within the quasi-one-dimensional approach. However, these authors generally employ different choking criteria and make different assumptions about the radial distribution of the tangential velocity component. Hence, there is little agreement among these investigators on the effects of swirl on nozzle performance with the result that no clearly superior quasi-one-dimensional theory exists for swirling nozzle flow. In addition, several of the investigators overspecified the problem in the course of their assumptions, and all of these studies confine themselves to certain classes of swirl velocity profiles, thereby limiting their generality. However, the most serious shortcoming of these theories is that they all ignore the radial velocity component and its derivatives. For the short, highly curved nozzles typical in ramjet applications, such theories are therefore inadequate. Indeed, the computations performed during the present investigation have demonstrated that over a large percentage of the nozzle flowfield, the radial velocity component is of the same order or larger than the tangential component.

In a series of reports Guderley et al. [22-25] have performed two-dimensional calculations for inviscid, swirling, converging-diverging nozzle flow. An approximate numerical solution for transonic throat flow with swirl was obtained to provide a starting line for the method of characteristics analysis of the supersonic diverging portion of the nozzle. Rao's method was employed to determine nozzle contours for maximum thrust in the

presence of swirl. This is the only known attempt at a two-dimensional analysis of swirling nozzle flow, and it is to be noted that only the supersonic portion of the flowfield is treated exactly. No consideration is given to the subsonic portion of the flow, and the transonic region is treated approximately.

Experimental measurements of swirling nozzle flow are extremely sparse. Both Norton, et al. [11] and Dunlap [26] used a spinning test apparatus to simulate spin-stabilized rockets. Norton measured the mass flow reduction due to spin while Dunlap's investigation concentrated on the flowfield in the simulated combustion chamber just upstream from the nozzle. In related publications, Batson and Sforzini [27] and Sforzini and Essing [28] experimentally investigated swirling nozzle flow by injecting cold gas tangentially to a cylindrical chamber wall and then letting it expand through single or multiple c-d nozzles. For the case of a single nozzle, limited mass flowrate, thrust, and velocity profile measurements were obtained. To date there are no known measurements available for swirled nozzle flow generated by fixed vane inserts, as occur in the ramjet application.

IV. THEORETICAL FORMULATION

As a result of the review of previous work, it is concluded that a computational method for analyzing two-dimensional swirling flow through nozzles is needed which correctly considers the entire subsonic-transonic-supersonic regions of the nozzle flowfield. For non-swirling flow, the most successful method for doing this has been the recently developed time-dependent technique, e.g. [29]. In this method the unsteady form of the governing equations is used to avoid the mixed, elliptic/hyperbolic nature of the steady flow equations. An initial value surface for the nozzle flowfield is guessed, the steady flow boundary conditions are applied, and the solution is advanced in time until it converges at the steady state. A disadvantage of this method is the long computing times which are sometimes required. If only the steady state solution is desired, several hundred time planes must be calculated before this limit is reached. On the other hand, its major advantages are its ability to analyze mixed flowfields and its obvious ability to handle truly transient flows.

Based on these considerations, the approach adopted in this research has been to modify a two-dimensional, time-dependent non-swirling nozzle flow code to account for the effects of swirl. The program chosen for use is a version of the widely-used and well-tested "VNAP" code [30-33] written by M. C. Cline at Los Alamos National Laboratory. This program has been demonstrated to accurately and efficiently calculate non-swirling nozzle flowfields.

In this analysis it is assumed that the nozzle contains the inviscid, non-heat conducting, axisymmetric flow of a thermally and calorically perfect gas. In this context, "axisymmetric" denotes that all derivatives in the tangential direction vanish while the tangential velocity component does not. Under these assumptions and using Cartesian coordinate notation where x and y are axial and radial coordinates and u, v , and w are the axial, radial, and tangential velocity components, respectively, the governing equations can be written in nonconservation form as:

$$\text{Continuity: } \rho_t + u\rho_x + v\rho_y + \rho u_x + \rho v_y + \frac{\rho v}{y} = 0 \quad (1)$$

Euler Equations:

$$\text{Axial: } u_t + uu_x + vu_y + \frac{P_x}{\rho} = 0 \quad (2)$$

$$\text{Radial: } v_t + uv_x + vv_y - \frac{w^2}{y} + \frac{P_y}{\rho} = 0 \quad (3)$$

$$\text{Tangential: } w_t + uw_x + vw_y + \frac{vw}{y} = 0 \quad (4)$$

$$\text{Energy: } (P_t + uP_x + vP_y) - a^2(\rho_t + u\rho_x + v\rho_y) = 0 \quad (5)$$

where P, ρ, a, t are pressure, density, speed of sound, and time, respectively, and the subscripts denote partial differentiation.

The inclusion of the swirl velocity component, w , requires the integration of an additional momentum equation, (4), as well as the appearance of the centripetal acceleration term, w^2/y , in the radial momentum equation (3). The result is a set of five coupled, nonlinear partial differential equations in the five dependent variables: u, v, w, P , and ρ . Knowing these variables and using the ideal gas equation of state, any other quantity of interest can be determined in the flowfield. It is to be noted that the tangential momentum equation can be integrated to:

$$\frac{D}{Dt}(yw) = 0 \quad (6)$$

i.e., the angular momentum of the flow is constant along pathlines. This

integrated form could be used to avoid the finite difference solution of equation (4). However, to obtain a solution which is consistent in time, pathlines would have to be projected from each gridpoint on the current time plane to the previous time plane and a bivariate spatial interpolation performed to determine w . Since there is no apparent numerical advantage to doing this, and since equation (6) will not apply, for example, if the code is extended to viscous flow, the tangential momentum equation has been solved using the same finite difference technique as for the other four equations.

Before solution, the governing equations and boundary conditions are further transformed from (x,y,t) coordinates to (ζ,η,τ) coordinates where the η coordinate is a dimensionless one which varies from zero at the nozzle centerline to unity at the wall. In this way the calculations are performed on a rectangular grid. For brevity the transformed equations are not included here.

V. NUMERICAL TECHNIQUE

The interior points in the flowfield are calculated using MacCormack's [34] second order accurate, explicit, predictor-corrector method. First order backward spatial differences are used for the predictor while first order forward spatial differences are employed on the corrector step. At the present time, this method is undoubtedly the most widely used technique for the time-dependent solution of the equations of compressible flow.

Even though the boundary points are far fewer in number than the interior points, it is well known that an accurate evaluation of them is a necessity in obtaining valid time-dependent solutions. In fact, several early attempts at using the time-dependent technique were plagued by convergence, stability, and/or accuracy problems due to inaccurate treatment of the boundary conditions. In the present work a second order reference-plane characteristics scheme is used at all boundary points. In this method, spatial derivatives along the boundaries are approximated by finite differences and are treated as source terms. The resulting unsteady, one-dimensional problem normal to the boundary is then solved by the method of characteristics.

At the nozzle wall, the characteristics analysis shows that three compatibility equations must be satisfied along the wall pathline and one compatibility condition must be satisfied along the right-running Mach line. In addition, for inviscid flow, the wall tangency condition must be applied. This, then, gives five simultaneous equations for the five unknowns u, v, w, P , and ρ which are solved at each wall point. The inclusion of swirl requires that an additional compatibility condition be integrated along the wall pathline.

For subsonic flow at the inlet, only one compatibility equation need be satisfied along the left-running Mach line. Therefore, four other flow properties must be specified at each inlet point. There has been some debate in the literature [35,36] on the "best" choice for these specified inlet conditions. However, it has been generally agreed that a good choice for non-swirling flow is the stagnation pressure, P_t , the stagnation temperature, T_t , and the meridian plane streamline angle, $\theta = \tan^{-1}(v/u)$. For swirling flow another property must be specified, and in this work the tangential velocity component, w , has been used. It is to be emphasized that these properties need not be constant across the inlet, but rather can vary arbitrarily. This is in direct contrast to the quasi-one-dimensional theories which generally make assumptions about the distributions of one or more of these properties. An iterative technique is required to determine the five dependent variables, u, v, w, P , and ρ , from the compatibility condition and the four specified properties.

For supersonic flow at the exit, five compatibility conditions must be simultaneously satisfied: three along the pathline and one each along the left and right-running Mach lines. As expected, therefore, supersonic flow at the exit is completely determined by upstream conditions. The solution of the five compatibility equations yields the five dependent variables of interest. The addition of swirl means that another compatibility equation must be integrated along the pathline.

Once the theoretical formulation for the time-dependent calculation of swirling nozzle flows was obtained, the required changes in the VNAP code were made with the resulting program being named "SNAP". Initial

results calculated using SNAP are described in the following section.

VI. RESULTS AND DISCUSSION

The nozzle geometry used in all of the calculations described below is shown in Fig. 1. It consists of a cylindrical inlet, a 35° conical convergent section, a 0.475 in. radius circular arc throat section, and an 18.5° conical divergent section. The inlet, throat, and exit radii are 1.401, 1.145, and 1.833 inches, respectively. This nozzle is similar to the baseline conical nozzle used by McDonnell-Douglas in their "Ramjet Nozzle Technology Program" supported by AFWAL/PORT.

In addition to a nozzle geometry, inlet conditions must also be specified before any calculations can be performed. As discussed in the previous section, the inlet properties to be specified were chosen as P_t , T_t , θ , and w . It was assumed that P_t , T_t , and $\theta = 0$ are constant across the inlet and that the w distribution is as given in Fig. 2. The rationale for this choice is that previous work [1] has shown that the "constant angle" swirler gives the best overall combustor performance, where the swirl angle is defined as $\phi = \tan^{-1}(w/u)$. Assuming that the swirl profile generated in the combustor inlet persists to the nozzle inlet and that the axial velocity distribution at the nozzle inlet is uniform, results in the w profile shown in Fig. 2. Note that w must vanish on the nozzle axis by symmetry and that a radius ratio of 0.2 was chosen as the location where w reaches its constant, maximum value.

Results have been obtained for the discharge coefficient, C_D , the vacuum stream thrust efficiency, η_N , and the nozzle flowfield as a function of the inlet swirl number, S_i , where:

$$C_D = \frac{\dot{m}}{\dot{m}_{\text{ideal}}} = \frac{2 \int_0^{R_e} T_{\text{out}} \rho u r dr}{(\rho^* u^* R_T^2)_{\text{ideal}}} \quad (7)$$

$$\eta_N = \frac{T_{\text{vs}}}{T_{\text{vs, ideal}}} = \frac{2 \int_0^{R_e} (P + \rho u^2) r dr}{[(P_e + \rho_e u_e^2) R_e^2]_{\text{ideal}}} \quad (8)$$

$$S_i = \frac{\int_0^{R_i} \rho u w r^2 dr}{R_i \int_0^{R_i} \rho u^2 r dr} \quad (9)$$

In this case the ideal conditions are defined as uniform, one-dimensional

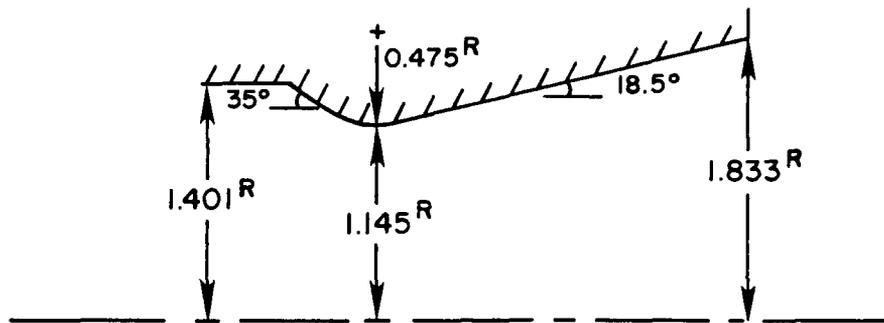
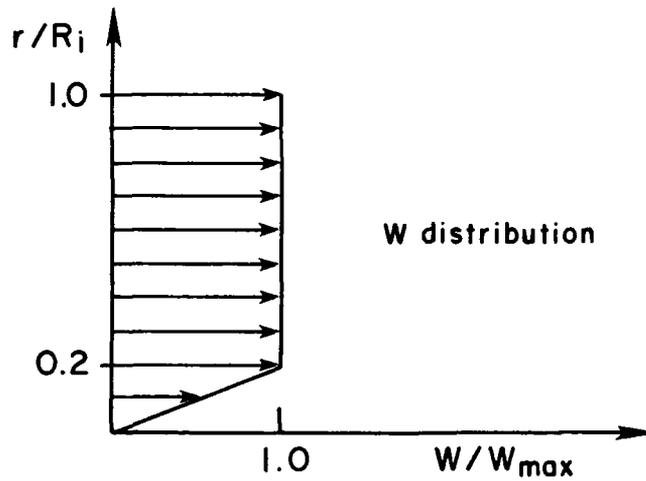


Fig. 1 Conical nozzle geometry (all dimensions in inches)



$P_t, T_t, \theta = 0$ assumed uniform across the inlet

Fig. 2 Assumed inlet conditions

values at the same stagnation conditions as the actual flow. Note that the inlet swirl number can be interpreted as the axial flux of angular momentum divided by the inlet radius times the axial flux of axial momentum and is therefore a measure of the level of swirl at the nozzle inlet. It should also be pointed out that this swirl number is not constant through the nozzle.

The calculations have been carried out for six swirl numbers on an 81×21 grid, and $\gamma = 1.4$ has been used in all cases. For computational efficiency each calculation was done in two parts: a 41×21 subsonic-transonic computation followed by a 41×21 supersonic calculation. Obviously a 41×21 grid will require less central memory than an 81×21 grid but, in addition, the supersonic calculations converged at least twice as fast as the subsonic-transonic computations. Each subsonic-transonic case had converged fairly well after 800 time planes and required approximately six minutes of CPU time on a CDC Cyber 74 while each supersonic case was extremely well converged after only 400 time planes and required about three minutes of computer time.

The results for the discharge coefficient and vacuum stream thrust efficiency as a function of the inlet swirl number are shown in Fig. 3. The effect of swirl at $S_i = 0.44$, the highest value investigated, is seen to result in approximately a 10.5% decrease in both the mass flowrate and vacuum stream thrust over the no-swirl case. However, the thrust per unit mass flow, i.e. specific impulse, is within 0.2% of being constant over the entire range investigated. This result suggests that for a constant mass flow device like a ramjet, the introduction of swirl entails essentially no loss in thrust but rather simply results in an increased combustor pressure. This increased pressure can also be avoided by appropriately enlarging the nozzle throat. Note that the curves in Fig. 3 become much steeper as S_i is increased, so that proportionately larger decrements in mass flow and thrust can be expected for swirl levels above that investigated here.

Mach number contour plots for the no-swirl case, $S_i = 0.0$, and for $S_i = 0.3$ are shown in Figs. 4 and 5. The flowfield shown in Fig. 4 is typical of that for an unswirled conical nozzle. Due to the relatively

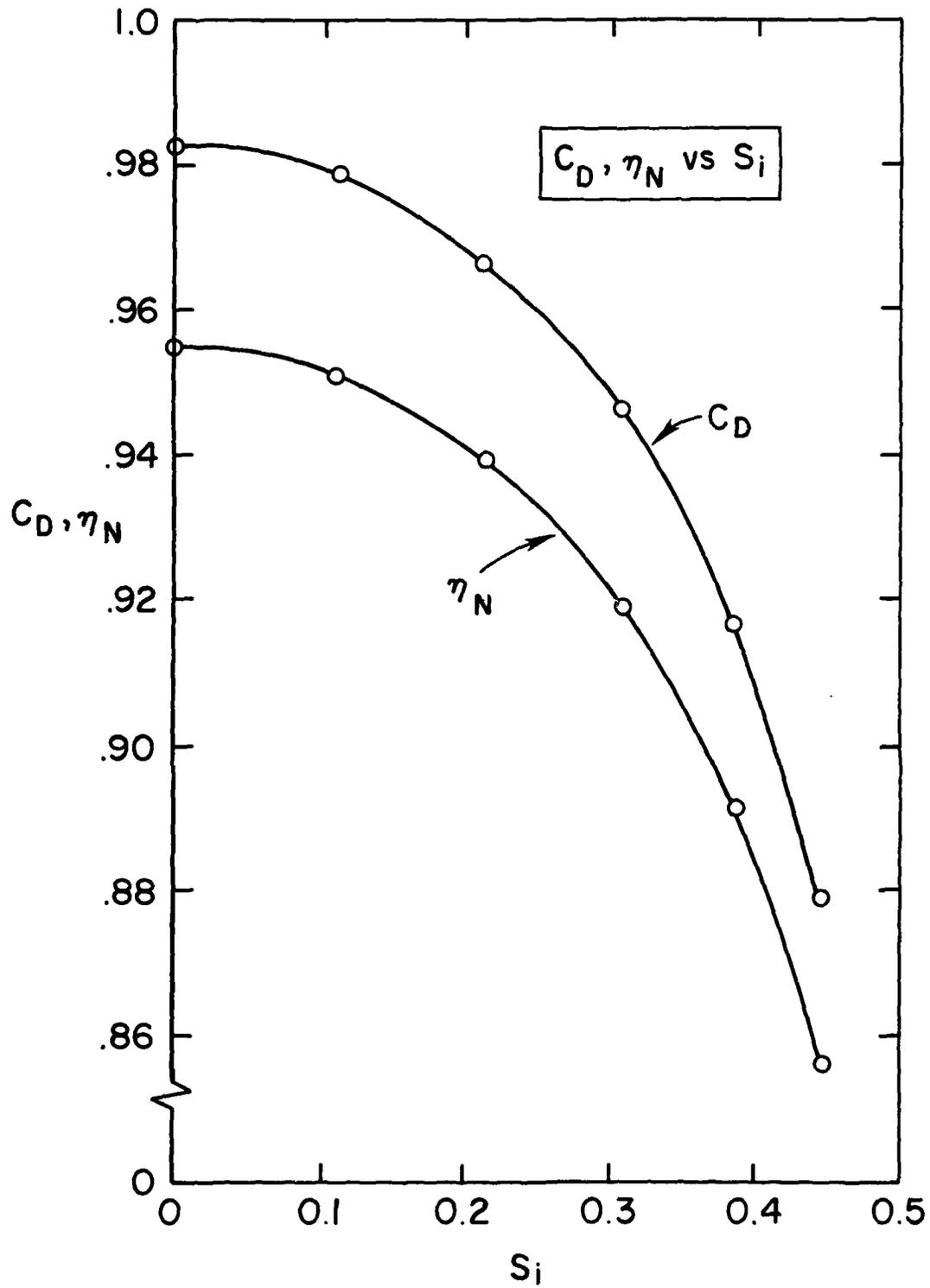


Fig. 3 Discharge coefficient and vacuum stream thrust nozzle efficiency as a function of inlet swirl number

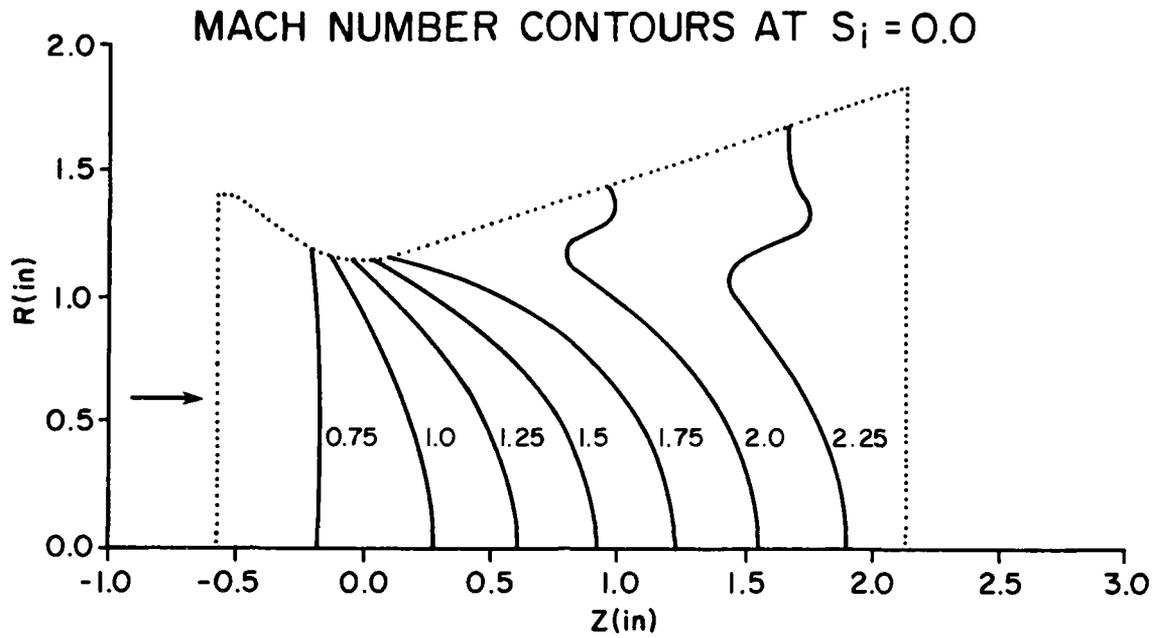


Fig. 4 Conical nozzle Mach number contours at $S_i = 0.0$

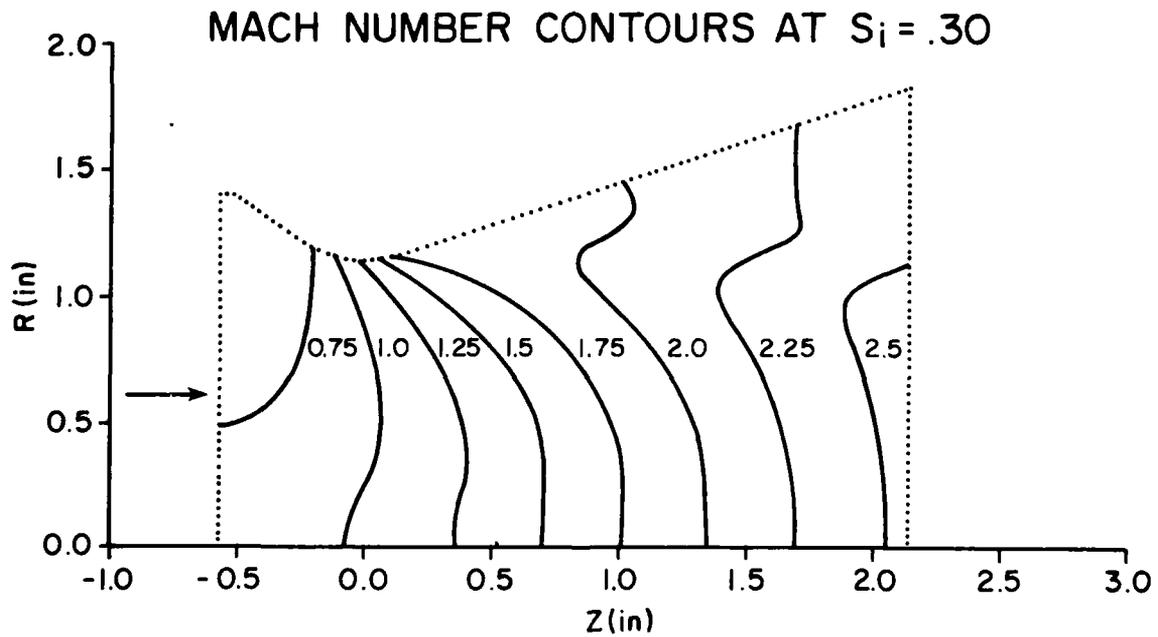


Fig. 5 Conical nozzle Mach number contours at $S_i = 0.3$

small wall radius of curvature, a strong expansion of the flow occurs in the throat wall region which greatly leads the expansion occurring along the nozzle centerline. The "wiggles" in the $M = 2.0$ and 2.25 contours are due to an oblique shock wave which originates at the tangency point between the circular arc and conical divergent wall sections. A discontinuity in the curvature of the wall contour occurs at this point, leading to a weak compression of the flow in this region. Comparing Figs. 4 and 5, the most notable effect of swirl is an upstream shifting of the Mach number contours near the nozzle centerline. In other words, the centerline axial velocity is much higher for a swirled nozzle flow than for the corresponding non-swirled flow under the assumed inlet conditions. A careful comparison of Figs. 4 and 5 also reveals that the axial velocity along the wall for $S_i = 0.3$ is slightly less than for $S_i = 0.0$. At the highest swirl number investigated, $S_i = 0.44$, a fairly large region of supersonic inflow is predicted to occur near the centerline at the inlet, as well as reverse flow along the wall near the inlet. Since the actual physical occurrence of these phenomena seems doubtful, it is hypothesized that the assumed inlet conditions are unrealistic at this level of swirl.

The evolution of the tangential velocity profile through the nozzle for $S_i = 0.3$ is shown in Fig. 6. At both the throat and exit planes, the w profile is similar in shape to the assumed inlet profile with a rounding of the sharp corner, followed by a region of nearly constant w , and a slight rise to the wall value. Based on the conservation of angular momentum, it is expected that the general level of the swirl velocity should be inversely proportional to the radius at a section. Referring to the dimensions given in Fig. 1, it is seen that the profiles presented in Fig. 6 confirm this expectation.

It is to be remembered that the inlet swirl velocity profile shown in Fig. 2 was chosen based on the desire to analyze the "constant angle" swirl distribution and the assumption that the axial velocity distribution at the nozzle inlet is relatively uniform. However, the results obtained in this study show that the inlet u distribution is very nonuniform with high values along the axis and very low values at the wall. For $S_i = 0.3$

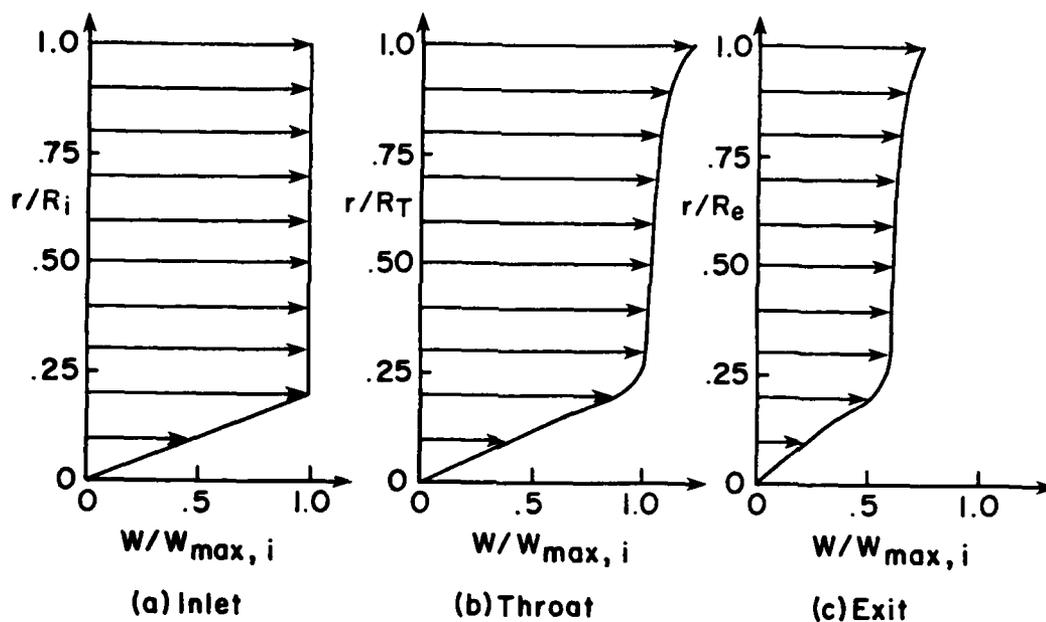


Fig. 6 Tangential velocity distributions at inlet, throat, and exit for $S_i = 0.3$

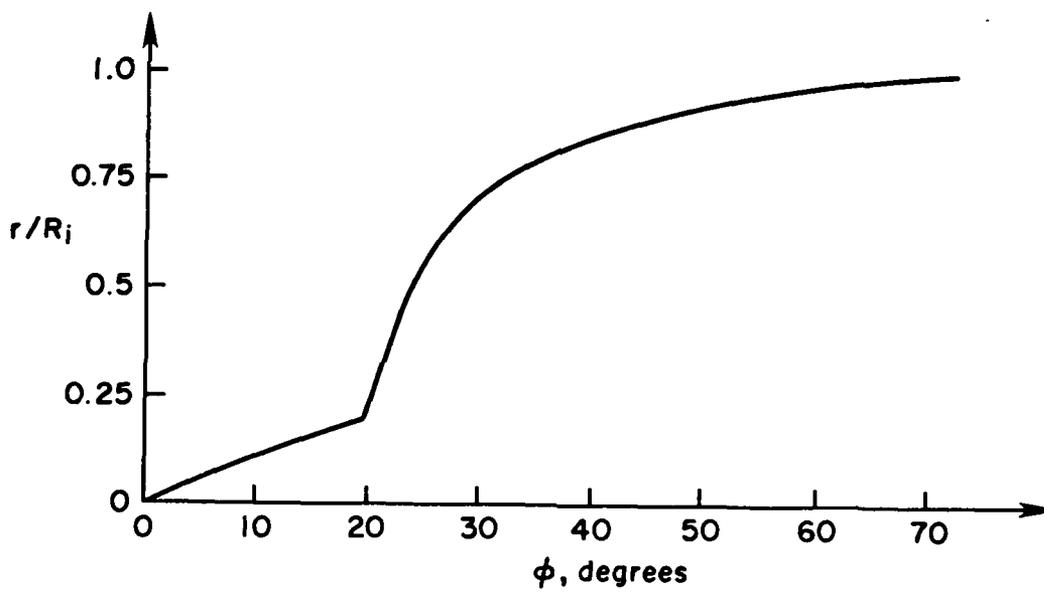


Fig. 7 Inlet swirl angle profile for $S_i = 0.3$

this results in the swirl angle profile at the inlet shown in Fig. 7. Obviously, this profile is a poor approximation to a constant angle distribution. In order to obtain a better approximation, a swirl velocity profile, w , which falls as the wall is approached would have to be assumed. An even better approach would be to re-formulate the inlet boundary conditions so that the specified properties are P_t , T_t , θ , and ϕ rather than P_t , T_t , θ , and w .

VII. CONCLUSIONS

As a result of this study, the following conclusions can be drawn:

- (1) A time-dependent finite difference technique for analyzing inviscid, axisymmetric, swirling flow in converging-diverging nozzles has been developed.
- (2) The first known fully two-dimensional swirling nozzle flow computations which treat the entire subsonic-transonic-supersonic regions have been obtained for a particular nozzle with a particular tangential velocity profile.
- (3) Based on this limited study, it is concluded that for $S_i \lesssim 0.4$, reductions in mass flow and vacuum stream thrust due to swirl are on the order of 10% or less.
- (4) The effect of swirl on the nozzle flowfield is to cause a large increase in the axial velocity near the centerline and a slight retardation near the wall as compared to the no-swirl case.
- (5) For the cases considered, little distortion of the swirl velocity profile shape occurs as the flow progresses through the nozzle. The magnitude of w at a particular axial station is generally inversely proportional to the wall radius at that section, as expected.
- (6) For large swirl, physically unrealistic inlet flow properties are obtained under the presently assumed inlet boundary conditions, viz. supersonic flow near the centerline with very large swirl angles and even reverse flow near the wall.

VIII. RECOMMENDATIONS

In order to complete and extend the work reported herein, the following further investigations are recommended:

- (1) Additional testing and validation of the code should be done, e.g. checking the constancy of angular momentum and stagnation temperature along streamlines, conservation of mass at each axial station, effects of grid refinement, etc.
- (2) The calculations should be extended to consider other c-d nozzle geometries and other tangential velocity distributions to determine their effects on C_D and η_N as functions of S_1 .
- (3) The code should be generalized so that exhaust plume calculations can be made for converging nozzles and centerbody calculations can be made for variable geometry nozzles.
- (4) Investigation of the inlet boundary conditions should be carried out including re-formulation in terms of specified (P_t, T_t, θ, ϕ) and (u, v, w, ρ) . The latter set has been proposed by Cline [31] as an alternative to the former, but its correctness has been questioned [35].
- (5) A series of nozzle experiments which include mass flow, thrust, and velocity profile measurements should be conducted both to verify the numerical calculations and to determine realistic inlet conditions for swirled and unswirled nozzle flows. No such measurements presently exist for fixed vane swirlers.
- (6) Long term goals are the inclusion of viscous effects, turbulence modeling, non-ideal gas, and/or reacting flow effects. In this way coupled analyses of the combustor/nozzle system could be performed as opposed to the present "patched" solution techniques. This method appears to provide an attractive alternative to the incompressible TEACH-based combustor analyses [37] which may be fundamentally inadequate for the high speed flow occurring in ramjet combustor/nozzle systems.

REFERENCES

1. Buckley, P.L., Craig, R.R., Davis, D.L., and Schwartzkopf, K.G., "The Design and Combustion Performance of Practical Swirlers for Integral Rocket/Ramjets," AIAA Paper No. 80-1119, 1980.
2. Binnie, A.M., "The Passage of a Perfect Fluid Through a Critical Cross-Section or Throat," Proceedings of the Royal Society, Series A, Vol. 197, 1949, pp. 545-555.
3. Mager, A., "Approximate Solution of Isentropic Swirling Flow Through a Nozzle," ARS Journal, Vol. 31, Aug. 1961, pp. 1140-1148.
4. Bastress, E.K., "Interior Ballistics of Spinning Solid-Propellant Rockets," Journal of Spacecraft and Rockets, Vol. 2, No. 3, May - June 1965, pp. 455-457.
5. Manda, L.J., "Spin Effects on Rocket Nozzle Performance," Journal of Spacecraft and Rockets, Vol. 3, No. 11, Nov. 1966, pp. 1695-1696.
6. King, M.K., "Comment on 'Spin Effects on Rocket Nozzle Performance'," Journal of Spacecraft and Rockets, Vol. 3, No. 12, Dec. 1966, pp. 1812-1813.
7. Manda, L.J., "Reply by Author to M. King," Journal of Spacecraft and Rockets, Vol. 3, No. 12, Dec. 1966, pp. 1813-1814.
8. King, W.S., "On Swirling Nozzle Flows," Journal of Spacecraft and Rockets, Vol. 4, No. 10, Oct. 1967, pp. 1404-1405.
9. Glick, R.L. and Kilgore, M.S., "Effect of Specific-Heat Ratio on Mass Flow for Swirling Nozzle Flow," Journal of Spacecraft and Rockets, Vol. 4, No. 8, Aug. 1967, pp. 1098-1099.
10. Lewellen, W.S., Burns, W.J., and Strickland, H.J., "Transonic Swirling Flow," AIAA Journal, Vol. 7, No. 7, July 1969, pp. 1290-1297.
11. Norton, D.J., Farquhar, B.W., and Hoffman, J.D., "An Analytical and Experimental Investigation of Swirling Flow in Nozzles," AIAA Journal, Vol. 7, No. 10, Oct. 1969, pp. 1992-2000.
12. Hsu, C.T., "Swirling Nozzle Flow Equations from Crocco's Relation," AIAA Journal, Vol. 9, No. 9, Sept. 1971, pp. 1866-1868.
13. Hsu, C.T., "Errata: 'Swirling Nozzle Flow Equations from Crocco's Relation'," AIAA Journal, Vol. 10, No. 3, March 1972, p. 368.
14. Hsu, C.T., "Mass Blocking of Swirling Flow in Nozzles," Journal of Spacecraft and Rockets, Vol. 8, No. 12, Dec. 1971, pp. 1232-1234.
15. Hsu, C.T. and DeJoode, A.D., "Inviscid Swirling Flows through a Choked Nozzle," AIAA Journal, Vol. 11, No. 11, Nov. 1973, pp. 1564-1566.
16. Wolfram, W.R. and Walker, W.F., "Swirling Flow of a Dissociated Gas," ASME Journal of Fluids Engineering, Vol. 97, March 1975, pp. 122-124.

17. Carpenter, P.W. and Johannesen, N.H., "An Extension of One-Dimensional Theory to Inviscid Swirling Flow through Choked Nozzles," *Aeronautical Quarterly*, Vol. 26, May 1975, pp. 71-87.
18. Carpenter, P.W., "A General One-Dimensional Theory of Compressible Inviscid Swirling Flows in Nozzles," *Aeronautical Quarterly*, Vol. 27, Aug. 1976, pp. 201-216.
19. Carpenter, P.W., "Effects of Swirl on the Subcritical Performance of Convergent Nozzles," *AIAA Journal*, Vol. 18, No. 5, May 1980, pp. 600-602.
20. Carpenter, P.W., "Supercritical Swirling Flows in Convergent Nozzles," *AIAA Journal*, Vol. 19, No. 5, May 1981, pp. 651-660.
21. Carpenter, P.W., "The Effects of Swirl on the Performance of Supercritical Convergent-Divergent Nozzles," *Aeronautical Quarterly*, Vol. 32, May 1981, pp. 126-152.
22. Guderley, K.G. and Tabak, D., "On the Determination of Optimum Supersonic Thrust Nozzles of Given Length for a Flow with Swirl: Theoretical Part," *Aerospace Research Laboratories Report No. ARL 66-0013*, Jan. 1966.
23. Guderley, K.G. and Breiter, M.C., "Approximation for Swirl Flows in the Vicinity of the Throat of a Laval Nozzle," *Aerospace Research Laboratories Report No. ARL 70-0009*, Jan. 1970.
24. Guderley, K.G. and Breiter, M.C., "On the Determination of Optimum Thrust Nozzles of a Given Length for a Flow with Swirl: Numerical Results," *Aerospace Research Laboratories Report No. ARL 70-0161*, Aug. 1970.
25. Guderley, K.G., Breiter, M.C., Bhutani, O.P., and Tabak, D., "Continuous and Discontinuous Solutions for Optimum Thrust Nozzles of Given Length," *Journal of Optimization Theory and Applications*, Vol. 12, Dec. 1973, pp. 588-628.
26. Dunlap, R., "An Investigation of the Swirling Flow in a Spinning End-Burning Rocket," *AIAA Journal*, Vol. 7, No. 12, Dec. 1969, pp. 2293-2300.
27. Batson, J.L. and Sforzini, R.H., "Swirling Flow through a Nozzle," *Journal of Spacecraft and Rockets*, Vol. 7, No. 2, Feb. 1970, pp. 159-163.
28. Sforzini, R.H. and Essing, J.E., "Swirling Flow through Multiple Nozzles," *Journal of Spacecraft and Rockets*, Vol. 7, No. 11, Nov. 1970, pp. 1366-1369.
29. Serra, R.A., "Determination of Internal Gas Flows by a Transient Numerical Technique," *AIAA Journal*, Vol. 10, No. 5, May 1972, pp. 603-611.
30. Cline, M.C., "VNAP: A Computer Program for Computation of Two-Dimensional, Time-Dependent, Compressible, Viscous, Internal Flow," *Los Alamos Scientific Laboratory Report No. LA-7326*, Nov. 1978.
31. Cline, M.C., "VNAP2: A Computer Program for Computation of Two-Dimensional, Time-Dependent, Compressible, Turbulent Flow," *Los Alamos National Laboratory Report No. LA-8872*, Aug. 1981.

32. Stiles, R.J. and Hoffman, J.D., "Analysis of Steady, Two-Dimensional Chemically Reacting Nonequilibrium Flow by an Unsteady, Asymptotically Consistent Technique, Volume I - Theoretical Development," AFWAL-TR-81-2127, Vol. I, Sept. 1982.
33. Stiles, R.J. and Hoffman, J.D., "Analysis of Steady, Two-Dimensional Chemically Reacting Nonequilibrium Flow by an Unsteady, Asymptotically Consistent Technique, Volume II - Computer Program Manual," AFWAL-TR-81-2127, Vol. II, Sept. 1982.
34. MacCormack, R.W., "The Effect of Viscosity in Hypervelocity Impact Cratering," AIAA Paper No. 69-354, April 1969.
35. Moretti, G., "Comment on 'Stability Aspects of Diverging Subsonic Flows'," AIAA Journal, Vol. 19, No. 5, May 1981, p. 669.
36. Cline, M.C., "Reply by Author to G. Moretti," AIAA Journal, Vol. 19, No. 5, May 1981, pp. 669-671.
37. Patankar, S.V. and Spalding, D.B., "A Calculation Procedure for Heat, Mass and Momentum Transfer in Three-Dimensional Parabolic Flows," International Journal of Heat and Mass Transfer, Vol. 15, 1972, pp. 1787-1806.

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 EVALUATION OF TWO NUCLEAR WEAPONS EFFECTS COMPUTER PROGRAMS

Prepared by:	Dr. John G. Eoll
Academic Rank:	Assistant Professor
Department and University:	Department of Physics, Lenoir-Rhyne College, Hickory, North Carolina
Graduate Student Research Assistant:	Mr. Robert Calvert, Jr.
Department and University:	Department of Mathematics & Computer Science, Emory University, Atlanta, Georgia
Research Location:	Air Force Weapons Laboratory Nuclear Criteria Group Secretariat and Nuclear Technology Engineering Division
USAF Research Colleagues:	Major Raymond Bell, NTEDA Lt John Ruble, NCGS
Date:	July 8, 1983
Contract No:	F49620-82-C-0035

AN EVALUATION OF TWO NUCLEAR WEAPONS EFFECTS COMPUTER PROGRAMS

by

John G. Eo11

Robert O. Calvert, Jr.

ABSTRACT

Work was performed on two FORTRAN programs used to model two separate nuclear weapons effects problems. The first code, the widely used HULL hydrodynamic code, was evaluated as to the possibility of including the effect of dust scattering in the radiation transport section of the code. It is our conclusion that considerable work needs to be done with HULL before the scattering effect can be included. Specific recommendations are described in detail.

The second code worked on was a family of FORTRAN programs used to model fallout problems. We combined the programs AFIT, BRAVO3, and SMEAR into a single interactive code named NCG capable of running on a PDP 11/24 minicomputer. Specific attention was paid to including a detailed wind field in the calculation of the distribution of fallout particles.

ACKNOWLEDGEMENTS

The authors wish to express their appreciation to the Air Force Office of Scientific Research, Air Force Systems Command, and the Southeastern Center for Electrical Engineering Education for providing us with the opportunity to participate in an important and interesting research experience.

We are also indebted to the host institute, the Air Force Weapons Laboratory, and to the many individuals who contributed to the success of our project. We specifically wish to thank Major Ray Bell of the Atmospheric Phenomenology Group of the Nuclear Technology Engineering Division (NTEDA) and Lieutenant John Ruble of the Nuclear Criteria Group Secretariat (NCGS) for providing us with much guidance and many helpful suggestions along the way.

We are also grateful to Lieutenants Jim Downey and Bob Connell for their frequent assistance and useful discussion.

Finally, we feel particularly indebted to John Ungvarsky, the Technical Coordinator, for making our stay at AFWL a pleasant one.

I. INTRODUCTION:

Airborne dust is important in at least two ways in the study of nuclear weapons effects. On the one hand, the dust may form from the condensation of fission products and weapon debris. In this form it is better known as "fallout;" although technically speaking, fallout refers to the radioactive dust only after it has settled to the ground. On the other hand, dust from the soil is lofted up into the air by the early effects of the bomb. In this case it is useful to discriminate between two regimes. In the first, the thermal pulse from the weapon causes dust to be injected into a layer which lies less than three meters from the ground. The presence of the dust in this layer then attenuates the subsequent blast shock wave which soon arrives. This is known as the thermal dust problem. The second regime is the dust which is lofted up into the stem by the afterwinds, which may reach hurricane force. This dust can become entrained in the rising fireball, and subsequently carried to great heights. In the case of a surface burst, the material from the crater can be thought of both as lofted dust and as an integral part of the fallout cloud. Besides the radioactivity which these particles emit, there is also the danger of erosion of aircraft parts due to the dust lofted to such high altitudes.

At the Air Force Weapons Laboratory (AFWL), two separate FORTRAN programs maintained by two different groups are used to investigate these two distinct dust problems. In the case of the thermal/lofted dust, a large multi-dimensional radiation hydrodynamics code named HULL is used. At AFWL, two versions of HULL are being maintained by

ACKNOWLEDGEMENTS

The authors wish to express their appreciation to the Air Force Office of Scientific Research, Air Force Systems Command, and the Southeastern Center for Electrical Engineering Education for providing us with the opportunity to participate in an important and interesting research experience.

We are also indebted to the host institute, the Air Force Weapons Laboratory, and to the many individuals who contributed to the success of our project. We specifically wish to thank Major Ray Bell of the Atmospheric Phenomenology Group of the Nuclear Technology Engineering Division (NTEDA) and Lieutenant John Ruble of the Nuclear Criteria Group Secretariat (NCGS) for providing us with much guidance and many helpful suggestions along the way.

We are also grateful to Lieutenants Jim Downey and Bob Connell for their frequent assistance and useful discussion.

Finally, we feel particularly indebted to John Ungvarsky, the Technical Coordinator, for making our stay at AFWL a pleasant one.

the Nuclear Technology Engineering Division (NTED), Atmospheric Phenomenology Group, under the direction of Major Raymond Bell. The first, EPHULL¹, contains subroutines which permit the modelling of elastic-plastic problems. This version also contains radiation diffusion routines which allow inclusion of non-equilibrium radiation flow in the problem. EPHULL has a serious deficiency; however, it is a scalar code with a correspondingly slow execution speed.

The newer version of HULL is known as Vector-HULL². This is a stripped-down version of HULL which contains neither the elastic-plastic ability, nor the radiation routines, but is fully vectorized for operation on the CRAY-1 computer.

The modelling of fallout effects is done by a family of small FORTRAN programs which originated at the Air Force Institute of Technology (AFIT), under the direction of Dr. Charles Bridgeman. Three of these codes were under study by our group:

1. AFIT³ - a program designed to calculate the activity which an aircraft encounters flying through the fallout cloud.
2. BRAVO⁴ - a program designed to calculate the locus of points which forms the curved axis of the fallout deposition ellipse, or "hotline" as it is known, for the particular case of the CASTLE-BRAVO shot.

3. SMEAR⁵ - a program designed to construct the fallout elliptical contour "footprint" from a statistical description of the fallout cloud.

II. OBJECTIVES OF THE RESEARCH EFFORT

The three programs described above for the fallout problem were each written with a different set of assumptions, and with different computational objectives. Our goal was to merge them into a single interactive program which will run on a small computer.

As regards the HULL code, our task was to investigate the possibility of installing a dust scattering mode in the radiation transport subroutines. The actual rewriting of HULL to include scattering was expected to be a long-range project.

III. PROBLEMS ENCOUNTERED WITH HULL

At the time our project began, the two HULL codes were undergoing some conversion. Vector-HULL was being converted from the CRAY Operating System (COS-a batch system), to the CRAY Time-Sharing System (CTSS). This conversion was not completed until the eighth week of our project, too late to enable us to use it.

The EPHULL code gave us difficulty from the beginning, because of an unusual processing system. The processing system, called SAIL⁶ includes the following features:

1. A source code editor - non-interactive batch only.
2. A library editor - containing a library of subroutines in a non-standard format.
3. A redimensioning feature - permitting the redimensioning of COMMON blocks, etc.
4. A file management system - based upon a previously used magnetic tape system. (This system can be run independently of SAIL.)

We found the SAIL system to be a serious barrier to code development, and we question its appropriateness in a time-sharing environment. In order to understand our objections, we describe the sequence of code development steps both under SAIL and under the more widely used UPDATE system. Table 1 describes the conventional code modification cycle. Table 2 describes our best guess as to the SAIL cycle, "best guess" because we never did successfully master it. One of us (JGE) has had more than ten years of experience modifying and running large FORTRAN programs both in astrophysics and weapons design, and yet the correct operation of SAIL eluded us.

Several additional problems were encountered using EPHULL, besides attempting to edit it. When we tried to run it using just a single material (air) the program crashed (or "sank" in the nautical parlance of the code). Neither would the code work with two materials and the equilibrium radiation diffusion mode set. This was a serious difficulty. We had been expecting a fully operational radiation hydro

TABLE 1
CONVENTIONAL CODE MODIFICATION

1	Create an UPDATE "ident" file ⁷ detailing the changes desired in the source code. An interactive editor such as TRIX or FRED is usually used for this purpose.
2	Run UPDATE in Quick mode, selectively modifying just those subroutines (decks) referenced by the ident changes. This run results in a partial source (compile file).
3	Compile the compile file.
4	Link (load) the new object file to the older object file, selectively replacing those subroutines which have been modified.
5	Execute the modified program, using a prior dump file.
6	Analyze the results.

TABLE 2
CODE MODIFICATION UNDER SAIL

1	Create an INPUT file for SAIL using special keywords to instruct SAIL as to the modification and how they are to be incorporated. Create a SAIL change file using TRIX.
2	Execute SAIL generating the library source code.
3	Compile the source code generated.
4	Use BUILD and/or LIB to modify the relocatable and absolute libraries.
5	Execute SAIL in the BOW/PLANK mode, generating a HULL source code.
6	Compile the HULL code.
7	Link the new object code to the HULL library files.
8	Execute HULL.
9	Analyze the results.

NOTE: If the vector version of HULL is being used, then additional or replacement steps are to be added involving the use of UPDATE and the problem POST, which accomplishes the extraction function of SAIL.

code. Instead, we discovered a processor system that, for the uninitiated, is difficult to work with, and a radiation hydro program which did not work.

The degradation in performance and efficiency which results from operating HULL under the SAIL system cannot be emphasized too strongly. At Los Alamos, probably some 25-30 hydro codes are presently in use (derivatives of ICE, ALE, YAQUI, LASNEX, BIG MAC, etc.). Add to this another dozen or so blast wave codes known to us from outside of Los Alamos, brings the total to 40 + hydro codes. HULL stands by itself in the exotic and arcane way in which it is maintained and operated. SAIL is a perfect example of being "too clever by half." Quite frankly, the SAIL system is a barrier to code accessibility, to cost efficient code development by contractors and/or staff, and to general and widespread use of HULL.

Note below the usual (and easy) sequence of operations by which most hydro codes are used:

1. Run a set-up code to initialize the mesh. This code is usually referred to as the "pre-processor."
2. Run the hydro code (the "processor").
3. Run a graphics code to obtain graphical output. This is usually referred to as the "post-processor." In fact, at Los Alamos there are several codes named POST (or CPOST, etc), which perform these

tasks. Note that the POST code used by HULL is actually part of the pre-processor. So much for having the name of a code denote its function!

IV. AN EVALUATION OF THE RADIATION ROUTINES IN HULL

It was not possible to get a radiation problem to run successfully on HULL even in the non-equilibrium mode. The radiation routines have not been exercised for some time, and were in COS compatible format. Lt Connell began the conversion of these routines into CTSS format for us at about the fifth week of our project. This was continued until the discovery that the opacity tables necessary for the radiation routines have apparently been lost. We cannot judge whether these tables were in the standard MAPLE format used by Los Alamos (and thus, available from Los Alamos) or not.

The radiation routines presently in EPHULL were originally installed by Alme, Westmoreland, and Fry⁸ in 1976. These routines calculated the radiation flux based upon a non-equilibrium diffusion model, using a flux limiter to limit the radiation flow in optically thin regions. An analysis of this model was performed by Fishbine⁹ in 1980. Apparently, very little has been done with the radiation routines since the Fishbine analysis.

Fishbine noted "several, quite serious coding errors," and discovered that "two subroutines needed for radiation work were ... missing from the HULL library." After correcting these deficiencies,

Fishbine went on to state that "this, however, did not result in a working code, suggesting that some additional errors remain to be uncovered."

Fishbine also criticized the test problem used by Alme, Westmoreland, and Fry, stating that "the combination of adhoc assumptions and the actual values of the mean-free paths chosen for the geometries investigated in (their) examples ..., and a selective use of the equation of state of air have, in effect, reduced (their) example to a case in which the hydrodynamics and radiation transport are decoupled. As a result, the non-equilibrium radiation algorithm (was) not really being tested."

Like us, Fishbine "started with the assumption that, aside from some possibly trivial coding errors, the DESHULL code was a working code--even for radiation problems. When initial analysis of the code showed that the coding errors were far from trivial, and when correction of these failed to yield a working code, an ab initio analysis of the underlying theory was carried out (instead)."

If anything, our conclusion is that the status of HULL's radiation routines have deteriorated even further since Fishbine's study.

As regards Fishbine's analysis of the Alme, Westmoreland, and Fry equations and algorithms, we agree, for the most part, with his conclusions. We feel, however, that it is necessary to emphasize the difficulties inherent in modelling radiation hydrodynamics in an Eulerian frame, as is done by HULL. Recall that the Eulerian frame is

a fixed laboratory frame, and the fluid moves with respect to it; but in the case of material opacity, the Lagrangian or local frame, at rest with respect to the fluid, is the natural frame to use. Attempting to impose these opacity terms in an Eulerian differencing scheme, necessitates the inclusion of additional factors and terms representing the effects of the Doppler shift and aberration, as well as fluid advection. All of these factors are missing from the HULL equations. Thus, these equations might be described as "primitive" set of transport equations. It is possible that for some limited class of problems these equations might be sufficient.

Keep the following in mind. The theory of hydrodynamics is quite simple. The principal difficulty in hydrodynamics is numerical (or calculational). When radiation transport is added to the problem, the difficulties double or triple (at least). It is the theory that is considerably more difficult, and there is little agreement as to which combination of radiation effects need to be included for a given problem¹⁰. When scattering is added on top of this, the difficulties increase by another factor of two or three. However, scattering is not a difficult theoretical problem. Like hydrodynamics, its main difficulty lies in the numerical calculations. Therefore, in a radiation hydro code with scattering, one should expect to expend equal amounts of programming time in the three areas: hydro, radiation, and scattering. HULL appears to have had hundreds of man-years of work done in the hydro, less than one man-year in the radiation, and none in

the scattering area. The quality of the product corresponds closely with this distribution of effort.

V. MODELLING FALLOUT PATTERNS: THE NCG CODE

The NCG code is a simplified merger of the three codes: AFIT, BRAVO3, and SMEAR. NCG is designed to run interactively on a small mini-computer such as the PDP 11/24.

We began by using the AFIT code as a base code. The AFIT code³ calculates the flythrough radiation dose to an aircrew caused by the radioactivity of the particles in the fallout cloud. It divides the fallout particles into 100 groups, based upon size. Each group has an equivalent amount of radioactivity. The descent of each particle group is computed by the code, particle by particle. Each particle starts at a different altitude depending approximately upon the condensation temperature of the material involved. In other words, particles which are more refractory (i.e., higher vaporization temperature) condense lower in the cloud and grow to larger size by the time the cloud reaches the stratosphere. AFIT uses a rather simple exponential model of the atmosphere, and uses the Davies-McDonald model to calculate the drag forces on the particle. According to this model, the drag force is composed of two forces: viscous drag and aerodynamic lift. Once the drag force is known, the terminal velocity of the particle is then calculated. Knowing the altitude of the aircraft, and the time it enters the fallout cloud, it is then a straightforward task to calculate the dose rate at the altitude of the aircraft. The actual

dose received by the plane depends upon the horizontal dispersion of the cloud, which is calculated according to a standard model (WSEG-10).

BRAV03, on the other hand, was designed to analyze the hotline of radioactivity produced by the CASTLE-BRAVO shot⁴. It computes a detailed atmospheric model based upon the National Oceanographic and Atmospheric Agency 1976 Standard (NOAA-76). The variable wind data of Dean and Olmstead are used to determine horizontal transport of fallout particles. The fallout is divided into 20 particle size groups. As in AFIT, the Davies-McDonald model is used to determine the terminal velocity of each particle group. The initial stabilization altitude is calculated in the same manner as AFIT.

In the NCG code, BRAV03 is a subroutine called by AFIT. This subroutine calculates a more accurate atmosphere, and includes the effect of wind on the particle transport. The 100 particle groups of AFIT were retained. An executive (main) program determines whether a flythrough problem is to be calculated or a hotline problem.

The third program, SMEAR, models the fallout in quite a different way. SMEAR calculates the dose rate at a given point on the ground (or determines dose contours). It assumes a dual-normal function to describe the horizontal dispersion of the fallout cloud, based upon the wind shear present (assumed to be constant).

We have several criticisms of the basic SMEAR model as described in Reference 5. First, by assuming a Dirac delta function for the initial vertical particle distribution, SMEAR is essentially assuming that all

particles start at the same height. This is contrary to the assumption of AFIT and BRAV03, and seems to conflict with the evidence. Second, the introduction of equation (10) as a new equation relating the grounding function, $j(t)$, to the activity-size distribution function, $A(r)$, seems to imply that only a given sized particle reaches the ground at a given time. If this is actually the assumption underlying equation (10), then it should be stated explicitly. We question the validity of this assumption, and therefore, equation (10). Third, an even more serious flaw comes about by assuming that the grounding function is linear in time as implied by equation (31). This seems to amount to some hypocrisy in view of the forceful criticism of the use of an exponential function for $g(t)$ by the canonical model (on pages 208-210). We strongly object to this linear approximation. None of the data seems to support it. Finally, equation (32) does not seem to follow from use of equation (31) as described in the text.

Besides these criticisms, there is an additional difficulty with SMEAR. It uses an integral method to determine the grounding function. Thus, it is not possible to extract the details of the particle transport itself, such as is necessary to determine flythrough dose or the effect of adding wind. The details of the transport are essentially embedded in the tables of Laurent coefficients used by the program. Because of these difficulties and criticisms, we did not incorporate the SMEAR algorithm in the NCG program, except for its contour generator.

VI. RECOMMENDATIONS

As regards the NCG code and the calculation of fallout patterns, we recommend that work continue on enhancing the code, particularly in two areas: reducing execution time, and providing the code with the ability to generate graphical output. In addition, the following areas should be studied for possible inclusion in NCG:

1. Time-varying wind.
2. Overlapping footprints due to multiple bursts.
3. The effect of vertical winds, updrafts, etc., in retarding the fallout descent.
4. The effect of rainout of fallout particles due to rainstorms.

We note, also, that currently most fallout codes are calibrated against just six nuclear tests: Johnie-Boy, Jangle-S, Small Boy, Koon, Zuni, and Bravo. Four of these tests took place before 1956. Aside from Bravo, which is the test that provides the most difficulty, the most recent test was Johnie-Boy in 1962, which was a fission-only test of very low yield. The design of nuclear weapons has changed a great deal since these tests. Some thought should be given to performing a detailed calculation of the fission products generated by the newer fission-fusion-fission weapons. This might alter not only the total amount of activity generated, but the activity-size distribution as well.

We also question the neglect of neutron activation products. Approximately three neutrons are emitted for each fission. One of these causes a fission with two fission products being formed. The fission products will emit approximately three gamma-rays each. The excess neutrons are captured by debris or lofted soil elements. These activated products will each emit a gamma-ray. Thus, neutron activation products will be responsible for two of eight delayed gammas emitted per fission, or 25 percent of the net fallout activity.

Finally, why have not the results of the Upshot-Knothole Harry test been factored into the fallout calculations? This test was designed to produce as much fallout as possible and has since been referred to as "Dirty" Harry. This should be the proof-test for any fallout code.

As regards HULL, we are convinced that it would be in the best interests of all concerned if the HULL source code were extracted from the suffocating clutch of SAIL. SAIL is an unnecessary obstacle to efficient code development and more widespread use of the HULL program. In a way, the question is similar to the debate over whether the U.S. should adopt the metric system or not. In our case, the rest of the world uses UPDATE; virtually, only HULL users use SAIL. Even if the two editing systems were equally attractive, the economic advantages (in software maintenance, training of staff, etc.) of switching to the more widely used commercial system would be too strong to resist; but, in fact, the two systems are not equally attractive. The UPDATE system of using end-of-line identifiers is clearly superior to the jumble of

in-line "skipto's" and "kepto's" of SAIL, which also violates ANSI standards. Old habits die hard, but this is one that should be forcibly retired.

Finally, NTEDA should assign a full-time staff member to care for and resuscitate the radiation routines in HULL. Our opinion is that Lt Connell is eminently qualified for this assignment both by virtue of his interest in radiation transport, and because of his knowledge of the codes. There is no simple or cheap way to get HULL to work on early fireball problems (i.e., radiation hydrodynamics problems). It is going to take a real commitment on the Laboratory's part to get the job done.

REFERENCES

1. Durette, R. E., and Matuska, D. A., "The HULL Code: Finite Difference Solution to the Equations of Continuum Mechanics," AFATL-TR-78-125, Air Force Armament Laboratory, Eglin AFB, FL, November 1978.
2. Booen, M. W., "The Operation of the Vectorized HULL Hydrodynamics Computer Code on the CRAY-1 Computer System," AFWL-NTE-TN-81-708, Air Force Weapons Laboratory, Kirtland AFB, NM July 1981.
3. Hickman, B., and Bridgeman, C. J., "Aircrew Ionizing Doses From Radioactive Dust Cloud Generated by Nuclear Burst," AFIT-GNE-PH-82-9, Air Force Institute of Technology, Wright-Patterson AFB, OH, March 1982.
4. Hopkins, A. T., "A Two Step Method to Treat Variable Winds in Fallout Smearing Codes," AFIT-GNE-PH-82M-10, Air Force Institute of Technology, Wright-Patterson AFB, OH, March 1982.
5. Bridgeman, C. J., and Bigelow, W. S., "A New Fallout Prediction Model," Health Physics, Vol 43, No. 2, pp. 205-218, 1981.
6. Gaby, L. P. "HULL System Guide," AFWL-TR-78-115, Air Force Weapons Laboratory, Kirtland AFB, NM, January 1979.
7. "UPDATE Reference Manual," SR-0013, CRAY Research, Inc., 1981.
8. Alme, M. L., Westmoreland, C., and Fry, M. A., "Non-Equilibrium Radiation for the HULL Code," AFWL-TR-76-244, Air Force Weapons Laboratory, Kirtland AFB, NM, October 1976.
9. Fishbone, B. H., "Preliminary Assessment of the Non-Equilibrium Radiation Treatment in the DESHULL Code," QSI-TR-79-49, Tuantum Systems, Inc., Albuquerque, NM, January 1980.
10. Eoll, J. G., "Note Concerning the Equations of Radiation Hydrodynamics," LA-UR-77-1048, Los Alamos National Laboratory, Los Alamos, NM, June 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

AXIAL VARIATION OF LOCAL HEAT FLUX ALONG THE

CONDENSER SECTION OF A DOUBLE-WALL ARTERY

HIGH CAPACITY HEAT PIPE

Prepared by: Dr. Amir Faghri*
Christine L. Rainey**

Academic Rank: *Associate Professor
**Research Assistant

Department and University: Department of Engineering
Wright State University

Research Location: Thermal Energy Lab, Aero Propulsion Lab
Wright Patterson AFB, Dayton, Ohio

USAF Research: Dr. T. Mahefkey

Date: August 28, 1983

Contract No.: F49620-82-C-0035

AXIAL VARIATION OF LOCAL HEAT FLUX ALONG THE
CONDENSER SECTION OF A DOUBLE-WALL ARTERY
HIGH CAPACITY HEAT PIPE

by

Amir Faghri
Christine L. Rainey

ABSTRACT

A new technique was developed to experimentally measure local heat flux and heat transfer coefficient along the condenser section of a heat pipe. This design consists of installing circular fins along the condenser section of a double-wall artery high capacity heat pipe. This method allows for overall heat flux as well as the local variation. In addition, a better cooling performance was achieved with this design compared to conventional cooling jackets. The trend in local heat flux and heat transfer coefficients are the same and show a high value at the end of the adiabatic section and decreases toward the end of the condenser section. The experimental result shows an overall energy balance. The general behavior of the individual fin agrees with the analytical result obtained from heat conduction analysis.

ACKNOWLEDGEMENT

The authors would like to thank the Air Force Wright Aeronautical Laboratories, Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for providing us with the opportunity to spend an interesting summer at the Thermal Energy Laboratory, Wright Patterson Air Force Base, Ohio.

We would like to thank Dr. Tom Mahefkey, Jerry Beam, and R. Ponnappan for their support and excellent working conditions. We are also grateful to Capt. Ruston Stewart for his help in this project. The technical assistance provided by Charlie Hall (AFWAL) and John Tennant (UES), throughout this project was invaluable.

I. Introduction

Since the invention of heat pipes in 1963 by Grover (4) a dramatic growth in the development and research in this area has taken place. Unfortunately the transport phenomena on condensation inside heat pipes, despite thousands of papers on the heat transfer phenomena is not well understood. Some of the questions still not answered are the following.

1. The general trend of heat flux as well as heat transfer coefficient for different wick structures.
2. Does the heat transfer (condensation) take place over part of the wick, whole wick surface or at the surface liquid meniscus?
3. What does the role of interfacial shear have on the condensation phenomena on conventional heat pipes?
4. What is the role of the surface tension on the condensation?

The majority of investigators have used a water circulating cooling jacket in experimental evaluation of the condenser performance. Using this conventional technique no information can be obtained on local values. Other cooling systems such as droplet/forced air evaporation or fins were used with heat pipes, but no cooling rate in the condenser was measured.

In the early stage of heat pipe development, Nusselt theory (5) was used to describe the condensation phenomena in heat pipes. Because of liquid flow in the conventional wick (porous media) or axially grooved wicks, as well as its interaction with the gas flow, this model is too simplified model to be used considering the fact that the heat pipe is also a closed system. Seban and Faghri (6) considered the reflux condensation in a closed two phase thermosyphon (gravity assisted wickless heat pipe) for both constant wall temperature and constant heat flux at the wall by accounting for the shear stress at the interface due to vapor friction and suction. Theoretical prediction was compared with experimental and shows a good agreement. The theory of condensation was also extended for axially grooved

heat pipes (7,8,9). In axially grooved heat pipe analysis, the following assumptions were made:

1. Zero-gravity condition.
2. Vapor mainly condenses in the land area between the grooves where the liquid film thickness is lowest.
3. Liquid condensed is sucked into grooves by capillary action and then flows along the groove to the evaporator section.

In both the gravity assisted wickless heat pipe and the axially grooved heat pipes the results were compared with overall heat transfer coefficients which were obtained by a cooling jacket.

The problem of heat pipes with porous wick with an annulus liquid flow between the wick and the solid pipe wall was considered by H. Hwang-Bo and W.E. Hilding (10). Not only the assumption made in solving the problems are not true, but because of the annulus liquid flow and reduction of total resistance, it is not of practical interest. An effort is now underway to analyze and make a parametric study of the condensation phenomena in the condenser section of a conventional heat pipe accounting for the porous media structure as well as the variation of the contact angle along the axial direction.

It is interesting to note that for axially grooved heat pipes the liquid thickness is decreasing along the axial direction toward the end of the condenser. Therefore the heat flux, or heat transfer, has an increasing effect in the same direction. The trend is completely the opposite for conventional heat pipes because the resistance thickness is fixed by the thickness of the wick; however, the effective thermal conductivity is decreasing toward the end of the condenser.

II. Objectives of Research Efforts

The main objective of this project was to investigate the question of the general trend of heat flux as well as the heat transfer coefficients along the condenser section and propose a practical technique

for obtaining this information for steady and transient heat pipes. We did not attempt to obtain results for various types of heat pipes. Instead our experimental result is limited to "double wall" heat pipe leaving the other cases for later investigation. It should be noted that the reason for choosing such a heat pipe (1,2) first was the availability in the Thermal Energy Lab, Aero Propulsion Laboratory, Wright-Patterson Air Force Base, Ohio, and secondly with this heat pipe design the effect of interfacial effect is basically insignificant.

Advanced spacecraft thermal management requirement motivates the development of high capacity heat pipes, and subsequently better physical understanding of the condensation phenomena in the heat pipe. This will lead to improved condenser designs. In general, there is no guarantee that the overall heat transfer coefficients obtained for one type of cooling system can be applied to other types without understanding the heat transfer mechanism. Thus, as a first step, the local heat flux was measured experimentally. It should be noted that the capillary limitation on heat transport rate $Q_{c,max}$ can be derived from the heat transport factor $(QL)_{c,max}$ if the heat flux distribution along the heat pipe is known. In conventional heat pipe analysis such as (3) the heat flux distribution is assumed uniform for simplicity and because there is no information available.

III. Experimental Design, Set Up and Performance

A. Heat Pipe Description

In 1982 R. Ponnappan and E.T. Mahefkey (1) reported on the design and performance of a double-wall artery high capacity heat pipe. This pipe incorporates two concentric copper tubes. The inner tube has axial external grooves, the outer tube has internal circumferential grooves, and finally, there is an interannular screen-mesh wick. The pipe itself is 1.2m long with 0.2m evaporator and condenser section.

There are several advantages to this design. First, from a

manufacturing point of view the external grooves on the inner tube are easier to fabricate than normal internal grooves. Also fitting the pipe together and inserting the screen is easier. Secondly, from a performance standpoint, the presence of the inner tube nearly eliminates the entrainment limit because the liquid and vapor flow interface are essentially eliminated. The inner tube also provides for a permanent contact of the screen with the grooves. In addition, the grooves in the outer pipe increase circumferential wetting and increase the film boiling limit. A cross sectional view of the pipe is shown in Figure 1.

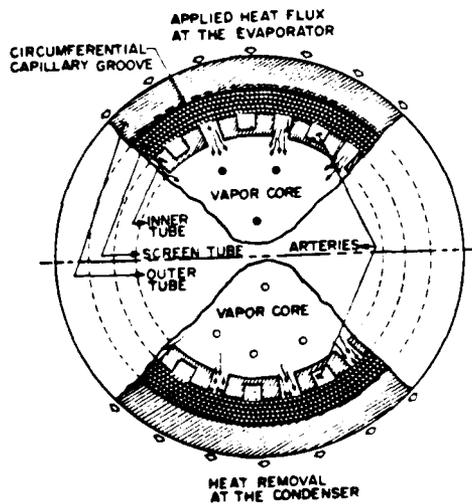


Fig. 1. Cross sectional view of double wall artery high capacity heat pipe.

B-1. Design of Cooling System

Since the overall experimental objective was to investigate the behavior of the local heat transfer coefficient along the length of the condenser section, we had to design a completely new instrumentation package to measure the local heat flux. Also, since most of the research efforts to date have used calorimetric water cooling systems, four out of five possible designs were based on efforts to better instrument existing or present state-of-the-art cooling systems.

These four water cooling systems basically measured the temperature gradient of the water at certain points along the condenser section. The first method was to use a water jacket with teflon posts to hold the thermocouple in place. The second method was to use instrumented circular fins inside of a water jacket. The third method was to use a sliding thermocouple harness inside a water jacket to allow for any number of readings along the section. The final water cooled design consisted of an instrumented ice bath around the condenser section. The first three designs were based on the following relationships:

$$Q_x = C_p \dot{m}_{in} (T_{out} - T_{in}) = h_x A (T_{wall} - T_{sat})$$

and

$$h_x = \frac{C_p \dot{m}_{in} (T_{out} - T_{in})}{A (T_{wall} - T_{sat})}$$

where T_{out} is the average fluid temperature at that cross section. The main disadvantage of these designs were twofold. First, only an approximate local heat transfer coefficient could be found. And secondly, the time required to fabricate each of these designs would have been prohibitive.

The final method, and the method chosen, was to use circular fins clamped to the heat pipe and instrumented to measure the temperature gradient along the radius at the base of the fin. The heat

flux for each fin was calculated using

$$q_x = K \left. \frac{dT}{dr} \right|_{r=0}$$

where K is the thermal conductivity, and $\left. \frac{dT}{dr} \right|_{r=0}$ is the temperature gradient at the base of the fin. This temperature gradient was also obtained analytically for a constant wall temperature. It was shown that this temperature gradient was linear up to .1" from the base of the fin for most metals and boundary conditions considered. Then the local heat flux and heat transfer coefficients are measured experimentally from the following equations respectively.

$$q_x = K \frac{(T_{\text{wall}} - T_u)}{\Delta r}$$
$$h_x = \frac{q_x}{T_{\text{wall}} - T_{\text{sat}}}$$

where T_{wall} is the wall temperature of the pipe, T_u is the temperature at .1" from the base of a fin along the radial direction of the fin, T_{sat} is the vapor temperature inside the heat pipe, and $\Delta r = .1"$.

There were several advantages to this design. First was the relative simplicity to manufacture the fins. Second was the ability to measure an actual local heat flux. Energy balance is the only method of checking the validity of the data. In other words the power entering the evaporation section should be approximately equal to the sum of the heat going out of each fin. It was also shown theoretically, using fin effectiveness, that the amount of power loss through the part which is not covered with fins was negligible. A diagram of this method is shown in Figure 2.

It is important to note also, at this time, that in order to accomplish the experimental objectives in the method proposed, we had to design a way to measure the saturation temperature of the working fluid. This was accomplished by using a steel sheathed

T type thermocouple and inserting it approximately 3/4 of the way from the end of the pipe through the condenser section. It is also interesting that the wall temperature at the middle of the adiabatic section was consistently within 0.1°C of the measured saturation temperature because of the long adiabatic section in this heat pipe.

B-2. Design Analysis for Fins

The next question to be answered is what material should the fins be made from, number of fins, and how the fins themselves are designed. The three metals considered were copper, 2024 aluminum, and steel. In choosing one metal over another, several criteria were kept in mind.

- a. Give the required cooling capacity for a given condenser length.
- b. Need a relatively large $\Delta T = (T_{\text{wall}} - T_u)$ for a short distance Δr .
- c. Is there a linear temperature profile close to the base of the fin?
- d. Is the metal easy to work with considering good contact, and the buckling problem?
- e. Quick and easy to manufacture.

In order to answer some of the above questions, analytical results for the temperature distribution, fin effectiveness and fin efficiency were obtained for a circular fin for constant wall temperature at the base of the fin and a constant surrounding heat transfer coefficient. The general differential equation describing the temperature distribution along a variable cross sectioned area fin is the following

$$\frac{d}{dx} \left(KA \frac{dT}{dx} \right) - h p (T - T_u) = 0 \quad (1)$$

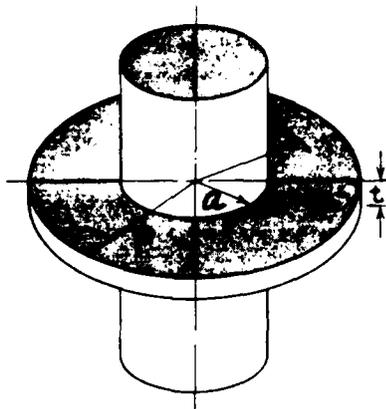


Fig. 4. A circular fin of constant thickness

The appropriate expression for the cross sectional area, A , and the perimeter, P , are

$$A = 2\pi r t$$

$$P = 4\pi r$$

when the above expressions are substituted into Equation (1) the applicable differential equation considering constant h and k with the appropriate boundary conditions are the following:

$$\frac{d^2 T}{dr^2} + \frac{1}{r} \frac{dT}{dr} - \frac{2h}{Kt} (T - T_\infty) = 0$$

at $r = a$ $T = T_0$ (wall temperature)

$r = b$ $\frac{dT}{dr} = 0$ (adiabatic)

Solution of the above differential equation for the boundary conditions given are the following (11).

$$\frac{(T-T_{\infty})}{T_o-T_{\infty}} = \frac{[I_o(\mu r) K_1(\mu b) + K_o(\mu v) I_1(\mu b)]}{[I_o(\mu a) K_1(\mu b) + K_o(\mu a) I_1(\mu b)]}$$

where

$$\mu = \left(\frac{2h}{kt}\right)^{1/2}$$

The effectiveness and efficiency of each fin are the following:

$$\begin{aligned} \gamma = \text{fin effectiveness} &= \frac{\text{heat loss with fin}}{\text{heat loss without fin}} \\ &= \frac{2}{\mu t} \left[\frac{I_1(\mu b) K_1(\mu a) - K_1(\mu b) I_1(\mu a)}{I_o(\mu a) K_1(\mu b) + K_o(\mu a) I_1(\mu b)} \right] \\ \eta = \text{fin efficiency} &= \frac{\text{actual heat transfer from the fin}}{\text{maximum possible heat transfer from the fin}} \end{aligned}$$

For design analysis, we considered an overall transport capacity limit of 1500 W-m for the present heat pipe (2) which was used for our investigation. Dividing this number by the total condenser section area, we calculated an overall heat flux to be 107.5 kw/m². This is the maximum amount of heat which is required to leave the condenser section with the proposed cooling system per unit area. Each metal was evaluated with fin size ranging from 0.0625" to 3 inches and thickness from 0.025" to 0.167 inches to determine the optimum fin radius.

Since all of the equations used were based on the assumption of one dimensional flow, this was obviously a very important design criteria. Physically the longer the fin is, relative to the width, the more accurate this assumption will be. On the other hand, however, there are practical considerations such as buckling limits and thermocouple attachments.

Based on these considerations and the fin efficiency calculations, we decided to use aluminum fins 0.125 inches thick and 3.0 inches in radius. This radius is defined as the distance between the pipe wall and the tip of the fin. We also assumed

that the effects of the steel clamp at the edge of the fin was negligible.

The final question relative to the design of the condenser section is, how many fins are required to cool the heat pipe effectively? Using the equation, $Q_{\text{max per fin}} = h A_{\text{SA}} (T_{\text{wall}} - T_{\infty})$, where A_{SA} is the total surface area of the fin, and knowing the fin efficiency, $\eta = 0.88$ as defined previously, you can solve for the actual heat flux into each individual fin by assuming $h = 2 \frac{\text{Btu}}{\text{hr ft}^2 \text{ } ^\circ\text{F}}$ and, $T_{\text{wall}} = 400^\circ\text{F}$ and $T_{\infty} = 68^\circ\text{F}$.

$$\begin{aligned} Q_{\text{max per fin}} &= 2 \frac{\text{Btu}}{\text{hr } ^\circ\text{F ft}^2} \times 0.526 \text{ ft}^2 \times (400^\circ\text{F} - 68^\circ\text{F}) \\ &= 349.3 \frac{\text{Btu}}{\text{hr}} \end{aligned}$$

$$\text{then } \eta = 0.88 = \frac{Q_{\text{actual}}}{Q_{\text{max}}} \rightarrow Q_{\text{actual}} = 307.4 \frac{\text{Btu}}{\text{hr}}$$

The above calculation assumes one dimensional flow and uniform heat loss across all of the fins. The heat loss from the primary surface, that area of the condenser section not covered by a fin, is negligible relative to the heat lost through the fins because the value of the effectiveness is about 326.

Now, using the assumed design total heat transfer capability of the pipe as mentioned previously, 1500 watts or 5100 Btu/hr, we can calculate the approximate number of fins required to cool the heat pipe.

$$\# \text{ of fins required} = \frac{5100 \text{ Btu/hr}}{307.4 \text{ Btu/hr}} = 16.6$$

It should also be noted that this is an approximate number and more or less fins may be required for different power inputs and with or without forced convection. For this reason we started with eight fins and went to fifteen. Another practical limitation is the spacing on the condenser section not allowing more than approximately fifteen fins.

These prototype fins were fabricated. Two used integral screw type clamps and the third used a steel ring clamp. We felt that the steel clamp provided the best and most uniform tightening and would provide the best contact with the pipe.

B-3. Instrumentation

Now that we've designed the cooling system and determined what material to use, there are several more questions that arise about how to instrument the fins to meet the experimental objectives. These questions are, what type of thermocouple to use, what size wire, where to place the thermocouples, and how to attach the thermocouples.

There were several factors that led to the decision to use J type (iron-constantan) thermocouples. The first, and foremost, reason is that Omega Engineering Corporation had only J type wire in stock in the size we needed. Type J was satisfactory in that its temperature envelope included our projected operating range, 0-200°C. Type T (copper-constantan) was the first choice because it has a narrower band and still included our range with lower errors in thermocouples.

The size of the thermocouple wire was very important because of the error in measurement. Because of the small distance between the thermocouple and the small projected temperature gradient the smallest wire, and the most accurate placement, was essential. For these reasons we decided to use .005 inch thermocouple wire and to place them at $r=0$, $r=.1$ ", $r=.15$ ". See Figure 3.

Initially we brazed several thermocouples and determined that the process was not satisfactory. This was because it was too time consuming, too tedious and too non-uniform. We then used a Dynatech Model 316 TIG welder. This method was very satisfactory resulting in uniform beads and the ability to keep the wire insulation right up to the bead to prevent extra junctions.

The original design was to attach the fins along the split wall

of the fin and put the head in a small indentation. The wires would run out the side of the fin and to the digital thermometer. The wall thermocouple would be glued to the fin and clamped between the fin and the wall. Several problems arose with this design, however. First, the beads were crushed sometimes to the point of failure. Second, the wire was flattened and hence weakened to the point of breaking. And third, the contact fin-to-fin and fin-to-pipe, was obviously poor. Initially we used Devcon F aluminum filled putty to attach the thermocouples. This method failed because the hardened putty caused gross contact problems. We also used "super glue", but still ran into problems with breaking the flattened wires.

Next we drilled each fin with .0225 inch holes. One was at .1 inch and .125 inches deep. The second hole was drilled through the side of the fin to the wall of the heat pipe. See Figure 3 for the orientation of the thermocouple wells. The thermocouples were dipped in Glyptal to provide extra insulation to prevent extra junctions. These were then held in place using either Dow Corning Silastic Silicon Rubber. This silicon rubber was also used to anchor the wires along the side of the fin. This aided immeasurably in handling the fins and wires. And finally, Dow Corning 340 heat sink material was used between the fin and the pipe to improve thermal contact.

And lastly, we discovered while calibrating the thermocouples using an icebath and a boiling water bath, that at least 100 feet of the thermocouple wire was bad, i.e. the temperature readings consistently ran higher than 0°C and lower than 100°C by about 10°C. Only twelve wires were satisfactory and were used for temperature measurements.

Another very serious problem with this whole design concept is that the accuracy of every temperature reading depends solely on the contact between the fin and the copper tubes. It is very important to note first that the pipe itself has inconsistencies

in its shape. These are due to the manufacturing process. Several actions were taken to account for and to alleviate this potential disastrous problem area. First, the inner diameter of each fin was milled to 0.870 inches. This is 0.005 inches smaller than the nominal outer diameter of the pipe. Second, a steel clamp was used to hold and clamp the fins to the tube. Third, each fin, when installed, was rotated in an attempt to get the best fit possible. And finally, we used Dow Corning 340 Heat Sink Silicon compound to further establish an assured thermal contact.

C. Experimental Set Up

The heat pipe experimental set up consists of three main components. These are the heat pipe itself, the power supply and monitoring equipment, and the instrumentation and data collection devices. The basic set up is the same as that reported by R. Ponnappan and E.T. Mahefkey (2). The only difference is in the design and instrumentation of the condenser section.

Each fin thermocouple is connected to a Fluke 2100A Digital Thermometer tied to a Fluke 2150A multi-point switch box. The thermocouples attached to the pipe itself were fed into a Instruments 8000 datalogger. This system provided for a real time check on the performance of the heat pipe and the ability to read the final temperature of the fins.

A large 20 inch box fan was used to provide a relatively uniform cooling air flow over the fins when required.

D. Experimental Procedure

Testing was initiated on the heat pipe as it was configured for previous experiments by R. Ponnappan. His experiments were duplicated to establish an experience base and to verify performance of the heat pipe. Procedures for equipment startup, data collection, and shutdown were established. Once these objectives were met, the heat pipe was reprocessed to eliminate the possibility of gas generation within the pipe. Experiments were

run again to confirm proper performance.

The heat pipe was then fitted with eight fins. Several tests were made to establish credibility of the fin thermocouple readings. Tests were performed with and without cooling air. In addition, in an attempt to increase the thermal contact, Dow Corning 340 Heat Sink Silicon Compound was put between the fin and the pipe.

After these tests were completed, nine additional uninstrumented fins were attached to the pipe. Tests were conducted with air and with and without heat sink material. In addition, in order to establish reliability in the thermocouple readings, the order of the instrumented fins were reversed. Also one set of experiments was conducted moving one fin to different locations along the condenser section. This was attempted to remove any possible error resulting from thermocouple error or spacing differences between the thermocouples.

IV. Results

The initial experiments were accomplished on heat pipe (2) exactly as left from experiments conducted by R. Ponnappan in late 1982 by a spray cooling system. As shown in Figure 4 our experiments closely matched previous results. Slight variances were observed due to the fact that we had no way of measuring the air or water flow exactly to duplicate previous conditions. As seen in the figure (4) we were able to attain a better spray cooling which gave us a near straight line on the graph. It can be seen, however, that the trends all correspond closely in shape. The only questionable areas in the temperature readings are at the end of the condenser section. This lower reading could be due to an increase in the spray cooling rate. Water was also observed to flow over the end of the pipe where the last thermocouple was located which would decrease the temperature. In addition, at lower spray cooling, previous trends were duplicated more exactly.

There were other possible explanations for this downward temperature reading. One possibility was that of internal noncondensable gas generation. For this reason we felt it necessary to reprocess the heat pipe before continuing the experiment. As can be seen in Figure 4 for the reprocessed heat pipe, the trends follow closely that of previous results, notably the condenser section temperature profile flattened out considerably and, in fact, reversed to an upward trend.

Throughout the remaining experiments the fin design was used as the cooling system, and several general trends were observed. First, the overall temperature profile across the pipe was somewhat higher than previously reported by spray cooling. However, the temperature profile is clearly more constant across the pipe. This is dramatically shown in Figure 4 where previous tests approached burn-out conditions. Using slow spray cooling by previous investigations (2) a ΔT of 62°C was approached for a power input of 900 watts. The most common failure in heat pipe is burn-out which is created by excessive power input at the evaporator section. It is brought about by the inability of the wick to feed sufficient liquid to the evaporator, and is characterized by a rapid rise in evaporator temperature compared to other regions of the heat pipe. Under normal spray cooling conditions the previous tests have a ΔT of 16°C which is still above the recommended ΔT . On the other hand, the fin design with air had a ΔT of only 3°C for the same power input. The spray cooling capacity is not as efficient as that of the fin design, for under normal cooling conditions a ΔT of 12°C was only seen at power inputs less than 750 watts. This figure also shows that perhaps with better fin contact, more fins, different fin material and a more efficient forced air system, there are possibilities for more efficient and better heat pipe performances.

The most important general trend is shown in Figure 5 for the variation of axial variation of local heat flux and heat transfer coefficients. From these graphs it can be seen that a general trend from high value to low value of the heat flux and heat

transfer coefficient exists. There is about a 10% drop in energy at the end of the condenser section compared to that at the beginning of the section. This seems to be a valuable conclusion since it existed for each power setting and shows a basic characteristic of the condenser section for this cooling method. It should also be noted that the total heat loss calculated from all the fins are within 5% of the energy input in the evaporator. This energy was shown both for eight and fifteen fins.

V. Recommendations

Several recommendations can be made with regard to the attachment of the fin to the pipe. Of course the least amount of resistance would occur if the fins were attached during the actual fabrication of the heat pipe. Molding the fins and pipe as a single piece of metal should be considered. Since the variability in the OD of the pipe itself created bad contact between the pipe and fin, it is recommended that the pipe itself (irrespective of the type of fin) be specially molded. If this were done then the routed fins would fit anywhere along the condenser section.

Some of our problems resulted from the small thermocouple wires we selected. By using duplex insulated wire a bending and breaking problem can be somewhat eliminated. This would also eliminate the problem of the thermocouple wires making contact with themselves and giving bad readings. Using a Silastic adhesive/sealant to attach the thermocouples tended to eliminate many of the problems one encounters using such fine wires.

It is of interest to study the effect of the surface tension within the heat pipe itself by the use of a nonfoaming surfactant. In prior studies different types of liquids have been used to study the effect of surface tension, but they also changed other properties of the working fluid. By using the nonfoaming surfactant, we are holding every property constant except reducing the surface tension.

More thermocouples should be instrumented on the fins so that the heat flux at the beginning and end of the condenser section can be evaluated. This will also allow us to evaluate the local heat transfer coefficient along the entire condenser section.

Research should be undertaken to study the conventional heat pipe and the thermosyphons (gravity assisted wickless heat pipe) with fins attached to the condenser section in order to study the axial variation in the local heat flux along the condenser section. This is a simpler and more classical approach to the problem.

It is also felt that a theoretical investigation of the local heat flux and heat transfer coefficient for the conventional heat pipe is needed. In this type of analysis, the effect of shear stress at the interface, axial variation of the radius of curvature of meniscus and the effect of porous medium should be accounted for.

REFERENCES

1. R. Ponnappan and E.T. Mahefkey, "Development of a Double Wall Artery High Capacity Heat Pipe," Paper No. AIAA-83-0318, AIAA 21st Aerospace Science Meeting, Reno, Nevada, Jan. 10-13, 1983.
2. R. Ponnappan and E.T. Mahefkey, "Development of a Double-Wall Artery High Capacity Heat Pipe," Paper No. AIAA-82-0906, AIAA/ASME 3rd Joint Thermo-Physics, Fluids, Plasma and Heat Transfer Conference, St. Louis, Missouri, June 7-11, 1982.
3. S.W. Chi, Heat Pipe Theory and Practice, (McGraw Hill, NY, 1976), P. 52-56.
4. G.M. Grover, T.P. Cotter and G.F. Erickson, "Structures of Very High Thermal Conductance," J. App. Phys. Vol. 35, p. 1990, 1964.
5. P. Dunn and D.A. Reay, Heat Pipes, (Pergamon, New York, 1976), P. 76-78.
6. R.A. Seban and A. Faghri, "Film Condensation in a Vertical Tube with a Closed Top," accepted for publication in the International Journal of Heat and Mass Transfer.
7. Y. Kamotani, "Analysis of Axially Grooved Heat Pipe Condensers," Paper No. AIAA-76-147, 1976.
8. V.A. Babenko, M.M. Levitan and D.K. Khrustalev, "J. Eng. Phys.", Vol. 40, P. 6, 1981.
9. L.I. Vasilev, L.P. Grakovich and D.K. Khrustalev, "Low-Temperature Axially Grooved Heat Pipes," Proceedings of the IV International Heat Pipe Conference, London, Sept. 1981.
10. H. Hwang-Bo and W.E. Hilding, "Optimization of a Heat Pipe with a Wick and Annulus Liquid Flow," J. Heat Transfer, May 1972.
11. H.S. Carslaw and J.C. Jaeger, Conduction of Heat in Solids, (Oxford University Press, Second Edition).

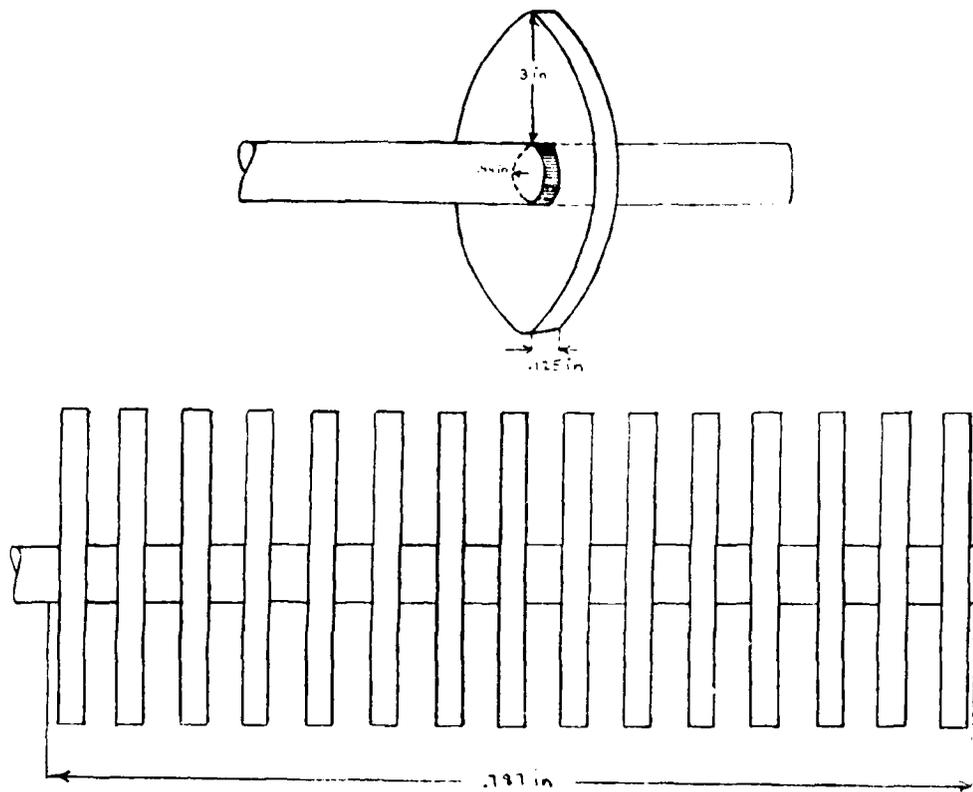


Figure 2 - Fin design for condenser section

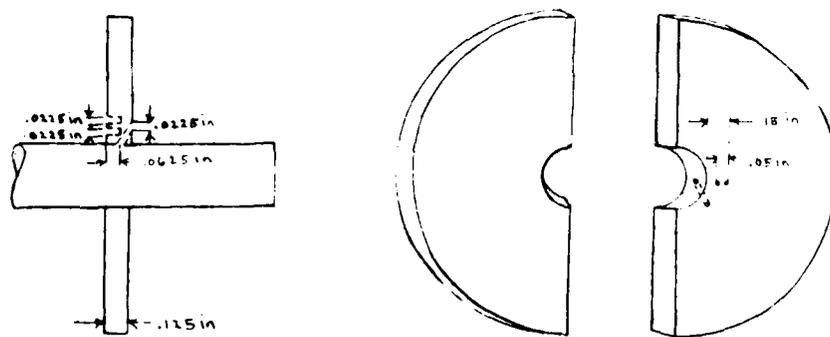


Figure 3 - Location of fin thermocouples

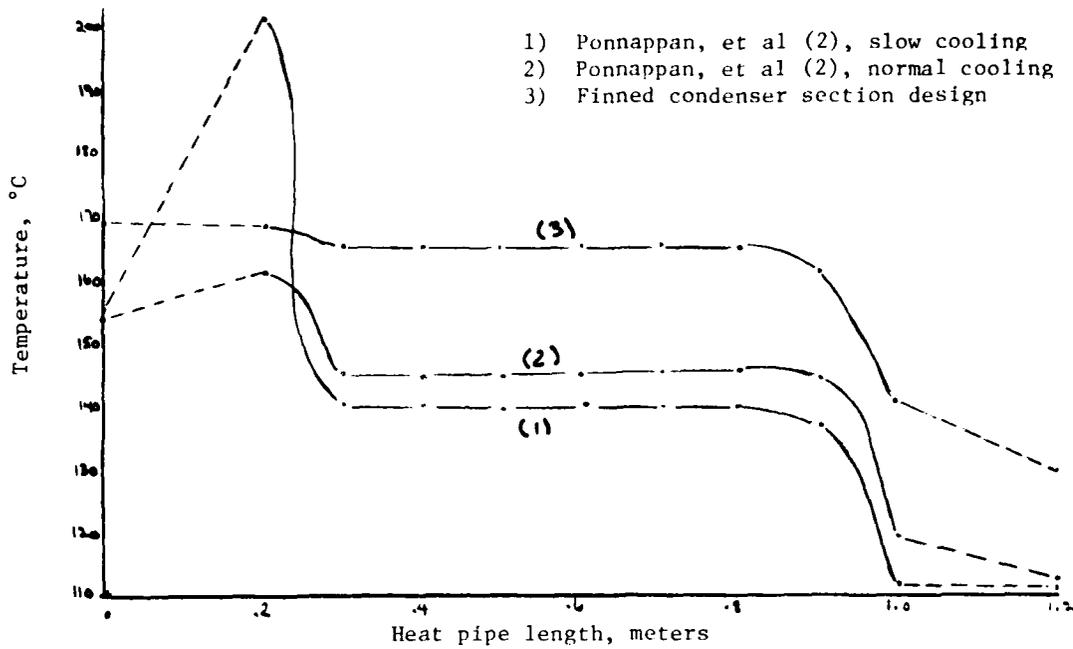


Figure 4A - Temperatures along heat pipe for 900 watts

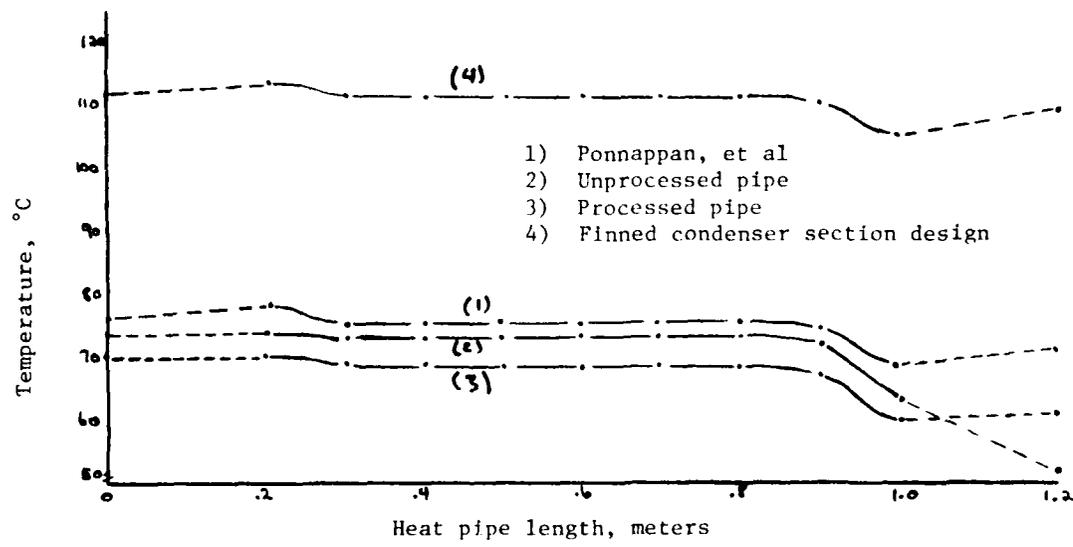


Figure 4B - Temperatures along heat pipe for 400 watts

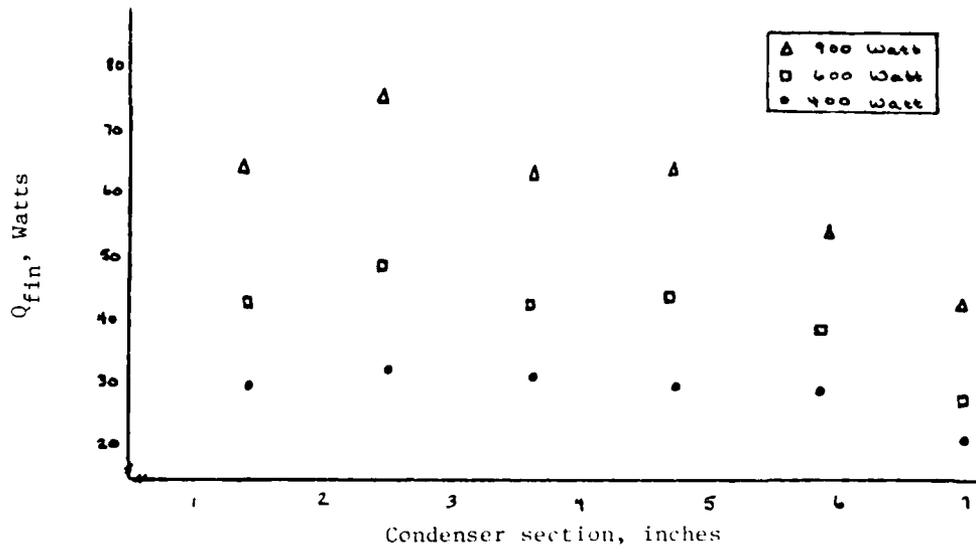


Figure 5A - Heat flux across the condenser section

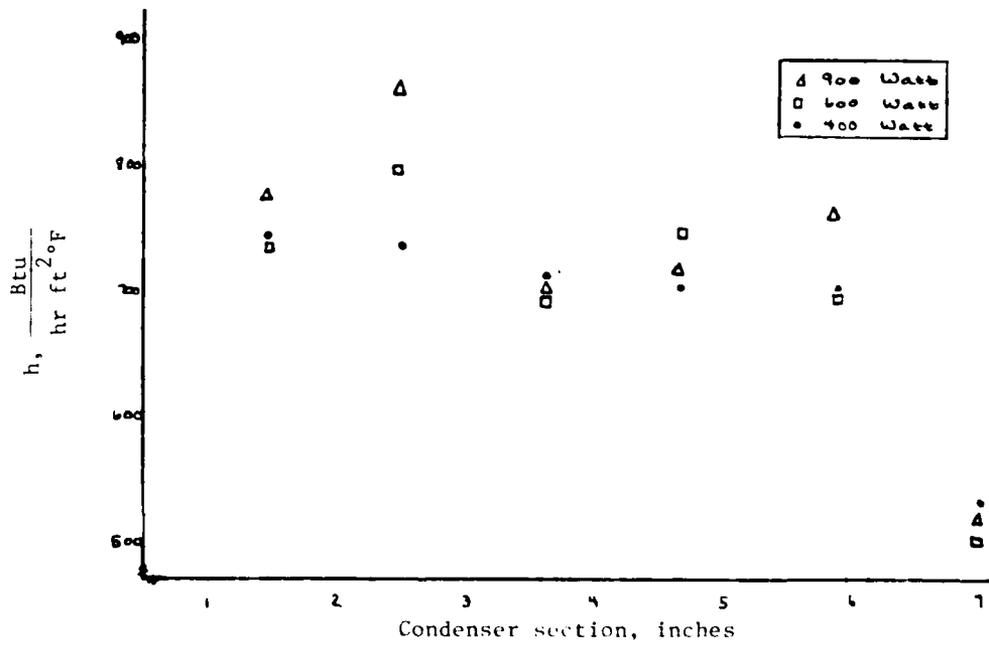


Figure 5B - Heat transfer coefficient across the condenser section

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
WINDSCREEN HAZE CHARACTERISTICS STUDIES

Prepared by:	Dr. Hans G. Fellner
Academic Rank:	Associate Professor
Department and University:	Department of Physics Slippery Rock University
Research Location:	Aerospace Medical Research Laboratory Human Engineering Division Integrated Visual Image Technology Section Wright Patterson Air Force Base
USAF Research:	Dr. H. Lee Task
Date:	September 15, 1983
Contract No.:	F49620-82-C-0035

WINDSCREEN HAZE CHARACTERISTICS STUDIES

by

Hans G. Fellner

ABSTRACT

Bird impact resistant aircraft transparencies manufactured from laminated plastics exhibit considerably poorer optical properties than the glass windshields they are replacing. After being in service on aircrafts their quality decreases even further as their surfaces become pitted and scratched. This study reports the first measurements of the angular dependence of the haze level in new and deteriorated F-111 aircraft windshield material samples. The measurements were used to successfully predict the loss in contrast (visibility) of a target when it was viewed through the windshield materials, which were oriented in several ways with respect to the design eye position.

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 productive summer in the Aerospace Medical Research Laboratory, Wright Patterson Air Force Base, Fairborn, Ohio.

He also extends his thanks to Lt. Colonel L. Genco, and particularly to Dr. H. Lee Task for suggesting the problem, taking out time for very helpful discussions, and providing unoppressive guidance.

I. Introduction and Objectives

It is the purpose of this project to quantify haze in F-111 airplane windshields. The surface of this windshield type is a section of a cone, and the interior consists of laminated sheets of polycarbonate, separated by flexible interlayers, for a total thickness of .8 inches. This construction makes the windshield bird impact resistant at high speeds. To insure durability, the outer layers are most often stretched acrylic. The thickness and the use of plastics significantly deteriorate the optical quality compared to conventional glass windows.

A transparent substance which is perfectly homogeneous (also nonexistent) would not scatter light, and as a consequence, be invisible. This condition is illustrated by a He-Ne laser beam traveling through the air in a dust-free room. It is virtually impossible to see that beam. However if inhomogeneities are introduced into the air, for example by blowing cigarette smoke or stirring up dust, then these inhomogeneities in the air will scatter light from the beam in all directions. Some of this scattered light will enter the observer's pupils; only then is the laser beam seen clearly.

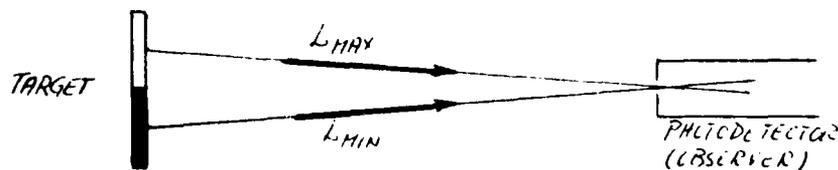
In general, gases and liquids scatter the least amount of light, but even highly purified and dust-free samples exhibit some scattering. It is due to the local variations in some parameters (such as density or temperature) of the substance. The blue color of the sky is due to this type of mechanism. In solids the presence of imperfections, impurities, and other irregularities gives rise to a comparatively large amount of light scattering. If the solid happens to be a windshield or other transparency, the light scattering is referred to as haze. The scattering of light from "point" irregularities within the volume of a transparency gives rise to volume haze. Physically we expect that this scattering behaves very similarly to the diffraction patterns due to minute circular

apertures, and the aggregate volume haze is due to the noncoherent superposition of the intensity patterns of point scatters at random locations, with random sizes. Furthermore, an initially smooth surface of a windshield becomes pitted and scratched with use. This gives rise to surface haze. We expect that the light scattering from scratches is similar to the diffraction patterns from "thin slits". Of course these line sources are again of random depth, width, orientation, and location. When an intense beam of light illuminates, in a darkened room, a section of the windshield, the "slits" due to surface deteriorations are not straight lines, but form parts of concentric circles within the area illuminated by the beam. The surface haze generated by both scratches and pits will eventually increase to such a level that it interferes with a pilot's ability to effectively operate the aircraft. Determining that level will be the ultimate aim of this project.

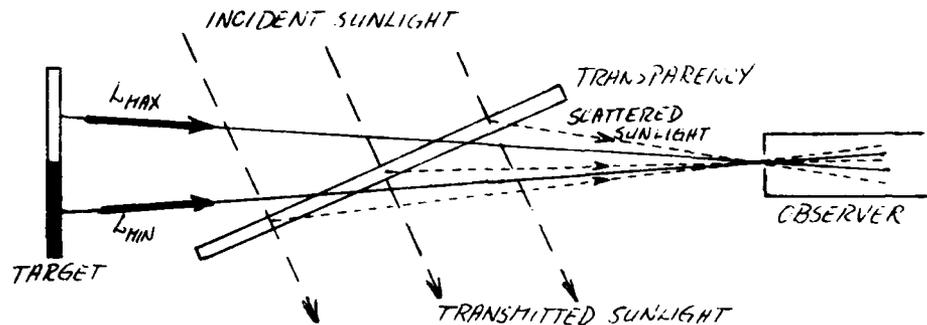
One possible way to describe the contrast (visibility) of a simple target composed of a black area adjacent to a white one is by the formula:

$$C = \frac{L_{\max} - L_{\min}}{L_{\max} + L_{\min}}$$

where L_{\max} is the luminance due to the white portion of the target and L_{\min} is the luminance due to the black portion of the target. Unfortunately this formula does not consider background illumination and thus neglects one very important aspect of the problem. To compensate, experimental conditions have to be arranged so that L_{\max} and L_{\min} conform to realistic lighting conditions encountered by pilots. It will be assumed that no degradation of the target occurs due to scattering by the atmosphere.



When a windshield is placed in the optical path between the target and the observer, the most important processes that can diminish the visibility of the target are (1) Not all the light emitted by target is transmitted by the windscreen, but some of the light is scattered, reflected and absorbed by the transparency. The entire process is characterized by t , the transmission coefficient of the windshield. The reflected portion is simply due to the difference in the index of refraction between the atmosphere and the windshield; but the scattering and absorption can be reduced by improving manufacturing techniques. (2a) A veiling luminance is introduced as the windscreen will scatter direct sunlight and other daytime illumination incident on it into the observer's eye, (2b) and primarily at night, a portion of illumination originating on the observer's side of the windshield is reflected back to his eye.

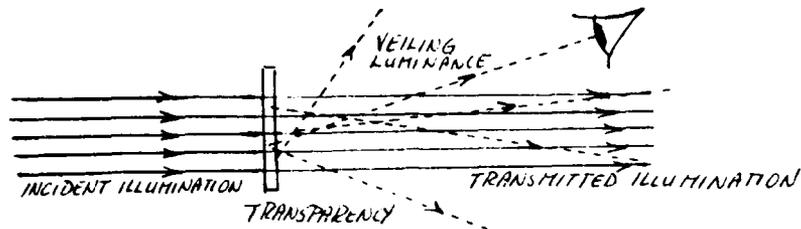


The definition for the contrast now becomes

$$C' = \frac{(tL_{\max} + \text{veiling luminance}) - (tL_{\min} + \text{veiling luminance})}{(tL_{\max} + \text{veiling luminance}) + (tL_{\min} + \text{veiling luminance})}$$

$$C' = \frac{tL_{\max} - tL_{\min}}{tL_{\max} + tL_{\min} + 2(\text{veiling luminance})} = \frac{L_{\max} - L_{\min}}{L_{\max} + L_{\min} + 2 \frac{k\phi I_{\text{inc}}}{t}}$$

Assuming that the veiling luminance is due to scattered sunlight only, we can write the veiling luminance as that fraction of the incident light which is scattered at the angle ϕ by the transparency.



Reflection may be treated identically, but will not be considered. By determining k_{ϕ} , the fraction of the incident beam scattered at a given angle, one obtains a direct measure of the haze, which in turn can be used to predict the loss in visibility.

II. THE EXPERIMENTAL DETERMINATION OF k_{ϕ}

Standard photometric light scattering techniques were used to determine k_{ϕ} . A collimated beam of white light illuminates a small area of the sample. To determine what fraction of the light in the incident beam has been scattered into a solid angle ω in a direction (ϕ, θ) , a luminance meter with a 1° acceptance angle and a spectral response corresponding to that of the standard human eye is placed, always at the same radial distance, at the angular position of (ϕ, θ) with respect to the illuminated area of the sample. The luminance in foot-lamberts may be read directly on the display of the luminance meter. The illumination incident on the sample is obtained by replacing the sample with a lambertian surface (e.g. BaSO_4 screen) which reflects the same amount of light into all directions. The luminance meter aimed at the lambertian surface will then directly record the incident illumination in foot-candles. Theoretically, for the measurement of the illumination the position of the light meter relative to the BaSO_4 screen is arbitrary, if the incident beam has a uniform illumination across its entire surface. From these two direct measurements the ratio of luminance to illumination is computed, which is k_{ϕ} .

All of the data in this study were taken with the polar angle θ equal to 90° , i.e. in a plane parallel to the floor. The incident

beam was usually normal to the illuminated area, and unless specified otherwise, this may be assumed as the experimental arrangement. However in actual use the angle between the incident beam and its surface may be far from 90° .

Since the distribution of light scattered by "points" and "slits" of random location and size in the Fraunhofer approximation is a non-coherent superposition of $\left(\frac{J_1(x)}{x}\right)^2$ and $\left(\frac{\sin x}{x}\right)^2$ functions, both of which are strongly peaked in the direction of the beam, it is desirable to obtain data as close to 0° and 180° as practicable. On the other hand, the pilot looks at the windshield at an angle of 68° . Our data were taken between 5° and 60° in the forward direction and between 120° and 175° in the backward direction, on both sides of the light beam.

A. THE LIGHT SOURCE

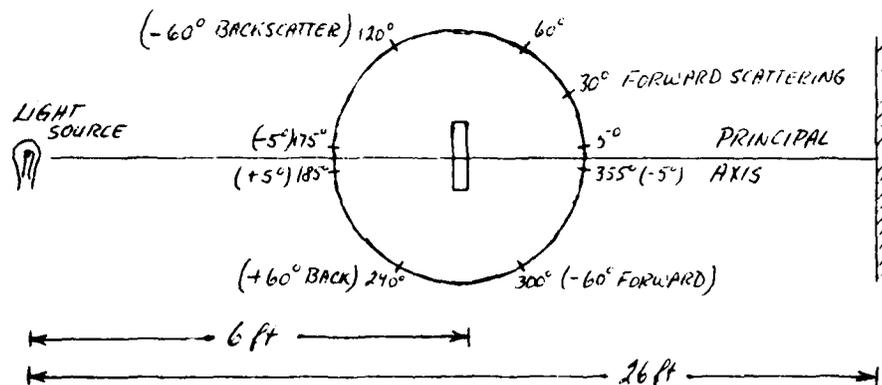
On short notice we found it difficult to obtain an intense, well collimated, white light source, whose illumination is constant across its entire beam. Several alternatives have been examined but we finally settled on a slide projector with a 500 watt lamp, 7 inch focal length lens, into which a slide that has a circular aperture of 3.5 mm at its center was inserted. The beam was well collimated, its divergence being less than 1° . With the intensity switch in the Lo position, a 350 ft-cd beam at the center is obtained. The diameter of the illuminated sample area was 6 cm. This is adequate for most measurements, however backscattering luminances, when the sample was inclined at 40° and larger, were less than 1 ft-lambert, thereby yielding only 2 digits on the 3 digit printout of the luminance meter. The major drawback is the nonuniform illumination over the area of the beam. At small scattering angles the luminance meter at a radial distance of 1m will collect light emitted from a roughly circular area of the sample with a diameter of 1.7 cm, but at large angles, e.g. 60° , it collects the light emitted by an elliptical area with major axis of 3.5 cm, minor axis 1.7 cm. Unless the illumination is uniform, the amount of light emitted by the additional area will not contribute

as much as that from the center, producing a fall off. This, of course, assumes that the same spot can be located for each measurement. To be able to compensate for the fall off we have measured the illumination of a BaSO_4 screen as a function of scattering angle, which should be independent of the scattering angle. The assumption is then made, that the luminance from a sample exhibits the same angular dependence as the illumination. A sheet of aluminum painted flat white gave results similar to the BaSO_4 .

B. ALIGNMENT PROCEDURE

At the approximate center of the laboratory a vertical axis has been established with a plumb bob to determine 2 of the 3 coordinates for the location of the sample. The plumb bob is suspended from the ceiling and can easily be raised and lowered. The sample height was chosen (an inconveniently low) 1.23m above the floor. A second point surrounded by concentric circles is marked directly on the wall at the far end of the room, also 1.23m above the floor. The light source is then positioned at the opposite end of the room so that the center of the beam passes simultaneously through the center of the sample position and the point on the wall. To aid with that, a small poster board with a 1.3cm diameter circular hole surrounded by concentric circles precisely located the sample position.

A 1.000m radius, circular scale has been marked on the floor with



its center underneath the sample position and 0° and 180° coinciding with points 1.23m underneath the principal axis from the light source through the sample position to the far end of the room.

Windshield materials of area less than 1 ft^2 were mounted in a vice that adjusts to any angle about the vertical; appropriate height adjustments could be made by a lab jack on which the vice rested. To assure an incident beam normal to the illuminated surface, the reflection of the beam was centered onto the light source. For curved samples this reflection is broadened vertically or horizontally to about 10° , depending on whether the cone axis is oriented horizontally or vertically. The sample had to be rotated about an axis parallel to the beam to accomplish that. A reasonably sensitive method for centering vertically extended reflections is to place a black screen with an aperture slightly larger than the beam diameter in front of the source. A horizontal index pointer at 1.23m above the floor is introduced into the beam between the source and the sample, and the shadow of its reflection is centered on the aperture. To center horizontally extended reflections onto the light source it was found convenient to use another plumb bob, which had been suspended from the ceiling above the principal axis between the light source and the sample position. Similarly, the shadow of the plumb bob string's reflection was centered on the aperture in the screen covering the light source.

This additional plumb bob was necessary anyhow to properly position complete windshields to obtain light scattering data at various points on the windshield. Because of their weight, the windshields were suspended from the boom of a hydraulic lift. Whenever the boom interfered with the primary plumb bob we used the secondary plumb line and measured back to the sample position. A nylon rope was attached to the holes in the metal frame of the windshield in such a way that they hung as vertically as possible when suspended freely. Whether the cylinder axis was horizontal or vertical the windshield was adjusted normal to the incident beam by repositioning 25 lb bags filled with shot, which were tied to the frame. Additional care had to be taken to move the crane and shot

bags toward or away from the light source until the center of the illuminated area was directly underneath the primary plumb bob. To rotate the windshield about an axis parallel to the beam, the point on the nylon rope in contact with the hook on the boom was readjusted until the reflection was horizontal or vertical. The string was then secured with masking tape in that position. Because the hydraulic valve of the crane leaked, the height of the boom had to be held constant by clamping an aluminum channel between the boom and the floor. At times the sagging of the boom shifted the sample slightly during a set of measurements. Half an hour settling time was allowed before beginning to take data.

C. DATA TAKING

One set of measurements consists of positioning the luminance meter at ± 5 , ± 7.5 , ± 10 , ± 12.5 , ± 15 , ± 20 , ± 30 , ± 40 , ± 50 , and $\pm 60^\circ$ for both forward and back scattering angles. The radial distance between the center of the sample and the luminance meter was held constant at 1.00m. For this a third plumb bob was suspended to mark the front surface of the objective lens of the Minolta luminance meter. The meter itself has been mounted on a reasonably sturdy photographer's tripod. The plumb bob just missed making contact with the floor at some positions, at other positions it hung several mm above it. This unevenness of the floor caused the luminance meter to sample regions at various heights within the illuminated area of the windshield. As the illumination across the beam was not uniform this built a small systematic error into the data.

The positioning procedure is tedious, since we attempted to keep the plumb bob to within 2mm of the marks on the floor and simultaneously aim at the center of the illuminated area. To improve on the accuracy in locating a 1.7cm diameter measuring field at the center of a 6cm diameter illuminated area, data were taken by first aiming at the shadow of the string of the secondary plumb bob. Due to the thickness of the windshields, two shadows were expected, however, the shadow on the side of the windshield away from the light source was

almost impossible to detect at angles much larger than 20° . Aiming only at the dominant shadow of the plumb line, instead of the middle between the two, causes too much of a parallax error. This rather cumbersome procedure of introducing the plumb bob, stabilizing it, positioning and aiming the luminance meter, removing the plumb bob from the beam, extinguishing all room lights, making and recording the observations, required at little more than two hours for the 40 data points per illuminated area. The results obtained, based solely on the experimenter's judgement as to where the center was, and those obtained by accurately aiming at the shadows were not all that different. Nevertheless, accurate aiming and positioning of the photometer is absolutely necessary. For instance, at the 5° position, with a high quality quartz flat as the sample, placing the photometer at the 4.5° mark and aiming $.8^\circ$ to the right of the center of the sample produces a result which differs by 150% from the 5.5° position and aiming $.8^\circ$ to the left of the sample's center.

The illumination has been recorded at least at the beginning and the end of each set of measurements, but more often it was measured four times. A 6cm x 6cm BaSO_4 screen was placed directly in front of the illuminated area and the luminance meter was consistently positioned at a radial distance of 1.00m at the 170° position. Although continuous monitoring would have been more desirable, our procedure was sufficient, since the illumination was found to be reasonably constant.

To avoid erroneous readings due to stray light, the room was kept dark, and absorbers and light shields were placed wherever needed. As the results with a single, clean quartz flat indicate, stray light was not a problem.

To measure the loss in visibility of a target when it is viewed through a windshield material, compared to when it is viewed directly, an apparatus was set up outside, with the sun as the light source. Windshield materials of surface area less than 1 ft^2 were mounted in a frame that enabled them to be oriented perpendicular to the incident sunlight. Entire windshields were held by a wooden frame in the same

position as when they are installed on a level airplane. The windshields faced directly into the sun, with the luminance meter at the cycloptic "design eye" position. Black velvet was draped over all sides and the bottom to eliminate stray light and reflections. A black and white target, each area about $8 \times 8 \text{ in}^2$, was located on the ground 11m from the luminance meter. The apparatus was continuously adjusted so that the photometer, target, sun, and boresight (or center of the fragment) were in the same plane.

With the windshield material removed from the holder, the luminance meter was aimed at the black, subsequently at the white area of the target. From the resulting luminances in ft-L the unobstructed contrast C could be computed. The transparency was then interposed between the target and luminance meter and the contrast C' was again measured by determining L_{max} and L_{min} .

To be able to predict C' from the light scattering data collected in the laboratory, the transmission coefficient of the windshield, the illumination at the windshield due to the sun, and the altitude of the sun had to be determined. The laboratory geometry is rotated by 90° about the principal axis compared to the outside geometry. So that the illumination remained constant for a set of data, the measurements were made on a cloudless day.

In a related set of experiments carried out indoors, the windshields were mounted in an adjustable frame that held them in the same position as that in which they are in the installed position on an aircraft. The luminance meter was again fixed at the "design eye" position and looked at the target through (0,0). A 450 watt aircraft landing light, 13 ft. away from the (0,0) position at 40° above the horizontal, constituted our artificial, but stationary sun. An 8 ft. high circular frame covered with black cloth served as background on which 7 white targets could be mounted. When the windshield was rotated, the luminance meter aimed at the targets at -25° , -15° , -5° , 0° , 5° , 15° and 25° . Another scan was then taken with the targets raised, so that the contrasts could be computed. Repeating this procedure with the windshield removed allowed us to determine the

contrasts of the unobstructed targets. When the luminance meter was refocused from the target to the windshield, data could be obtained to predict the contrast loss with the windshield in place compared to the unobstructed view. Two light levels on the targets were utilized.

The transmission coefficients for various geometrical arrangements were measured by first measuring the luminance of a diffuse light source for a particular geometrical arrangement and then measuring the luminance of the same source again, but with the windshield between the meter and the light source.

III. MEASUREMENTS COMPLETED

During the 10 week period we measured the angular dependence of the amount of light scattering by

(A) a 1 ft² "commercial grade" sample of windshield material. Care had been taken to preserve it in the same "new" condition in which the manufacturer had furnished it. It had a yellow tint, was laminated, nearly flat, and .75 inches thick. Surface haze was assumed to be minimal.

(B) a 1 ft² "select grade" sample of windshield material. Care had also been taken to preserve it in the same condition in which the manufacturer had furnished it. It had a yellow tint, its radius of curvature was 78cm and it was .625 inches thick (1/8 inch less than the companion commercial grade). Surface haze was assumed to be minimal.

(C) three "coupons" labeled A, C and D, approximately 5 x 5 in² which had been cut from discarded F-111 windshields. Sample A was greyish, with a radius of curvature of 92cm, samples C and D had a yellow tint. The visible surface scratches on all three samples were predominantly parallel to the cone axis.

(D) three F-111 windshields removed from airplanes for reasons other than haze. Two of these windshields were rejected for unacceptable levels of birefringence, the other for distortion.

(E) three F-111 windshields which had been removed from airplanes primarily for being excessively hazy.

Measurements on windshield fragments (items A, B and C above) were taken in their approximate centers. On the windshields rejected for reasons other than haze the measurements were taken on areas located in the pattern of the dots of the five side of dice, whereas on the windshields rejected for haze the pattern was that of a cross, the center area coinciding with the boresight. For the commercial and select grade sample materials, as well as sample A, data were obtained with the cylinder axis both horizontal and vertical, and the center of curvature both toward and away from the light source. For samples C and D, and the three hazy windshields the center of curvature was always away from the light source, but the cylinder axis was oriented both horizontally and vertically. Finally, for the measurements on the three windshields rejected for reasons other than haze, the cone axis was always kept vertical and the center of curvature away from the light source.

Visibility measurements of a target with the sun as the light source were also taken on the three windshields rejected for haze, the commercial and select grade fragments, and the three coupons. A related set of measurements on eight new F-111 windshields was performed indoors with a 450 watt aircraft landing light at 40° as the light source. These data primarily establish a baseline so that the extent of the surface deterioration may be quantified after the windshields have been in service for a certain amount of time. Graphs of most available results are included in the appendix.

IV. CONCLUSIONS

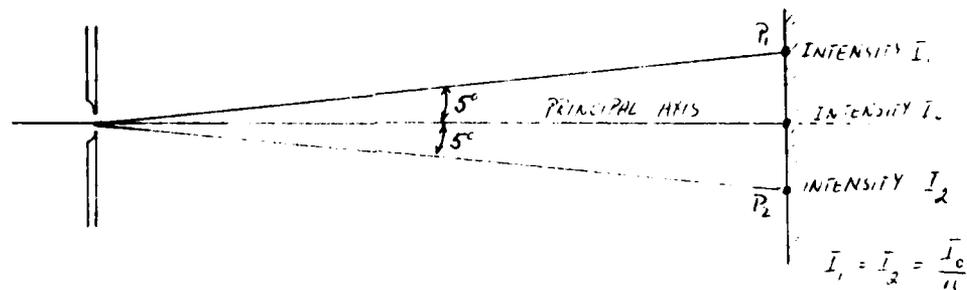
1. The light scattered by bird impact resistant windshield materials is peaked very sharply in directions parallel and antiparallel to a light beam incident normally on the material. At small angles the scattering from "hazed" materials is 3-6 times larger

than that from undegraded materials.

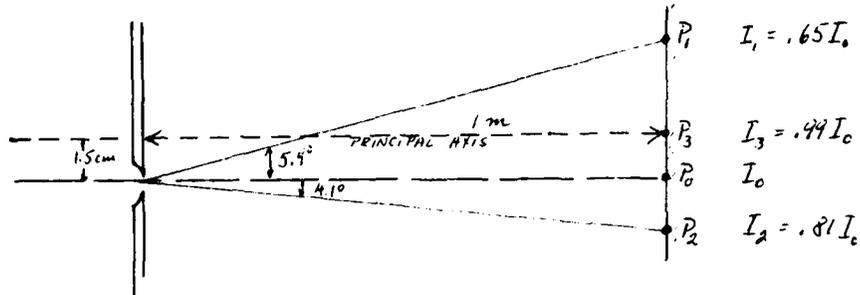
A mathematical expression which describes the relation of the forward scattering as a function of angle (with a correlation coefficient larger than .98) is $L = a;^{-n}$ ($1.7 < n < 2$). However, both the coefficient and exponent varied for different samples and for different areas on the same sample. Worse yet, the same function did not fit back scattering data as well. The correlation coefficient dropped to about .8 ($r^2 = .7$). Curve fitting backscattering data to the mathematical expression $L = a + b;^{-1}$ produced results with r^2 also equal to .7. This is somewhat of a puzzle since we would expect the same angular distribution for light diffracted from either a reflection or transmission grating. It will be an interesting exercise to noncoherently add $(\frac{J_1(x)}{x})^2$ and $(\frac{J_0(x)}{x})^2$ functions for point and "circular line" scatterers of random location and size to determine if the assumed model is realistic.

2. Only for normal incidence is the backscattering reasonably similar to the forward scattering. If the normal to the plane is at angles larger than 10° above the incident beam the luminance due to backscattered light is constant with scattering angle. We have no plausible explanation.

3. The data we have obtained were frequently plagued by an asymmetry between the amount of light scattered into a solid angle at $(+\phi, 0)$ compared to the same solid angle at $(-\phi, 0)$. It is not difficult to imagine processes that would cause that. For instance the nonuniformities on one side of the illuminated area could scatter more light than those on the other. Let us construct a model which attributes all the excess scatter to single slit. If this slit were

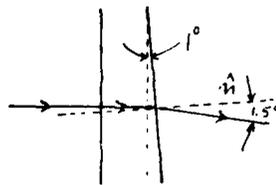


located on the principal axis we expect equal intensities at the observation points P_1 and P_2 . However, if the slit in the sample area is moved 1.5cm to one side of the principal axis, the luminances at P_1 and P_2 will no longer be equal. Choosing $\lambda = 5000\text{\AA}$, the slit width $d = 0.1\text{mm}$, and the distance between the slit and observation plane equal to 1.00m then at P_1 the intensity will be .65 of the



intensity at the center, whereas at P_2 the intensity will be .81 of the intensity at the center.

Wedginess in the windshield will produce a similar effect. For a



local apex angle of 1° the diffraction pattern will be deviated by 1.5° with respect to the principal axis if an index of refraction of 1.5 is assumed. It, however, is disturbing that the data were predominantly too large on the same side of the apparatus.

We are confident that for uniformly scattering samples with parallel surfaces the reproducibility of the data is about 10%, even through the asymmetry between right and left hand scattering was as high as 35% on occasion. Considering that these measurements are the first of their kind and our equipment consisted of a luminance meter, string, duct tape, a slide projector, a hydraulic crane, plumb bobs, 25 lb shot bags, a vice, an index pointer, lab jacks, and angle finders we have taken the accuracy to the absolute limit of the

available apparatus.

4. It makes no difference in the amount of scattered light whether the center of curvature of the windshield points toward or away from the light source.

5. Almost all windshield materials we examined that had been in service on aircraft exhibit much more scattering, in an observation plane parallel to the floor, when the cone axis is perpendicular to the floor compared to the arrangement in which the cone axis parallel to the floor. That is, they seem to behave very much like diffraction gratings with rulings predominantly parallel to the cone axis. On the wall opposite the slide projector one observes in essence the Fraunhofer diffraction pattern due to vertically oriented slits when the cone axis is approximately vertical. That pattern rotates by 90° when the cone axis is oriented horizontally. But how the cone axis is oriented on the F-111 windshields is not entirely clear. For instance, when the windshields are suspended so that the reflection of the incident beam is accurately horizontal, the dominant diffraction line observed on the wall is usually at some (small) angle to the horizontal. In addition, secondary patterns were present. The "rulings" could be due to abrasions by particles hitting the windshield during flight, abrasions due to cleaning, or to stress cracking. On the select and commercial grade material the data are not sensitive enough for us to claim a difference.

6. We have used light scattering data gathered in the laboratory to predict measured visibility losses due to the haze of windshields. Even though the predicted losses are somewhat less than the actually measured losses in visibility, agreement is acceptable and tends to follow the experimental curves. This project merits continuation simply to refine these measurements. They certainly are of interest to pilots. See appendix.

Because in the laboratory the incident beam was normal to the illuminated surface, the surfaces of the fragments were kept normal to the sun's rays and with the passage of time the surface of these windshield materials had to be tilted from the vertical at an angle

equal to the altitude of the sun. We only had the facilities for entire windshields to be mounted in the installed position, consequently only a single measurement per entire windshield was feasible with the sun's rays at 90° to the windshield.

7. The current industry standard to determine haze levels is the Gardner Haze Meter. The data obtained from measurements developed during this project and measurements we took with the Gardner Haze Meter follow the same trends. Our methods simply furnish much more detailed information and more adequately suit the geometrical and physical requirements of performing measurements on airplane windshields.

V. RECOMMENDATIONS

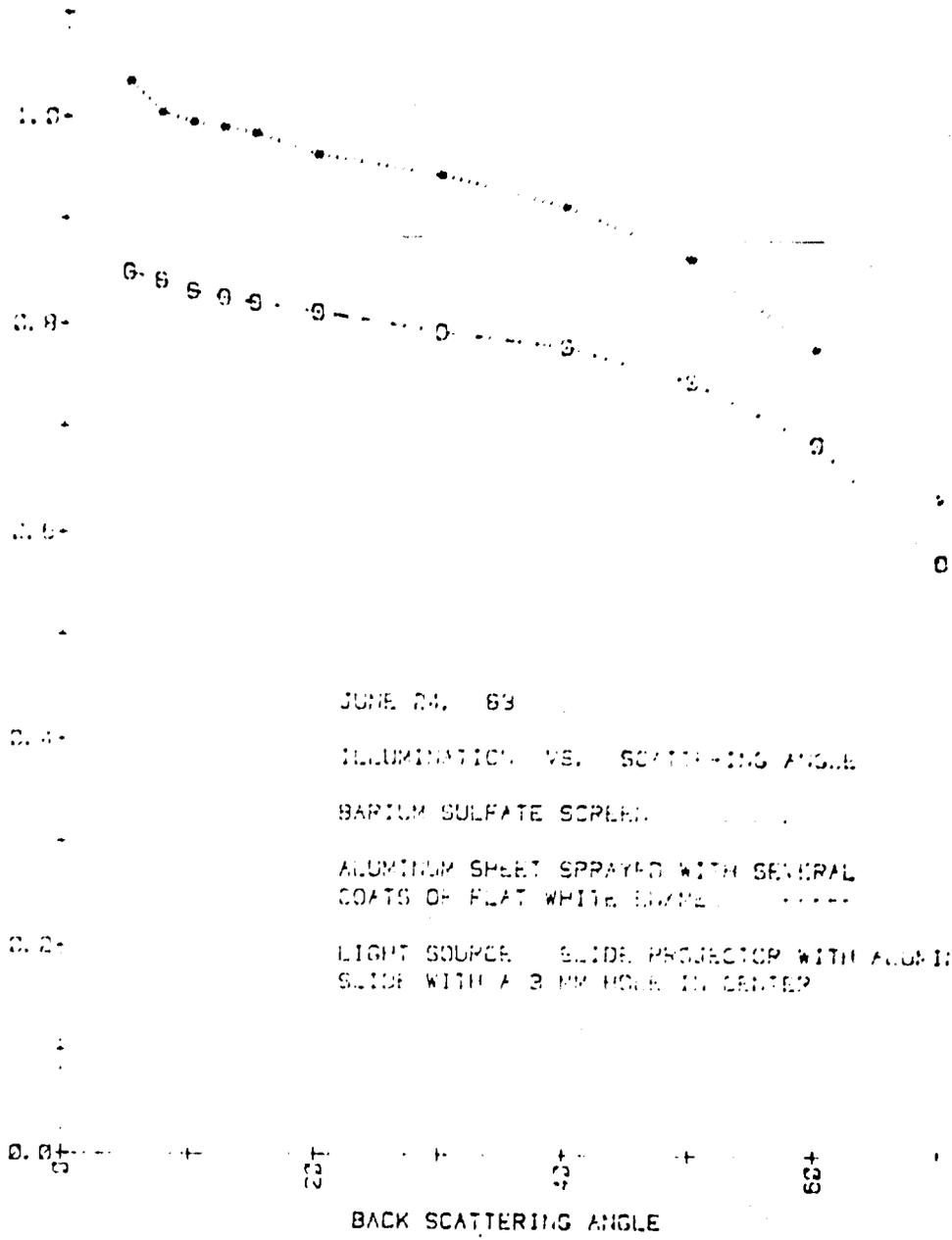
1. Continue and extend the study initiated this summer. A description of a sturdier, more convenient, and automated apparatus will be included in the migrant application the author intends to submit. Now that the measurement techniques are developed and general trends in the data are known, the design of that apparatus should be rather reliable and final. Among other things we suggest that the luminance meter be mounted on a 20-24 inch machinist's index table and the windshields be suspended from a lead screw mounted in the apex of a metal spider. Solving the problems outlined in the conclusion section of this report will take some time. It will be worthwhile to examine the results of buffing the windshields with various lapping materials and waxes or to examine the effect of different types of light sources with different intensities.

An approach different from ours would be to replace the luminance meter with the 1° acceptance angle by a photomultiplier that examines the illuminated sample area through a suitable pinhole. This way the light scattered from the entire illuminated area is gathered, but shielding becomes much more important. We chose our method since that corresponds more closely to the conditions under which pilots make

observations.

2. It is our opinion that to construct a portable haze meter based on reflected, scattered light from a windshield, a narrow pencil of light should be incident normally to the windshield. The backscattered light can then be collected simultaneously by 4-6 detectors located at some small angle, e.g. 20° , with respect to the incident beam. The detectors will have to be shielded properly from stray light. What the acceptable levels of scattered luminances, with reference to a standard are, remains to be determined.

SCATTERED ILLUMINATION



JUNE 24, 69

ILLUMINATION VS. SCATTERING ANGLE

BARIUM SULFATE SCREEN

ALUMINUM SHEET SPRAYED WITH SEVERAL
COATS OF FLAT WHITE EMPEL

LIGHT SOURCE SLIDE PROJECTOR WITH ALUMINUM
SLIDE WITH A 3 MM HOLE IN CENTER

June 24, 83

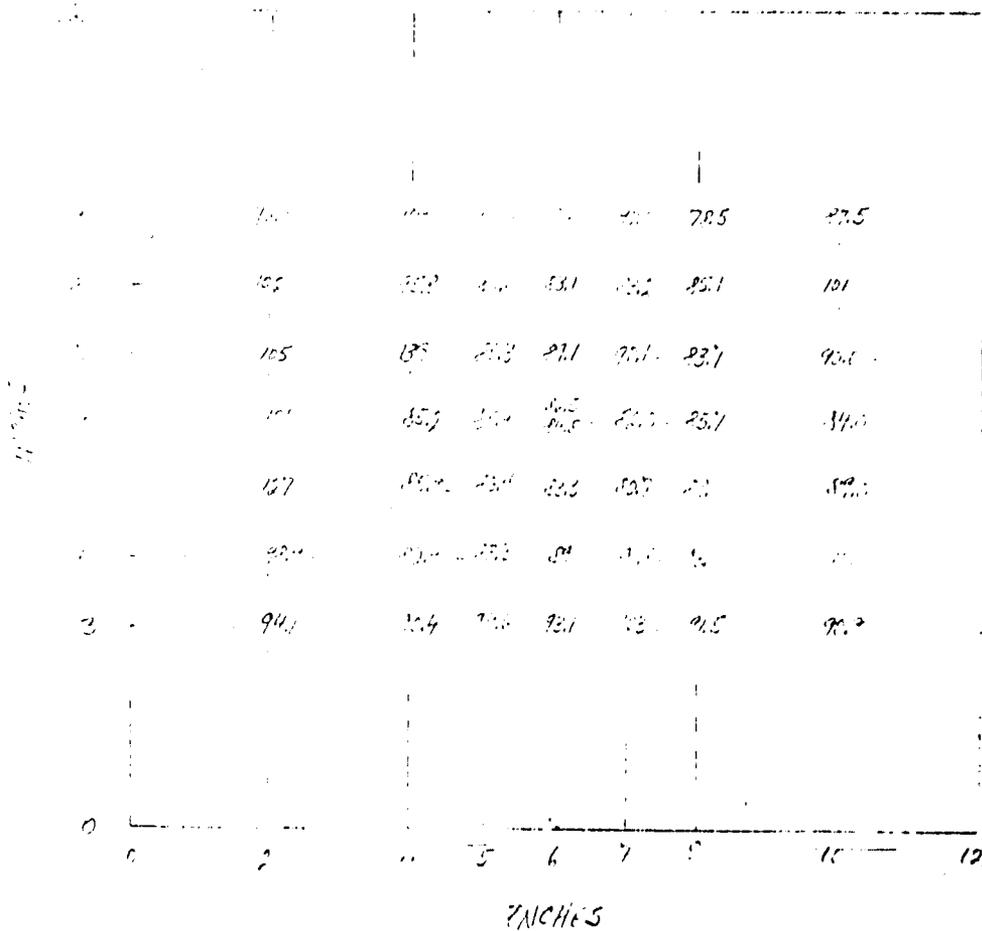
SCATTERED POINTS AS A FUNCTION OF SAMPLE POSITION
 INCIDENT RADIATION: 1000 Hz

SCATTERED RADIATION: 1000 Hz

120000 Hz

SAMPLE AVERAGE: (92.21, 11.4) Hz

(92.21, 12.9%)



JUNE 24, 63

SPECTRAL INTENSITY AS A FUNCTION OF SAMPLE POSITION

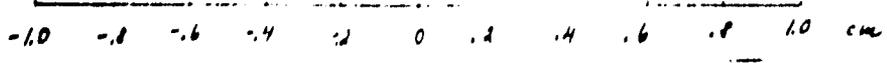
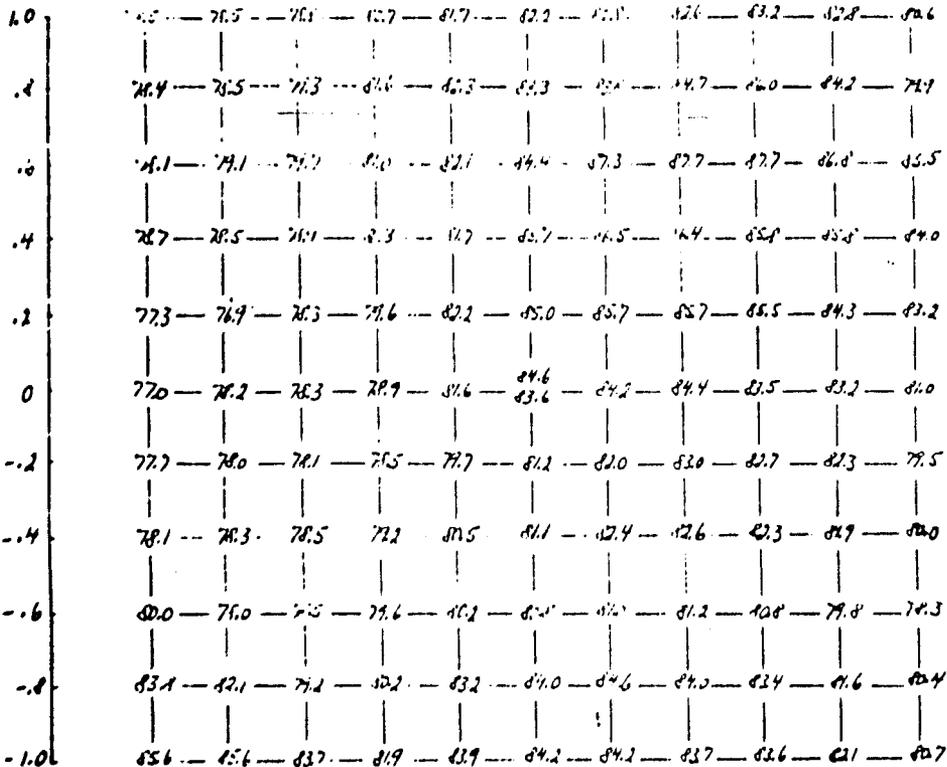
INCIDENT ILLUMINATION: 450 ft-cd

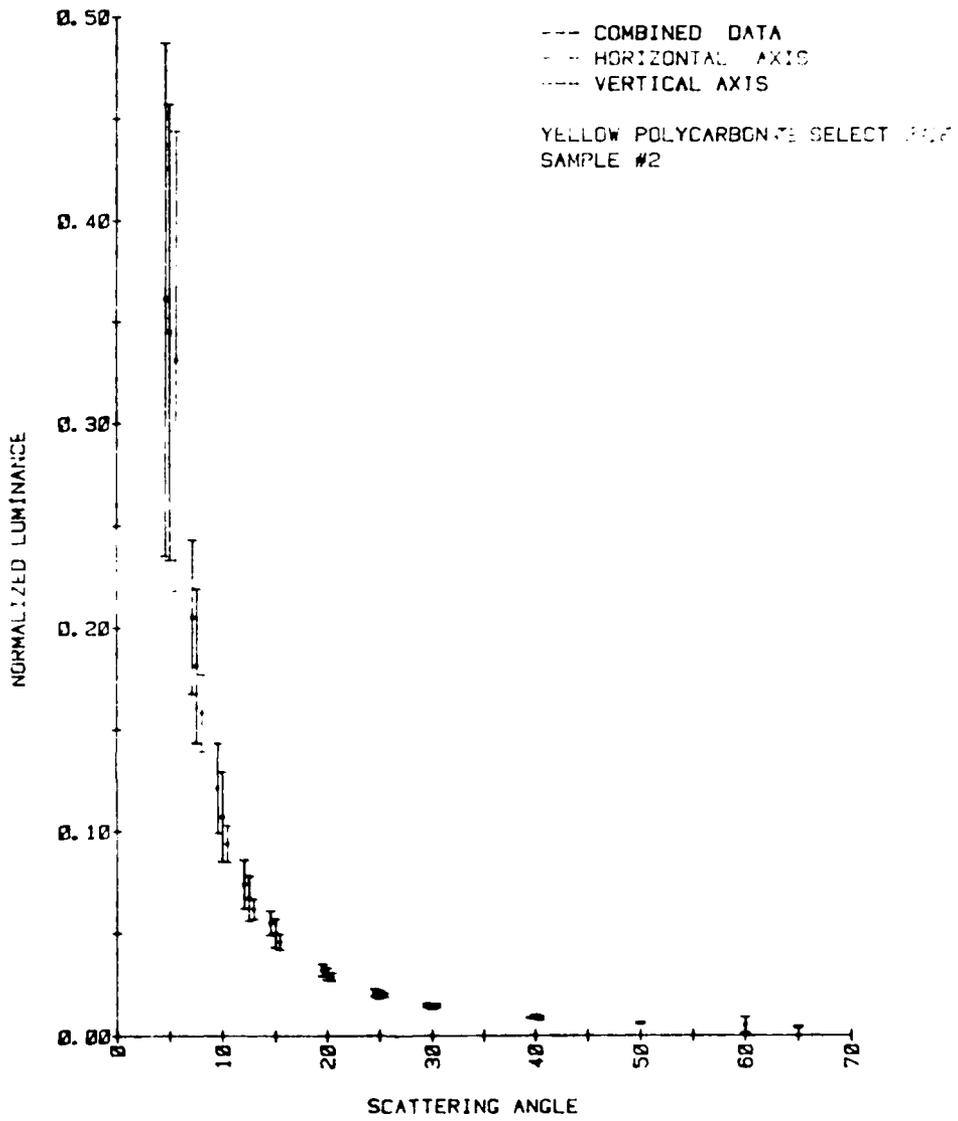
LUMINOUS INTENS. (ACCEPTANCE ANGLE 10) FILMED AT +5

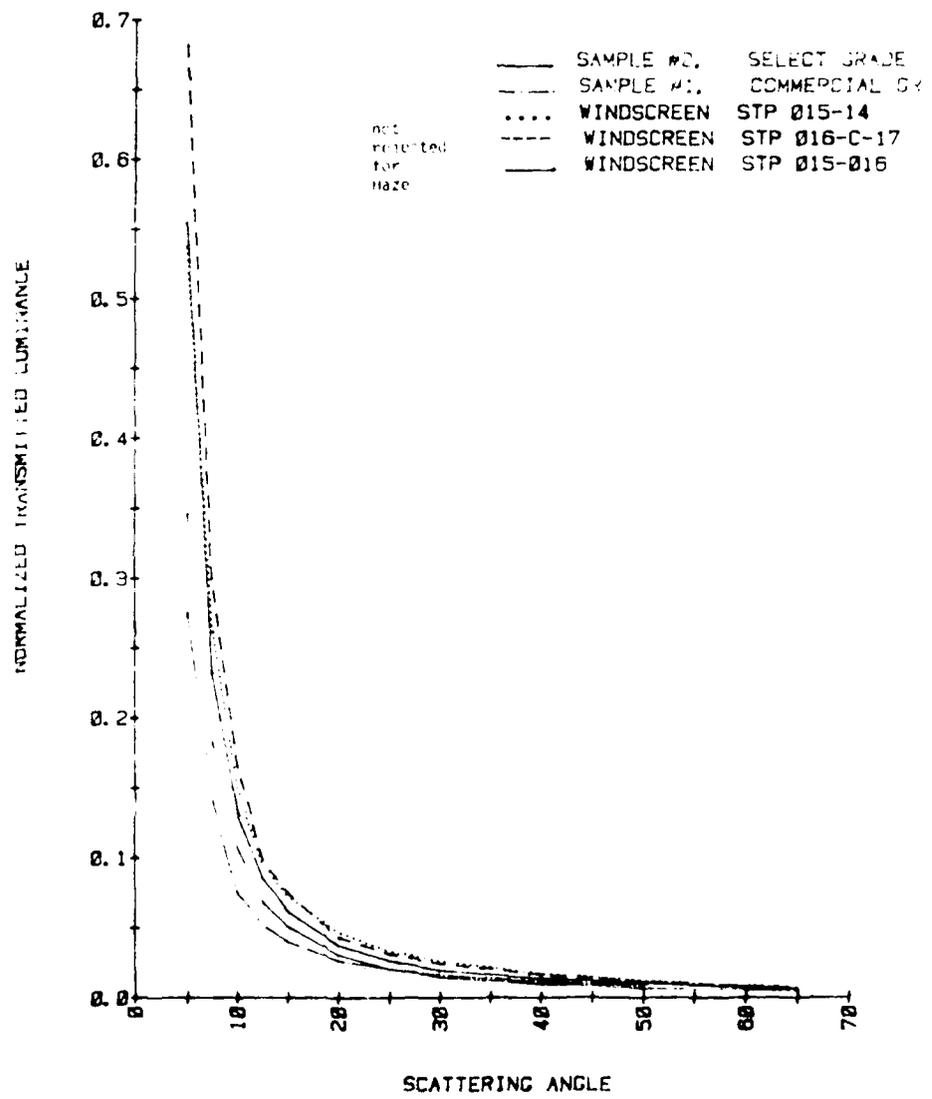
RECORDING INSTRUMENT: 2000

SAMPLE NUMBER: 108 ± 2.7 (100 ± 3%)

cm





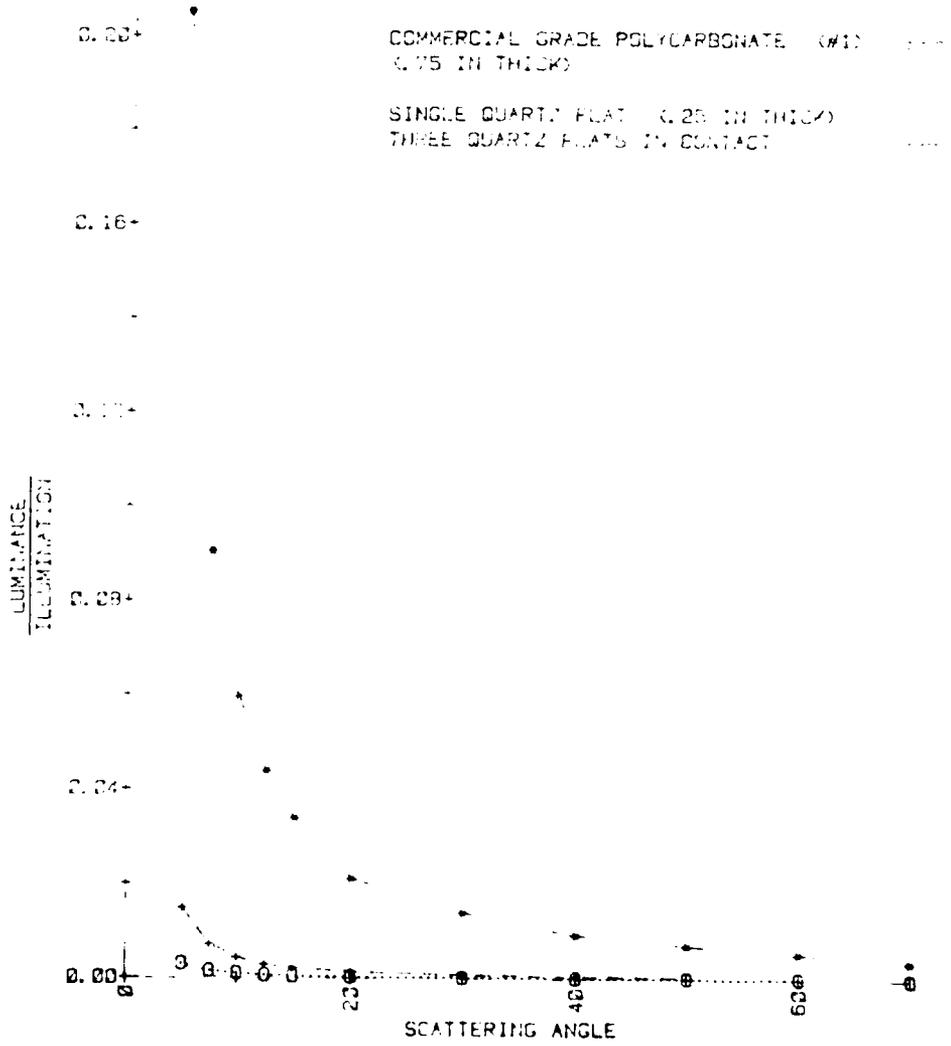


JUNE 24, 83

FORWARD SCATTERING VS. ANGLE

COMMERCIAL GRADE POLYCARBONATE (WIDENED)
(.075 IN THICK)

SINGLE QUARTZ PLAT (.25 IN THICK)
THREE QUARTZ PLATS IN CONTACT

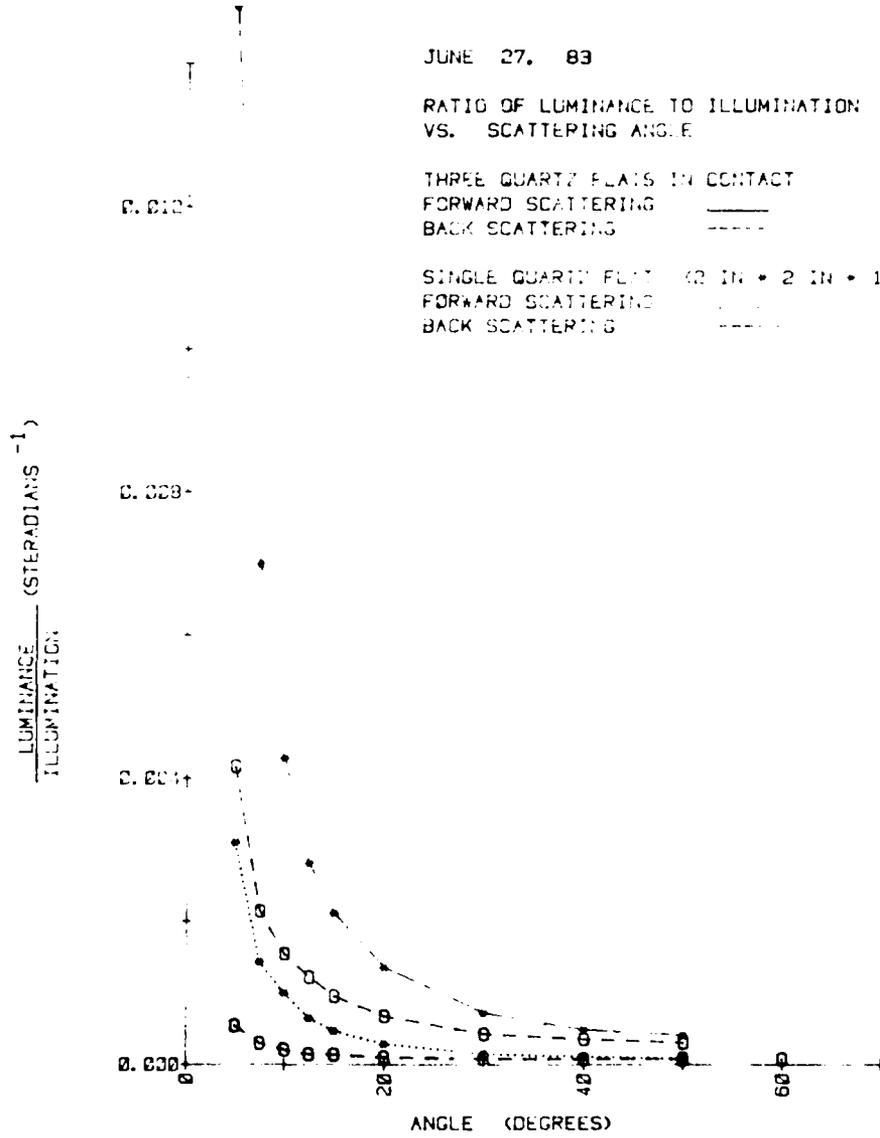


JUNE 27, 83

RATIO OF LUMINANCE TO ILLUMINATION
VS. SCATTERING ANGLE

THREE QUARTZ PLATS IN CONTACT
FORWARD SCATTERING ———
BACK SCATTERING - - - - -

SINGLE QUARTZ FLAT 40 IN • 2 IN • 1/4 IN
FORWARD SCATTERING ———
BACK SCATTERING - - - - -



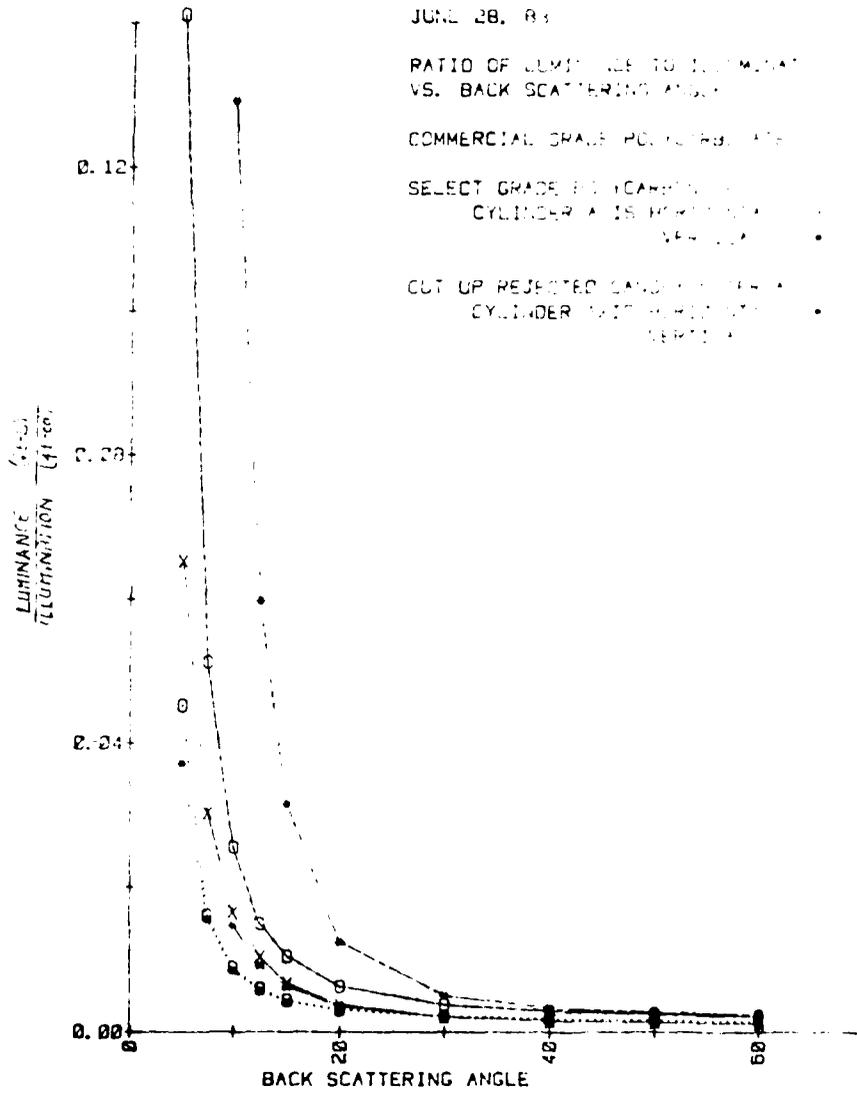
JUNE 28, 1943

RATIO OF LUMINANCE TO ILLUMINATION
VS. BACK SCATTERING ANGLE

COMMERCIAL GRADE POLYESTER FIBER •

SELECT GRADE POLYESTER FIBER
CYLINDER AXIS HORIZONTAL ○
CYLINDER AXIS VERTICAL x

CUT UP REJECTED POLYESTER FIBER
CYLINDER AXIS HORIZONTAL ●
CYLINDER AXIS VERTICAL ▲



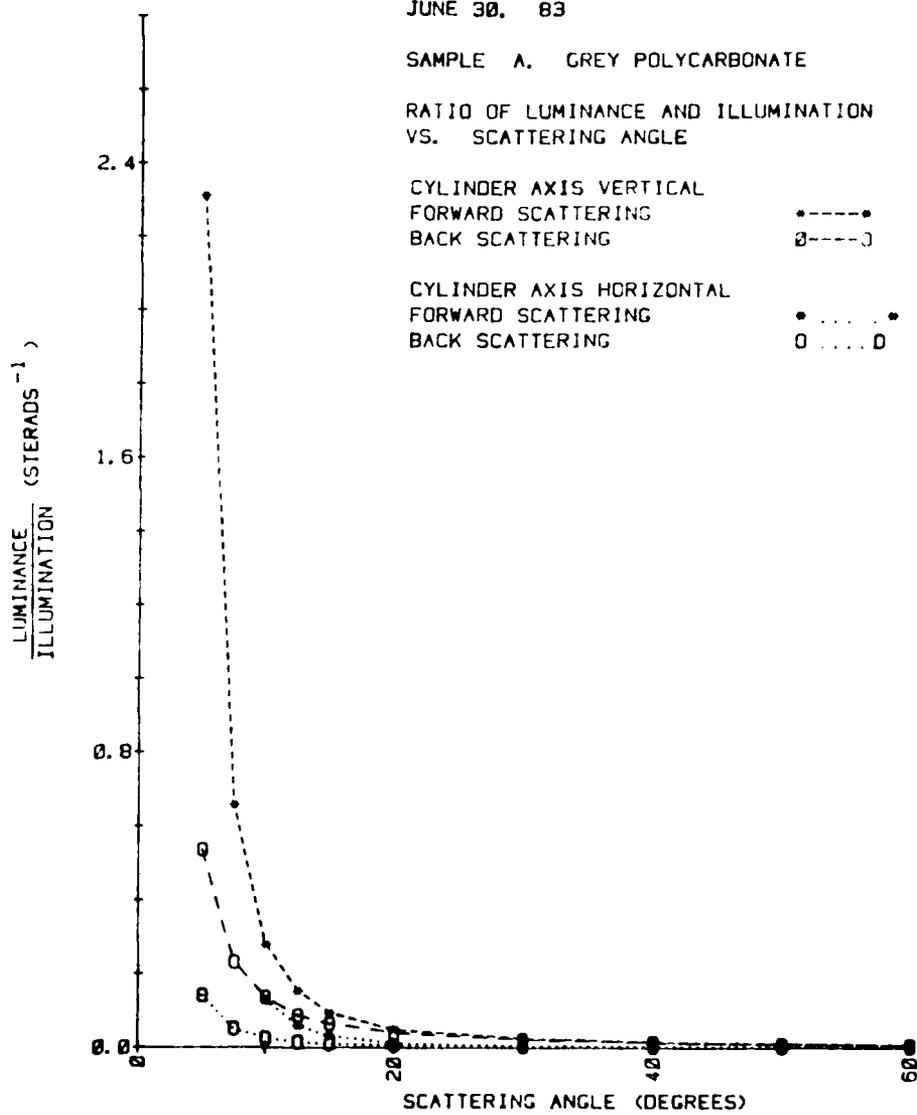
JUNE 30, 83

SAMPLE A. GREY POLYCARBONATE

RATIO OF LUMINANCE AND ILLUMINATION
VS. SCATTERING ANGLE

CYLINDER AXIS VERTICAL
FORWARD SCATTERING •-----•
BACK SCATTERING ◊-----◊

CYLINDER AXIS HORIZONTAL
FORWARD SCATTERING •.....•
BACK SCATTERING ◊.....◊



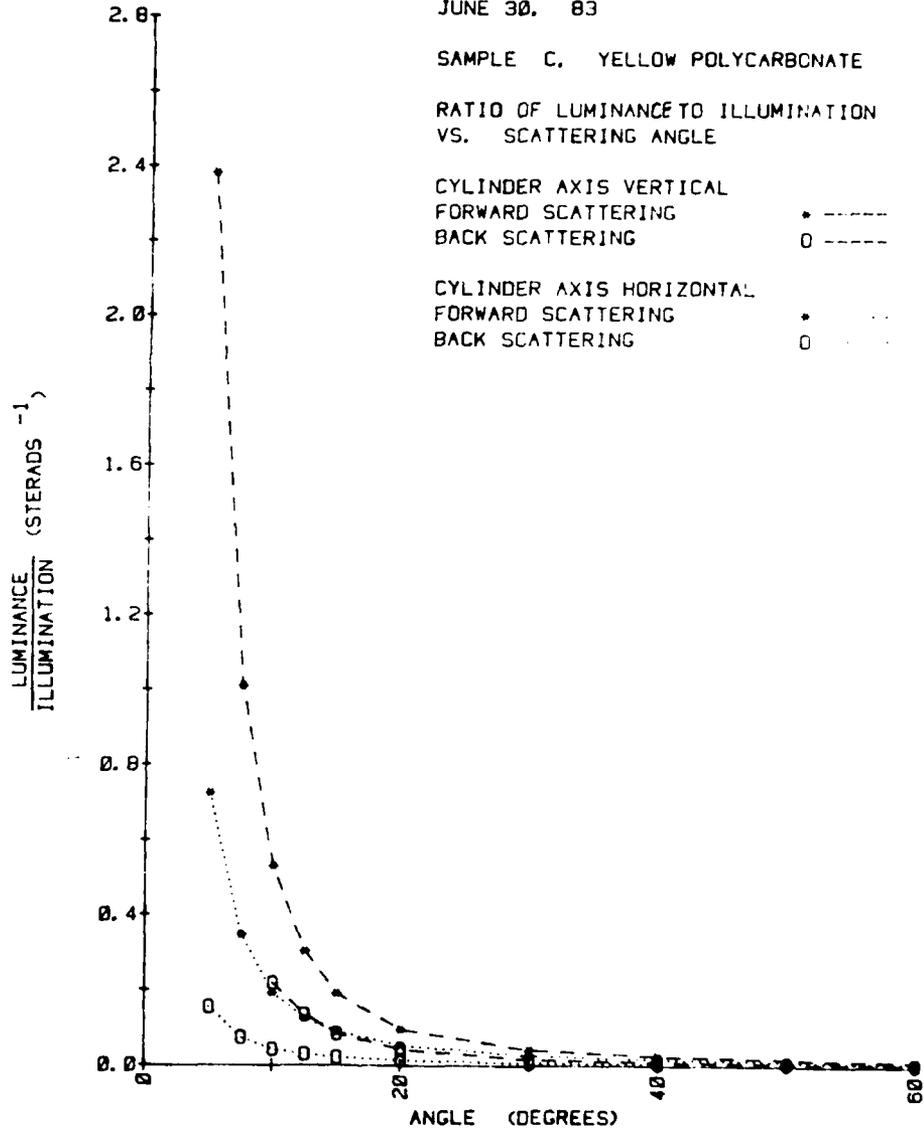
JUNE 30, 83

SAMPLE C. YELLOW POLYCARBONATE

RATIO OF LUMINANCE TO ILLUMINATION
VS. SCATTERING ANGLE

CYLINDER AXIS VERTICAL
FORWARD SCATTERING * - - - -
BACK SCATTERING 0 - - - -

CYLINDER AXIS HORIZONTAL
FORWARD SCATTERING * - - - -
BACK SCATTERING 0 - - - -



JUNE 30. 83

SAMPLE D. YELLOW POLYCARBONATE

RATIO OF LUMINANCE TO ILLUMINATION
VS. SCATTERING ANGLE

CYLINDER AXIS VERTICAL

FORWARD SCATTERING

BACK SCATTERING

•----

o----

CYLINDER AXIS HORIZONTAL

FORWARD SCATTERING

BACK SCATTERING

*-----

o-----

CYLINDER AXIS AT +45 DEGREES LOOKING

TOWARD LIGHT SOURCE

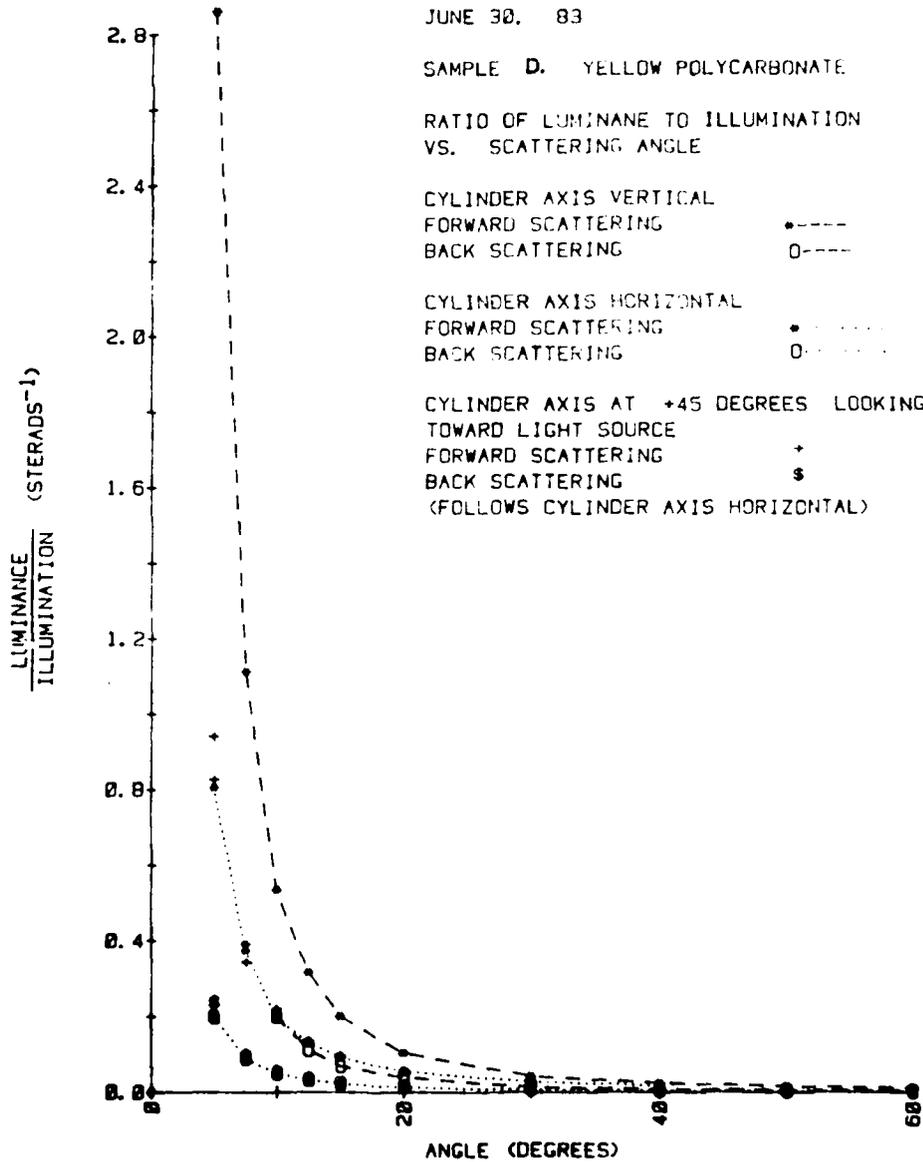
FORWARD SCATTERING

BACK SCATTERING

+

\$

(FOLLOWS CYLINDER AXIS HORIZONTAL)

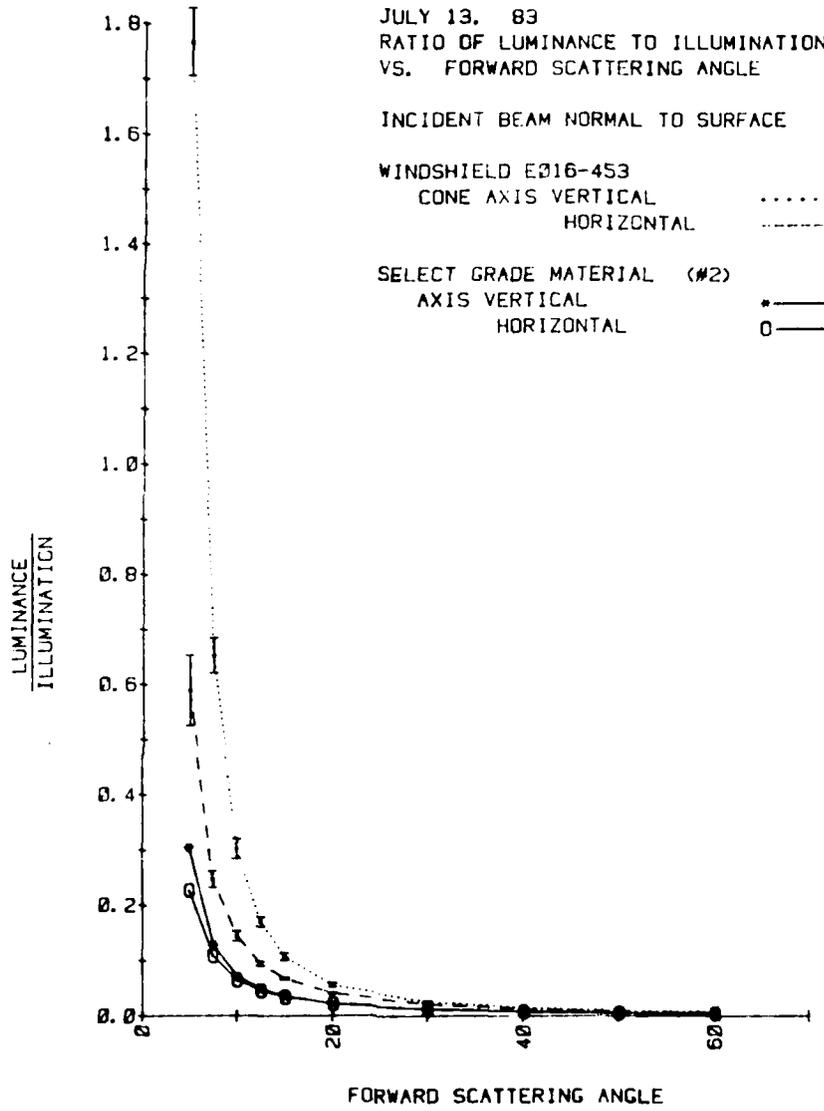


JULY 13, 83
RATIO OF LUMINANCE TO ILLUMINATION
VS. FORWARD SCATTERING ANGLE

INCIDENT BEAM NORMAL TO SURFACE

WINDSHIELD E016-453
CONE AXIS VERTICAL
HORIZONTAL - - - - -

SELECT GRADE MATERIAL (#2)
AXIS VERTICAL * - - -
HORIZONTAL O - - -



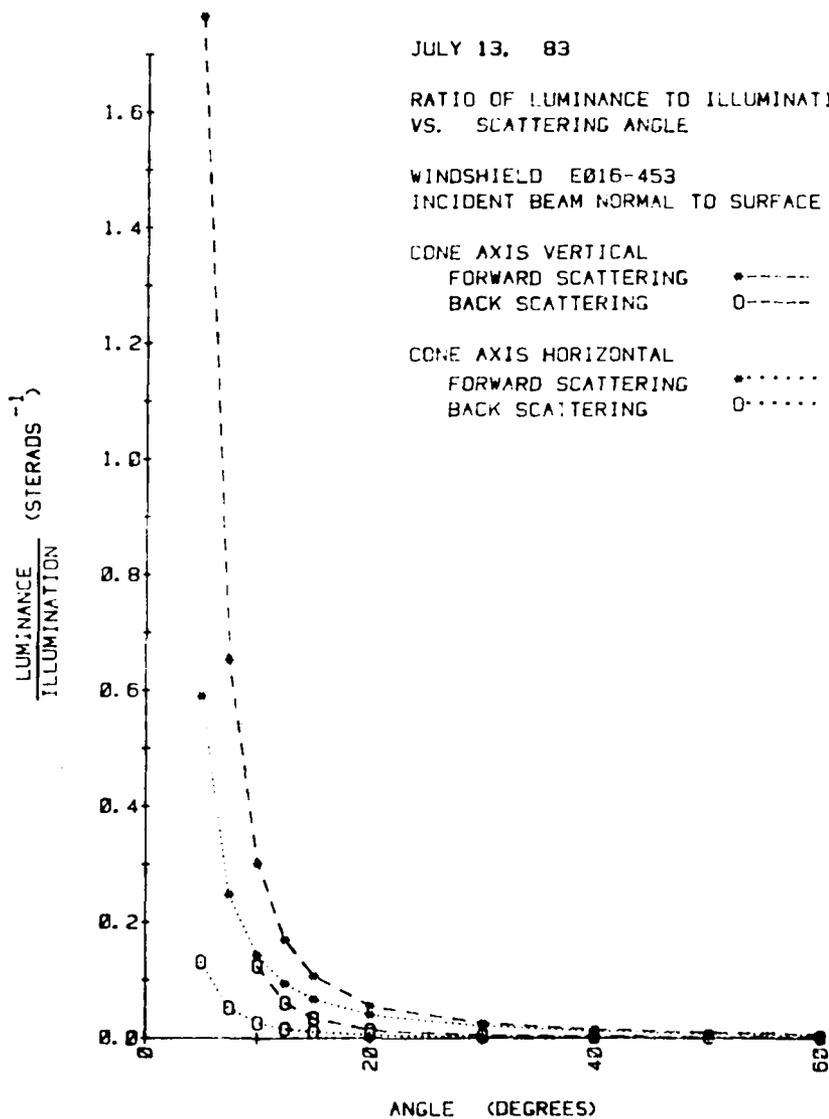
JULY 13, 83

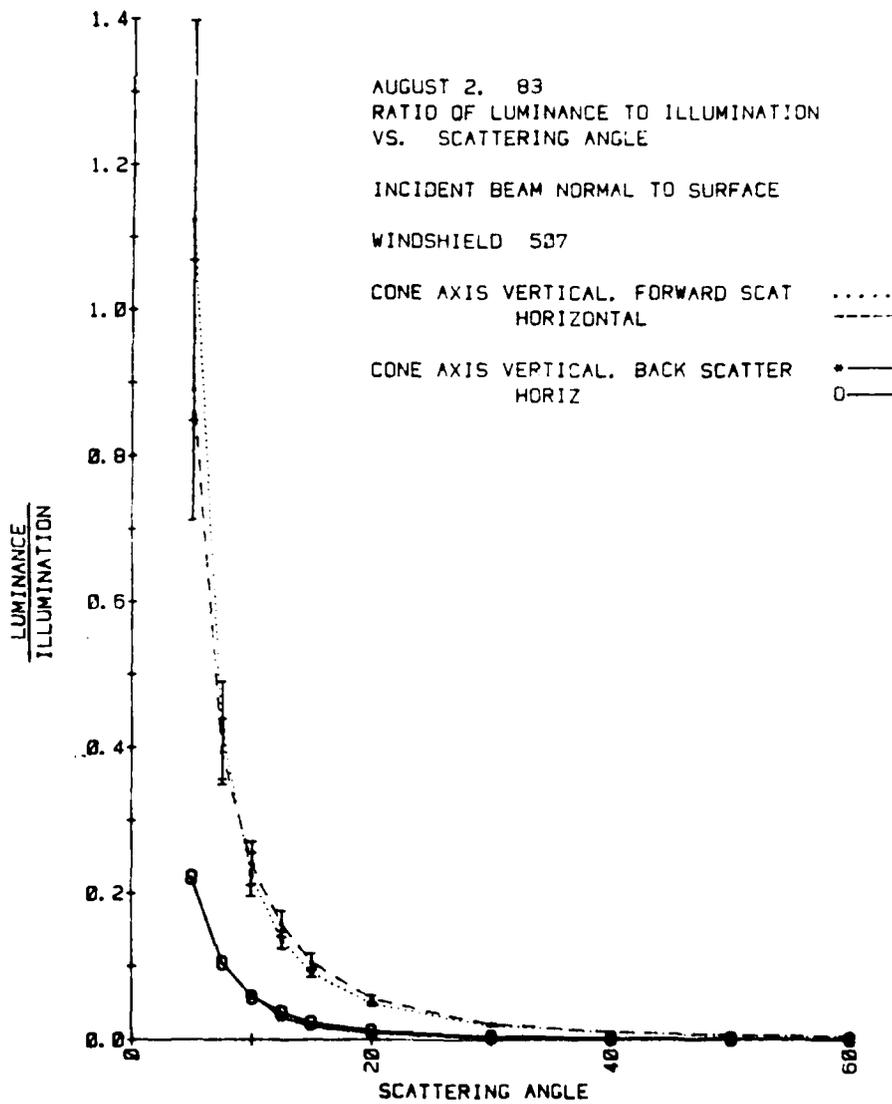
RATIO OF LUMINANCE TO ILLUMINATION
VS. SCATTERING ANGLE

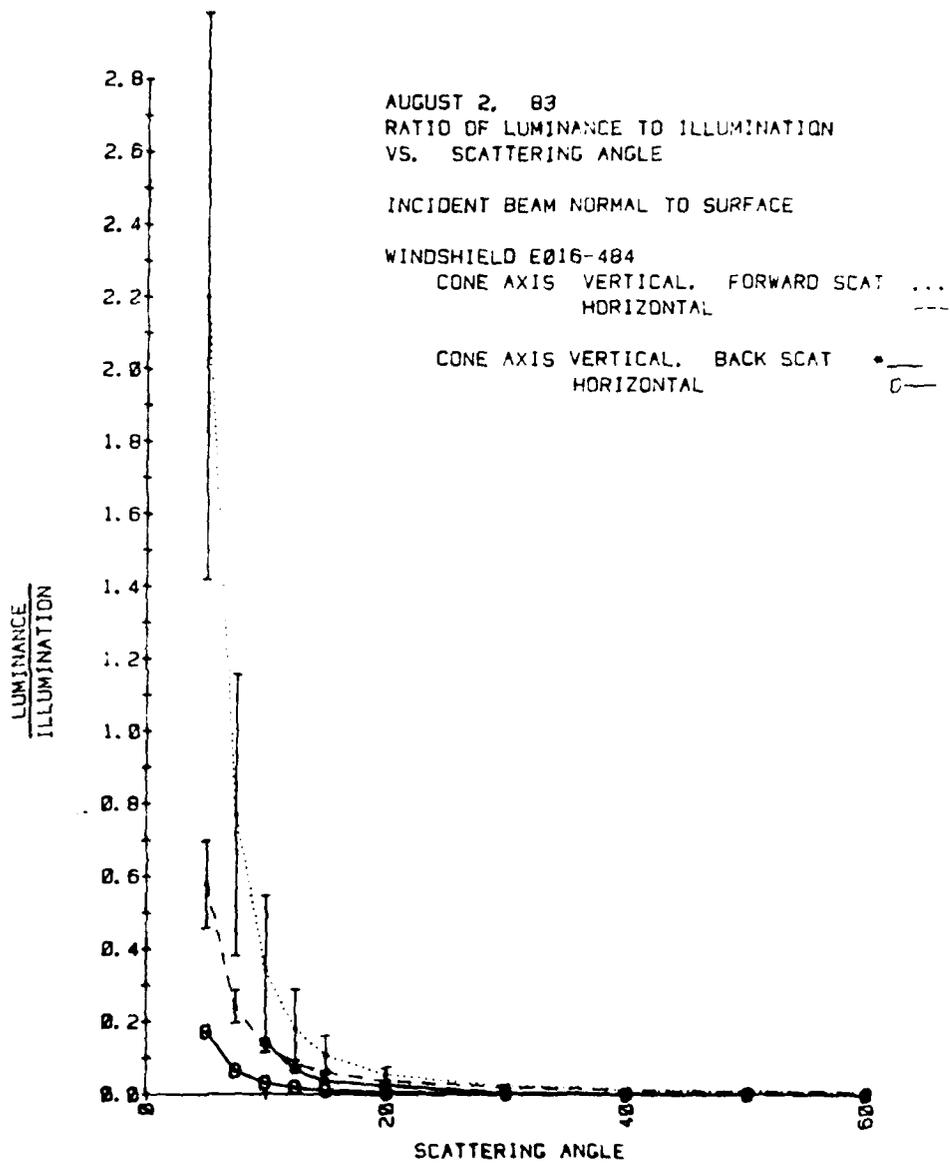
WINDSHIELD E016-453
INCIDENT BEAM NORMAL TO SURFACE

CONE AXIS VERTICAL
FORWARD SCATTERING ●-----
BACK SCATTERING ○-----

CONE AXIS HORIZONTAL
FORWARD SCATTERING ●.....
BACK SCATTERING ○.....







JUNE 22, 83

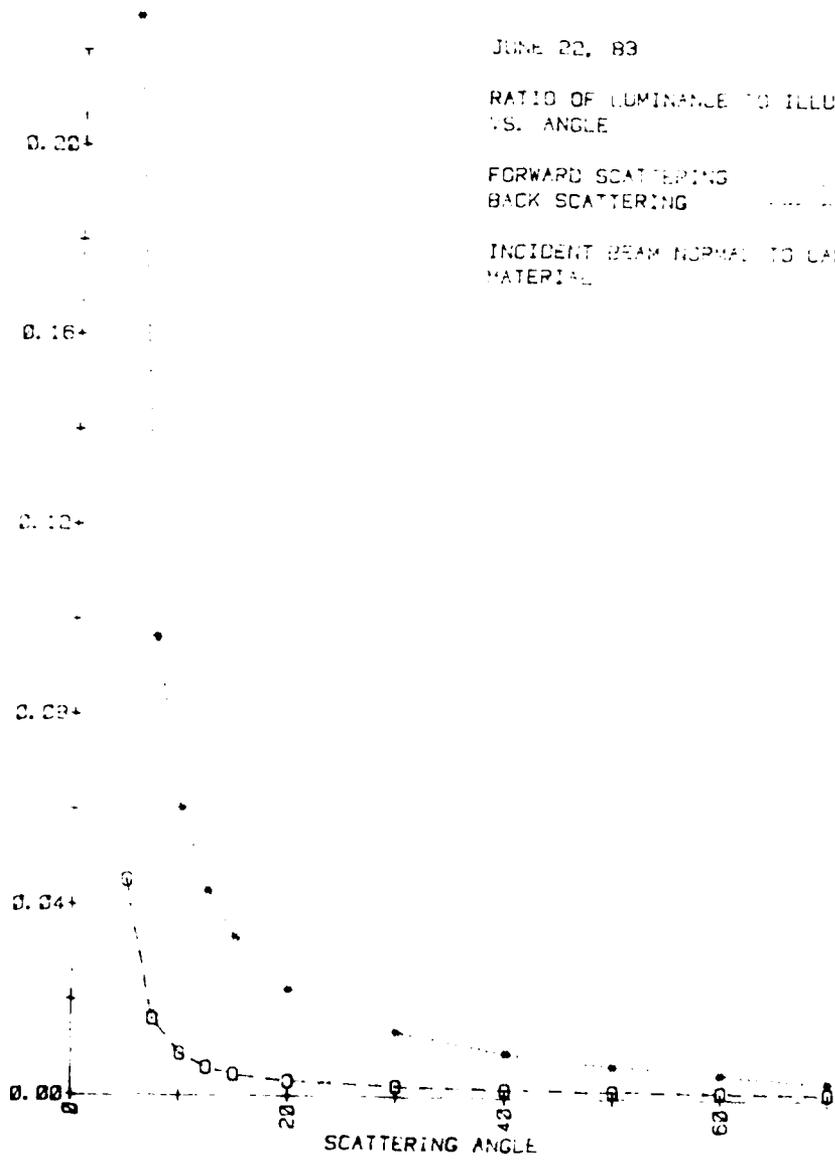
RATIO OF LUMINANCE TO ILLUMIN.
VS. ANGLE

FORWARD SCATTERING

BACK SCATTERING

INCIDENT BEAM NORMAL TO LAMP
MATERIAL

LUMINANCE
ILLUMINATION



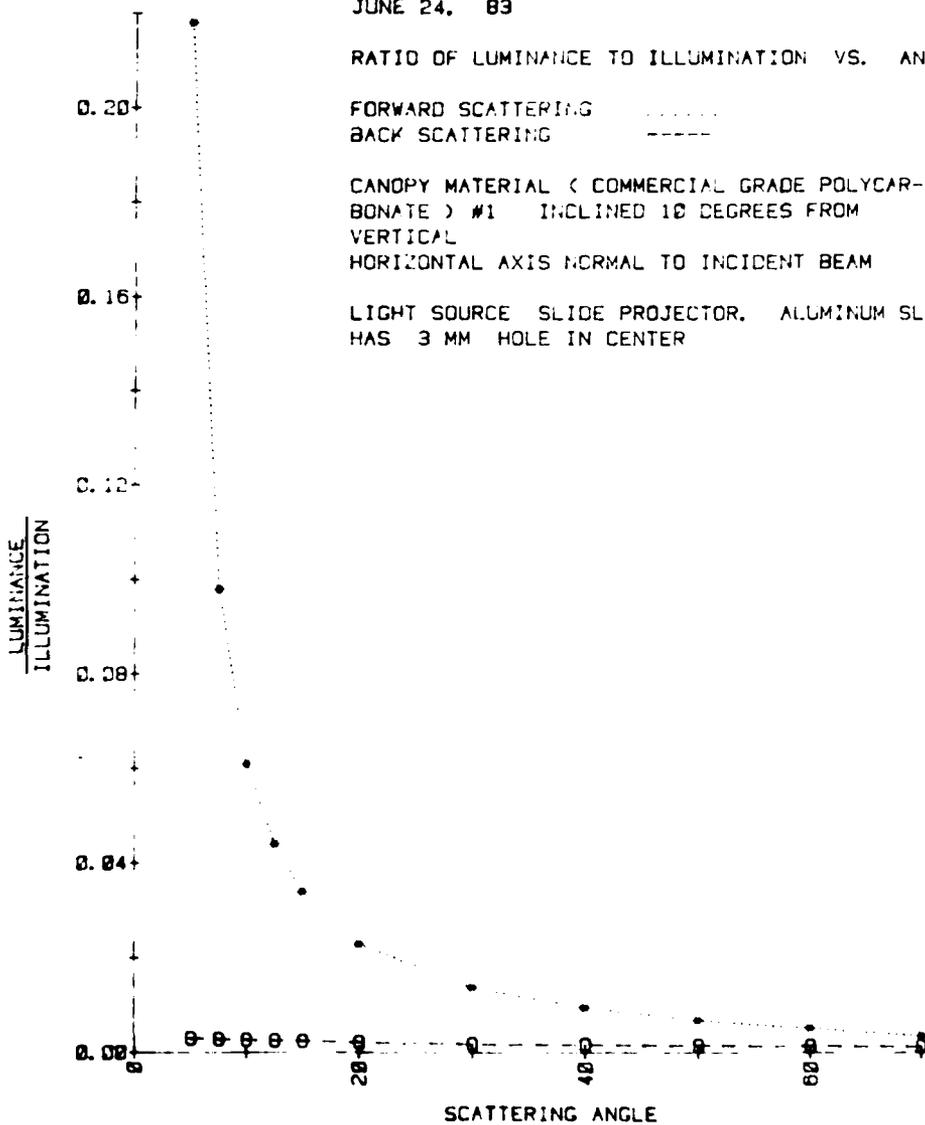
JUNE 24, 83

RATIO OF LUMINANCE TO ILLUMINATION VS. ANGLE

FORWARD SCATTERING
BACK SCATTERING -----

CANOPY MATERIAL (COMMERCIAL GRADE POLYCAR-
BONATE) #1 INCLINED 10 DEGREES FROM
VERTICAL
HORIZONTAL AXIS NORMAL TO INCIDENT BEAM

LIGHT SOURCE SLIDE PROJECTOR. ALUMINUM SLIDE
HAS 3 MM HOLE IN CENTER

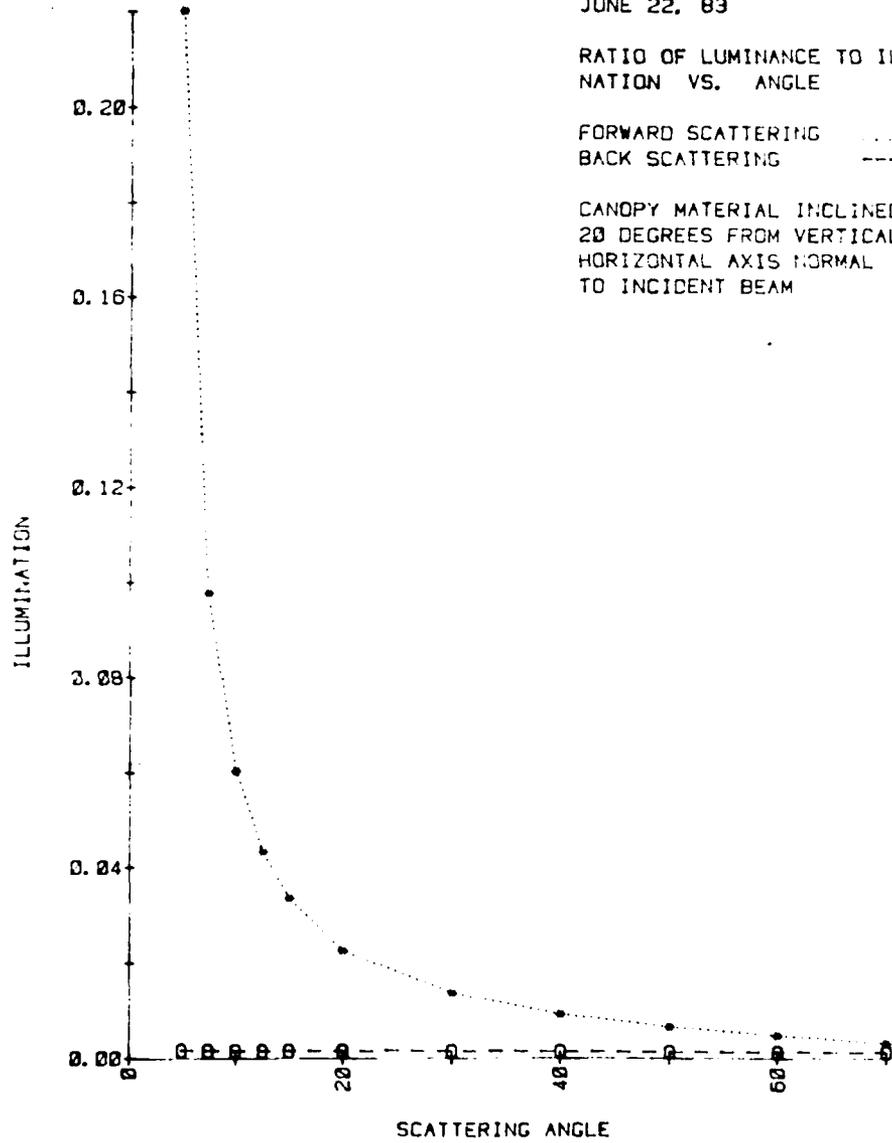


JUNE 22, 83

RATIO OF LUMINANCE TO ILLUMINATION VS. ANGLE

FORWARD SCATTERING
BACK SCATTERING -----

CANOPY MATERIAL INCLINED
20 DEGREES FROM VERTICAL
HORIZONTAL AXIS NORMAL
TO INCIDENT BEAM



JUNE 22, 83

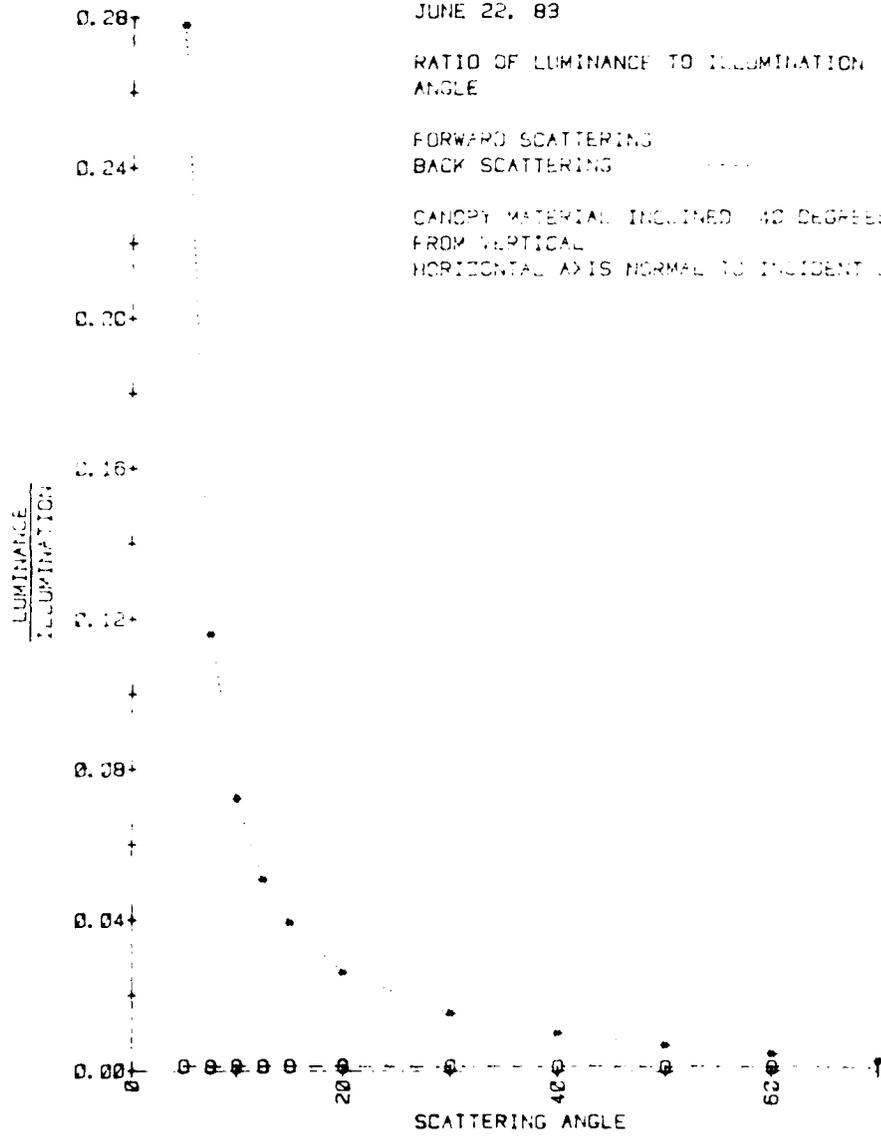
RATIO OF LUMINANCE TO ILLUMINATION VS. ANGLE

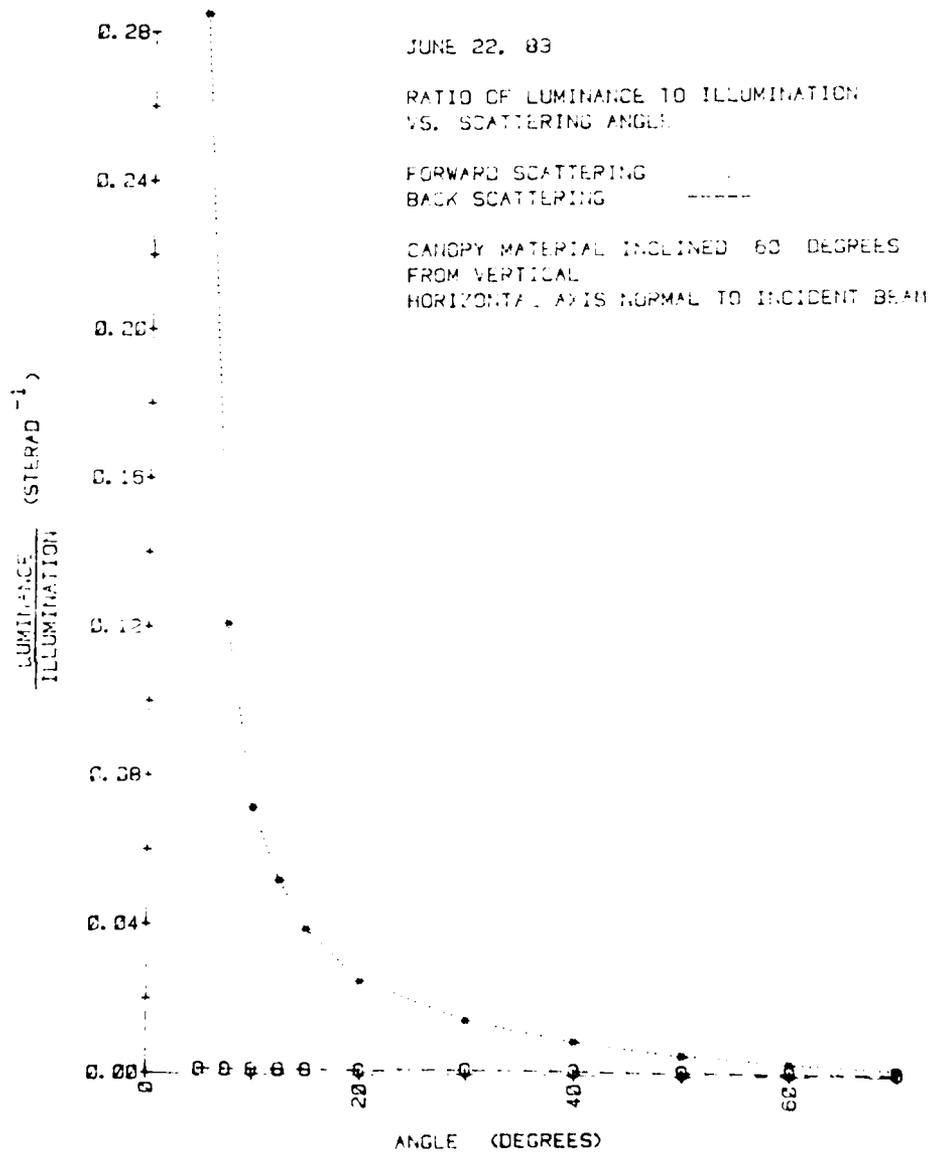
FORWARD SCATTERING

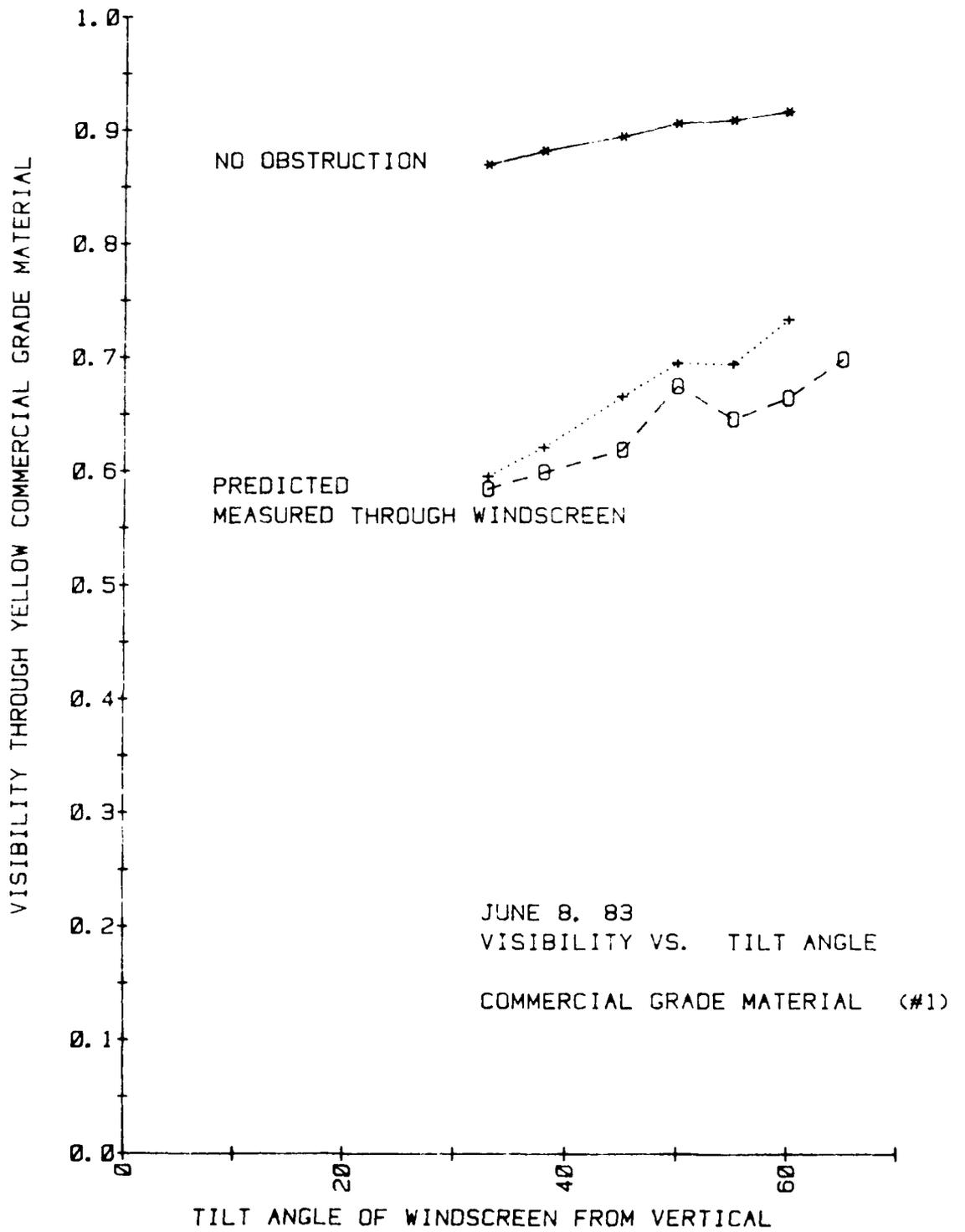
BACK SCATTERING

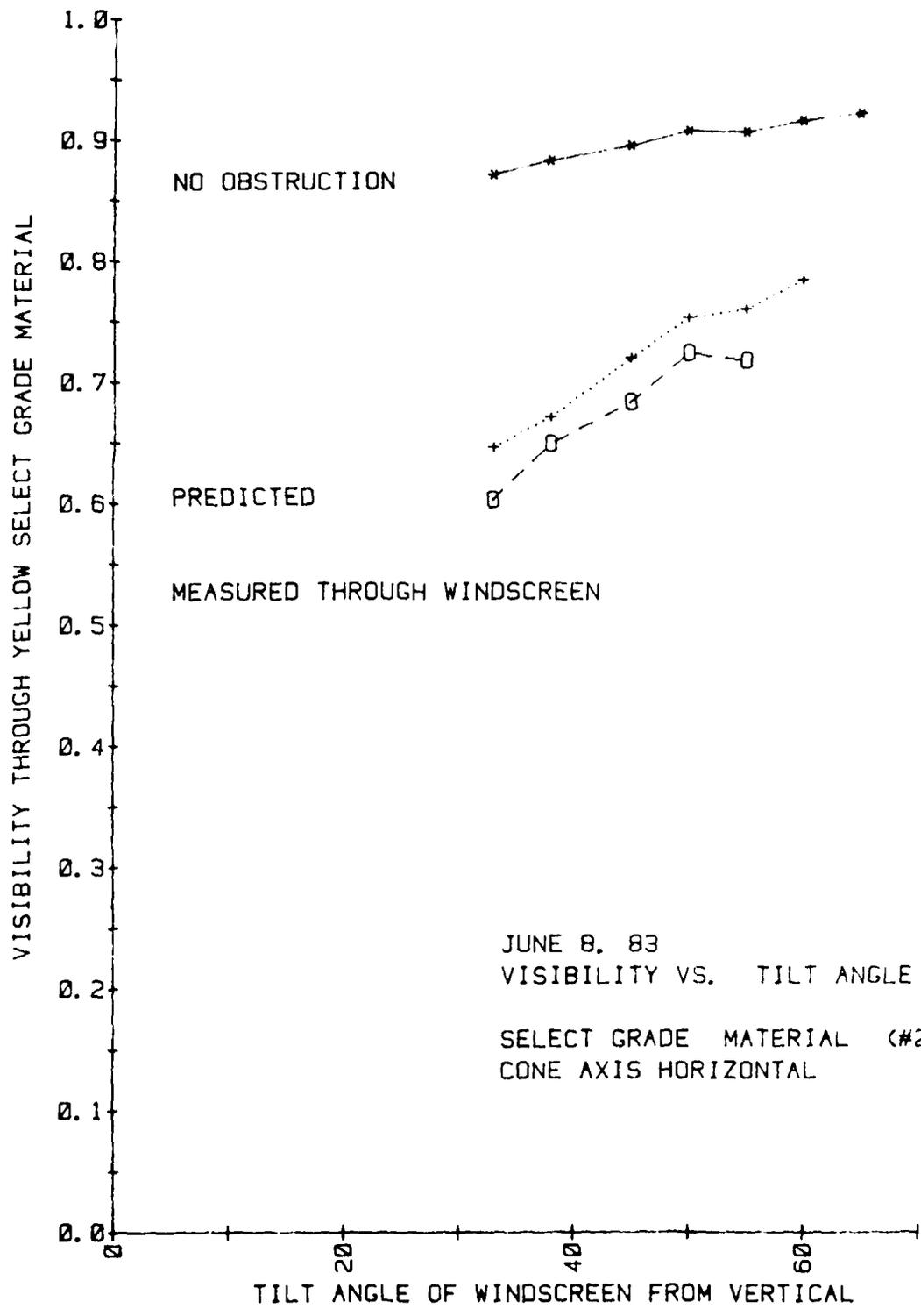
CANOPY MATERIAL INCLINED 40 DEGREES FROM VERTICAL

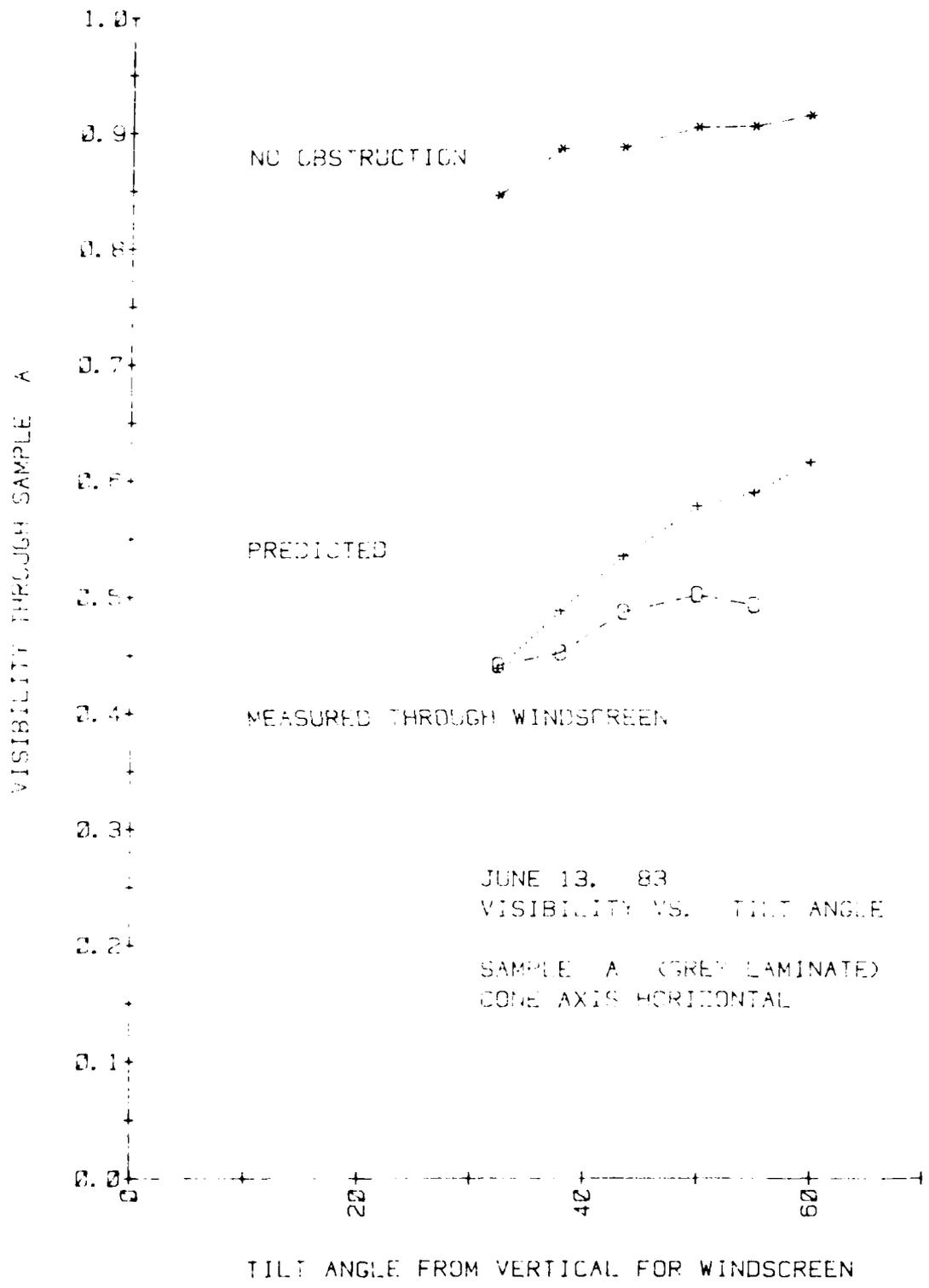
HORIZONTAL AXIS NORMAL TO INCIDENT BEAM

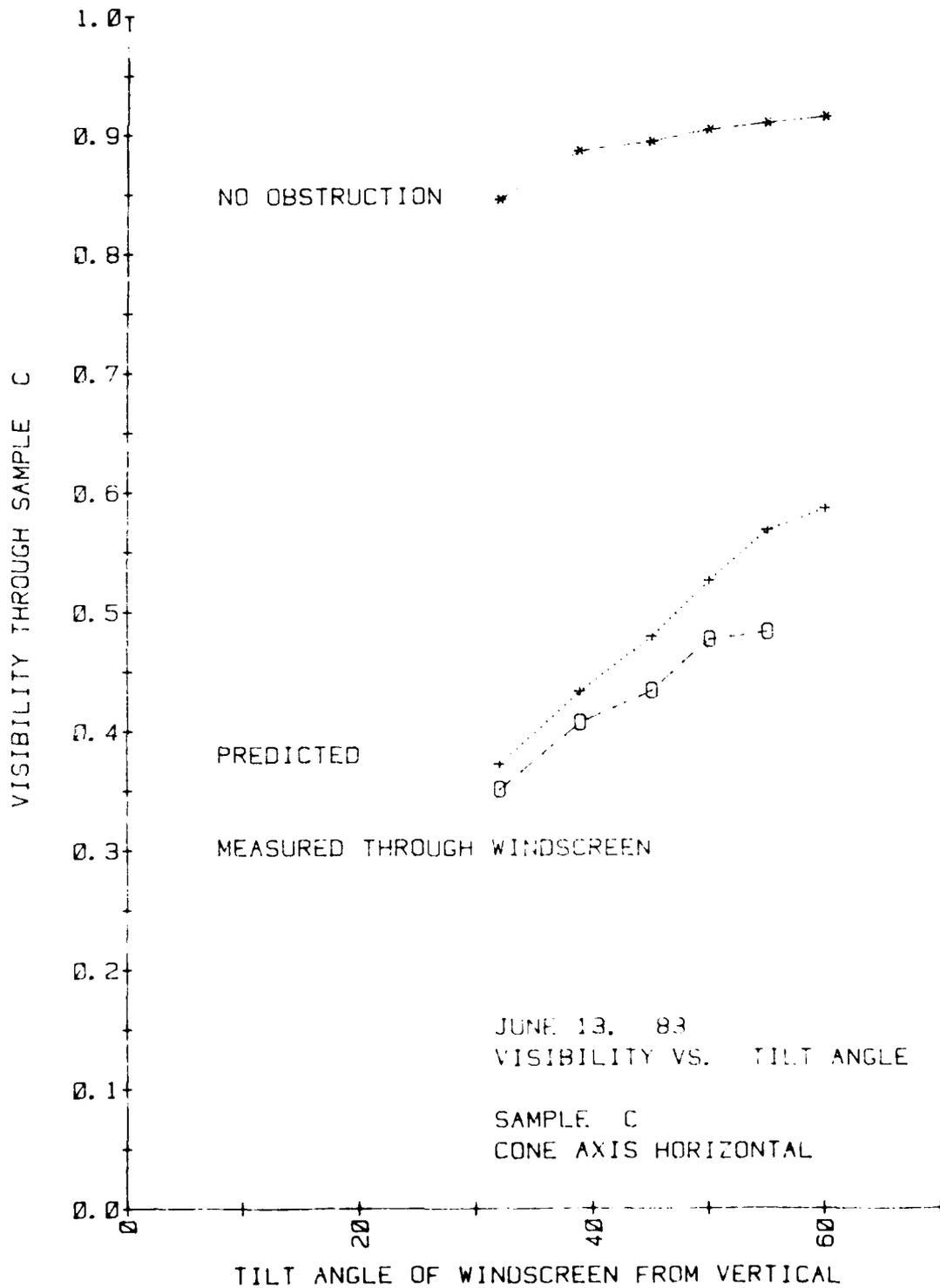


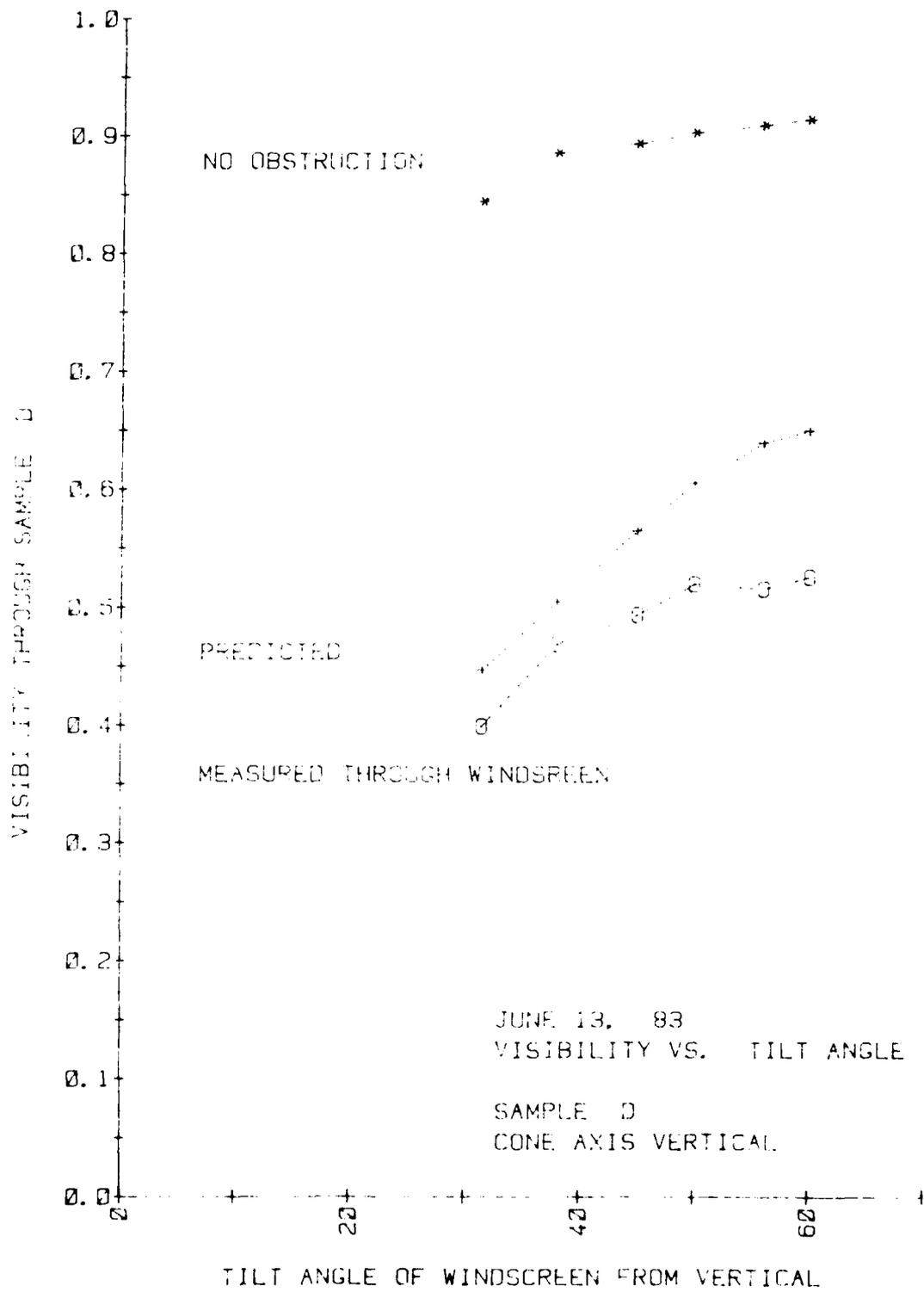




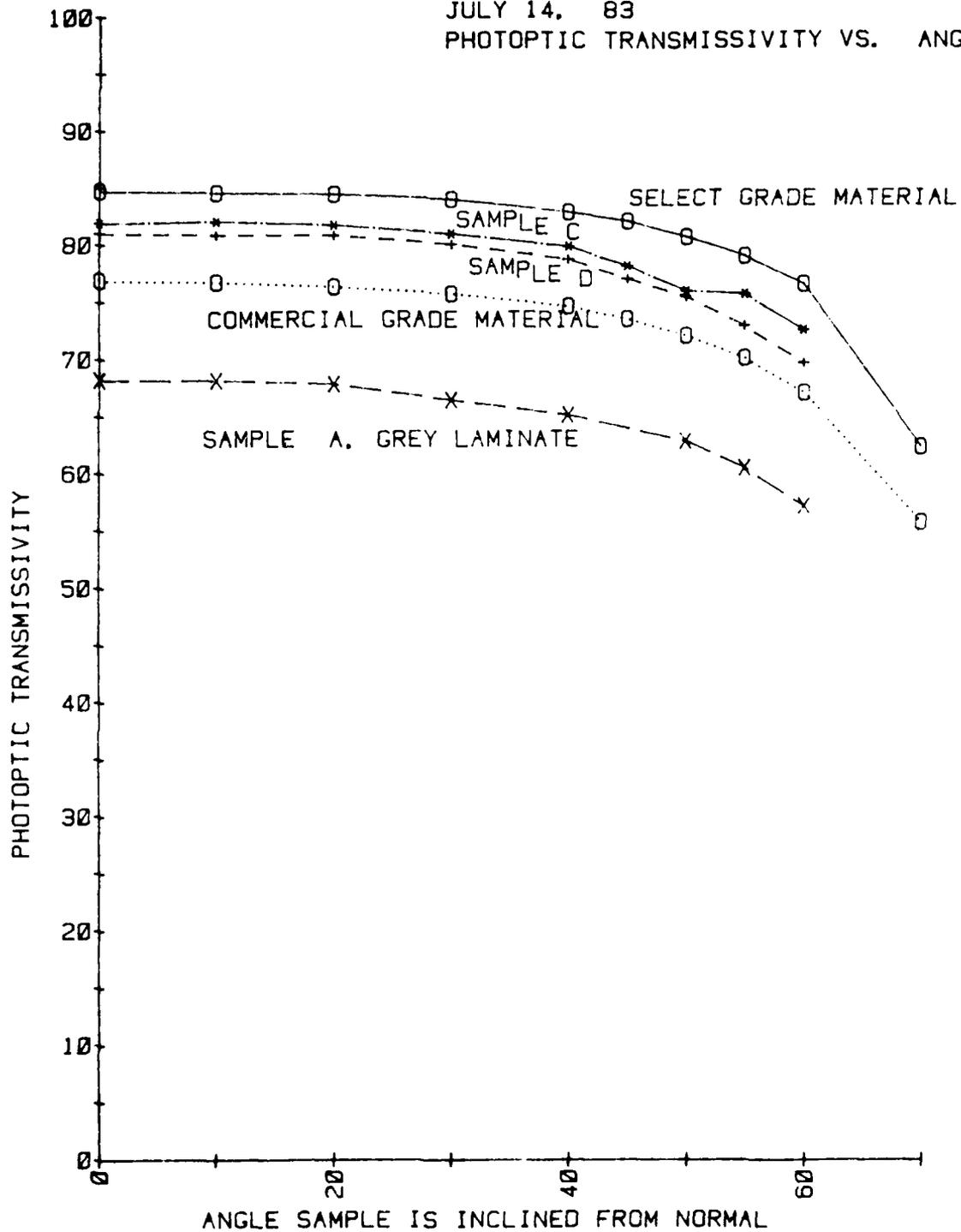








JULY 14, 83
PHOTOPTIC TRANSMISSIVITY VS. ANGLE



Item	QTY	UNIT	PRICE	TOTAL	DESCRIPTION
1	100	EA	0.077	7.70	...
2	100	EA	0.077	7.70	...
3	100	EA	0.077	7.70	...
4	100	EA	0.077	7.70	...
5	100	EA	0.077	7.70	...
6	100	EA	0.077	7.70	...
7	100	EA	0.077	7.70	...
8	100	EA	0.077	7.70	...
9	100	EA	0.077	7.70	...
10	100	EA	0.077	7.70	...
11	100	EA	0.077	7.70	...
12	100	EA	0.077	7.70	...
13	100	EA	0.077	7.70	...
14	100	EA	0.077	7.70	...
15	100	EA	0.077	7.70	...
16	100	EA	0.077	7.70	...
17	100	EA	0.077	7.70	...
18	100	EA	0.077	7.70	...
19	100	EA	0.077	7.70	...
20	100	EA	0.077	7.70	...
21	100	EA	0.077	7.70	...
22	100	EA	0.077	7.70	...
23	100	EA	0.077	7.70	...
24	100	EA	0.077	7.70	...
25	100	EA	0.077	7.70	...
26	100	EA	0.077	7.70	...
27	100	EA	0.077	7.70	...
28	100	EA	0.077	7.70	...
29	100	EA	0.077	7.70	...
30	100	EA	0.077	7.70	...
31	100	EA	0.077	7.70	...
32	100	EA	0.077	7.70	...
33	100	EA	0.077	7.70	...
34	100	EA	0.077	7.70	...
35	100	EA	0.077	7.70	...
36	100	EA	0.077	7.70	...
37	100	EA	0.077	7.70	...
38	100	EA	0.077	7.70	...
39	100	EA	0.077	7.70	...
40	100	EA	0.077	7.70	...
41	100	EA	0.077	7.70	...
42	100	EA	0.077	7.70	...
43	100	EA	0.077	7.70	...
44	100	EA	0.077	7.70	...
45	100	EA	0.077	7.70	...
46	100	EA	0.077	7.70	...
47	100	EA	0.077	7.70	...
48	100	EA	0.077	7.70	...
49	100	EA	0.077	7.70	...
50	100	EA	0.077	7.70	...
51	100	EA	0.077	7.70	...
52	100	EA	0.077	7.70	...
53	100	EA	0.077	7.70	...
54	100	EA	0.077	7.70	...
55	100	EA	0.077	7.70	...
56	100	EA	0.077	7.70	...
57	100	EA	0.077	7.70	...
58	100	EA	0.077	7.70	...
59	100	EA	0.077	7.70	...
60	100	EA	0.077	7.70	...
61	100	EA	0.077	7.70	...
62	100	EA	0.077	7.70	...
63	100	EA	0.077	7.70	...
64	100	EA	0.077	7.70	...
65	100	EA	0.077	7.70	...
66	100	EA	0.077	7.70	...
67	100	EA	0.077	7.70	...
68	100	EA	0.077	7.70	...
69	100	EA	0.077	7.70	...
70	100	EA	0.077	7.70	...
71	100	EA	0.077	7.70	...
72	100	EA	0.077	7.70	...
73	100	EA	0.077	7.70	...
74	100	EA	0.077	7.70	...
75	100	EA	0.077	7.70	...
76	100	EA	0.077	7.70	...
77	100	EA	0.077	7.70	...
78	100	EA	0.077	7.70	...
79	100	EA	0.077	7.70	...
80	100	EA	0.077	7.70	...
81	100	EA	0.077	7.70	...
82	100	EA	0.077	7.70	...
83	100	EA	0.077	7.70	...
84	100	EA	0.077	7.70	...
85	100	EA	0.077	7.70	...
86	100	EA	0.077	7.70	...
87	100	EA	0.077	7.70	...
88	100	EA	0.077	7.70	...
89	100	EA	0.077	7.70	...
90	100	EA	0.077	7.70	...
91	100	EA	0.077	7.70	...
92	100	EA	0.077	7.70	...
93	100	EA	0.077	7.70	...
94	100	EA	0.077	7.70	...
95	100	EA	0.077	7.70	...
96	100	EA	0.077	7.70	...
97	100	EA	0.077	7.70	...
98	100	EA	0.077	7.70	...
99	100	EA	0.077	7.70	...
100	100	EA	0.077	7.70	...

2001 August 28
 200200 00 500 PMS
 200000 000000 000
 200000 000000 000
 200000 000000 000

Account	Balance	Debit	Credit	Balance
101	100.00			100.00
102	200.00			200.00
103	300.00			300.00
104	400.00			400.00
105	500.00			500.00
106	600.00			600.00
107	700.00			700.00
108	800.00			800.00
109	900.00			900.00
110	1000.00			1000.00

Account	Balance	Debit	Credit	Balance
111	1100.00			1100.00
112	1200.00			1200.00
113	1300.00			1300.00
114	1400.00			1400.00
115	1500.00			1500.00
116	1600.00			1600.00
117	1700.00			1700.00
118	1800.00			1800.00
119	1900.00			1900.00
120	2000.00			2000.00

Station	Time	Lat	Long	Alt	Temp	Hum	Wind	Dir	Speed	Pressure	Clouds	Remarks
33	10:00	35° 15' N	122° 45' W	107	10.0	100	10	100	10	1010	100	10
32	10:15	35° 15' N	122° 45' W	107	10.0	100	10	100	10	1010	100	10
31	10:30	35° 15' N	122° 45' W	107	10.0	100	10	100	10	1010	100	10
30	10:45	35° 15' N	122° 45' W	107	10.0	100	10	100	10	1010	100	10
29	11:00	35° 15' N	122° 45' W	107	10.0	100	10	100	10	1010	100	10
28	11:15	35° 15' N	122° 45' W	107	10.0	100	10	100	10	1010	100	10
27	11:30	35° 15' N	122° 45' W	107	10.0	100	10	100	10	1010	100	10
26	11:45	35° 15' N	122° 45' W	107	10.0	100	10	100	10	1010	100	10
25	12:00	35° 15' N	122° 45' W	107	10.0	100	10	100	10	1010	100	10
24	12:15	35° 15' N	122° 45' W	107	10.0	100	10	100	10	1010	100	10
23	12:30	35° 15' N	122° 45' W	107	10.0	100	10	100	10	1010	100	10
22	12:45	35° 15' N	122° 45' W	107	10.0	100	10	100	10	1010	100	10
21	13:00	35° 15' N	122° 45' W	107	10.0	100	10	100	10	1010	100	10
20	13:15	35° 15' N	122° 45' W	107	10.0	100	10	100	10	1010	100	10
19	13:30	35° 15' N	122° 45' W	107	10.0	100	10	100	10	1010	100	10
18	13:45	35° 15' N	122° 45' W	107	10.0	100	10	100	10	1010	100	10
17	14:00	35° 15' N	122° 45' W	107	10.0	100	10	100	10	1010	100	10
16	14:15	35° 15' N	122° 45' W	107	10.0	100	10	100	10	1010	100	10
15	14:30	35° 15' N	122° 45' W	107	10.0	100	10	100	10	1010	100	10
14	14:45	35° 15' N	122° 45' W	107	10.0	100	10	100	10	1010	100	10
13	15:00	35° 15' N	122° 45' W	107	10.0	100	10	100	10	1010	100	10
12	15:15	35° 15' N	122° 45' W	107	10.0	100	10	100	10	1010	100	10
11	15:30	35° 15' N	122° 45' W	107	10.0	100	10	100	10	1010	100	10
10	15:45	35° 15' N	122° 45' W	107	10.0	100	10	100	10	1010	100	10
9	16:00	35° 15' N	122° 45' W	107	10.0	100	10	100	10	1010	100	10
8	16:15	35° 15' N	122° 45' W	107	10.0	100	10	100	10	1010	100	10
7	16:30	35° 15' N	122° 45' W	107	10.0	100	10	100	10	1010	100	10
6	16:45	35° 15' N	122° 45' W	107	10.0	100	10	100	10	1010	100	10
5	17:00	35° 15' N	122° 45' W	107	10.0	100	10	100	10	1010	100	10
4	17:15	35° 15' N	122° 45' W	107	10.0	100	10	100	10	1010	100	10
3	17:30	35° 15' N	122° 45' W	107	10.0	100	10	100	10	1010	100	10
2	17:45	35° 15' N	122° 45' W	107	10.0	100	10	100	10	1010	100	10
1	18:00	35° 15' N	122° 45' W	107	10.0	100	10	100	10	1010	100	10

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

MIREM AND MISSION PHASING

Prepared by: Robert D. Foley

Academic Rank: Assistant Professor

Department and University: Department of Industrial Engineering and Operations Research
Virginia Polytechnic Institute and State University

Research Location: Air Force Human Resources Laboratory,
Logistics Research Branch

USAF Research James McManus

Date: October 13, 1983

Contract No.: F49620-82-C-0035

MIREM AND MISSION PHASING

by

Robert D. Foley

ABSTRACT

MIREM was developed to analyze the reliability of the ICNIA avionics system during single-phase missions. In this paper, we develop a way of extending MIREM's capability to analyzing multi-phase missions. During the development, several questions about the existing implementation of MIREM arose. Specifically, MIREM uses two approximations, at least one of which works poorly on some test problems. Thus, it is recommended that MIREM's accuracy on realistic problems be evaluated. A method for testing MIREM is developed which could be incorporated into MIREM if MIREM is insufficiently accurate on realistic problems.

ACKNOWLEDGEMENTS

I would like to thank the Human Resources Lab for giving me the opportunity to spend the summer working in the Lab. In particular, I would like to thank Jim McManus for his help and Mike Veatch of TASC for explaining parts of MIREM. In addition, I must thank the Air Force Systems Command, the Air Force Office of Scientific Research, and the Southeastern Center for Electrical Engineering Education for giving me the opportunity to work at Wright-Patterson Air Force Base.

I. INTROUCTION:

The Air Force is currently developing an integrated communication, navigation, and identification avionics (ICNIA) system to be used in tactical aircraft. One of the most attractive features of ICNIA is fault-tolerance or graceful degradation. For the ICNIA system, fault-tolerance means that even though some of the avionics components have failed, the system can reconfigure itself in order to perform all of the critical avionics functions. In other words, the system can degrade gracefully by preserving the ability to handle all of the critical functions in spite of diminished abilities. The design of the ICNIA system greatly affects the ability of the system to degrade gracefully. Hence, design engineers to be informed of the amount of fault-tolerance in proposed designs for the ICNIA. MIREM (Mission Reliability Model) was developed as a tool to inform design engineers of the extent of fault-tolerance in proposed designs. Specifically, MIREM has the ability to determine the reliability of the avionics for a single phase mission, i.e., a mission in which the set of critical avionics functions does not change during the mission. I was requested to determine whether MIREM could be enhanced to handle mission phasing, i.e., to handle the fact that different avionics functions are critical during different phases of a mission. In the first part of this paper, some of the possible drawbacks of the current version of MIREM are analyzed. Most of these possible drawbacks are not a problem with the possible exception of some approximations used in computing the reliability. On some artificial cases, the approximations perform poorly. On realistic cases, the approximations should perform better -

however, this needs further investigation. The approximations are discussed in Section IV.

The next two sections describe methods for determining the reliability of multi-phase missions. The first method involves little effort and yields upper and lower bounds to the reliability, but it requires MIREM to be able to accurately analyze single-phase missions. The second method accurately analyzes multi-phase missions; however, it involves a large amount of computation and should prove to be computationally infeasible on realistic problems. The best approach appears to be using the first method provided that MIREM can analyze single-phase missions accurately. Hence, we develop an alternate method for accurately analyzing single-phase missions which should be computationally feasible yet does not involve any approximations. This alternate method can be used to determine if MIREM is sufficiently accurate on realistic problems. If MIREM is not sufficiently accurate, MIREM can be modified to use the compromise alternate method.

II. OBJECTIVES OF THE RESEARCH EFFORT.

The objective of the research effort was to determine if the reliability of the avionics system could be determined for multi-phase missions. In particular, it was hoped that MIREM could be enhanced to analyze multi-phase missions. In order to obtain this objective, the existing version of MIREM needed to be studied in some detail.

III. SOME POSSIBLE DRAWBACKS WITH THE EXISTING MIREM:

In this section, we review some of the possible shortcomings of MIREM as described in [1], which we assume the reader is familiar with.

Overall MIREM seems to be a well thought out approach to estimating the reliability of the ICNIA system. Some possible difficulties are:

- a) exponentiality assumption
- b) independence assumption
- c) imperfect switching
- d) interconnection failures
- e) undetected failures by BIT
- f) detected "failures" by BIT
- g) approximations in MIREM
- h) absence of mission phasing

We will briefly discuss all of the possible shortcomings, but we will be most concerned with g) and h).

a) exponentiality. The distribution of the lifetime of each component is assumed to be exponentially distributed. This assumption should be and probably has been tested. If the lifetimes are exponentially distributed, some reference should be made to studies supporting this assumption. If the lifetimes appear to be IFR (Increasing Failure Rate) or DFR (Decreasing Failure Rate), it should be determined whether the exponentiality assumption causes the system reliability to be over or underestimated. The exponentiality assumption does not seem to be a serious drawback.

b) independence. The component lifetimes are assumed to be independent. Assumptions a) and b) greatly increase the possibility of obtaining analytical solutions. If there are studies justifying the independence assumption, they should be cited. In reality, the lifetimes are dependent, and probably positively correlated. For

example, if one component is damaged by a projectile or a power surge, other components are also likely to be damaged. Positive correlations should lead to overestimating the system reliability. It is conjectured that the dependence between components is sufficiently small so that the independence assumption is not a serious drawback.

c) imperfect switching. MIREM assumes that the switches work perfectly. Rumor has it they don't. The seriousness of assuming perfect switching is directly related to the extent that the switching is imperfect.

d) interconnection failures. MIREM assumes that the wiring, etc. interconnecting the components doesn't fail. Similar comments to c).

e) undetected failures by BIT. BIT (Built In Test equipment) determines whether each component has failed. MIREM assumes that BIT detects all failures. In reality some failures go undetected. Again the seriousness of the assumption depends on how often the assumption is violated.

f) detected "failures" by BIT. Note that BIT may also claim that a component has failed when the component is fine. To get a rough idea of the magnitude of this problem, some recent studies have shown that approximately 30% of the electrical equipment identified as faulty and removed from the aircraft pass the retesting procedure at the intermediate repair shop and receive no corrective maintenance.

g) approximation in MIREM. On pp. 63-64 of the Appendix in [1], two approximations are used while computing the system reliability. It is not clear that the approximations are sufficiently accurate. Note that a) through f) are simply assumptions upon which the model is based.

However the approximation is not a problem with the model, it is a problem with the analysis of the model. We will discuss the approximation more thoroughly in Section IV.

h) absence of mission phasing. The absence of mission phasing causes problem in determining the system reliability for a mission. Different functions are critical at different phases of a mission as shown in Figure 1.

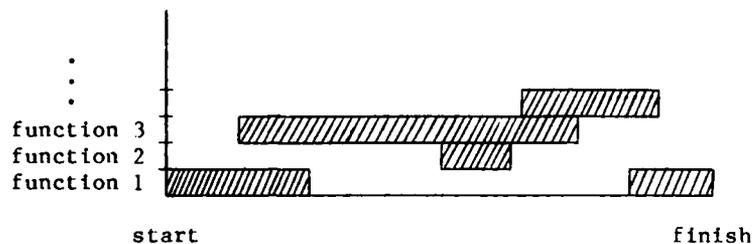


Fig. 1. Figure showing which functions are critical and when.

MIREM assumes that critical functions are critical throughout all phases of the mission. Basically this corresponds to extending each of the black bars in Figure 1 from the start until the finish of the mission. Clearly this will cause the reliability to be underestimated MIREM alleviates this problem slightly on page 61 of [1] by treating contending pools as noncontending if the functions are not required simultaneously. For example, suppose all of the functions in Figure 2 need to use a single preprocessor from a particular pool of preprocessors.

Since none of the functions are critical at the same time, the pool of preprocessors can be classified as noncontending. Thus the pool needs to have only one working preprocessor in order to support these five functions. Now suppose that function 2 is critical from time zero

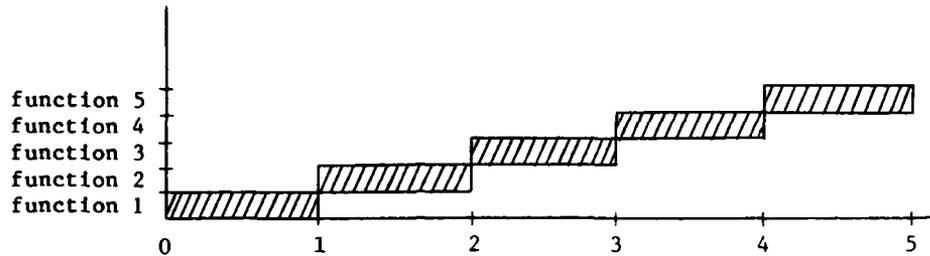


Fig.2. Example of a contending pool treated as noncontending.

until time 2. Since functions 1 and 2 are both critical from time zero until time 1, the pool must be treated as contending. Since the pool is contending, MIREM will require at least five working preprocessors from time zero until time 5 to support these five functions. In reality, the system needs only two preprocessors working from time 0 until time 1, and only 1 preprocessor working from time 1 until time 5. Thus MIREM is suited only for determining the reliability of the system during a single phase of a mission. In Section V, we introduce a quick-and-dirty method for extending MIREM's capability to a multi-phase mission assuming that MIREM can accurately determine the reliability of a single phase. Before that we look at how accurately MIREM analyzes a single phase.

III. MIREM's APPROXIMATIONS:

If one accepts the assumptions upon which MIREM is based, it would seem that one must accept the results MIREM yields, at least for single phase missions. However, there are several approximations used in MIREM which makes it questionable whether the results of MIREM should be trusted. The approximations cause MIREM to overestimate the system reliability. The magnitude of the errors may be significant.

The approximations are described on pp. 63-64 of [1]. The authors wish to obtain

$$p(\underline{x}^1, \underline{x}^2) = P\{UP^{1+2}(C) | \underline{x}^1 = \underline{x}^1, \underline{x}^2 = \underline{x}^2\}$$

which is the probability that the contending pools of the two chains can handle all of the critical functions conditioned on knowing which functions can be handled by which chains with respect to the noncontending pools. This can be rewritten as

$$p(\underline{x}^1, \underline{x}^2) = P\{\exists \underline{y}, \text{ s.t. } 1 - \underline{x}^2 < \underline{y} < \underline{x}^1; r_i(\underline{y}) < C_i^1, r_i(1 - \underline{y}) < C_i^2 \\ \text{for } i = 1, \dots, n\} \quad (2)$$

where i indexes the n pairs of contending pools. Equation (2) corresponds to equations (A-7) of the appendix to [1].

The approximation used in [1] for $p(\underline{x}^1, \underline{x}^2)$ can be derived as follows:

$$p(\underline{x}^1, \underline{x}^2) < \prod_{i=1}^n P\{\exists \underline{y}, \text{ s.t. } 1 - \underline{x}^2 < \underline{y} < \underline{x}^1; r_i(\underline{y}) < C_i^1, r_i(1 - \underline{y}) < C_i^2\} \quad (3)$$

and the r.h.s. of (3) is

$$< \prod_{i=1}^n P\{r_i(1 - \underline{x}^j) < C_j^1 \text{ for } j = 1, 2; r_i(1) < C_i^1 + C_i^2, \\ \max_{j \in CF} u_{ij} < \max\{C_i^1, C_i^2\}\}. \quad (4)$$

Equation (4) corresponds to (A-8) of the appendix to [1]. Note that there is a typo in equations (A-8a), (A-8b), and (A-9): (\underline{x}^1) should be replaced by $(1 - \underline{x}^2)$, and (\underline{x}^2) should be replaced by $(1 - \underline{x}^1)$.

The error incurred in going from (2) to (3) is that instead of having a single allocation of functions that must simultaneously satisfy all pairs of contending pools, we only require that each pair of pools

has some feasible allocation. For example, suppose we have two functions, two chains, and two contending pairs of pools with the number of working components as shown in Figure 1.

Suppose that the first function requires 3 components and the second function requires 1 component from each pool in the chain that the function is assigned to. Note that the example in Figure 1 cannot

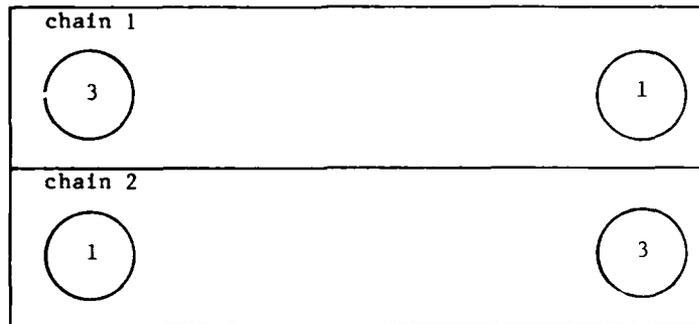


Fig. 3. An example with two chains and two pairs of contending pools. support these two functions since function 1 cannot be assigned to either chain. However MIREM would conclude that the system can support these two functions. According to (3), MIREM looks at each pair of pools in isolation. In the first pair of pools, MIREM says that function 1 can be assigned to chain 1 and function 2 to chain 2. Then MIREM looks at the second pair of pools and says that function 1 can be assigned to chain 2 and function 2 to chain 1. Since both functions can be supported at each pair of pools when considered in isolation, MIREM erroneously concludes that both functions can be supported. Clearly, looking at pairs of pools in isolations results in overestimating the system reliability.

The second approximation occurs while going from (3) to (4). The probability that there exists a feasible allocation at the i th pair of pools is approximated by (4). As the authors point out, this is only a necessary condition; the r.h.s. of (4) overestimates the r.h.s. of (3). To see this, consider the first pair of pools in the example in Figure 3, but assume that each of the functions needed two components. Clearly there is no feasible allocation of the two functions even when a pair of pools is considered in isolation. However if $\underline{x}^1 = \underline{x}^2 = \underline{1}$, we can assign one function to one chain and the other function to the other chain, and all of the equations in (A-8) are satisfied (don't forget to correct the typo in (A-8) mentioned earlier). Thus MIREM would erroneously conclude that the system in the example could support the two functions.

The error in the first approximation gets worse as the number of contending pools increases and as $\underline{x}^1 + \underline{x}^2$ approaches $\underline{2}$. The error in the first approximation occurs only if there is more than one pool. The error in the second approximation occurs only if some of the u_{ij} are greater than one. Of the two errors the first is more serious. The first error seems to be sufficiently serious based on the following example to warrant further investigation.

Each pool in the example has one component. The reliability of each component is .9. There is a single function which needs one component from each pool in the chain that it is allocated to. In Table 1, we compare the true system reliability with the results that MIREM would produce for various sizes of n . MIREM produces the correct answer for one pair of pools. However, even with two pairs of pools, MIREM is

not close to the true answer. As the number of pools increases MIREM becomes worse and worse. All of the error is due to the first approximation. Thus it is unclear whether the existing version of MIREM accurately determines the reliability of realistic systems with contending pairs of pools. In Section VI, we present a direct approach to computing the system reliability. Unfortunately, the direct approach will probably be computationally infeasible on realistic sized problems.

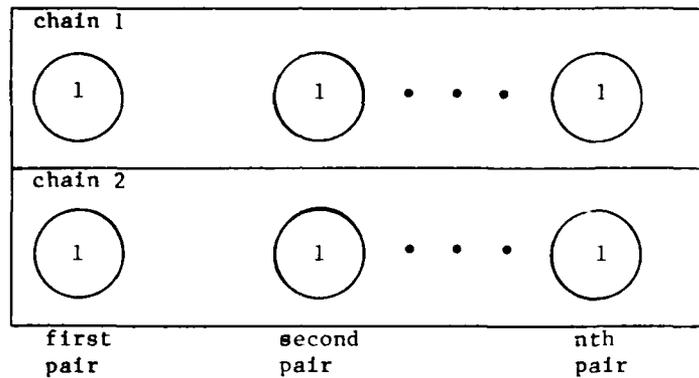


Fig. 4. Example with n pairs of contending pools.

<u>n</u>	<u>MIREM system reliability</u>	<u>True system reliability</u>
1	.9900	.9900
2	.9801	.9639
3	.9703	.9266
4	.9606	.8817
5	.9506	.8323
10	.9044	.5758
15	.8601	.3694
20	.8179	.2284

V. A QUICK-AND-DIRTY EXTENSION OF MIREM TO MISSION PHASING:

In this section, we propose a simple method for extending MIREM's capability from single phase to multi-phase missions under the assumption that MIREM accurately analyzes a single phase mission. Based on the previous section, this assumption is a little shaky. It is hoped that the modification to MIREM proposed in Section VII will allow MIREM to accurately analyze a single phase mission.

Define a mission phase to be an interval of time in which the same set of functions are critical. Then the entire mission can be divided into m phases. Phase 1 is the first interval, i.e., take-off, then phase 2 follows, etc., as shown in Figure 5.

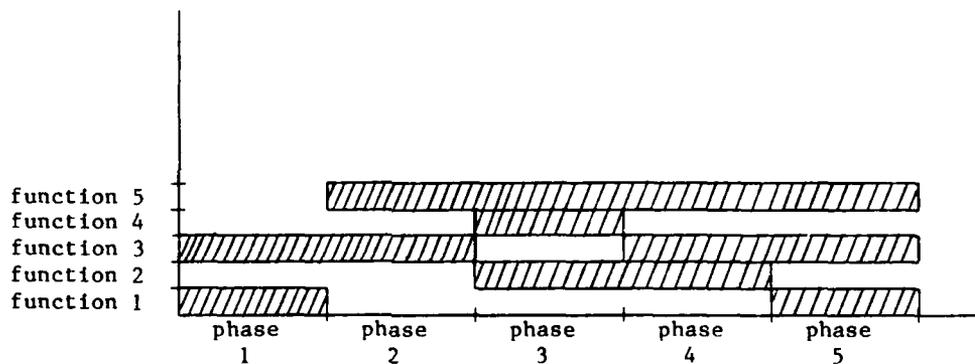


Fig. 5. Graph showing which functions are critical and when during a five phase mission.

Let PH_i denote the event that the i th phase is successful. The Weapon System Reliability (WSR) is

$$\begin{aligned} WSR &= P\{PH_1, PH_2, \dots, PH_m\} \\ &= P\{PH_1\}P\{PH_2|PH_1\} \cdots P\{PH_m|PH_1, \dots, PH_{m-1}\} \end{aligned}$$

Let F_k denote the set of critical functions during phase k . In the example shown in Figure 5,

$$F_1 = \{1,3\} \text{ and } F_4 = \{2,3,5\}.$$

Corresponding to each phase, we define two other mission scenarios: one long, and one short. Let PH_k^L be the event that a single phase mission is successfully completed where the mission has the same critical functions as the k th phase of the original mission and the length of the long mission is the length of time from the start of the original mission until the end of phase k . Let PH_k^S be the same event except that the length of the mission is exactly the same as the length of phase k .

We claim that

$$P\{PH_k^L\} < P\{PH_k | PH_{k-1}, \dots, PH_1\} < P\{PH_k^S\}. \quad (5)$$

Using (5), we can get upper and lower bounds on the system reliability:

$$P\{PH_1^L\}P\{PH_2^L\} \cdots P\{PH_m^L\} < WSR < P\{PH_1^S\}P\{PH_2^S\} \cdots P\{PH_m^S\}. \quad (6)$$

The second inequality in (5) can be justified intuitively as follows. The probability that phase k is successful given the previous phases were successful must be smaller than

$$P\{PH_k^S\} \quad (7)$$

since (7) corresponds to the probability that phase k is successful given that we fix all broken components at the start of phase k .

The first inequality in (5) can be justified by realizing that

$$P\{PH_k^l\}$$

is the probability that phase k is successful without knowing whether the previous phases were successful or not. But conditioning on the previous phases being successful increases the probability that phase k is successful because the components are more likely to be working if we know the previous phases were successful. A formal proof could be given but does not seem necessary.

Thus we have upper and lower bounds on the reliability of a multi-phase mission based on the reliability of a single phase mission. It is conjectured that the lower bound is quite good. In some cases the lower bound can be improved. If there exists some phase i and a later phase j with all of the critical functions in phase i being critical in phase j, i.e.,

$$F_i \subset F_j,$$

then $P\{PH_i^l\}$ can be deleted from the l.h.s. of (6). Thus in the example shown in Figure 5, we originally had

$$P\{PH_1^l\}P\{PH_2^l\} \cdots P\{PH_5^l\} < WSR < P\{PH_1^s\}P\{PH_2^s\} \cdots P\{PH_5^s\}.$$

Now, since

$$F_1 \supset F_5 \quad \text{and} \quad F_2 \supset F_4,$$

we have

$$P\{PH_4^l\}P\{PH_5^l\} < WSR < P\{PH_1^s\}P\{PH_2^s\} \cdots P\{PH_5^s\}.$$

VI. A DIRECT APPROACH:

In this section, we present a direct approach to accurately computing the reliability of a multi-phase or single phase mission. The disadvantage of this method is that it becomes computationally intractable for real sized problems. In Section VII, we propose an alternate method which is hopefully both accurate and computationally tractable. The direct approach does have several advantages. In particular, it provides a check on more sophisticated algorithms and their implementations.

Let \underline{C}_t be the vector whose components describe the number of working components in each pool. \underline{C}_t is a Markov process. Let E denote the state space and P_t the transition matrix for $\{\underline{C}_t; t > 0\}$. Assume that the mission has m phases and define

$$\varphi_k(i) = \begin{cases} 1 & \text{if all critical functions of phase} \\ & \text{k can be handled in state i,} \\ 0 & \text{otherwise.} \end{cases}$$

Let t_k denote the length of phase k and $\gamma_k(i)$ denote the probability of successfully completing phases k through m inclusive given that the state at the beginning of phase k was i. Then

$$\gamma_m(i) = \sum_{j \in E} P_{t_m}(i,j) \varphi_m(j)$$

and

$$\gamma_k(i) = \sum_{j \in E} P_{t_m}(i,j) \varphi_k(j) \gamma_{k+1}(j).$$

Thus the γ 's can be iteratively computed since P_t and φ_k are easily determined.

The major disadvantage with this method is that it quickly becomes computationally intractable.

VII. ALTERNATIVE METHOD:

The purpose of this paper is to present a method for analyzing single-phase missions which is both accurate and computationally feasible on realistic problems. Hopefully the method can either be used to show that MIREM is sufficiently accurate or be incorporated into MIREM and allow MIREM to accurately determine the system reliability.

MIREM seems to work well except when computing

$$P\{UP^{1+2}(C) | \underline{X}_1 = \underline{x}^1, \underline{X}_2 = \underline{x}^2\} \quad (8)$$

which appeared as equation (A-5) in the appendix of [1]. Hence, we will concentrate on developing a method to compute (8). Assume that we have two parallel chains as described in the appendix to [1]. We will be concerned only with the contending pools. Assume that there are n pairs of pools numbered $1, \dots, n$, and a set of critical functions. Each critical function must be allocated to one of the two chains although some critical functions may be restricted to one of the chains. An allocation completely specifies which critical functions are allocated to which chains. A possible allocation satisfies the constraints at the noncontending pools; however, a possible allocation may or may not satisfy the constraints at the sharing or contending. Assume that there are p possible allocations numbered $1, \dots, p$. Now a feasible allocation is a possible allocation that satisfies all of the constraints at the contending pools. To compute (1), we simply need to determine

$$P\{\text{there exists a feasible allocation}\}. \quad (9)$$

Let A_k denote the event that allocation k is feasible. Equation (9) can be restated as

$$P\{A_1 \text{ or } A_2 \text{ or } \dots \text{ or } A_p\} \quad (10)$$

which can be rewritten as

$$\begin{aligned} &P\{A_1\} + P\{A_2\} + \dots + P\{A_p\} \\ &- P\{A_1A_2\} - P\{A_1A_3\} - \dots - P\{A_{p-1}A_p\} \\ &+ P\{A_1A_2A_3\} + \dots + P\{A_{p-1}A_{p-2}A_{p-3}\} \\ &\quad \vdots \\ &+ (-1)^{p+1}P\{A_1A_2\dots A_p\}. \end{aligned} \quad (11)$$

Thus we need to be able to determine the values of the terms in (11).

Let C_i be an ordered pair representing the state of the i th pair of contending pools at the end of the mission. Since the pools within a pair are mutually independent, we have

$$P\{C_i > (s,t)\} = G_i(s)H_i(t).$$

The functions G_i and H_i are easy to determine since each is the complementary distribution for a binomial random variable.

Let $g_i(k)$ denote the number of working components needed at the chain 1 pool of the i th pair of pools by allocation k , and let $h_i(k)$ denote the corresponding quantity for the chain 2 pool. Then the probability of allocation k being feasible can be computed as

$$\begin{aligned}
P\{A_k\} &= P\{C_1 > (g_1(k), h_1(k)), \dots, C_n > (g_n(k), h_n(k))\} \\
&= \prod_{i=1}^n P\{C_i > (g_i(k), h_i(k))\} \\
&= \prod_{i=1}^n G_i(g_i(k))H_i(h_i(k)).
\end{aligned}$$

More complicated terms such as $P\{A_j A_k\}$ can be evaluated as follows:

$$\begin{aligned}
P\{A_j A_k\} &= P\{C_1 > (g_1(j) \vee g_1(k), h_1(j) \vee h_1(k)), \dots, \\
&C_n > (g_n(j) \vee g_n(k), h_n(j) \vee h_n(k))\} \\
&= \prod_{i=1}^n G_i(g_i(j) \vee g_i(k))H_i(h_i(j) \vee h_i(k)).
\end{aligned}$$

where " \vee " denotes "maximum of." In general, we have

$$P\{A_{k_1}, \dots, A_{k_m}\} = \prod_{i=1}^n G_i(\bigvee_{j=1}^m g_i(k_j))H_i(\bigvee_{j=1}^m h_i(k_j)). \quad (12)$$

The method can be outlined as follows:

- 1) determine the functions G_i , H_i , g_i , and h_i .
- 2) Evaluate (11) using (12) to evaluate each of the terms.

The alternate method does not use any approximations; hence, it determines the correct answer. The method has the same advantage as the existing version of MIREM since it analyzes each pair of pools separately and then combines the information. However, the alternate method involves more computational effort since more information must be computed at each pair of pools and the combination of the results is more involved. In the alternate method, the computational effort grows with the number of possible allocations which in turn grows roughly exponentially with the number of critical functions. However, for the number of critical functions needed in realistic problems, the method

appears to be computationally feasible. The computation effort may be reduced by several means if necessary. For example, (4) can be truncated at various points yielding either upper or lower bounds to the system reliability. Thus the method appears to satisfy both criteria. The alternate method yields the correct answer, and appears to be computationally tractable.

VIII. RECOMMENDATIONS:

The original task was to determine a method for enhancing MIREM to permit analyzing multi-phase missions. The method outlined in Section IV appears promising; however, this method relies on MIREM accurately analyzing a single-phase mission. Section III shows that MIREM can have trouble with some artificial problems. Hence it should be investigated whether MIREM accurately analyzes realistic problems. If MIREM is sufficiently accurate on realistic problems, the MIREM can be left unchanged. However, if MIREM performs poorly, MIREM should be modified to use the alternate method described in Section V. Thus, the important point needing investigation is MIREM's accuracy on realistic problems. The alternate method should be implemented in order to have a standard against which MIREM can be tested.

- [1] Veatch, M. H., Calvo, A. B., and Myers, J. B. Logistics engineering design techniques for fault-tolerant avionics. The Analytic Sciences Corporation. TIM-4128-1.

ADDENDUM

A draft of this document was discussed by the developers of MIREM and myself. In Section III, I have given two examples showing where MIREM makes approximations. Due to some other features, MIREM does not make approximations for the two specific examples I gave and is able to accurately analyze the two examples. However, the approximations described are representative of the type of approximations occurring in slightly more complicated examples. Rather than construct more complicated examples, it is better to understand the spirit of the approximation in the simple examples. The developers of MIREM are confident that MIREM will be sufficiently accurate on realistic cases. This should be verified by checking it against the alternative method described in Section VI. If MIREM performs poorly, the compromise method should be incorporated into MIREM. If MIREM performs well, it might be preferable not to incorporate the alternative method since it is computationally more work than the existing version.

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

HEMP VULNERABILITY/SURVIVABILITY OF COMPUTER NETWORKS

Prepared by : Dr Eddie R Fowler
Academic Rank : Associate Professor
Department and University : Department of Electrical Engineering
Kansas State University
Research Location : Air Force Weapons Laboratory, Nuclear Technol-
ogy Directorate, Space/C3/Reentry Systems Div-
ision, Applications Branch
USAF Research : Dr Babu Singaraju
Date : 8 August 1983
Contract No. : F49620-82-C-0035

HEMP VULNERABILITY/SURVIVABILITY OF COMPUTER NETWORKS

by

Dr Eddie R Fowler

ABSTRACT

This report presents an analytical approach for quantifying the message traffic through-put degradation of a computer network in a high altitude electromagnetic pulse (HEMP) environment. The ARPANET/DDN is characterized to provide information as to which attributes to incorporate in a computer network simulation model. Plots of the HEMP generated E-field are presented. An E-field strength contour map superimposed on orthogonal projection of the USA for one height of burst (HOB) is presented. Finally suggestions for further research in this area are presented.

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 Weapons Laboratory (AFWL), Kirtland AFB, N.M. He would also like to acknowledge the AFWL Nuclear Technology Directorate for providing an interesting research problem, a well stocked Technical Library, and the use of the NTC Division's VAX 11-780.

He also wishes to thank Jacque Rodriguez, a NTC Stay-in-School, for his programming, plotting, and engineering assistance and support.

Finally, he wishes to thank his wife, Patricia Ann, for spending many late nights typing, proof reading, and collating this and the unabridged report.

1.0 Introduction And Research Motivation. The USAF operational capability is dependent upon its C4 systems and computer networks have become an integral part of these C4 systems. Thus to process and communicate C3I information during nuclear hostilities, these computer networks must be high altitude electromagnetic pulse (HEMP) survivable. To be HEMP survivable, computer networks must be designed, constructed, and operated according to standards based on quantified through-put degradation for all possible threat and friendly generated HEMP. Presently this degradation has not been quantified empirically because of the ban on all exoatmospheric nuclear tests. Thus analytical techniques using simulation modeling is the only feasible approach to quantify through-put degradation. Based on this, there is motivation for a research effort to develop an analytical technique that evaluates and quantifies computer network through-put degradation resulting from HEMP.

1.1 Research Objectives. The present research objectives are long term and short term. The long term research objective is to develop an analytical technique that utilizes a simulation model to quantify the through-put degradation of a computer network in a HEMP environment. Presently the primary C3I computer network candidate for analysis is the packet switched Defense Data Network (DDN) whose design philosophy is being evolved from that of the Advanced Research Project Agency Network (ARPANET). The short term objectives have been to: develop the necessary background understanding by accomplishing a HEMP literature survey, and gain a familiarization of EMP phenomenology and of the ARPANET/DDN

packet switched network characteristics; and, develop a description of an analytical technique and simulation model that can be used to quantify the through-put degradation of a computer network.

1.2 Previous and Present Work. Several agencies/organizations have been interested and involved in research and development of EMP phenomenology and the susceptibility/vulnerability and survivability of communication networks illuminated by EMP. Also several authors have been working with the design and analysis of computer communication networks. This section will briefly discuss pertinent work and documentation associated with EMP phenomenology, computer network analysis, and HEMP effects on networks.

1.2.1 EMP Phenomenology. EMP phenomenology has been a topic of major interest for the past 20 years at both the U S Air Force Weapons Laboratory (AFWL) and the Defense Nuclear Agency (DNA). Early work at AFWL led to a basic classified EMP handbook (69) that contains descriptions of EMP phenomenology associated with system and weapon environment and environment confidence. A second AFWL design handbook (63) provides major concerns for system protection against EMP effects, with recommended practices and procedures. There is also a time domain and frequency domain mathematical description of electric field strength. The AFWL Nuclear Criteria Group Secretariat (NCGS) has published the latest approved version (71) of these mathematical descriptions. DNA has likewise collated descriptions of EMP phenomenology and published these data in six volumes (70). The DNA volumes contain: information on system evaluation and system design that are hardened to EMP; formulas for calculating EMP coupled currents and voltages in cables; information for assessing equipment and component susceptibility to

EMP and damage due to EMP; information on nuclear induced EMP into systems in flight; a 3400 entry EMP bibliography; and, Information on computer codes for analyzing EMP effects problems. Bell Laboratories is the one civilian industry that has published an EMP handbook (4). This document contains a description of the EMP phenomenology concerning the generation and characteristics of the EM field, its intensity contour map, EMP coupling and shielding of structures and facilities, and component susceptibility. The latest AFWL document (40) is a thorough discussion on EMP interaction principles, techniques, and reference data. The combined information in the above documents provide a rather complete discussion on EMP phenomenology.

1.2.2 Network Characterization. Computer network characterizations and associated through-put degradation analytical techniques have been a topic of research in isolation from EMP considerations. However this work is directly applicable to through-put degradation resulting from nuclear effects. There has been a flood of interest in analyzing the packet-switched ARPANET. It has produced ample research in such areas as: adaptive routing and flow control (3,6,8,10,12,13,17,18,25,26,30,38,44,46,47,48,52,56,62,64); computer-to-computer protocol (8,17,18,26,34,35,36,44,46,56,61,62); communication processor hardware and software design (8,13,17,18,26,44,46,56,62); line protocol (8,17,18,26,44,46,56,61,62); network reliability (8,13,17,18,25,26,44,46,56,62,65,66); network topological design (3,6,8,12,17,18,25,26,30,44,46,48,56,62,65); and, packet switching concepts (8,13,17,18,25,26,44,46,56,62,66). The link transmission media of the ARPANET/DDN are designed and operated in accordance with AT&T policy(5). Any analytical techniques and/or computer network simulation models should be based on the previous work

documented in the above references.

1.2.3 HEMP effects. HEMP effects on computer and communication networks has been a Department of Defense (DOD) concern since the nineteen-fifties. The communications network that supported the anti-ballistic missile system, SAFEGUARD, was analyzed for EMP survivability and the resulting documentation (72) presents EMP coupling, with resulting voltage/current transient, predictions and equipment/subsystem EMP susceptibility. This study also included the AT&T long haul lines in the communications sub-system analysis. DNA has conducted studies (73,74,75) on the nuclear susceptibility and survivability of the Automatic Digital Network (AUTODIN), Automatic Voice Network (AUTOVON), and the communications facilities for the Pacific Command's C4 system. These studies have analyzed and assessed the nuclear EMP effects on the C4 system's landlines, wireless and satellite links, and terminal facilities. The Defense Communications Agency (DCA), Strategic Air Command (SAC), and USAF Air Staff have pursued programs to analyze and assess the nuclear vulnerability and survivability of communications/computer command and control networks which each will use during hostilities. The most complex of these is the Defense Data Network (DDN) which is evolving from the ARPANET, World Wide Military Command and Control System (WWMCCS) Intercomputer Network (WIN), and Movement Information Network (MINET). The nuclear vulnerability and survivability analysis of the DDN is presently being conducted under the supervision of the DDN Program Office (20, 29, 58) DCA has used the network simulation model (54), GRINDER/ROUT, in some of its studies and suggested that it might be useful in a nuclear vulnerability and survivability study of DDN. Likewise the Air Staff has used the SIMSTAR model (14) and AFWL the PNAC model (49). Both

are functioning models and candidates for any analytical study of communications/computer networks. The civilian contractor company, XRI, McLean, VA, is studying the nuclear vulnerability and survivability of the DDN node facility equipments. XRI has just initiated work on this study. Bell Laboratories, Holmdel, NJ, has recently organized a Nuclear Weapons Effects Group to study the EMP effects on the Bell System long haul trunk lines and the node facilities and equipments. It must be emphasized that this group is not a continuation of any previous EMP study. Due to the enormous facility size requirement for EMP testing, the only known testing presently is that being done by Harry Diamond Labs on buried cables. However the cable length used in this testing still limited to less than 10 km. An unabridged report (68) of this survey contains an appendix with all of the above references with a synopsis of each.

2.0 General Discussion. The research objective of developing an analysis technique and simulation model to assess nuclear vulnerability and survivability of a packet switched computer network requires characterizing in writing such a network, e.g., ARPANET/DDN. To get a general idea on how the DDN works and therefore the number of network characteristics that must be considered for a simulation model to characterize the DDN, one must first have a good understanding of how a packet-switched network functions.

2.1 ARPANET/DDN Characterization. A characterization study was done of the ARPANET/DDN and is included in the unabridged report (68). The following topics are discussed in that report; messages, Interface Message Processor (IMP), packet descriptions, positive acknowledgement (ACK), host, source and destination IMP, error detection and retransmission,

packet delay, ALLOCATE, sliding window protocol, and routing algorithm. Figures that show the geographical location of the ARPANET and DDN nodes and their logical interconnections are given in (68).

2.2 AT&T Dependency. Both the ARPANET and the DDN use AT&T leased lines (20). Therefore, for purposes of computer simulation, a database containing the exact geographical location and orientation of each AT&T link is necessary for degradation purposes. Figures showing maps of the 1972 AT&T long haul lines are given in (68). After referring to these figures, one immediately notes the extreme dependency of the DDN on the AT&T lines. The vulnerability/survivability of the AT&T lines is of critical importance to the vulnerability/survivability of the DDN.

3.1 Model Survey. There have been simulation models previously developed which analyze packet-switched computer networks, but none will fully characterize the ARPANET/DDN. The four simulation models surveyed were the Propagation Network Analysis Code (PNAC), the ROUT module of GRINDER (Graphical Interactive Network Designer), SIMSTAR, and a connectivity program. These are discussed in detail in the unabridged report (68).

3.2 Network Connectivity Program. A connectivity program was generated during the course of this research in order to get a better understanding of the problems associated with the modeling of a computer network. In this simulation model the maximum number of nodes allowed is eight, all nodes are fixed, and the links may be either directed or bidirectional. Two of the problems identified in generating a program to characterize a network with a large number of nodes are computer run time and memory. Both of these factors are discussed in detail in the unabridged report (68) and the computer code

for the connectivity program is also included.

3.3 Network Simulation Model Description. The development of any simulation model should be in a top down structured manner. Since the purpose of a simulation model is to analyze some system aspect with respect to a measure of effectiveness (MOE). This MOE must be well defined because it is the basis/foundation of the system's simulation model attributes. Therefore candidate MOE and attributes for a packet-switched computer network simulation are given below.

3.3.1 Measures of Effectiveness (MOE). A list of candidate MOE that quantify some aspect of through-put degradation are: connectivity of input/output (I/O) devices; transmission delay per bit error rate; availability of I/O devices; and, through-put of message traffic, which can be a combination of all of the above plus other MOE.

3.3.2 Model Attributes. Based on the listed MOE above, candidate model attributes are: analog and/or digital transmission signal format; nodes; links; message traffic; incident times; link degradation; node degradation; and, process degradation. These are discussed in detail in the unabridged report (68).

3.4 HEMP Effects Analysis. The EMP attributes discussed above that are of interest to this report are the induced voltages and currents that are coupled into the network links and nodes from the incident EM field. This EM field can be characterized by the electric field (E-field), in both the time domain and frequency domain, at each observer (coupling) location (link/node) within the illuminated area. This EMP coupling characterization is necessary to determine the voltage or current induced in the links and nodes. This coupled voltage/current is input into the network simulation model to determine if there is any resulting

memory upset or component burnout of link, node, and/or facility equipments. And upset or burnout determination is the ultimate objective of HEMP effects analysis, because it is used in the designated MOE function to determine through-put degradation. Thus a detailed discussion concerning the characterization of the E-field, resulting EMP coupling to the links and nodes of a computer network, and candidate HEMP effects attributes are presented in the unabridged report (68).

4.0 Analytical And Simulation Problems. The analytical techniques, presented in (68), have a number of problems that need individual attention before they will be resolved. Some of these problems are inherent to the scope of the task. They are: parameters uncertainties; AT&T proprietary information; simulation model run time; magnetohydrodynamics (MHD) effects; MOE development; and, computational techniques. These are discussed in detail in (68). However at the risk of sounding flippant, this anonymous quote is appropriate, "With enough infusion of money, any problem can be solved!"

5.0 Suggested Research Problems. The research problems that have become apparent from this study are: network simulation model development; computational techniques; MOE development; and, MHD effects. These research problems are discussed in detail in (68).

6.0 Summary. A summary of accomplishments reported in this document and the unabridged report (68) are: a literature survey; technical personnel survey; E-field computations and plots; E-field strength contour maps; computer network characterization; connectivity program; and, vulnerability/survivability analytical techniques.

UNCLASSIFIED BIBLIOGRAPHY

1. Aguet, Michel, Mircea Ianovici and Chung-Chi Lin; "Transient Electromagnetic Field Coupling to Long Shielded Cables," IEEE Transactions on Electromagnetic Compatibility, Vol. EMC-22, No. 4, November 1980; IEEE Press, New York, New York
2. ARPANET Directory; SRI International for the Defense Communications Agency; March 1982; NIC 49000
3. Balabanian, Norman and Theodore A. Bickart; Electrical Network Theory; John Wiley & Sons, Inc.; New York, New York, c. 1969
4. Bell Laboratories; EMP Engineering and Design Principles; Bell Telephone Laboratories, Inc.; Whippany, New Jersey; c. 1975
5. Bell Laboratories; Engineering Operations in the Bell System; Western Electric Company, Inc.; Indianapolis, Indiana; c. 1977
6. Bellman, Richard, Kenneth L. Cooke and Jo Ann Lockett; Algorithms, Graphs and Computers; Academic Press, Inc.; New York, New York; c. 1970
7. Bevensee, R.N., James N. Brittingham, F.T. Deadrick, Theodore H. Lehman, Edmund D. Miller and Andrew J. Poggio; "Computer Codes for EMP Coupling," IEEE Transactions on Electromagnetic Compatability, Vol. EMC-20, No. 1, February 1978; IEEE Press; New York, New York
8. Ed. by Blanc, Robert P. and Ira W. Cotton; Computer Networking; IEEE Press; New York, New York; c. 1976
9. Broad, William J.; "The Chaos Factor," Science 83, January/February 1983
10. Cantor, David G. and Mario Gerla; "Optimal Routing in a Packet-Switched Computer Network," IEEE Transactions on Computers, Vol. C-23, No. 10, October 1974; IEEE Press, New York, New York
11. Casey, Kendall F. and Edward F. Vance; "EMP Coupling Through Cable Shields," IEEE Transactions on Electromagnetic Compatibility, Vol. EMC-20, No. 1, February 1978; IEEE Press; New York, New York

12. Christofides, Nicos; Graph Theory: An Algorithmic Approach; Academic Press, Inc.; New York, New York; c. 1975
13. Cole, Gerald D.; "Performance Measurements on the ARPA Computer Network," Large-Scale Networks: Theory and Design; IEEE Press; New York, New York; c. 1976
14. Conrad, J.; Documentation of the SIMSTAR Multi-Message C3I Simulator, Vol. I-VIII; IRT Corporation for HQ USAF: August 1982; IRT 6442-029
15. Corrigan, Michael L.; Defense Data Network Protocols; The Defense Communications Agency; IEEE Press; New York, New York; c. 1982
16. Corrigan, Michael, Steven F. Holmgren, Robert K. Miller, and David C. Wood; DDN Interfacing Strategy; The MITRE Corp. for the Defense Communications Agency; IEEE Press; New York, New York; c. 1982
17. Davies, Donald W. and Derek L.A. Barber; Communication Networks for Computers; John Wiley & Sons, Inc.; London, England; c. 1973
18. Davies, D.W., D.L.A. Barber, W.L. Price, and C.M. Solomonides; Computer Networks and Their Protocols; John Wiley & Sons Ltd.; Chichester, England; c. 1979
19. A brochure on the Defense Data Network; The Defense Communications Agency
20. Defense Data Network Program Plan; The MITRE Corp. for the Defense Communications Agency: January 1982, Rev. May 1982
21. Ed. by Feinler, Elizabeth; ARPANET Resource Handbook; SRI International for the Defense Communications Agency: February 1980; NIC-47500
22. Ed. by Feinler, Elizabeth and Jonathan Postel; ARPANET Protocol Handbook; SRI International for the Defense Communications Agency: January 1978; NIC-7104-Rev-2
23. Ferrari, Domenico; Computer Systems Performance Evaluation; Prentice-Hall, Inc.; Englewood Cliffs, New Jersey; c. 1978
24. Frank, Howard and Ivan T. Frisch; Communication, Transmission, and Transportation Networks; Addison-Wesley Publishing Co.; Reading, Massachusetts; c. 1971
25. Frank, H., R.E. Kahn and L. Kleinrock; "Computer Communications Network Design: Experience with Theory and Practice," Large-Scale Networks: Theory and Design; IEEE Press; New York, New York; c. 1976
26. Ed. by Green, Jr., Paul E. and Robert W. Lucky; Computer

- Communications; IEEE Press; New York, New York; c. 1974
27. Hackett, General Sir John, Air Chief Marshal Sir John Barracough, Sir Bernard Burrow, Brigadier Kenneth Hunt, Vice-Admiral Sir Ian McGeoch, Norman Macrae, and Major-General John Strawson; The Third World War: August 1985; Macmillian Publishing Company, Inc.; New York, New York; c. 1978
 28. Harris, Dr. Thomas C., Capt. Peter V. Abene, Wayne W. Grindle, Darryl W. Henry, Dennis C. Morris, Maj. Glynn E. Parker and Jeffery Mayersohn; Development of the MILNET; The Defense Communications Agency; IEEE Press; New York, New York; c. 1982
 29. LTC Heidi, B. Heiden and Howard C. Duffield; Defense Data Network; The Defense Communications Agency; IEEE Press; New York, New York; c. 1982
 30. Henley, Ernest J. and R.A. Williams; Graph Theory in Modern Engineering; Academic Press, Inc.; New York, New York; c. 1973
 31. Herman, James G. and Susan L. Berstein; Monitoring, Control, and Management of the Defense Data Network; Bolt Beranek and Newman, Inc. for the Defense Communications Agency; IEEE Press; New York, New York; c. 1982
 32. IEEE Spectrum: October 1982; IEEE Press; New York, New York
 33. An Information Briefing on the Defense Data Network (DDN); The Defense Communications Agency
 34. Internet Mail Protocols; SRI International for the Defense Communications Agency: November 1982
 35. Internet Protocol Implementation Guide; SRI International for the Defense Communications Agency: August 1982
 36. Internet Protocol Transition Workbook; SRI International for the Defense Communications Agency: March 1982
 37. Kent, Stephen T., Peter J. Sevik and James G. Herman; Personal Authentication System for Access Control to the Defense Data Network; Bolt Beranek and Newman, Inc. for the Defense Communications Agency; IEEE Press; New York, New York; c. 1982
 38. Kleinrock, Leonard; Queueing Systems Volume II: Computer Applications; John Wiley & Sons, Inc.; New York, New York; c. 1976
 39. Kulikovskiy, Andrey Gennadievich and Grigoriy Aleksandrovich Lyubimov; Magnetohydrodynamics; Addison-Wesley Publishing Company, Inc.; Reading, Massachusetts, c. 1965
 40. Ed. by Lee, K.S.H.; EMP Interaction: Principles, Techniques and Reference Data (A Complete Concatenation of Technology from the

EMP Interaction Notes) EMP Interaction 2-1; Dikewood Industries, Inc. for the Air Force Weapons Laboratory (NTMT): December 1980; AFWL-TR-80-402

41. Lerner, Eric. J.; "Electromagnetic Pulses: Potential Crippler," IEEE Spectrum, May 1981; IEEE Press; New York, New York
42. Longmire, Conrad L.; "On the Electromagnetic Pulse Produced by Nuclear Explosions," IEEE Transactions on Electromagnetic Compatibility, Vol. EMC-20, No. 1, February 1978; IEEE Press; New York, New York
43. A letter from Longmire, Conrad L. to Col. Robert Francis, Director Air Force Weapons Laboratory: June 4, 1982
44. Martin, James; Telecommunications and the Computer, 2nd ed.; Prentice-Hall, Inc.; Englewood Cliffs, New Jersey; c. 1976
45. Ed. by Mather, Norman W. and George W. Sutton; Engineering Aspects of Magnetohydrodynamics; Gordon and Breach Science Publishers; New York, New York; c. 1964
46. McQuillan, John M. and Vinton G. Cerf; Tutorial: A Practical View of Computer Communications Protocols; IEEE Press; New York, New York; c. 1978
47. McQuillan, John M., Ira Richer and Eric C. Rosen; "The New Routing Algorithm for the ARPANET," IEEE Transactions on Communications, Vol. COM-28, No. 5, May 1980; IEEE Press; New York, New York
48. Minieka, Edward; Optimization Algorithms for Networks and Graphs; Marcel Dekker, Inc.; New York, New York; c. 1978
49. Nichols, Wesley G., Dr. Gary D. Cable, John F. Erhart and D. Lynn Kellum; Satellite Communication System Performance Assessment- Propagation Network Analysis Code (PNAC) User's Manual; Computer Sciences Corp. for the Defense Communications Agency: July 1982; 4489-C-001 (Rev. 2)
50. Nuclear Survivability Organization; EMP Assesment Methodology Program Relationship Between TEMPS and HEMPS Responses in C3 Facilities; Boeing Aerospace Co. for the Defense Nuclear Agency: October 1981; DNA 5918F
51. Price, Gary H.; "The Electromagnetic Pulse From Nuclear Detonations," Reviews of Geophysics and Space Physics, Vol. 12, No. 3, August 1974
52. Read, Ronald C.; Graph Theory and Computing; Academic Press, Inc.; New York, New York; c. 1972
53. Ricketts, L.W., J.E. Bridges and J. Miletta; EMP Radiation and Protective Techniques; John Wiley & Sons, Inc.; New York, New York; c. 1976

54. ROUT User's Guide (Grinder); Network Analysis Corp. for the Defense Communications Agency: June 1977
55. Rudie, Norman J.; Principles and Techniques of Radiation Hardening, Vol. 3, 2nd ed.; Western Periodicals Co.; North Hollywood, California; c. 1980
56. Schwartz, Mischa; Computer-Communication Network Design and Analysis; Prentice-Hall, Inc.; Englewood Cliffs, New Jersey; c. 1977
57. Selvaggi, Philip S.; The Department of Defense Data Protocol Standardization Program; The Defense Communications Agency; IEEE Press; New York, New York; c. 1982
58. Sevcik, Peter J., Graeme J. Williams and Bruce L. Hitson; Defense Data Network Survivability; Bolt Beranek and Newman, Inc. for the Defense Communications Agency; IEEE Press; New York, New York; c. 1982
59. Shercliff, J.A.; A Textbook of Magnetohydrodynamics; Pergamon Press; Oxford, England; c. 1965
60. Shreider, Yu. A.; The Monte Carlo Method; Pergamon Press Ltd.; Oxford, England; c. 1966
61. Specifications for the Interconnection of a Host and an IMP (Report no. 1822); Bolt Beranek and Newman, Inc. for the Defense Communications Agency: December 1981
62. Tannenbaum, Andrew S.; Computer Networks; Prentice-Hall, Inc.; Englewood Cliffs, New Jersey; c. 1981
63. Trybus, Paul; Ground Based Systems EMP Design Handbook; Mission Research Corp. for the Air Force Weapons Lab: February 1982; AFWL-NTYCC-TN-82-2
64. Tucker, Alan; Applied Combinatorics; John Wiley & Sons, Inc.; New York, New York; c. 1980
65. Wilkov, Robert S.; "Analysis and Design of Reliable Computer Networks," IEEE Transactions on Communications, Vol. COM-20, June 1972; IEEE Press; New York, New York
66. Williams, Larry J.; briefing on Survivability Considerations in Computer-to-Computer Communications for the Air Force Weapons Lab: January 1983
67. Wittwer, Leon A., Gregory H. Canavan and James E. Brau; CHEMP: A Code for Calculation of High-Altitude EMP; Air Force Weapons Lab: July 1974; AFWL-TR-74-49
68. Fowler, Eddie R., and Russell D. Thomas; HEMP Vulnerability/Sur-

vivability Of Computer Networks; Air Force Weapons Lab; AFWL/NTCA;
29 July 1983; Unpublished

CLASSIFIED BIBLIOGRAPHY

69. Schlegel, G.K., M.A. Messier, W.A. Radasky, and W.C. Hart, Electromagnetic Pulse Environmental Handbook (U), AFWL EMP Phenomenology I-1, January 1972, Kirtland AFB, NM, (SRD-CNDWI)
70. Sweton, J.F., DNA EMP (Electromagnetic Pulse) Handbook (U), 5 July 1979, (CONFIDENTIAL)
Volume 1 - Design Principles (U), DNA 2114H-1
Volume 2 - Coupling Analysis (U), DNA 2114H-2
Volume 3 - Component Response and Test Methods (U), DNA 2114H-3(rev)
Volume 4 - Environment and Applications (U), DNA 2114H-4
Volume 5 - Resources (U), DNA 2114H-5
Volume 6 - Computer Codes (U), DNA 2114H-6
71. Nuclear Criteria Group Secretariat Report, Air Force High Altitude Electromagnetic Pulse Criterion (U), NCGS 80-6 (Rev), September 1980 - April 1981, Kirtland AFB, NM, (SECRET)
72. Final Report - SAFEGUARD Communications Agency EMP Effects, Analysis, and Test Program (U), M 74-220, The MITRE Corporation, Bedford, MA, 30 September 1974. (SECRET)
M 74-220, Vol.1 - Summary (U), 30 September 1974. (SECRET)
M 74-220, Vol.2 - Transient Predictions and Coupling Program Analysis (U), 20 September 1974. (SECRET)
M 74-220, Vol.3 - SAFTCS Equipment/Subsystems EMP Susceptibility (U), 15 August 1974. (CONFIDENTIAL)
M 74-220, Vol.4 - SAFTCS Equipment/Subsystem EMP Susceptibility (Appendices) (U), 15 August 1974. (CONFIDENTIAL)
M 74-220, Vol.5 - Transient Predictions and Coupling Program Analysis (Appendices) (U), 15 August 1974. (SECRET)
73. Baxter, J.A., D.L. Rogers, and W.H. Roell, Project APACHE (U), Pacific AUTODIN System Performance in a HAEMP Environment (U), DNA 4012T-A-2, 30 September 1978, BDM Corporation, McLean, VA. (SECRET)
74. Schmidt, R., J.A. Baxter, J. Anderson, L. Bayliff, M. Rhoden, Project APACHE, Pacific AUTOVON System Performance in an HAEMP Environment (U), DNA 4012T-B02, 31 March 1978, BDM Corporation, McLean, VA. (CONFIDENTIAL)
75. Communication Facility EMP Assessment (U), Boeing Aerospace Co., Seattle, WA
NAVCAMS WESTPAC, Guam (U), DNA 4005T-GBG, 1 December 1979. (CONFIDENTIAL)

NAVCAMS EASTPAC, Hawaii (U), DNA 4006T-HBS, 30 November 1979. (CONFIDENTIAL)
Anderson AFB, Guam (U), DNA 4005T-GBC, 30 March 1979. (CONFIDENTIAL)
Barrigada, Guam (U), DNA 4005T-GBE, 30 November 1979. (CONFIDENTIAL)
Hickam AFB, Hawaii (U), DNA 4006T-HBE, 30 November 1979. (CONFIDENTIAL)
Makalapa, Hawaii (U), DNA 4006T-HBL, 30 November 1979. (CONFIDENTIAL)
Pearl Harbor, Hawaii (U), DNA 4006T-HBO, 31 August 1979. (CONFIDENTIAL)
Lualualei, Hawaii (U), DNA 4006T-HBJ, 31 August 1979. (CONFIDENTIAL)
Mount Kaala, Hawaii (U), DNA 4006T-HBU, 30 March 1979. (CONFIDENTIAL)
Camp Smith, Hawaii (U), DNA 4006T-HBB, 30 November 1979. (CONFIDENTIAL)

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 APPROACH TO THE DESIGN OF AN ADAPTIVE
SPREAD SPECTRUM MODEM

Prepared by: Dr. Victor S. Frost
Academic Rank: Assistant Professor
Department: Electrical Engineering Department
Univeristy: University of Kansas
Research Location: RADC, Griffiss AFB, Rome NY
USAF Research Colleague: Mr. John Graniero
Date: August 5, 1983
Contract No. F49620-82-C-0035

AN APPROACH TO THE DESIGN OF AN ADAPTIVE
SPREAD SPECTRUM MODEM

Abstract

The capability to reprogram spread spectrum (S.S.) modems on a short term basis to reflect changes in the environment will soon exist. An adaptive S.S. modem (ASSM) would insure maximum instantaneous link performance and graceful degradation of the overall system performance. This report outlines an approach to the design of an ASSM. The major design issues are identified and a research plan to address these issues is presented.

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 Rome Air Development Center, Griffiss AFB New York. The hospitality of the DCCD Section and The Communications Transmission Branch is gratefully acknowledged.

The author would also like to express his thanks to Mr. John Graniero, Mr. John Patti and Lt Michael Weir for their collaboration and guidance during the course of the summer research program.

1.0 Introduction

For many years spread spectrum (S.S.) systems have provided the solution to military communications problems(1-5). As technological advancements lower the cost of spread spectrum systems their use will become more ubiquitous. S.S. modems (modulator/demodulator) are used for jamming protection, multiple access, multipath protection, ranging, and to avoid detection by unauthorized listeners. The primary applications of S.S. systems of interest here are antijamming (AJ) and low probability of intercept (LPI). S.S. systems achieve AJ and LPI capabilities by pseudo-randomly distributing the transmitted energy over a wide range of alternatives. These alternatives include carrier frequency (known as frequency hopping, FH), phase (known as direct sequence, DS or pseudonoise, PN) and time (known as burst or time hopping, TH). Error correcting codes and diversity can be used in addition to pseudo-random transmissions to further increase the AJ and LPI capability of the modem.

Traditionally modems have been designed to operate in one environment. In most designs worst case assumptions regarding the environment are made resulting in a fixed system configuration, i.e., fixed levels of FH, PN, TH, error control coding, and diversity. In the design process trade-offs between important system parameters are made to accommodate the fixed scenario. For example, to achieve a specified level of AJ in a fixed radio frequency (RF) bandwidth limits on the link data rate must be imposed. However, the environment in which these modems must operate is highly volatile and the design assumption and trade-offs may not always match the environment resulting in a link performance that may not always be optimum. The link performance can be measured in terms of information rate, intercept range or other criteria as will be discussed later. Therefore, there is a need to develop adaptive S.S. modems (ASSM). An adaptive S.S. modem would estimate the environment and modify its configuration to optimize the link performance. A network of ASSMs would accommodate time-varying user loads and allow for the reallocation of network resources to counteract link degradations.

The capability to modify a S.S. modem's configuration as a function of time will exist with the production of the VHSIC AJ dedicated-function-chip-

set currently under development at Hughes Aircraft Co.(6) A programmable common equipment core which can accommodate both AJ, LPI, and error control features is the basis for the Hughes design. (See Reference 6 for a detailed description of the Hughes VHSIC programmable modem). In the near future it will be technically feasible to adapt a modem configuration to reflect changes in the environment.

The research questions that need to be addressed are how can the modem estimate the environment and then given knowledge of the environment, how can the modem select a new configuration. These problems become more difficult because S.S. modems do not operate alone. They are part of a larger network so each link of the network must coordinate its transmission modes. Further, the feasibility of adapting the modem configuration does not imply that such a capability could be justified in terms of an improved overall link performance. The purpose of this report is (1) to identify some potential advantages and disadvantages of an adaptive S.S. modem (ASSM), (2) to propose an approach to the design of such a modem, (3) to discuss the major issues involved in implementing the proposed design, and (4) to develop a research plan for addressing the major issues in the design of an adaptive S.S. modem.

In the following section the objectives and scope of this report will be defined. The advantages and disadvantages of an adaptive S.S. modem will be identified next. An approach to the design of an ASSM will be presented in section 4. Based on the approach the major issues which must be addressed in the design of an ASSM were identified and are discussed in section 5. This report concludes with the presentation of a research plan for determining the feasibility of an ASSM.

2.0 Objectives.

The specific objectives of this study were:

- 1) To define the concept of an ASSM.

2) To identify the potential improvement in system performance using ASSM.

3) To identify potential problems with the ASSM concept.

4) To develop a design approach for an ASSM.

5) To define the major questions which must be resolved before ASSM can be developed.

6) To outline a research plan for addressing these questions.

It is expected that an ASSM will operate as part of a network. Using adaptive links in a network poses problems in the coordination of each link's transmission mode. For example, which user sets the transmission mode and how is that information to be distributed among all the network users. Similar problems have been encountered in the development of adaptive routing algorithms for computer communications(7, 8, 9, 10) networks. The scope of this discussion will be limited to the analysis of one link, i.e. the networking problem will not be addressed. The networking problem is important and interesting; however, if an adaptive link connecting two users is not feasible then obviously there is no need to consider a network.

3.0 The Potential Benefits and Problems of an Adaptive Spread Spectrum Modem.

Military S.S. systems are designed to meet set AJ and LPI specifications. Given that current systems operate with fixed waveforms and error control coding trade-offs between the AJ, LPI, data rate, capabilities of the system must be made. For example, to achieve maximum AJ you want to transmit the maximum power possible, however, for LPI the minimum transmitter power is the optimum choice. Thus, adaptive control of the transmitter power in an intercept environment, i.e., no jamming, would be desirable while if the system is operating in a jamming environment then the maximum power should be transmitted. Consider the data rate trade-off. Both LPI and AJ capabilities

of a system are enhanced by reducing the data rate. Clearly a reduction in the data rate degrades the link performance. In a network environment the network data rate is reduced by removing users from the system. However, it is highly desirable to design the system to degrade gracefully, i.e., if the link can not support one data rate, then communications should be possible at a lower rate with degraded quality as opposed to not being able to use the link at all. Suppose at data rate r_1 the voice quality is such that the listener can recognize the talker and clearly understand the signal, at data rate r_2 the listener can understand the signal but does not recognize talker, at r_3 the talker has a limited vocabulary and synthetic speech is used. Clearly the data rate can be changed as a function of time to insure that some communications capability exists in many different environments. Only in the worst signal environment would data rate r_3 be used. If the link can adapt its data rate as a function of the environment then the user is assured that his instantaneous performance is maximized.

The above examples are intended to illustrate some of the benefits that can be derived from an ASSM. Summarizing an ASSM would provide: 1) maximum instantaneous link performance (2) graceful degradation of the system performance and 3) the ability to change the relative importance of different performance measures e.g. AJ or LPI as a function of the scenario.

The potential problems with an ASSM are significant. As previously discussed, there is a question as to how the ASSM can be connected in a network. Estimating the environment is not a trivial task. The development of receiver structures to estimate the characteristics of different jamming signals is needed. Given a receiver for intercepting and analyzing the jamming signals the technical question becomes can the estimation be accomplished fast enough to help the link performance. Knowing the signal environment an algorithm must then be developed to select the optimum link configuration. Further, would the addition of this extra receiver structure drive the cost, power, or space requirements beyond practical limits. Finally, if an ASSM is feasible, does the improvement in the average link performance gained through its deployment justify the extra cost incurred.

The potential benefits of ASSM definitely justifies further study. Even though some aspect of an ASSM might be currently unrealizable advancements in technology could be expected to overcome those problems.

4.0. An Approach to the Design of an Adaptive Spread Spectrum Modem.

An ASSM offers the possibility of changing many modem parameters e.g. power, diversity, or error control coding as a function of time. An approach is needed to specify how each parameter should change as a function of the observed environment. Such an approach is outlined below.

The initial steps in the approach are to define, 1) the modem parameters of interest, 2) the classes of jamming signals of interest, e.g. tone jammers, 3) the important characteristics of each class of jamming signal, e.g. power and frequency. Next modem performance measures, e.g. AJ margin are defined. The definition of S.S. modem performance measures is an important problem by itself. For each class of jamming signals a relationship between the modem variables, the jamming signal characteristics and the modem performance measures is established. The relative importance of each performance measure is then defined and an overall link quality metric found. A baseline link quality is established. A sensitivity study is then used to define the adaptive algorithm, i.e. given that the jammer class is known as well as its characteristics, the modem variable which will cause the biggest change in the link quality is adjusted so that the baseline link quality is maintained. If the baseline quality can not be achieved, then the next most sensitive modem variable is adjusted. After all the modem variables are exhausted then the relative importance of each performance measure can be changed and the procedure repeated. For example, a modem might initially weight LPI more than AJ then after the jamming reaches some critical level AJ would become dominant. The algorithms will be stored in the modem as a lookup table; the sensitivity analysis and variable selection will be done offline. It will be assumed that we can infer the modem performance from the jammer parameters; it will not be necessary to actually measure the modem performance.

The above procedure can be described mathematically as follows. Define a vector of modem variables V as

$$V = [v_1 \dots v_m]$$

where

$v_i = i$ th modem variable.

A list of potential modem variables is shown in Table I.

- chip rate
- hop rate
- transmitted power
- data rate
- level of diversity
- level of error control coding
- level of source coding
- time hop patterns
- duty cycle
- length of sync symbols
- acquisition algorithm
- code tracking algorithm
- number of adaptive array elements
- length of adaptive filters

Table I
Potential Modem Variables

A matrix can be used to define the jamming environment. The rows of the matrix correspond to classes of jamming signals while columns define the

characteristics of each class; that is, the matrix J has the form of

$$J = \begin{bmatrix} \gamma_{11} & \dots & \gamma_{1K} \\ \gamma_{21} & \dots & \gamma_{2K} \end{bmatrix}$$

where

γ_{ij} = jth characteristic of the ith jammer class

Not all classes of jammers have the same number of parameters so those elements of the matrix not needed can be set to zero. Suppose we are only considering single tone and partial band jammers, then the matrix J could have the form

$$J = \begin{bmatrix} \gamma_{11} & \gamma_{12} & 0 \\ \gamma_{21} & \gamma_{22} & \gamma_{23} \end{bmatrix}$$

where

γ_{11} = power in the tone jammer

γ_{12} = frequency of the tone

γ_{21} = power in the partial band jammer

γ_{22} = center frequency of the partial band jammer

γ_{23} = bandwidth of the partial band jammer

Table II contains a list of potential jammer classes and their characteristics.

Jammer Class	Characteristics
Single tone	power, frequency
Multiple tone	power, number, frequencies
Partial band noise	power, center frequency, bandwidth, noise type
Swept	power, center frequency, bandwidth, sweep rate
Pulsed	power, center frequency duty cycle

Table II
Potential Jammer Classes and Their Characteristics

The next step in the design process is the specification of a vector of modem performance measures, M , as

$$M = [M_1 \dots M_L]$$

where

$$M_i = i^{\text{th}} \text{ performance measure}$$

Many different performance measures can be suggested⁽¹¹⁾ for S.S. systems. A potential list is shown in Table III.

Communications Range
Intercept Range
Data Rate
Bandwidth
Time-to-achieve Synchronization

Table III
Potential S.S. System Performance Measures

Each performance measure is a function of both the modem variables, and the jamming environment J, i.e.

$$M = f(\gamma, J)$$

This function must be found so that the appropriate trade-offs can be made. The relationship between J, γ , and M is in general very complex. There are some cases which can be solved analytically, however, it is expected that computer simulations will be needed to completely define this function.

The overall performance of a S.S. modem is difficult to assess because of the diversity of system parameters as well as the wide range of operating environments and user requirements. However, some individual performance measures can be easily identified, e.g. communications range, intercept range, and data rate. The difficulty occurs in weighting the importance of each metric. That is, some metrics are more significant depending on the scenario. An overall link quality can be defined as a function of the performance vector as

$$Q = g(M).$$

It seems reasonable to assume the function $g(M)$ to be linear so

$$Q = \sum_{i=1}^L w_i m_i$$

where

w_i = importance weighting on the
ith performance metric

The vector $W = [w_1 \cdot \cdot \cdot w_L]$ provides the mechanism for assigning a number to the overall link quality. Note that W may also be a function of time, i.e. as the environment is changed the importance of different matrices is also likely to change, i.e. $W(J)$.

Now we can write the overall link performance as

$$Q = \alpha(\mu, W) = \alpha(\mu, J, W(J))$$

The goal of the ASSM is to maintain Q at some specified level as J changes. The adaptive algorithm must be designed to meet this goal efficiently. Note that the performance metrics are not directly measured, rather they are predicted from the jamming environment and the modem variables. One such algorithm would consist of performing a sensitivity analysis on the link quality Q relative to the modem vector V and then assigning the new V by choosing those variables which yield the greatest change in Q .

This design procedure is best illustrated using a simple example. The scenario consists of a single tone jammer at the S.S. system's carrier frequency. The S.S. system uses only PN modulation. The adaptable modem variables are the bit rate and transmitter power, i.e.

$$V = [\mu_1, \mu_2]$$

where

μ_1 = transmitter power

μ_2 = bit rate.

The jamming matrix is given by a single number, the jamming power so

$$J = [\gamma]$$

γ = Jamming power

The performance measures are the square of the communicating range for fixed probability of error, the square of the intercept range and the data rate i.e.

$$M = [m_1, m_2, m_3]$$

$$m_1 = d_R^2$$

d_R = communications range

$$m_2 = d_I^2$$

d_I = intercept range

$$m_3 = r_d$$

v_d = data rate

Notice that each performance measure must be scaled before the importance weightings are applied. One convenient scaling is to normalize each metric by its minimum value that is

$$m_j = m_j / m_{j\min}$$

To establish these minimum values the range of the modem and jamming parameters must be known. The range of the β_i 's will be known while the range of the γ_j 's must be assumed. Also note that this form of normalization removes all multiplicative constants. For this example let the range of these variables be

$$N_1 = 1, 100$$

$$N_2 = 1, 10$$

$$\gamma = 1, 1000$$

The performance measures can now be written as a function of the modem variables and the jammer characteristics as

$$m_1 = \frac{10^4 \omega_1 N_1}{N_2 \gamma}$$

$$m_2 = N_1$$

$$m_3 = N_2$$

The overall link performance is then

$$Q = \frac{10^4 \omega_1 N_1}{N_2 \gamma} - \omega_2 N_1 + \omega_3 N_2$$

Notice that weighting on the intercept range is negative, this insures that as the intercept range (m_2) increases the link performance measure decreases.

Next assume that the LPI function should be initially favored so we set

$$\begin{aligned}w_1 &= .25 \\w_2 &= .5 \\w_3 &= .25\end{aligned}$$

and the link quality becomes

$$Q = \frac{2.5 \times 10^3 \omega_1}{\omega_2 \gamma} - .5 \omega_1 + .25 \omega_2$$

A baseline link performance can be found by setting $\omega_1 = 1$, $\omega_2 = 1$ and $\gamma = 1000$, this performance level is 2.25.

The information now available allows us to specify how the modem should change its variables as the jamming power is increased. Through a sensitivity analysis, it is found that the bit rate, ω_1 should be changed first, i.e. Q is most sensitive to ω_1 . Figure 1 shows how the bit rate should be changed in response to increasing jammer power with the transmitter power fixed at 1. It is this function that would be stored in the modem and the adaptive algorithm would simply measure the jammer power then set the bit rate. This process would probably be done in steps rather than continuously.

As the jammer power exceeds 1000 the link performance can no longer be maintained at 2.25. This might indicate that the jammer is directing all of his resources directly at the receiver and thus weighting LPI more heavily is no longer valid. In this case the weighting vector should be changed and the procedure repeated. A new function would be stored to tell the modem how to adjust the transmitter power as γ increased further.

In this section a procedure for designing an ASSM was presented. A mathematical framework was developed to describe the procedure. A simple example was presented to illustrate the approach. Based on the framework

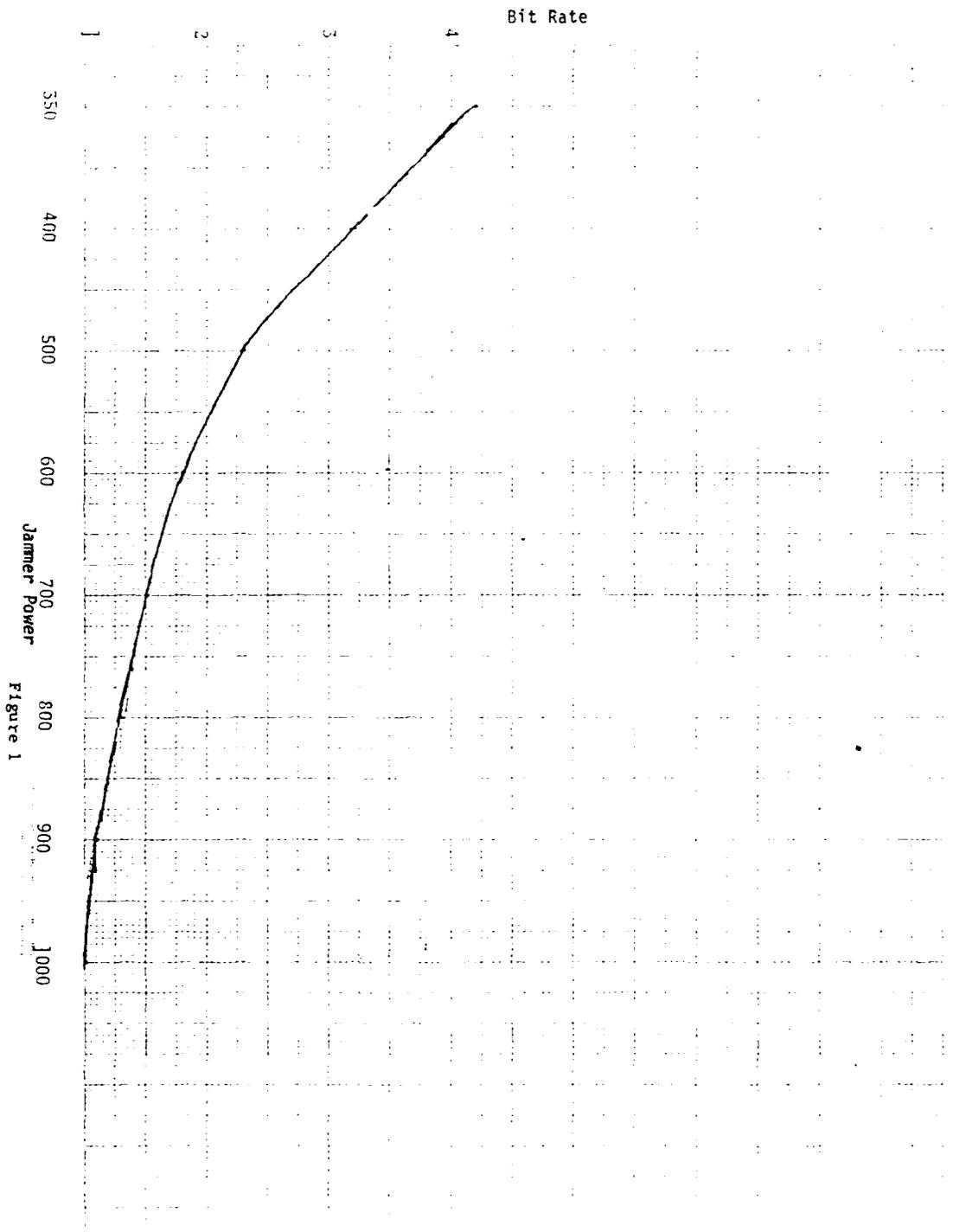


Figure 1

developed here the major issues which must be addressed in applying the design procedure have been identified and are discussed in the next section.

5.0 Adaptive Spread Spectrum Modem Design Issues.

The purpose of this section is to clearly define the major points of concerns regarding the development of an ASSM.

1) What does short term adaptability buy?

The point of concern here lies in the possibility that a prudent S.S. modem design can be developed which will achieve the desired link performance given all possible scenarios. The overhead and cost associated with an ASSM would then become unnecessary. It is known that to maintain LPI power control is essential so some level of adaption will be needed. Whether power control can be best handled manually or automatically is unknown.

2) What are appropriate S.S. modem performance measures?

There are a wide variety of performance metrics for S.S. modems. These range from voice quality which is subjective to probability of a bit error which is quantitative. The set of metrics which are appropriate for ASSM design is unknown. Further, the question of S.S. modem performance evaluation is unresolved in general.

3) What classes of jammers need to be considered and what are their critical characteristics.

There are many different kinds of jammers, each with different parameters. To sufficiently characterize the environment for an ASSM which classes can be combined. Or put another way should the environment be perceived not as a composite of independent jammers, rather as one signal with certain properties.

4) How can the characteristics of the environment be estimated?

Jamming signals operate well above the noise so conventional spectrum analyzers or compressive receivers can be used for detection. Once detected what receiver structures are needed to estimate the parameters of the jammers? The development of an intercept receiver for jamming signals with feature extraction capabilities is essential to the ASSM concept. Many of the components of an intercept receiver are contained within the VHSIC chip set under development by Hughes. It would be of interest to investigate the capability of the chip set as an intercept receiver for jamming signals. Independent of the ASSM design, a network of inexpensive intercept receivers for jamming signals would provide an important global view of the jamming environment.

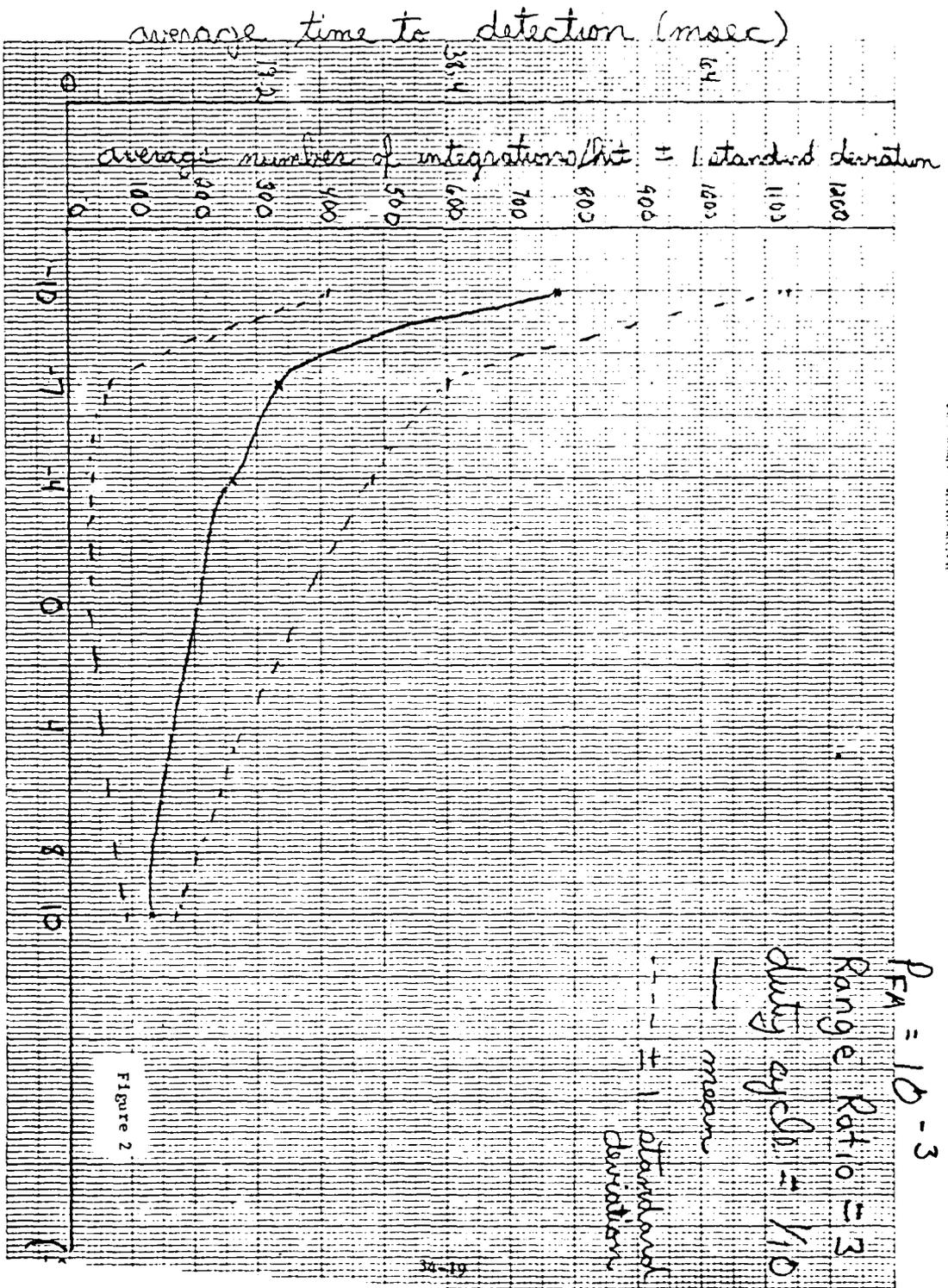
5) How can the relationship between the modem variables, jamming parameters and the performance metrics be found?

This will be the most difficult aspect of ASSM design but these relationships will have applications beyond the ASSM. Understanding these relationships is essential to prudent S.S. modem design in general. Theoretical analysis can provide these relationships in some cases. For example, the case of multiple tone jamming has been analyzed for direct sequent systems(12). However, the complexity of the S.S. waveforms will hinder analytic analysis and in some cases totally preclude the theoretical approach. In these cases computer simulation will be used to obtain the needed relationships. For example, figure 2 shows the result of a simulation study where a time-to-intercept was related to the S.S. modem's input (S/N) for the case of a scanning radiometer intercept receiver. A complete description of this simulation study is presented in Appendix A.

6) How can the overall link performance be defined?

Within the context of the developed approach this issue becomes a question of how to select the importance weightings, W.

In the following section a research plan is presented for addressing these issues.



6.0 Recommendations.

This section describes a study which addresses the issues raised in the previous section. Refer to figure 3 for the relationships between the various tasks.

6.1 Adaptive Spread Spectrum Modem Feasibility Study.

A set of tasks are outlined below whose goal is the determination of the feasibility of an ASSM.

6.1.1 Review Spread Spectrum Performance Measures.

There are many performance metrics used for S.S. modems. Some of these are redundant while others are not applicable to ASSM design. The purpose of this task is to catalog all S.S. performance measures and to develop any relations between them. A common framework for S.S. performance metrics will be the result of this task.

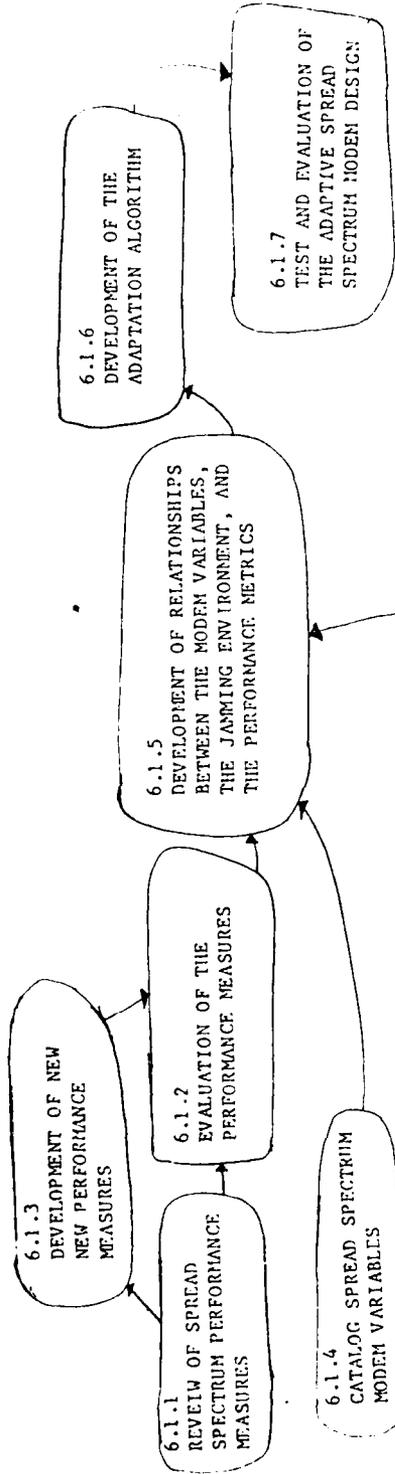
6.1.2 Evaluation of Performance Metrics

The goal of this task is the determination of the suitability of the S.S. performance metrics for ASSM design. It is expected that some performance metrics will naturally fit into ASSM design while others will not. The result of this task will be a list of performance metrics for use in ASSM design.

6.1.3 Development of New Performance Metrics.

From the review and evaluation of existing S.S. performance metrics it is possible that new metrics can be developed. The goal of this task will be to propose new performance metrics for S.S. modems as needed. These metrics will be suitable for use in ASSM design.

6.1 ADAPTIVE SPREAD SPECTRUM MODEN FEASIBILITY STUDY



6.2 DEVELOPMENT OF EFFICIENT RECEIVER STRUCTURES FOR THE ANALYSIS OF JAMMING SIGNALS

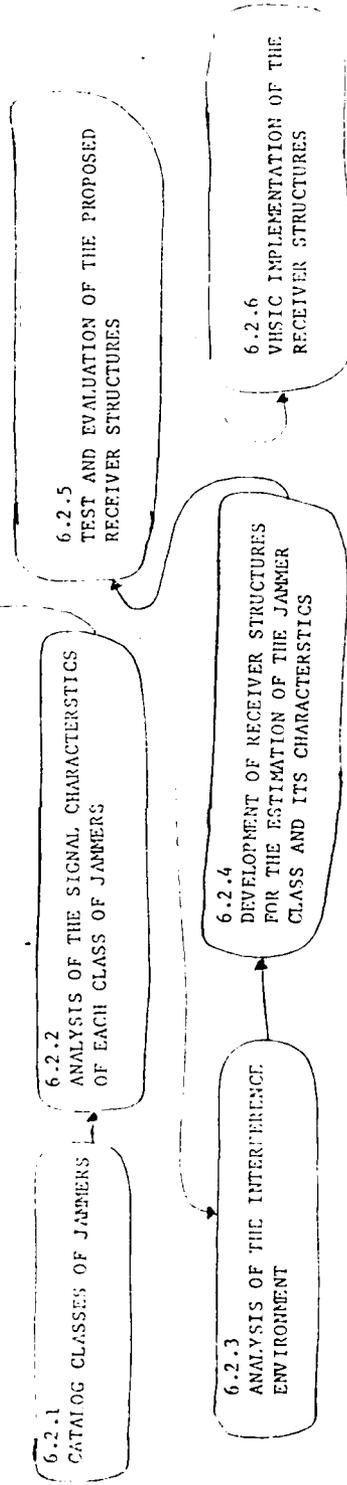


Figure 3

6.1.4 Catalog Spread Spectrum Modem Variables.

The purpose of this task is to compile a list of S.S. modem variables that can be used in ASSM design. This list will be analyzed and only those suitable to ASSM design retained. The result of this task will be the list of S.S. modem variables that will be used in the ASSM design.

6.1.5 Development of Relationships Between the Modem Variables, the Jamming Environment and the Performance Metrics.

The goal of this task is to develop the relationships needed to infer the S.S. modem performance given the modem configuration and the jamming environment. For simple scenarios these techniques exist. For more complex environments, e.g. combinations of different jammers, computer simulation will be used to obtain the desired relationships. The result of this task will be a set of design curves relating the jamming environment to the modem performance given a set of modem variables. (See figure 2 for an example).

6.1.6 Development of the Adaption Algorithm.

Given the design curves developed in the previous task an adaption algorithm will be developed. In this task the procedure for updating the modem configuration will be developed. As discussed previously this development will require a sensitivity analysis. The result of this task will be a set of look up tables in which each set of jamming characteristics indicates the appropriate set of modem parameters.

6.1.7 Test and Evaluation of the Adaptive Spread Spectrum Modem Design.

This task has two goals: first, to test the ASSM algorithms to verify they perform as expected; second, to evaluate the benefit of ASSM. That is, the improvement in overall link performance gained through the use of an ASSM over a non-adaptive S.S. modem will be determined. The result of this task will be a recommendation as to the feasibility of the ASSM concept.

6.1.8 Required Facilities and Expertise.

To successfully complete this feasibility study, expertise is required in spread spectrum systems and their operating environment. Further, several aspects of this study rely on computer simulation of S.S. systems and jamming environments. Thus a communications system simulation capability with emphasis on S.S. systems is needed. Assuming the above requirements are net minor, effort should be devoted to the development of additional analysis and simulation tools.

6.2 Development of Efficient Receiver Structures for the Analysis of Jamming Signals.

A set of tasks is outlined below whose goal is the development of efficient techniques to estimate the jamming environment in which the ASSM must operate. The emphasis of these tasks is on the exploitation of VHSIC technology for estimating the jamming environment.

6.2.1 Catalog Classes of Jammers.

The goal of this task is to analyze different jammers and to separate them into classes. Each class will generate signals with common characteristics.

6.2.2 Analysis of the Signal Characteristics of Each Class of Jammers.

The purpose of this task is to identify the characteristics of each class of jammers. It is these characteristics which the ASSM must estimate. The product of this task will be a list of signal characteristics for each class of jammers.

6.2.3 Analysis of The Interference Environment.

Another way to approach the problem of analysis of jamming signals for ASSM is to focus on the interference environment as a whole instead of attempting

to pick out individual jammers and their parameters. The goal of this task is to determine characteristics of the interference environment which can be used to infer modem performance. For example, a complete power spectral density of interference environment might be sufficient to infer the modem performance. The result of this task will be a characterization of the interference environment sufficient to predict the S.S. modem's performance. A similar analysis for line-of-sight microwave links has recently been performed(13).

6.2. 4 Development of Receiver Structure for the Estimation of the Jammer Class and its Characteristics.

The purpose of this task is to develop signal process structures suitable for estimating 1) the type of jammers present in the environments and 2) the parameters of those jammers. Standard intercept techniques will be considered. These receiver structures should take advantage of the fact that the jamming signals will be operating well above the noise floor. The receiver structures will be developed with capabilities of the Hughes VHSIC programmable modem in mind. The result of this task will be a set of receiver structures.

6.2.5 Test and Evaluation of the Proposed Receiver Structures.

The goal of this task is to evaluate the performance of the proposed receiver structures. This evaluation will be performed using analytical techniques where appropriate. However, because of the complexity of the environment it is expected that computer simulation will be needed to completely test these receivers. The result of this task will be verification of the performance of each receiver structure.

6.2.6 VHSIC Implementation of the Receiver Structures.

The purpose of this task is to evaluate the feasibility of implementing the developed receiver structures using the Hughes VHSIC programmable modem. It is expected that this programmable modem will have the flexibility necessary to implement some of the proposed receiver structures. The product of this

task will be a VHSIC implementation of receiver structures for analyzing jamming signals.

6.2.7 Required Facilities and Expertise.

To successfully complete the VHSIC design of receivers for this analysis of jamming signals expertise is in the area of intercept receivers and the nature of jamming signals. Again computer simulation will be needed to evaluate the proposed designs, so experience in the simulation of intercept receivers and jamming signals is also needed. Familiarity with the capabilities of the Hughes VHSIC programmable modem is also required. It is expected that new tools in the form of new simulation software will be needed to complete these tasks.

7.0 Conclusions.

The possibility of adapting the parameters of an S.S. modem to changing jamming environments has been explored. The possible advantages and problems with an adaptive S.S. modem have been addressed. An approach to the design of an ASSM was presented and the major design issues discussed. Finally a research plan was proposed for determining the feasibility of an ASSM. Even though significant difficulties can be foreseen in the development of an ASSM, the opportunity to insure optimum link performance in a very wide range of environments justifies its study. Further, many of the issues which must be addressed in this development also apply to other aspects of S.S. systems design and analysis.

REFERENCES

- (1) R. Dixon, Spread Spectrum System, Wiley, New York, 1976.
- (2) R. Dixon, Spread Spectrum Techniques, IEEE Press, 1976.
- (3) IEEE Trans. on Communications - Special Issue on Spread Spectrum Communication, COM-25, August 1977.
- (4) IEEE Trans. on Communications - Special Issue on Spread Spectrum Communication, COM-30, May 1982.
- (5) J. K. Homes, Coherent Spread Spectrum Systems, John Wiley, New York, 1982.
- (6) A. N. Chester et al, "VHSIC, Phase I" Vol I, Interim Technical Report, CDRL A002, Hughes Aircraft Co., Nov 1982.
- (7) A. J. Tanenbaum, "Computer Networks", Prentice-Hall, Englewood NJ 1981.
- (8) D. J. Baker et al, "Distributed Network Reconfiguration in Response to Jamming at HF", MILCOM '82, Boston MA, Oct 17-20, 1982, paper 23.2.
- (9) W. Tan, "Adaptive Protocol for Worsening Communication Conditions", MILCOM '82, Boston MA, Oct 17-20, 1982, paper 16.3
- (10) K. Mase and M. Imase, "An Adaptive Capacity Allocation Scheme in Telephone Networks", IEEE Trans. on Communications, Vol COM-30 No. 2, Feb 1982 pp 354-358.
- (11) G. Huang et al, "A Spread Spectrum Analysis and Design Laboratory System", MILCOM '82, Boston MA, Oct 17-20, 1982, paper 3.1.
- (12) L. B. Milstein et al, "The effect of Multiple Tone Interfering Signals on a Direct Sequence Spread Spectrum Communication System", IEEE Trans. on Comm, Vol. COM-30, No. 3, March 1982, pp. 436-445.
- (13) R. E. Bozek, "ECM Signal Discrimination and Identification Techniques", Final Technical Report, PAR Technology Corporation, PAR Report # 83-18, Jan 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

QUANTIZATION ERROR ANALYSIS FOR THE DEANZA IMAGE PROCESSOR

Prepared by:	Dr. Patrick Garrett
Academic Rank:	Adjunct Professor
Department and University:	Electrical Engineering University of Cincinnati
Research Location:	Air Force Wright Aeronautical Laboratory, Advanced Systems Research Group, AAAT-3, Avionics Laboratory, WPAFB, OH 45433
USAF Research:	LtC B. R. Altschuler
Date:	August 19, 1983
Contract No:	F49620-82-C-0035

Quantization Error Analysis For The DeAnza Image Processor

Patrick Garrett

ABSTRACT

The prevailing time-amplitude step-interpolator A/D quantization process offers a quantitative basis for relating the DeAnza image processor system specifications to image accuracy through derivation of a quantization levels/noise relationship. It is shown that significant noise is generated by system quantization which is deterministic, and that pixel detection and measurement are influenced accordingly. Independent frame averaging is shown to be equivalent to increasing the number of quantization levels available from higher sample rates with a corresponding reduction in quantization noise. The DeAnza obtains 1.7 samples per pixel for a 512 x 512 pixel frame which is inadequate for reliable threshold detection until four frames are averaged to provide 1-bit resolution. Thirty frame averages are determined to provide 4-bit resolution which is necessary to resolve the 10 grey levels from an image-intensifier equipped Vidicon. Further, averaged resolution cannot exceed the A/D converter wordlength because of data quantization truncation at the LSB value, but for the DeAnza resolution is limited to 7 bits following 256 averages maximum for a 512 x 512 pixel frame. This analysis demonstrates one element of a total system error budget whose investigation is recommended.

ACKNOWLEDGEMENT

The author wishes to thank LtC Bruce R. Altschuler for providing encouragement and perspectives helpful in pursuing tasks while associated with his Advanced Systems Research Group for the summer. I am equally indebted to Dr. Donald Moon for providing the opportunity for this summer appointment to the AAAT Branch, AFWAL, at Wright- Patterson Air Force Base. Many helpful discussions were also exchanged with Capt Kevin Rose of the same organization.

The author also wishes to acknowledge the Air Force Systems Command, the Air Force Office of Scientific Research, and the Southeastern Center for Electrical Engineering Education for the support which made this summer appointment possible.

1.0 INTRODUCTION

A goal in the design of signal processing systems is to achieve transparency with regard to the hardware elements of the system. The process of A/D conversion in a sampled-data system is characterized by the step-interpolator representation of Figure 1. Developing an understanding of this process in the context of the DeAnza image processing system, primarily evident as additive quantization noise, provides the limits within which data can be meaningfully interpreted. This was chosen for analysis because it represents a significant error source in a major system component, and demonstrates the value for one element of a total system error budget.

2.0 OBJECTIVES

The requirement for modularization in signal processing systems frequently results in cascaded implementations each with individual input/output subsystems. This has demonstrated the need for quantitative accountability to enable improved system identification and design optimization. Required is the development of comprehensive system and component models with methods for combining them which advance the understanding of computer data acquisition- conversion- recovery functions.

This specific investigation is limited to analysis of the data quantization process and the influence of sampled-data parameter selection upon it. Benefits from extension of this task to a total system error analysis include the determination of accurate system characterizations useful for the specification and audit of computer interfaced systems,

and indication where system improvement is most productive in achieving the overall performance of interest.

3.0 DETECTION VERSUS MEASUREMENT

The detection of a signal in noise is concerned with the selection of a threshold value which minimizes both false alarm and false rest probabilities (signal and noise confusion). Closed-form expressions for these probabilities are well known, such as those of Neyman-Pearson, and provide useful decision criteria concerning the presence or absence of the signal. The process of estimation is more complex than detection because more information must be acquired in decisions between many alternatives to measure the value of the amplitude or phase of a signal in noise. In practice, this requires a substantially higher signal-to-noise ratio than detection with results predictable, for example, by Bayes' Rule.

Both detection and estimation theory are historically limited to classical communications applications where noise is external and generally uncontrollable. In the case of digital image processing, however, substantial noise is also generated by the system quantization process and will influence results depending upon selection of the sampled-data parameters. This deterministic relationship applies whether detection or measurement of signal parameters is of interest and may be viewed as a matter of degree. Consequently, an understanding of the quantization process is essential for prediction of system performance.

4.0 QUANTIZATION DEFINITIONS

Spatial resolution is a measure of the number of individual quantities which can be separately determined, and amplitude resolution a measure of the smallest amplitude value to which a quantity can be determined. In a data-conversion system the latter is influenced by the effect of the conversion period T on amplitude quantization. Quantization of a sampled analog signal involves the assignment of a finite number of amplitude levels corresponding to discrete values of voltage increasing from zero to some full-scale value V_{FS} . For an A/D converter having an n -bit binary wordlength, there exist 2^n assignable levels between 0 and V_{FS} each of spacing $q = V_{FS}/2^n$ volts. Although an n -bit A/D wordlength is necessary for n -bit resolution, it is not sufficient. This is a consequence of each specific application and the dynamic signal conditions which the hardware encounters. Actual A/D converter output resolution is determined by the following considerations.

The step-interpolator is the way data are handled internally in digital processors, whereby the present sample is current data until a new sample is acquired. Equation (1) illustrates that a peak-to-peak signal will acquire a ΔV amplitude quantization during each conversion period T , where the input signal rate-of-change dV/dt is proportional to its bandwidth BW . Equation (2) is an extension of this quantization relationship and describes the actual number of uniformly spaced quantization levels $V_{FS}/\Delta V$ realized between zero and the full-scale

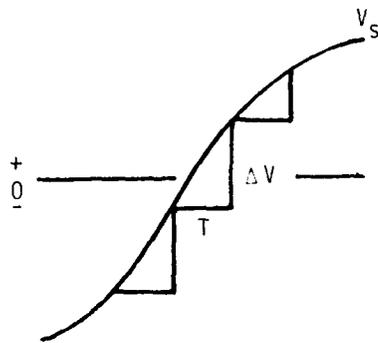


FIGURE 1. Step-Interpolator Signal Representation

$$\Delta V = T \frac{dV_s}{dt} \quad (1)$$

$$= T \frac{d}{dt} \frac{1}{2} V_s \sin 2\pi BWt \Big|_{t=0}$$

$$= T \pi BW V_s$$

$$2^n = \frac{V_{FS}}{T \pi BW V_s} \quad (2)$$

$$\text{SER}_{\text{dB}} = 10 \log \left(\frac{2^n q / 2 \sqrt{2}}{q / \sqrt{12}} \right)^2 \quad (3)$$

$$= 6.02n + 1.78$$

signal amplitude value V_{FS} in a binary encoded system. Resolution in binary-bits n follows from the argument of 2^n .

With the prevailing uniform quantizing algorithm and an input signal value between any two levels, which spans q volts, then the specific value is encoded as the digital equivalent of the closest level with a quantization error which can range up to $+ q/2$ volts. This error constitutes an irreducible quantization noise of $q^2/12$ watts added to the digital output signal, which may be characterized as a normal probability density function similar to Gaussian noise. A signal-to-error ratio (SER) is defined by equation (3) which expresses this quantization noise in terms of output power dB, and describes the maximum achievable output quality for an A/D converter with a noiseless input signal. This rms-signal-power-to-rms-noise-power ratio and its relationship to quantization is tabulated in Table 1 for several binary resolution values. Significant are the equivalents between quantization SER in dB describing digital image quality, and the average number of amplitude levels describing the accuracy with which the data may be interpreted in digital form. Together these quantities offer a quantitative basis for relating system specifications to image accuracy.

Binary Bits (n)	Amplitude Levels (2^n)	Amplitude Error (% FS)	Output SER (dB)
1	2	50.0	8
2	4	25.0	14
3	8	12.5	20
4	16	6.2	26
5	32	3.1	32
6	64	1.6	38
7	128	0.8	44
8	256	0.4	50

TABLE 1. Binary Quantization Relationships

5.0 DEANZA APPLICATIONS

When sweep rates are coordinated with Vidicon granularity the minimum pixel may be defined as 130 nanoseconds in width. The DeAnza samples this video data with a conversion period T of 78 nanoseconds and acquires 1.7 uniformly spaced samples per pixel for a 512 x 512 pixel frame. Although each sample is encoded with an 8-bit A/D wordlength, the relationships of equation (2) provide an average signal quantization of only $2^n = 1.7/\pi$ levels per pixel between 0 and V_{FS} as described in Table 2. This is so because of the high rate-of-change of the video signal, which in fact must be sampled 256π times its bandwidth to achieve true 8-bit resolution. However, $1.7/\pi$ levels is inadequate for detection purposes and must be raised to at least 2 levels, or 1 binary bit, for reliable threshold determination.

If the DeAnza were partitioned into 256 x 256 contiguous pixels per frame, then $7/\pi$ levels would be available per pixel and the minimum requirement for reliable pixel detection met. However, this improvement is at the expense of spatial resolution. Clearly, increased resolution is required to provide reduced image uncertainty, and methods for achieving this are of interest.

Digital signal averaging provides a finite geometric series in the time domain which is the equivalent of sinusoidal interpolation in the frequency domain. Data averaging can effectively improve signal quality if the necessary a posteriori signal processing delay is permissible.

Independent frame averaging provided by the DeAnza is equivalent to increasing the average number of quantization levels which are available from a higher sample rate. This improvement is evident as a corresponding reduction in quantization noise whose value may be determined by substitution of equation (2) into (3). For example, four averages of a 512 x 512-pixel frame provides 1-bit resolution and is the minimum necessary for reliable detection.

Averaging thirty 512x512-pixel frames provides a per pixel resolution of 4 bits from 16 average quantization levels (30 times 1.77 levels per pixel). This enables a 26 dB image quality from Table 1 which is adequate to resolve the 10 grey levels available from an image intensifier equipped Vidicon. Note that this process is entirely deterministic instead of stochastic, and that signal averaging improves image accuracy by a factor equal to the sum of the specific values of equation (2) for the number K of pixel/frames averaged as illustrated in the model of Figure 2. Further, the averaged resolution cannot exceed the A/D converter wordlength because of truncation of data quantization at its LSB value. For an 8-bit A/D converter the LSB is proportional to 0.4% of V_{FS} . Therefore, an K equal to 467 512x512-pixel frames is required for the limiting resolution of 8 bits. However, since the DeAnza is limited to 256 averages the maximum achievable resolution is 7 bits.

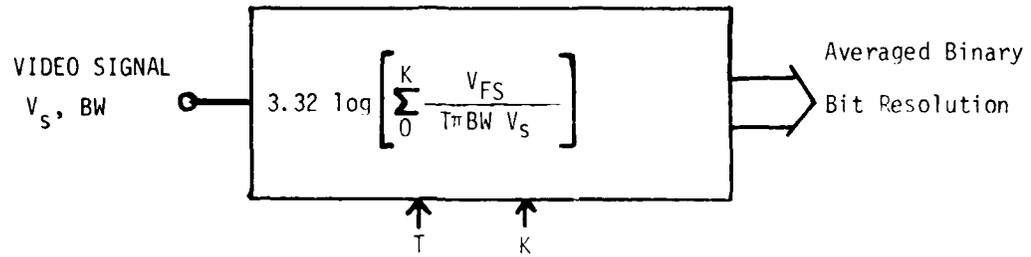


FIGURE 2. DeAnza Quantization - Averaging Model

PIXELS	WIDTH	BW	SER	LEVELS	RESOLU.	4 AVE	30 AVE	256 AVE
512x512	0.13 μ s	7.8MHz	-4dB	1.7/ π	1/4 bit	1 bit	4 bits	7 bits
256x256	0.52 μ s	1.9MHz	8dB	7/ π	1 bit	3 bits	6 bits	8 bits
128x128	2.0 μ s	0.5MHz	20dB	24/ π	3 bits	5 bits	8 bits	8 bits

TABLE 2. DeAnza Quantization Per Pixel

6.0 RECOMMENDATIONS

Improvement in quantization can also be effected by an appropriate transformation of the input signal using a preprocessor. Even spacing between amplitude levels is not necessarily ideal for the requirements of image processing. For example, input signal compression prior to A/D conversion increases the number of quantizing levels at lower signal amplitudes, resulting in a decrease in quantization noise at these levels. This is a practical option for signal statistics whose amplitude frequency of occurrence is not uniform over the 0 to V_{FS} input range, but instead concentrated at specific amplitude values.

Future system designs could benefit from such nonlinear quantization as illustrated in Figure 3, which compares a conventional linear with a Gaussian characteristic. Benefits include maintaining output SER nearly constant over a 100/1 input signal range if the match to signal statistics is reasonably accurate. This corresponds to approximately a 2-bit resolution improvement without averaging. Implementation would be by an analog logarithmic function prior to A/D conversion. However, an antilog operation is required to restore signal linearity following signal processing.

Table 3 describes a recommended partitioning of modeling and analysis methods for total system error determination. Each element included in this error budget is represented by an appropriate model, as demonstrated for quantization, which by value of its contribution also

indicates where performance improvement is beneficial. The principal concerns of these analyses are: (1) signal quality and its necessary upgrading due to interference, (2) electronic device errors, and (3) the merits of various system design options resulting from choices among available alternatives.

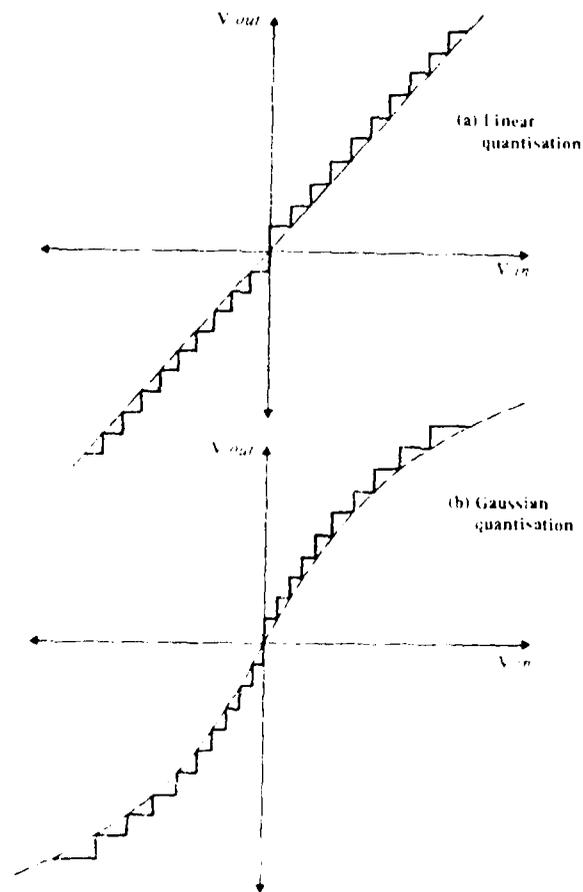


FIGURE 3. Linear and Gaussian Quantization

Subsystem	Model	Error	Description
Acquisition	Analog Components	ϵ_a	Sensor, amplifier, filter, multiplexer errors.
	Signal Quality	ϵ_b	Error following signal conditioning upgrading.
Conversion	Digital Components	ϵ_c	Sample hold, A/D, computer errors.
	Sampling	ϵ_d	Quantization, aliasing, aperture, sin X/X errors.
Recovery	Output Components	ϵ_e	D/A, output interpolator errors.
	Reconstruction	ϵ_f	Resolution following output signal smoothing.
Uncertainty	Combined RMS Total	ϵ_{FS}	One standard deviation σ

Table 3. Total System and Device Error Budget

REFERENCES

1. K.G. Beauchamp and C.K. Yuen, Data Acquisition for Signal Analysis, Allen and Unwin Publishers, London, c. 1980
2. P.H. Garrett, Analog I/O Design, Reston Publishing Company, Virginia, 1981
3. P.H. Garrett, Analog Systems for Minicomputers and Microprocessors, Reston Publishing Company, Virginia, c. 1978
4. B.M. Gordon, The Analogic Data-Conversion Systems Digest, Analogic Corporation, c. 1977
5. D.F. Hoeschele, Analog-to-Digital Digital-to-Analog Conversion Techniques, John Wiley, New York, c. 1968
6. E. Zuch, Data Acquisition and Conversion Handbook, Datel-Intersil Corporation, c. 1979

PILOT WORKLOAD AND G-STRESS

by

Dr. Richard T. Gill

ABSTRACT

The purpose of this study was to assess the effects of G-stress, in particular +Gz, on pilot workload. However, due to time constraints the actual experiment has not yet been completed. It will be conducted in 3 phases (static training, dynamic training, and data collection) and will employ two different techniques (primary task performance and subjective ratings) to assess the impact of G-stress on pilot workload. In the static training phase, subjects will practice solving two-dimensional mazes in a normal +1Gz environment. For the dynamic training phase, subjects will continue to practice solving two-dimensional mazes but in a +Gz environment ranging from +1Gz to +6Gz in the AFAMRL human centrifuge, the Dynamic Environmental Simulator. The data collection phase will be a repeat of the dynamic training phase with an additional experimental condition wherein the subject will be provided with the optimal maze solution and merely have to trace the cursor through the solution path. Both the time required to solve a maze and subjective ratings will then be used to assess the effect of +Gz on pilot workload. In addition, comparisons of maze solution times, both with and without the solution provided, between the various levels of +Gz, will provide a measure of the effects of G-stress on the pilot's cognitive abilities.

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

PILOT WORKLOAD AND G-STRESS

Prepared by: Dr. Richard T. Gill

Academic Rank: Assistant Professor

Department and University: Systems Engineering Department
Wright State University

Research Location: Air Force Aerospace Medical Research Laboratory
Biodynamics & Bioengineering Division
Acceleration Effects Branch

USAF Research: Mr. William Albery

Date: 10 August 1983

Contract No: F49620-82-0035

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 worthwhile and interesting summer at the Air Force Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio. I would like to acknowledge the laboratory, in particular the Acceleration Effects Branch for its hospitality and stimulating work environment. Finally, I would like to thank Bill Albery, Robert Van Patten, Sharon Ward, and Tom Eggemeier for their collaboration and endless support, and Rachel Birtle for her secretarial support.

1. INTRODUCTION

Technological advances in recent years have considerably increased the complexity of fighter aircraft cockpits. The number of individual parameters which the pilot must monitor and control has increased dramatically. A fundamental consequence of these advances has been to significantly alter the pilot's role from primarily a skilled manual control operator, to being more of an executive manager or decision maker.

Such alterations in the pilot's tasks have created an additional constraint or problem for design engineers, namely pilot workload. As pilot workload increases, not only does he become fatigued more readily, but his performance begins to deteriorate. Excessive pilot workload can result in some piloting tasks not even being performed with the potential consequences being catastrophic. Clearly, pilot workload is a crucial parameter which must be addressed in the design of modern fighter aircraft.

Concurrent with the advances in aircraft avionics have been major advances in propulsion, aerodynamics, and airframe materials. As a result, the modern fighter aircraft is capable of maneuvers with considerably higher G-levels and G-onset rates. For example, the F-4 is rated at approximately 7.5 G's with an onset rate of approximately 2-4 G's per second; while the F-16 is capable of 9.3 G's and 6 G's per second. These excessive G-levels and G-onset rates have been cited as the cause of numerous F-16 accidents due to pilot loss of consciousness. This problem will become even more acute with the next generation fighters which are expected to exceed the G capabilities of the F-16.

Although the increased G environment of modern fighter aircraft may appear to be independent of the pilot workload problem, it probably is not. First of all, in order to maintain an adequate field of vision and/or consciousness at higher G levels, pilots must perform an M-1 or L-1 straining maneuver. This coordinated muscular activity is an additional piloting task which requires some mental attention thus the potential for increasing pilot workload. In addition, the G-induced, reduced blood flow to the brain could impair higher level mental functioning or cognitive activity. This in turn would decrease the pilot's mental processing efficiency, thereby making it more difficult for him to adequately complete all the necessary piloting tasks.

Pilot workload and pilot G-stress are two very important issues in the design of modern fighter aircraft. Although the physiological effects of G-stress have been studied for years by both the aerospace and medical communities, little is known about the psychological effects or costs associated with G-stress. An issue of immediate concern is the potential interaction of G-stress with pilot workload.

II. OBJECTIVES

The objective of this research effort was to develop an experimental design to empirically assess the effects of acceleration on pilot workload. Since a preliminary literature review had revealed that little or no research had been conducted in this general topic area, the primary goal was to determine if G-stress does, in fact, significantly effect pilot workload, and if so, to establish the relationship between G-stress and workload. It was decided to limit this initial investigation to +Gz only, since it is the most severe acceleration environment in which pilots must function.

In pursuit of this primary goal of assessing the relationship between +Gz stress and pilot workload, the following subgoals were identified:

- (1) Adapt a technique recently developed at AFAMRL/HE, two- dimensional maze solving, as the primary means of assessing workload.
- (2) Investigate alternative workload assessment techniques with respect to the feasibility of using them in a high G environment and in conjunction with the maze solving technique.
- (3) Determine the experimental G profile, including maximum G- level and durations, to be used.
- (4) Develop the experimental protocol including the details of the experimental design.
- (5) Conduct pilot study, as necessary, to finalize the experimental design.

III. APPROACH AND RESULTS

G-Stress: A literature review on the effects of acceleration was conducted. Since this is such a broad topic area, the literature review was restricted to research involved in: (1) determining human tolerance limits to high-sustained G; (2) assessing the effects of G-stress on pilot workload; and (3) determining the effects of G-stress on the human's higher mental functioning or cognitive abilities. No attempt was made to review the literature with respect to the physiological effects of G-stress. Although such

effects are extremely important, and represent a significant proportion of the literature, it is not directly relevant to this research effort.

As noted above, one of the objectives of the literature review was to determine human tolerance to G-stress. More specifically, tolerance to +Gz was desired since this investigation would only involve acceleration in that direction. There has been considerable research effort focused on this specific area. For example, a 1974 paper entitled "Man at High Sustained +Gz Acceleration: A Review" (1), cited over thirty references; a more general review by Collyer (2) referenced over sixty sources. Even with such extensive research, absolute human tolerance limits cannot be specified.

There are numerous contaminating variables which prohibit specifying absolute human tolerance limits to +Gz acceleration. Obviously both within and between subject variance is a major problem. Additional contaminating variables identified in the literature were: (1) seat tilt back angle [i.e. 0°, 13°, 30°, etc. (3,4,5)]; (2) subject posture [i.e. feet down vs feet horizontal (5,6,7)]; (3) use of G- protection [G-suit, M-1 or L-1 maneuver, positive pressure breathing, etc. (6,8)]; (4) degree of subject experience or training at high sustained G (1,9,10); and (5) how tolerance is operationally defined [peripheral light loss, central light loss, loss of consciousness, etc. (5,6,9)].

Since the specification of a reliable +Gz tolerance vs exposure duration curve was infeasible, an alternative approach was taken. As noted earlier, one of the goals of this research effort was to use two-dimensional maze solving as a technique for assessing pilot workload (the details of this task will be discussed later). Previous research (11,12) had shown typical solution times to be approximately one-minute in a nominal 1G environment. It was then estimated that subjects should have a minimum of two minutes solution time available to them while under the influence of high sustained G. Thus the approach taken was to identify the maximum likely G-level which subjects could endure for two minutes with the G protection techniques of the modern fighter aircraft (i.e. 30° seatback angle, G-suit, and M-1 or L-1).

A detailed review of all the specific papers that influenced the determination of the maximum G-level that could be used in this research would not be fruitful. Rather, the interested reader is referred to the most relevant ones (1,2,4,5,7,8,9,13,14,15). Typical tolerance limits reported ranged from +5Gz for 240 seconds (14) to +9Gz for 45 seconds (15). It was concluded that only very highly trained subjects would be able to reliably

sustain +7Gz for two minutes, thus, +6Gz was selected as the upper limit for this research.

The second and third objectives of the acceleration literature review were to address the effects of +Gz on pilot workload and more specifically its impact on the pilot's cognitive abilities. No literature was found that specifically addressed the issue of G-stress on pilot workload. In fact, very little work has been done which addresses the effects of G-stress on the pilot's cognitive abilities. In Collyer's very thorough 1973 review (2), he cited several studies (16,17,18,19,20) which suggested increased G-stress has a negative effect on pilot's cognitive abilities. However, he concluded that knowledge in this area was quite incomplete, and little has been done since them. Given the increased cognitive demands and G-stress being placed on the pilot's of modern fighter aircraft, this area demands thorough investigation.

Before leaving the topic of G-stress, two other issues should be addressed. First, several studies (7,21,22) noted the problem of vertigo experienced by centrifuge subjects and how it would induce a spurious decrease in their initial performance at a given G-level. This effect would be minimized if the G-onset rate was limited to .75G/sec or less. Secondly, it was noted that many subjects suffered from anxiety/ apprehension just prior to a high G run (22,23), thereby decreasing their subsequent performance. However, after several runs their anxiety would subside and their performance would improve. Thus it is desirable to allow for some training runs at high G prior to any data collection.

Workload: The workload literature was reviewed with the focus of the effort on two specific topics. First of all, what modifications, if any, would be needed to use the maze solving paradigm (11,12) as a measure of pilot workload. Secondly, identify any other workload measurement technique that would be used in conjunction with the maze solving paradigm to assess pilot workload in a high G environment.

As noted in the literature (24,25), there is no one definition of workload. Rather it is a hypothetical construct that must be operationally defined in each specific instance. The basic notion of workload is that the human is a limited capacity processor. That is, the human is capable of processing only a limited amount of information and/or performing a limited number of tasks in any given time period. If the human's channel capacity is exceeded, then clearly the workload is too high.

To fully comprehend the workload measurement problem, one must first become acquainted with the fundamentals of human information processing. Recent research (26) has suggested that human attention or processing resources are comprised of three mutually exclusive categories, perceptual, cognitive, and response resources. If the subject's tasks demands exceed his capabilities in any one of these three areas, then he would be unable to adequately perform the tasks due to excessive workload. If a workload measurement technique is to be reliable, it must be sensitive to variations in all three areas.

In the maze solving technique developed by Ward and her colleagues (11,12), subjects are presented with an unfamiliar two-dimensional maze via a CRT and are required to move a dot thru from one side to the other as rapidly as possible. The dot moves at a constant speed and direction, but subjects can change the direction with discrete central inputs via a trim tab button on a joystick. The subject's score is computed as the ratio of the optimum solution time to his actual solution time.

Clearly, this task demands processing resources from all three major categories, and they should be a reliable workload measurement technique. There is, however, one way in which it might be improved. That is to replace the discrete control with a continuous control thereby demanding more response resources. However, there are several problems associated with such a change: (1) the maze solving task has been validated in a nominal +1G environment with the discrete control input but not for the continuous control mode; (2) motor coordination is very difficult under high G and excessive response demands can obscure more subtle changes in the perceptual and cognitive areas or even make the task impossible to perform; and (3) it would require extensive software modification which probably could not be completed within the ten week time period of this research effort. Thus for this investigation it was decided to keep the task as is with the discrete mode of operation.

The second objective of the workload literature review was to identify any other potential workload assessment techniques that could be used in conjunction with the maze-solving paradigm under sustained high G. In a comprehensive reivev of the workload literature including over 400 references, Wierwille and Williges (25) identified twenty-eight specific techniques for assessing operator workload. In the same paper, they also presented a method for selecting the most appropriate technique for any given context. Following

their guidelines, it was decided that a subjective assessment would be the most viable alternative.

Recently, Eggemeier, Reid, and their colleagues at the Air Force Aerospace Medical Research Laboratory have developed a generic subjective technique called the subjective workload assessment technique or SWAT (27,28,29,30). In short, it combines subjective ratings on three different levels, via the mathematical technique of conjoint measurement, to produce an interval scale of workload. It has been shown to be both a sensitive and reliable measure of workload. One significant advantage of SWAT is that it is a relatively simple and unobtrusive technique which could be easily implemented jointly with the maze solving technique in the high G environment. Thus, it was decided to use SWAT as an alternative workload measure.

To complete the discussion of workload, a common misnomer in the literature should be noted. Generally speaking if an experimental manipulation, for example changing from manually to vocally activated switches, resulted in improved operator performance it was deduced that such a manipulation decreased the operator's workload; alternatively, if an experimental manipulation (i.e. increased vibration) resulted in a decrement in operator performance it was concluded that operator's workload had increased. Strictly speaking, however, operator workload did not change in either of these cases. Rather, it was a change in the operator's processing efficiency or ability to produce work, or more simply a change in his channel capacity.

This misnomer does not invalidate such previous research, nor does it impair the use of the various workload measurement techniques that have been developed. But it does emphasize the need for care and precision in interpreting experimental results. For example, in this study two alternative workload measurement techniques will be employed to investigate the effects of G-stress. It is anticipated that increased G-stress will result in an "apparent" increase in pilot workload, yet the correct interpretation would be a reduction in pilot channel capacity.

One final clarification concerning the concept of workload should be made. In this particular experiment, the subject's task of maze solving will remain unchanged throughout, thus his overt or "external" workload will remain constant. However, a consequence of increased G- stress is an increased difficulty in performing motor-coordination tasks as well as the necessity to perform an M-1 or L-1 straining maneuver. These increased demands on the

pilot can be thought of as an increase in his "internal" workload. In short, it is anticipated that G-stress will cause a decrease in pilot performance. This decrement will be induced by either or both an increase in the internal workload an/or a decrease in processing efficiency. Currently, workload measurement techniques cannot discriminate between either of these sources.

In summary, to be consistent with the referenced literature, the generic term of workload has been used throughout this paper. However, the correct interpretation of workload is that it is a combination of both internal and external workload, as well as processing efficiency. Or, alternately, workload can be thought of as being universally related to amount of unused processing resources. That is, as the amount of available or unused processing resources decrease, the workload, by definition, has increased.

Experimental Design: As noted previously, due to the scheduling backlog on AFAMRL's dynamic environmental simulator (DES), the actual experiment has not yet been completed. Thus, only the experimental design and its rationale will be presented. Recall, the objective of this study is to analyze the effects of +Gz stress on pilot workload. It will be conducted in three phases, Phase I - Static Training, Phase II - Dynamic Training, and Phase III - Data Collection. Pilot workload will be measured using both primary task performance, via the two-dimensional maze solving task, and subjective ratings via SWAT.

The DES facility will be utilized to provide the G-stress with a modified ACES II or F-16 seat. Flight suits, standard issue helmets, oxygen masks, and anti-g suits will also be used.

There are two tasks which will be accomplished in the static training phase. Subjects will perform a card sort or ordering of the subjective ratings that will be used in Phase III. In addition, subjects will practice solving two dimensional mazes similar to the ones which will be used as the primary task in Phase III. All work conducted in Phase I will be in a normal +1 Gz environment.

Each subject will participate in four one-hour training sessions with each session occurring on a different day. The purpose of the first session is to perform a card sort on the subjective ratings to be used in the SWAT portion of Phase III. The subjects will be provided with a deck of 27 cards placed in random order. Each card will represent one of the 27 possible combinations of three categories (time load, mental effort load, and psychological stress load) with each category at three different levels (slow,

medium, and high). The subject's task will be to sort these cards so that all 27 combinations are rank-ordered with respect to the degree of subjective workload imposed by each. These rank-orderings will then be used to develop an interval scale of workload for evaluating the subjective ratings to be obtained in Phase III.

In the remaining three static training sessions, subjects will practice solving two-dimensional mazes. Figure 1 depicts a typical maze that will be used throughout the experiment. All mazes will consist of the basic 10 x 10 grid as shown, but will differ in the placement of the maze barriers. For each trial, a given maze will be displayed on a CRT with a dot at the entrance of the maze. The subject's task will be to move the dot thru the maze as rapidly as possible. The dot will move at a constant speed; the subject can change its direction (left vs right or up vs down) by moving the trim tab button, on a joystick controller, in the appropriate direction.

The trial will conclude as soon as the dot has been successfully guided thru the maze to the goal. The shortest possible solution time divided by the actual time to complete the maze will be used as the measure of performance. After solving the maze, subjects will then be asked to do a SWAT rating on the maze. Once the SWAT rating is completed, they will be provided feedback concerning their maze solving performance. If the subject fails to complete the maze within two minutes (typical solution times are approximately one minute or less), the trial will terminate and a message will be displayed indicating that the subject ran out of time. The subject's score will then be computed as the ratio of the shortest possible solution time, multiplied by the percent of the maze solved, and then divided by the actual solution time (two minutes).

The purpose of the dynamic training phase, which will be conducted entirely on the DES, is to reduce the experimental variance in Phase III by permitting subjects to practice maze solving while under G-stress. Each subject will practice in two daily sessions of approximately one hour each. The specific G-profile, number of runs/sessions, duration of each run, etc. will be identical to those which will be used in the data collection phase and are discussed in detail in the section below entitled "Phase III - Data Collection". Different mazes will be used in each of the three phases to prohibit learning effects from the subjects becoming familiar with any particular maze.

The data collection phase will also be conducted on the DES. Each subject will participate in five daily sessions of approximately one hour each; each session will be comprised of eight runs. A typical run will be comprised of four parts:

(1) Positive Onset: Starting at a baseline level of +1.5 Gz, the subject's Gz level will increase (or decrease) at the rate of .25Gz/sec until the desired level of Gz for that run is attained. Four levels of +Gz (1,3,5, and 6) will be employed. The slow onset rate and the baseline level of 1.5Gz were chosen to minimize the problems associated with vertigo.

(2) Test: Once the desired level of +Gz is attained, the subject will be presented with a maze, via the CRT monitor, and required to solve it as rapidly as possible. The maximum time allowed for solving the maze will be two minutes.

(3) Negative Onset: Immediately after the successful completion of the maze or the two minutes time limit, whichever occurs first, the subject's Gz level will decrease (or increase) at the rate of .25Gz/sec until the baseline level of +1.5Gz is attained.

(4) Rest: The subject will then rest at the +1.5Gz level for a minimum of one minute before initiating the next run. However, this rest period can be extended for as long as desired by either the subject or the medical monitor. During this rest period, they will be required to rate their perceived workload during the previous run via SWAT. This technique consists of rating each category (time, mental effort, and psychological stress load) as being either low, medium, or high. After the SWAT rating is completed, the subject will be informed of his maze solving score.

During each daily session, subjects will receive two runs at each of the four Gz levels for a total of eight runs. Subjects will also be provided with an additional mandatory three-minute rest period after the first four runs are completed. The order of the runs and the specific maze presented will be counterbalanced across subjects and +Gz levels.

The first four daily sessions of Phase III will be identical to that of Phase II. On the fifth day, each maze will have the optimum solution path identified on the CRT. Thus the subject's only task will be to maneuver the dot along the path. Comparisons of solution times between mazes with and without the solution path being shown will give a direct measure of the cognitive effects of G-stress.

Figure 1 summarizes the experimental design of Phase III. (Recall, the same basic design will also be employed in Phase II). The design construction/rationale is as follows:

(a) A balanced 4 x 4 Latin Square was used to determine the order of presentation of the G-levels. Thus each G-level (represented by the numbers 1 thru 4) will be preceded and followed by every other G-level an equal number of times. This counter balancing will eliminate the effects of training and more importantly fatigue. The same Latin Square will then be repeated, after ten-minute break, to complete each eight-run daily session.

(b) Eight different mazes (represented by the letters A thru H) were employed. One maze was randomly assigned to each run with the following constraints:

- (1) Each maze could only be used once within any given daily session.
- (2) Each maze had to be used exactly four times at each G-level.
- (3) The average run-number for each maze, over all four daily sessions, had to be approximately equivalent.

Finally, a minimum of four and a maximum of eight subjects will be used. Each subject will be presented with each one of the four daily sessions, but in a different, randomly determined order.

IV. RECOMMENDATIONS

It is quite evident that technological advances in fighter aircraft have produced a work environment in which human capabilities can and often are exceeded either in terms of workload or G-stress. Although both of these topic areas have been addressed quite extensively by the research community, little if any effort has been focused on the interaction of these two crucial flight parameters. Clearly, this is a problem which warrants immediate and thorough investigation.

The primary purpose of the proposed experiment (which is currently underway) is to test the hypothesis that +Gz stress does in fact increase pilot workload. Or more correctly, +Gz stress will decrease the processing resources which are available for the various piloting tasks. Assuming the hypothesis is supported, the next step should be to quantify this effect. That is, to determine the pilot's available channel capacity as a function of his G-stress. In addition, the effect on pilot workload of G-stress along other axes should also be addressed as well as the effects of high G onset rates.² The maze solving technique appears to be a very promising method for

assessing workload. There are, however, several areas which warrant further investigation. One is to address the previously mentioned change from discrete control to continuous control. Additionally, the development of an algorithm that would quantify the number of bits of information required to solve any given maze would be invaluable since it would allow direct quantification (i.e. number of bits/second) of human processing capabilities. Finally, efforts should be made to reduce maze solution times (for example, by increasing dot speed or decreasing maze size and complexity) so that this technique can provide a more instantaneous measure of workload.

REFERENCES

1. Burton, R.R., S.D. Leverett, Jr. and E.D. Michaelson. Man at high sustained +Gz acceleration: a review. *Aerospace Medicine* 45(10): 1115-1136. 1974.
2. Collyer, S.C. Testing psychomotor performance during sustained acceleration. *School of Aerospace Medicine, SAM-TR-73-52*. 1973.
3. McElreath, K.W. and M.D. Clader. Pilot tracking performance as a function of g-stress and seat back angle. *AMRL-TR-76-107*. 1977.
4. Barer, A.S., T.A. Sokolova, V.M. Tardov, and Y.P. Yashin. Effect of prolonged +Gz acceleration on human performance. *Kosmicheskaya Biologiya i Aviakosmicheskaya Meditsina, Vol 1, No 3, 37-40*, 1980.
5. Hyde, A.S. and H.W. Raab. A summary of human tolerance to prolonged acceleration. *AMRL-TR-65-36*. 1965.
6. Grether, W.F. Acceleration and human performance. *Aerospace Med.* 42(11): 1157-1166. 1971.
7. Creer, B.Y., J.D. Stewart, J.G. Douvillier. Influence of sustained accelerations on certain pilot performance capabilities. *Aerosp. Med* 33:1086-1093. 1962.
8. Clarke, N.P., A.S. Hyde, N.S. Cherniak. A preliminary report of human response to rearward facing reentry acceleration. *WADC Technical Note 59-109 (AD 231651)*. 1959.
9. Chambers, R.M. The psychology of space flight and centrifuge training. *Journ of British Interplanetary Society. Vol 21, 232-273*. 1968.
10. Tesch, P.A., H. Hjort and U.I. Balldin. Effects of strength training on G tolerance, *Aviation Space & Environ Med, 54(8): 675-782*.
11. Poturalski, R.J. and S.L. Ward. Maze solving as a tool for measuring performance under stress. *Annual Mtg of AsMA*. 1982.
12. Ward, S.L. and R.J. Poturalski. Changes in maze solving errors due to stress. *Annual Mtg of Human Factors Society*. 1982.
13. Creer, B.Y., H.A. Smedal, R.C. Wingrove. Centrifuge study of pilot tolerance to acceleration and the effects of acceleration on pilot performance. *NASA-TN-D-337. Sci. Techn. Info. Div. NASA, D.C.* 1960.
14. Miller, H., M.B. Riley, S. Bondurant, and E.P. Hiatt. The duration of tolerance to positive acceleration. *Journal of Aviation Medicine. Vol 30, 360-366*.

15. Parkhurst, M.J., S.D. Leverett, S.J. Shubrooks. Human tolerance to high, sustained +Gz acceleration. *Aerosp. Med.* Vol 43, No 7. 1972.
16. Comrey, A.L. et al. The effect of increased positive radial acceleration upon perceptual speed ability. *J. Aviat. Med* 22: 60-69. 1951.
17. Frankenhaeuser, M. Effects of prolonged gravitational stress on performance. *Acta Psychol* 14:92-108. 1958.
18. Chambers, R.M. Long term acceleration and centrifuge simulation studies. N63-19313. *Aviation Medical Accel Lab, US NADC, PA.* 1963.
19. Ross, B.M. and R.M. Chambers. Effects of transverse g-stress on running memory. *Percept Mot Skills* 24:423-435. 1967.
20. Ross, B.M., R.M. Chambers, R.R. Thompson. Effects of transverse acceleration on performance of two running matching memory (RMM) tasks. NADC-MA-6309. 1963.
21. Smedal, H.A., B.Y. Creer, R.C. Wingrove. Ability of pilots to perform a control task in various sustained acceleration fields. *Aerosp Med* 31:901-906. 1968.
22. Little, V.Z., S.D. Leverett, B.O. Hartman. Psychomotor and physiologic changes during accelerations of 5, 7, and 9 +Gx. *Aerosp Med* 39:1190-1197. 1968.
23. Webb, M.G. Jr. Some effects of acceleration on human subjects. *J. Aviat Med* 29:879-884. 1958.
24. Gartner, W.B. and M.R. Murphy. Concepts of workload. AGARDOGRAPH No. 246, Survey of Methods to Assess Workload, AGARD-AG-246, ADA078319, 1979.
25. Wierwille, W.W. and R.C. Williges. Survey and analysis of operator workload assessment techniques. *Naval Air Test Center, S-78-101.* 1978.
26. Wickens, C.D. Processing resources in attention. Chapter in *Varieties of Attention*, edited by Parasuraman, R.J. Beatty and R. Davis, 1983.
27. Reid, G.B. et al. Application of conjoint measurement to workload scale development. *Annual Mtg of Human Factors Society.* 1981.
28. Reid, G.B. et al. Development of multi-dimensional subjective measures of workload. *IEEE Conf on Cybernetics & Society.* 1981.
29. Eggemeier, F.T. et al. Subjective workload assessment in a memory update task. *Ann Mtg of Human Factors Society.* 1982.
30. Eggemeier, F.T. et al. The effects of variations in task loading on subjective workload rating scales. *IEEE Nat Aerosp & Elect Conf.* 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

IDENTIFICATION OF RAPID EYE MOVEMENTS BY COMPUTER

DURING DISCRETE TRACKING TASKS

Prepared by: Dr. John P. Giolma
Academic Rank: Assistant Professor
Department and University: Department of Engineering Science
Trinity University

and

Prepared by: James Evans Lyne
USAF-SCEEE Graduate Student
Summer Support Recipient
Vanderbilt University School of Medicine
Research Location: USAF School of Aerospace Medicine
Neurosciences Function
USAF Research: Dr. J. W. Wolfe
Date: 23 August 1983
Contract No.: F49620-82-C-0035

IDENTIFICATION OF RAPID EYE MOVEMENTS BY COMPUTER
DURING DISCRETE TRACKING TASKS

by

John P. Giolma and J. Evans Lyne

The identification of rapid eye movements (saccades) and the extraction of performance parameters by digital computer is described. Algorithms written in Fortran are described that use signal noise estimates, main sequence relationships and the digitally derived eye velocity waveform to locate saccades of a half degree and larger. Recommendations are provided for future applications under other tracking conditions and for the improvement of the infra-red eye movement recorder.

I. INTRODUCTION

Oculomotor responses of humans to vestibular stimulation and to the displacement of visual targets is of interest to neurophysiologists who seek to understand basic oculomotor function and to the Air Force in the development of additional measures with which to predict pilot performance. One aspect of the visual response to these stimuli consists of rapid eye movements called "fast phases", "saccades" or simply "rapid eye movements". This type of movement consists of the most rapid change in position of which the eye is capable and serves to keep the fovea on the desired portion of the visual field.¹ Extraction of information concerning the performance of the eye in executing rapid movements is a tedious task if performed manually, and the resultant measures of velocity and acceleration are questionable.

II. OBJECTIVES

The principal objective of this project was the development of an automated technique for the location of rapid eye movements or saccades from sampled measurements of human eye position. The algorithms developed to locate saccades were specifically intended to identify rapid movements that occurred in response to discrete stimuli, i.e., the sudden jump in position of a spot of light in the subject's field of view. These discrete stimuli were generated by displacing a spot of light more rapidly than the eye could follow. The subject fixated on the light, attempted to follow its change in position and thus generated saccades.

A second objective of the project was to provide sufficient generality in the software to permit its application to saccade

identification under other tracking conditions (e.g., tracking a target that moves smoothly and saccades elicited by optokinetic or vestibular stimuli).

A third objective of the project was to provide computer algorithms to extract measures for various parameters of rapid eye movement including latency, duration, amplitude, maximum velocity, maximum acceleration and deceleration, and overshoot. These algorithms have been implemented on the laboratory PDP-11/34 at the Neurosciences Function, NGNS, USAF School of Aerospace Medicine (SAM). The output calculations are provided in tabular and graphical (histogram) form for analysis. These routines will be used by the Neurosciences Function Laboratory in studies of the physiology of normal human oculomotor function.

Finally, operating characteristics of the existing eye movement recorder system were examined. These characteristics will be considered for future development of location algorithms, and form the basis for recommendations concerning improvements to the measurement system.

III. CHARACTERISTICS OF THE MEASUREMENT SYSTEM

The system which was used for visual stimulation and for data acquisition consisted of a Digital Equipment Corporation PDP-11/34 minicomputer with A/D and D/A capabilities, an array of light emitting diodes (LEDs) and an infrared reflective eye movement recorder. The computer provided the desired stimuli by controlling the sequence of LEDs turned on for the observer. The voltage output of the eye movement recorder was also sampled by the computer. Nine LEDs were positioned in a linear array placed in the horizontal plane with respect

to the observer, providing visual targets of five degree increments ranging from -20 to +20 degrees from the center. For the current project, the sequence of target (LED) movements consisted of short (2 to 3 seconds) intervals with the center position diode on, after which the center diode was turned off and one of the off-center diodes turned on for .75 seconds. During this interval, the output of the eye movement recorder was sampled by a 12 bit A/D at 1000 samples per second. The sampling rate was chosen to permit the application of digital filtering techniques if required, although about half that rate should be adequate for the signal bandwidth and the algorithms employed in the present work.

The eye movement recorder consisted of infrared sources and detectors mounted on eyeglasses that used the reflective properties of the sclera and limbus to derive horizontal eye position.² In bench tests, the system had a design 3db bandwidth of 150 Hz, was linear to 0.1 degree in the range of -25 to +25 degrees and could be aligned in practice to better than one degree.

Computer generated plots of sampled eye movements appear in Figure 1. Position was calibrated by static measurements of sampled signals from each eye following a gaze to the left of 10 degrees and to the right of 10 degrees. The derivative plots were obtained using a two-point central difference approximation with a step size of four to provide some filtering consistent with the expected frequency content of both saccadic movements and the 1 kHz sampling rate.^{3,4,5} The four-point step size provides a 3db bandwidth of 56 Hz for velocity and 40 Hz for acceleration.

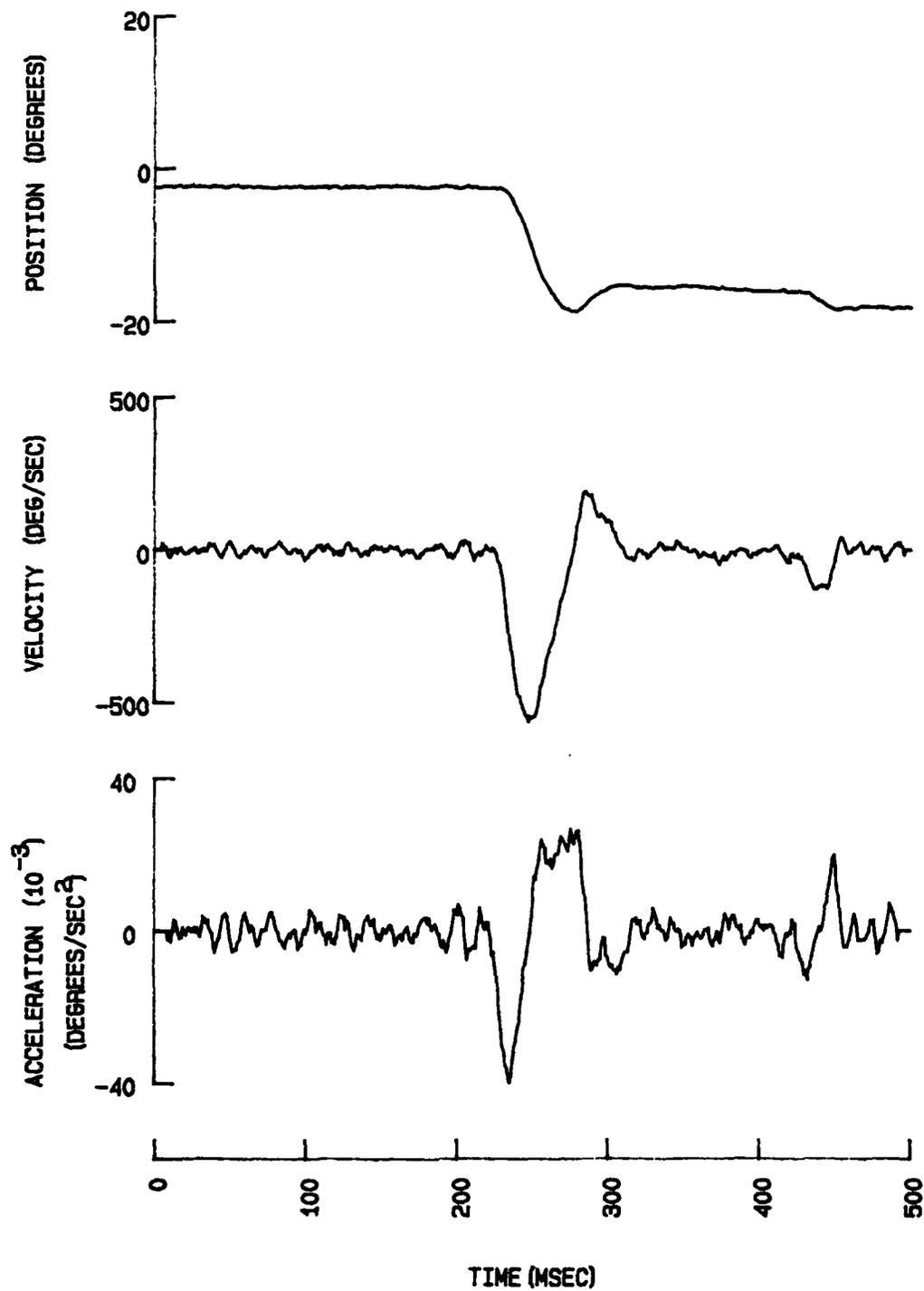


Figure 1: Computer generated plots of eye position, velocity and acceleration obtained during a response to a 15 degree stimulus to the subject's left.

To determine noise levels of the eye movement recorder, the system output was sampled both with and without a human subject. A frequency averaged digital Fourier transform was generated for each condition. In the "no subject" case, (Figure 2a) the system noise was predominately random (i.e., it assumed the shape of the system transfer function) with a single coherent peak at 60 Hz. The 60 Hz peak corresponded to millivolt levels of signal at the resolution limits of the A/D. The origin of the noise was determined to exist primarily in the infrared emitter/detector end of the recorder.

In order to characterize the "oculomotor noise" levels, eye position signals were collected when a subject was fixating a stationary target. Small eye movements that were present during fixation (microsaccades and tremor) can be expected to display the amplitude levels (minutes of arc) and frequency range (up to 150 Hz) that are present in the plots of Figures 1 and 2.⁶ In Figure 2b (subject present), the low frequency region of the spectrum changed markedly due to the presence of small saccades and to occasional slow drift movements of the eye. Total rms values of the position signals with the subject not present indicated an average value of 8 millivolts rms (system noise alone) while the subject present condition resulted in 20 millivolts rms (system noise plus "oculomotor noise"). If calibrated to eye position, these levels correspond to .066 degrees (4 minutes) for system noise and to .16 degrees (10 minutes) for system plus oculomotor noise. Measurements of positional noise for five subjects ranged from .05 to 0.2 degrees rms. The limits that this noise placed on the ability of the system to locate saccades of small amplitude will be discussed in the following section.

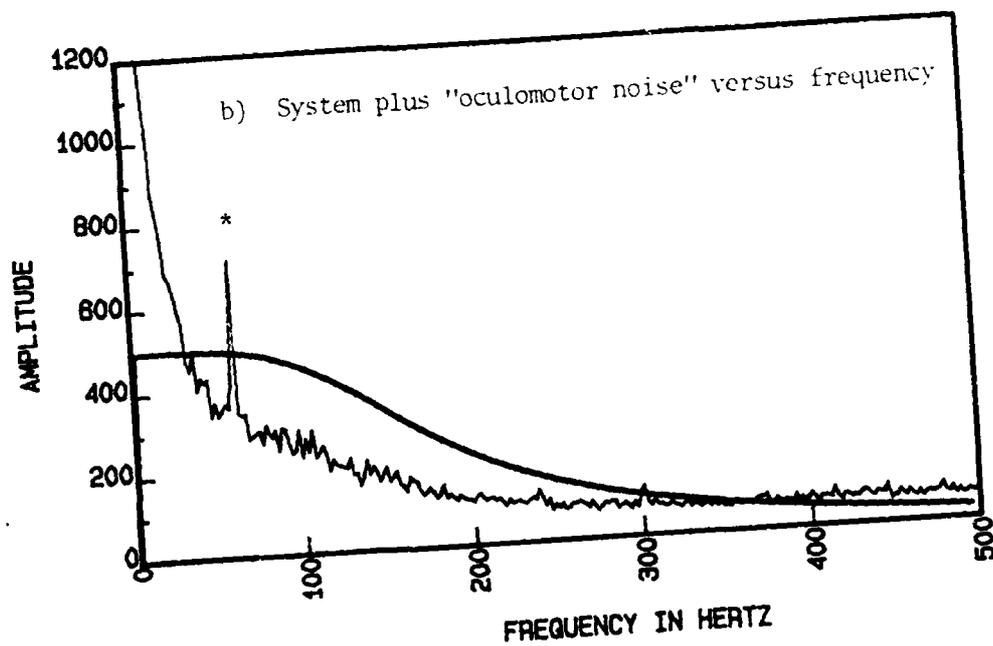
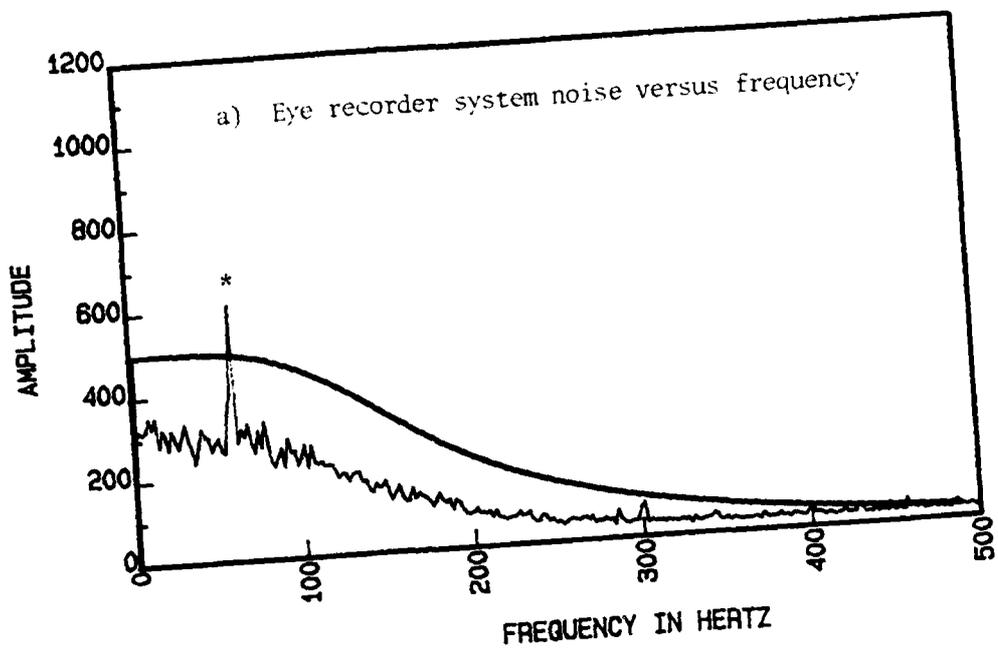


Figure 2: Computer generated amplitude spectrum plots of recorder system alone (a), and with fixating subject present (b). The small peak (*) present in both plots is at power line frequency and corresponds to millivolt (rms) amplitudes. Solid line in both plots is the calculated frequency response of the recorder filters.

IV. DEVELOPMENT OF THE SACCADE IDENTIFICATION ALGORITHM

The curves of Figure 1 suggest that saccades might be identified from the characteristic shape of the position curve (a ramp-step change of position) from the peak of the velocity curve or from the peak and valley of the acceleration curve. Some investigators have reported techniques for saccade location based on the peak of the velocity curve^{7,8,9} while others have used acceleration and matched filter techniques to identify saccades.¹⁰ Tole and King¹⁰ define a matched filter which was designed to be sensitive to the acceleration characteristics of a saccadic motion. The filter output was compared to a threshold derived from noise estimates. The matched filter approach may require separate filters for saccades of differing amplitudes since there is a roughly linear relationship between relationship between saccade amplitude and duration.¹¹

From a signal-to-noise ratio point of view, the position waveform is the best choice. In the curves of Figure 1, the signal-to-noise ratio defined by peak signal to rms noise declined by about 10db for each derivative. For more generalized applicability, we considered that the derivative served as a filter to separate slow or pursuit eye movements from the fast components due to saccades. Based on the signal-to-noise characteristics of the eye movement recorder system, a velocity peak detector was selected as part of the location algorithm used in the present work.

For large voluntary saccades, the intercept of plots of duration versus amplitude appeared to be in the range of 20 to 30 milliseconds.^{9,12} Main sequence relations indicated that durations of this magnitude would be appropriate for rapid eye movements of about 0.5 degrees with

peak velocities of about 50 degrees per second. Since the positional noise levels for the eye recorder system and oculomotor system were in the range of 0.05 to 0.2 degrees rms, we could not expect to reliably locate movements of that order of magnitude, or movements whose main sequence duration would fall in the corresponding time range (much less than 20 milliseconds). Due to this limitation, a minimum duration criterion of 15 milliseconds was selected as a part of the identification criteria. We chose to define duration as the time between zero crossings that bound the local peak in the velocity curve.

Initial observation of laboratory data resulted in a fixed threshold of 80 degrees per second coupled with the 15 millisecond duration criterion. This approach worked well with noisy EOG recordings but did not identify saccades of the order of one degree or less with the infrared recorder. The variable threshold suggested by Tole and Young¹⁰ led to the application of a variable velocity threshold for saccade identification based upon measurements obtained during steady fixation. Rms noise determination were obtained from five volunteers to determine possible threshold levels for peak detection in the velocity signal. The rms values were obtained during short (100 millisecond) periods of stable fixation following each discrete stimulus. These values ranged from 9 to 20 degrees per second.

The threshold value used in this work was obtained by determining the rms velocity noise level as described and choosing the lower of the values: 80 degrees per second or three times the rms value. Using this criterion, saccades above about 0.5 degrees were found reliably. We then observed that periods immediately following large

saccades, especially those with large overshoots, contained higher levels of "oculomotor noise" than did the steady fixation periods. Therefore, a lower bound was placed on the velocity threshold (30 degrees per second) to reduce the incidence of false selections while maintaining the appropriate amplitude, velocity, and duration relationships of the main sequence.

In summary, saccades were identified by:

- 1) defining a variable velocity threshold;
- 2) searching through the velocity signal for peaks that exceeded the threshold; and
- 3) requiring a minimum duration of 15 milliseconds.

V. PERFORMANCE OF THE ALGORITHM DURING DISCRETE TRACKING

The routines for identification of saccades were tested using laboratory personnel as subjects. Multiple sessions were conducted with some subjects. The responses to each discrete stimulus were reviewed using a graphics terminal (VT11). Saccades identified by the algorithm were then compared to those identified as saccades by the reviewer. Of the 640 stimuli in the tests (either eye counted separately) the reviewer identified 720 large saccades (greater than one degree) and 356 small saccades (less than one degree). All large saccades were identified by both reviewer and algorithm. Based on small saccades, the algorithm missed 13.5% (48 false negatives) and incorrectly identified 19 events (5.3% false positives). Almost all saccades missed were missed only for one eye and were 0.5 degree or less in magnitude. In addition, almost all erroneous selections were in the time period following a large saccade with overshoot and

subsequent increase in "oculomotor noise". Significant improvement in performance might be achieved by using the locating of a saccade in the record from one eye to assist in locating a saccadic response in the other eye by reducing thresholds. Figure 3 illustrates the concepts of thresholding, rms noise calculation, and selection errors.

VI. MEASURES OF EYE MOVEMENT PERFORMANCE DURING SACCADES:
GRAPHICAL OUTPUT

An emphasis was placed on latency, duration, amplitude, maximum velocity, and maximum acceleration and deceleration in the saccadic location algorithm because of the desire to calculate these oculomotor parameters in normal humans. Amplitude was the net position change from onset to the steady state following the response, while the maximal velocity, acceleration and deceleration were absolute values found within the saccadic boundaries. Overshoot was defined for all saccades as the net position change between the end of the rapid movement (zero crossing in velocity) and the next zero crossing in the velocity. This approach searched for the next saccade only beyond the end of the overshoot even though the recovery from overshoot may itself meet the main sequence criteria for a saccade.⁹

Visual stimuli consisting of 10 degree discrete displacements of the target to the left and right of center were presented to the five volunteers used for the development of the algorithms. The saccadic parameters were calculated following the signal acquisition and were output in a variety of forms: raw calculations for each saccade found, means and standard deviations for the primary (first) response, and the histogram form shown in Figure 4. This last approach

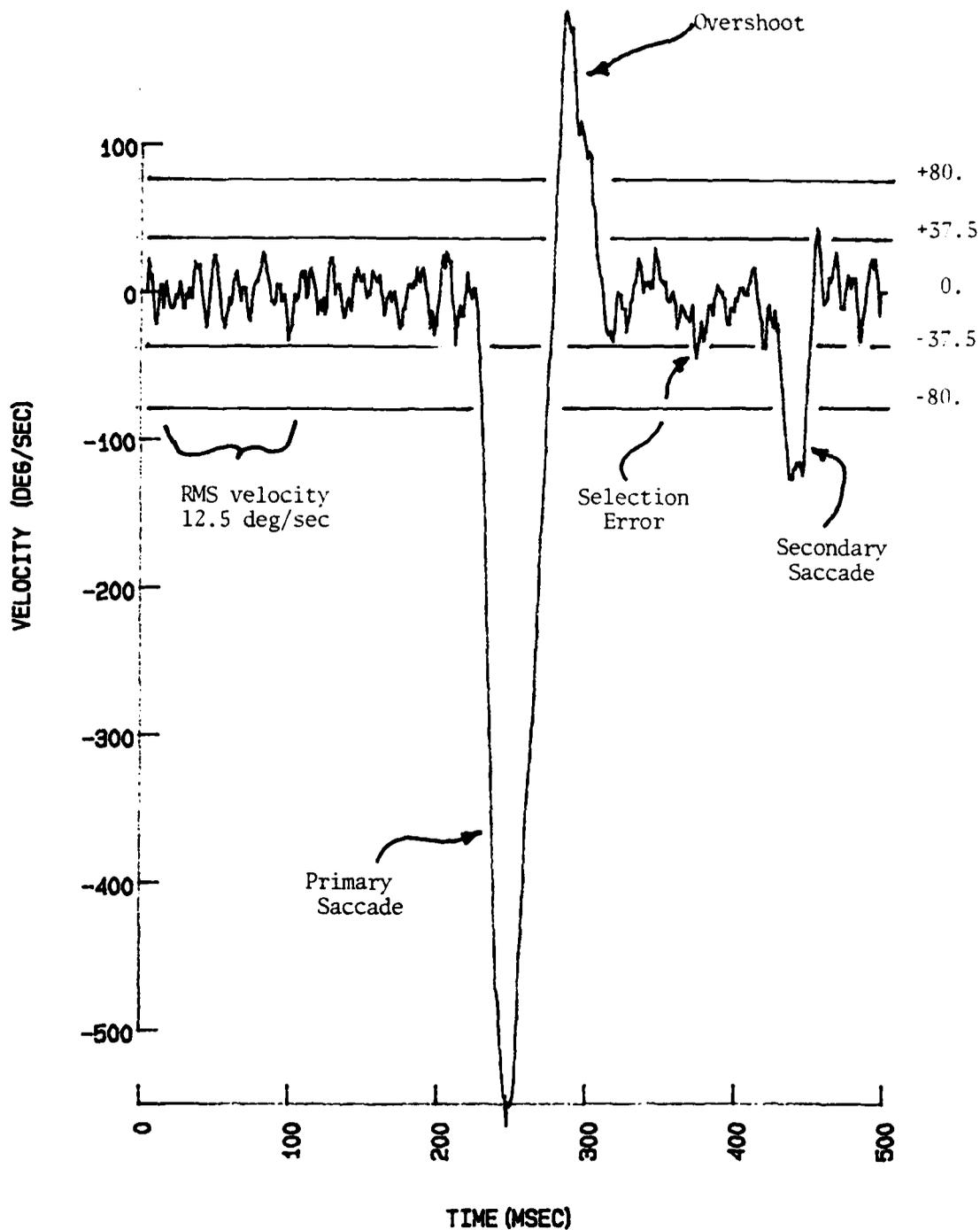


Figure 3: Computer generated plot of eye velocity during response to 15 degree stimulus. Velocity threshold for peak detection is indicated with horizontal lines. RMS derived threshold is indicated and is calculated from the first 100 msec. of the response (see text). Waveform is from same response as Figure 1.

LR	MEAN	58.3
	ST. DEV.	8.5
	N	24
LL	MEAN	57.8
	ST. DEV.	7.8
	N	24

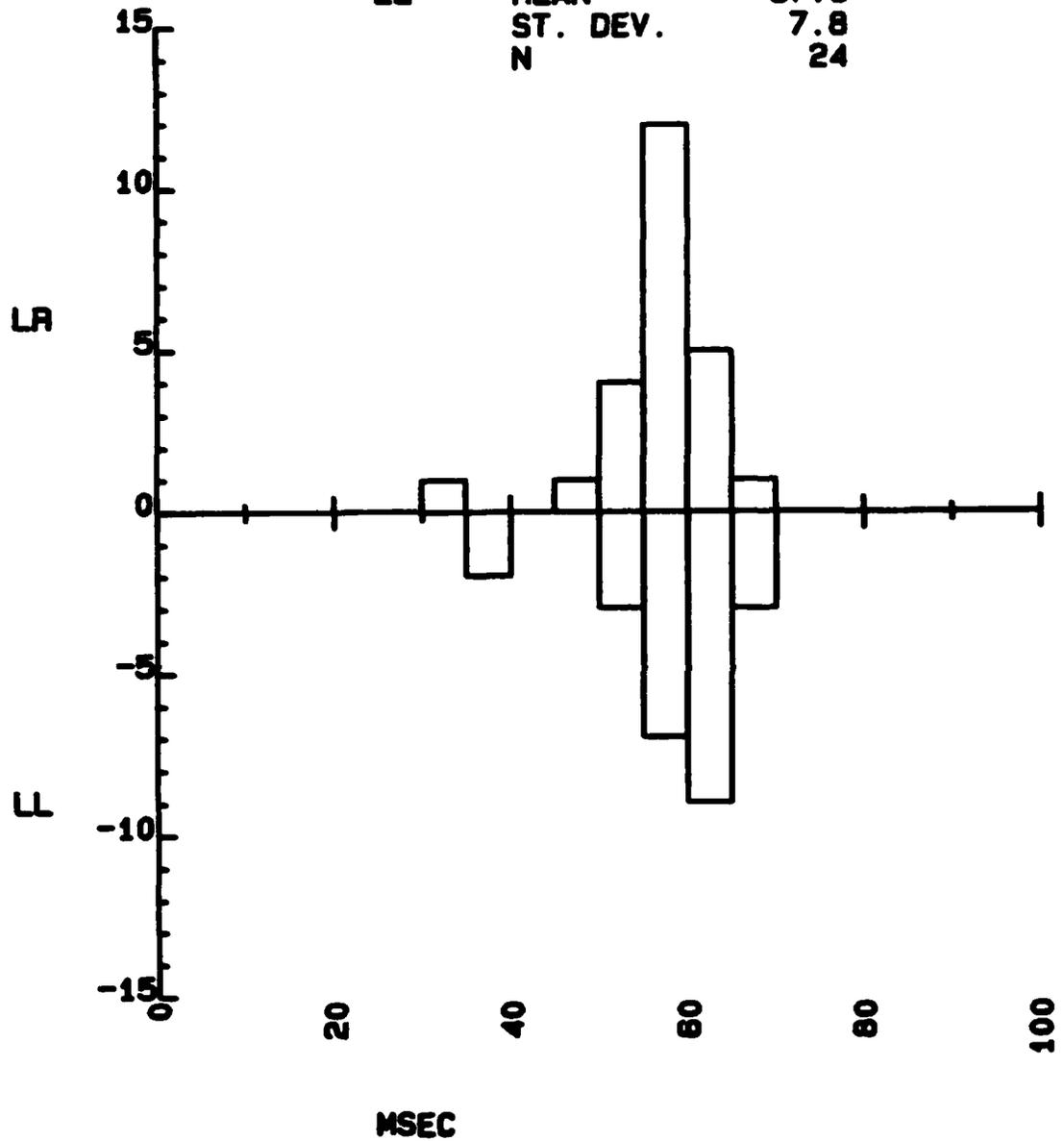


Figure 4: Histogram of latency for primary saccades from responses to discrete ten degree stimuli. LL indicates the left eye moving to the left and LR the left eye moving to the right.

was felt to be especially useful in developing techniques for comparing normal human oculomotor function under monocular and binocular tracking conditions.

VII. RECOMMENDATIONS

This work has provided a technique for the automated identification of rapid eye movements that occur during a discrete tracking task. In order to apply the algorithms to other oculomotor tasks, such as sinusoidal target movements or optokinetic stimuli, the algorithms must be preceded by appropriate manipulations of the eye movement signals to separate the slow pursuit movements from the rapid movements. In some cases, the digital derivative alone will provide this, but in others further work may be required to successfully identify saccades and correctly obtain performance measurements. When no period of fixation occurs during the oculomotor tasks, the noise estimates used for threshold detection should be determined under fixation conditions before the tracking task begins.

At present, this laboratory (Neurosciences Function, NGNS USAFSAM) has in place a series of tests that extract useful information from the smooth pursuit capabilities of subjects. Our algorithms could be used in discrete tracking tasks, sinusoidal tracking tasks and optokinetic stimuli and when combined with the present tests would provide a significant advance in the understanding of human oculomotor and vestibular function. The combined battery could then be used to provide measures of pilot performance for the Air Force and to assist in explaining the 'space sickness' syndrome experienced by some astronauts. I would propose the further development of the saccade

1

identification techniques to include sinusoidal tracking and optokinetic stimuli which will provide a data base of normal human oculomotor performance.

The infrared eye movement recorder developed in this laboratory offers a significant improvement in signal to noise ratio and in drift to traditional EOG methods of eye movement recording. Experience gained in using the system during this work indicates that further improvements are desirable and possible in both the noise and drift characteristics. Studies should be performed to isolate the principal noise sources in the emitter and detector components of the system. Further refinement of the synchronous demodulator, filters and amplifiers should essentially remove the drift in the output signal, which now is as large as one degree in a four minute recording session. Finally, system modifications are necessary to increase the signal level available to the A/D converter while maintaining an appropriate level for the strip chart recorders often used in conjunction with the computer.

REFERENCES

1. A. T. Bahill and L. Stark, "The Trajectories of Saccadic Eye Movements," Sci. Amer., 240:108-117, 1979.
2. E. J. Engelken, K. W. Stevens, J. W. Wolfe and J. T. Yates, "A Limbus Sensing Eye Movement Recorder," SAM TR 1983 (in preparation).
3. A. T. Bahill, J. S. Kallman and J. E. Lieberman, "Frequency Limitations for the Two-Point Central Difference Derivative Algorithm," Biological Cybernetics, 45:1-4, 1982.
4. A. T. Bahill and J. D. McDonald, "Frequency Limitations and Optimal Step Size for the Two-Point Central Difference Derivative Algorithm with Application to Human Eye Movement Data," IEEE Trans. Biomed. Eng. 30:3, 191-194, March 1983.
5. D. A. Robinson, "The Mechanics of Human Saccadic Eye Movements," J. Physiol., 174:245-264, 1964.
6. R. Grant, "Neurophysiology of the Retina," in The Eye, Vol. 2, p. 626 Academic Press, 1962.
7. E. J. Engelken and J. W. Wolfe, "Analog Signal Processing of Eye Movements for On-Line Digital Computer Analysis," Physiol. and Beh. 18:175-58, 1977.
8. R. M. Jell, G. T. Turnipseed and F. E. Geudry, Jr., "Digital Analysis of the Voluntary Head Movement-Induced Vestibulo-Ocular Reflex, with Saccade Extraction," DTIC Report AD-A097 680, NAMRL-1273, 1980.
9. A. T. Bahill, A. Brockenbrough and B. T. Troost, "Variability and Development of a Normative Data Base for Saccadic Eye Movements," Invest. Ophthalmol. Vis. Sci., 21:116-125, July 1981.
10. R. W. Baloh, A. W. Sills, W. E. Kumley and V. Honrubia, "Quantitative Measurements of Saccadic Amplitude, Duration and Velocity," Neurology, 25:1065-1070, November 1975.
11. J. R. Tole and L. R. Young, "Digital Filters for Saccade Extraction and Fixation Detection," in Eye Movements: Cognition and Visual Perception, Lawrence Erlbaum Assoc., 1981.
12. A. T. Bahill, M. R. Clark and L. Stark, "The Main Sequence, A Tool for Studying Human Eye Movements," Math. Biosci., 24:191-204, 1975.
13. M. R. Clark and L. Stark, "Control of Human Eye Movements: III. Dynamic Characteristics of the Eye Tracking Mechanism," Math. Biosci., 20:239-265, 1974.

1983 USAF-SCREE 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 Evaluation of the Measurement System Used by
the Leadership and Management Development Center for
the Assessment of Its Consulting Efforts

Prepared by	Samuel B. Green
Academic Rank	Associate Professor
Department and University	Department of Psychology Auburn University
Research Location	Leadership and Management Development Center, Directorate of Research and Analysis, Maxwell Air Force Base, Alabama
USAF Research	Maj Lawrence O. Short, Ph.D. Capt Janice M. Hightower, M.S.
Date	Sept. 12, 1983
Contract No.	F49620-82-C-0035

An Evaluation of the Measurement System Used by
the Leadership and Management Development Center for
the Assessment of Its Consulting Efforts

Samuel B. Green

ABSTRACT

The purpose of the present research was to evaluate the measurement system employed by the Leadership and Management Development Center (LMDC) for the assessment of their consulting efforts. A number of changes were recommended after reviewing the literature pertinent to the measurement problems encountered by LMDC. Recommendations were made regarding the theory underlying the measures, their psychometric properties, and proper methods of interpretation. Throughout the report, emphasis was placed on the analysis of organizational units rather than individuals within the units. Finally, some suggestions were made concerning possible future research.

ACKNOWLEDGEMENTS

I would like to thank everyone at the Leadership and Management Development Center at Maxwell Air Force Base for making my summer a productive, enjoyable one. Special thanks are extended to Major Larry Short and Captain Janice Hightower for their help during the course of this project. Finally, appreciation is extended to the Air Force Systems Command, Air Force Office of Scientific Research, who made this summer research a rewarding experience.

I. INTRODUCTION

Consultative services are offered by the Leadership and Management Development Center (LMDC). The goals of these services are to improve the leadership and management skills of Air Force personnel and, in the long run, to enhance combat effectiveness through increased motivation and productivity. The consultative process used to achieve these goals involves five steps.

(1) A major unit commander or agency chief provides a written request for the services.

(2) LMDC consultants make a first visit to the unit or agency to measure characteristics of that organization. A primary measure is the Organizational Assessment Package Survey (OAP), an attitudinal questionnaire completed by a stratified sample or all personnel in the unit or agency. Additional measures include (a) the Executive's Impressions Questionnaire (EIQ), containing open-ended questions answered by senior supervisors as designated by LMDC consultants; (b) Supervisor's Impressions Questionnaire (SIQ), containing open-ended questions answered by all other supervisors; and (c) Objective Data Performance Indicators (ODPI), which includes file data such as judicial and disciplinary summaries, sortie rates, and inspection reports.

(3) The LMDC consultants determine the strengths and

weaknesses of the organization based on the data collected in the prior step.

(4) When the consultants return to the base on their second visit, they provide feedback to supervisors about the data and offer assistance to improve leadership and management skills. Specifically, a supervisor obtains a written report that gives summary analyses of the OAP's completed by the respondents whom he/she supervises. Also on the basis of the collected data, consultants select individuals who could benefit from a more intensive consulting effort. The techniques used in this effort are tailored to the needs of the individuals. Finally, workshops and/or seminars may be scheduled on particular topics in leadership and management.

(5) On the third and final visit, data comparable to those collected in step 2 are again obtained to provide an update on the unit and assist LMDC in retaining the consultative process.

II. OBJECTIVES

The major purpose for the data-collection steps (steps 2 and 5) is to provide feedback to supervisors about their units so that weaknesses in the organization can be remediated. However, a secondary goal is to give feedback to LMDC about the effectiveness of the consultative service so

that it may be improved. The distinction between these two purposes for data collection is important. The adequacy of the data to fulfill one goal is conceptually distinct from the adequacy of the data to fulfill the other. The objective of this report is to evaluate procedures used by LMDC to meet these two goals. However, this brief report can neither evaluate all procedures used by the LMDC nor consider all issues in the evaluation of any single procedure. Consequently, those aspects of the measurement process which seemed to present the greatest problems to LMDC were discussed in greater depth. Admittedly, almost all issues received much less attention than they deserve.

III. MEASUREMENT OF THE ORGANIZATION

The LMDC measurement process may be divided arbitrarily into five components: (1) organizational measures in general, (2) OAP Survey, (3) choice of respondents, (4) data analysis of pre-measure, and (5) data analysis of post-measure. Each of these components will be reviewed below.

ORGANIZATIONAL MEASURES IN GENERAL

LMDC consultants must have available to them an accurate view of an organization in order to make interventions that might lead to positive changes. Such a picture can be obtained only if the measurement process is guided by a

well-defined, comprehensive organizational model. An appropriate model should include multiple facets and clearly define the relationships among facets and between facets and behavior. Within such a framework, the goal of measurement should be to assess each facet precisely. In order to achieve this goal, each measure of a facet should yield reliable information. In addition, multiple measures using varying methods should be available to assess each facet. For example, a unit's productivity may be determined by (1) asking the individuals in the unit about their productivity, (2) asking the supervisor about the unit's performance, and/or (3) computing indices of productivity based on file data. Each of these measures has strengths and weaknesses. Only by obtaining all three types of measures (and perhaps others) can the facet be fairly evaluated.

Of course, all of these measures are useless unless the consultants actually review the results and have available a frame of reference to make an interpretation of them. The consultants are more likely to do so if these data are presented to them in an easy-to-digest manner. For example, a busy consultant is more likely to review and understand a computer-generated profile of scores that presents how a unit is doing in comparison with similar units than 10 individuals' responses to 15 open-ended questions.

LMDC has attempted to develop a set of measures that assess the multidimensional nature of the organization, including multiple measures of some facets, and to package the measurement results so that the consultants may find them beneficial. For example, the OAP assesses the perceptions of the workers on a number of aspects related to work, such as organizational climate, work characteristics, supervisor characteristics, and job satisfaction. This information is obtained by asking close-ended questions on a self-report measure (OAP), and, to some extent, by asking open-ended questions of supervisors (EIQ and SIQ) and by consultants' conducting interviews with selected supervisors during the second visit (step 4). For another example, performance is assessed using both self-report and file-data measures. LMDC also has made some of the information easily accessible to the consultants. The OAP data are computer-scored and the results for any work unit are compared with those of similar units from bases previously consulted. The consultants receive computer printouts of these analyses.

Although the measurement system employed by LMDC has a number of strengths, it could be improved. LMDC seems to be using two different organizational theories: (1) a contingency model of leadership and (2) a systems model of organizations. The contingency model was advocated by Hendrix (1976) and served as the basis for the development of

the OAP. On the other hand, a general systems model with its components -- input, process, and output-- is advocated currently. For example, in all the recent descriptions of the OAP, its factors are grouped on the basis of these three components. Even though this literature suggests that the specific systems model employed is an adaptation of McGrath's model, in practice some generic form seems to be in use by consultants. A well-defined theory is needed to produce better integration among measures, to insure that all important facets are being assessed, and to have greater continuity between the consultative and the evaluative efforts within LMDC.

Without a theory, it is difficult to evaluate whether the present set of measures adequately assesses a base's organization. However, some comments that are more limited in scope can be made. Because the data from the OAP are computer analyzed while the results from the other measures are not, LMDC tends to emphasize the findings from the OAP. An attempt should be made to enter data other than the OAP on the computer. For example, Feild (1982) suggested some performance indices that could be routinely collected. These indices may differ somewhat from unit to unit depending on the function of the units. However, if the same measures could always be obtained on units with the comparable functions, the data could be computer scored and a data base created for them. Similarly, a change in format

on the EIQ and SIQ would permit the resulting data to be stored on the computer. Some of the open-ended questions included on these questionnaires for supervisors could be rewritten so that a limited-response format like the one on the OAP could be employed. With such a system, a consultant could more easily determine, for example, not only how individuals within a work unit view their own productivity, but also how supervisors evaluate their performance, as well as their effectiveness as assessed with file data. The results would be of interest not only if they were in agreement, but also if there was a lack of congruence among them.

If the the EIQ and the SIQ were revised to include exclusively close-ended questions, some information would be lost; therefore, a revision of these instruments might still contain some open-ended questions. Alternatively, the OAP, EIQ, and SIQ could include only close-ended questions. Based on the results from these measures, consultants could determine which supervisors should receive only standard feedback of the data from these measures and which supervisors need a more intensive consulting effort. There is no need to obtain any additional data from the former group of supervisors. However, for the latter group, the information that could not be acquired from the close-ended questions might then be obtained during interviews with these supervisors.

OAP SURVEY

Given the emphasis on the OAP by IMDC, it is important to consider the quality of this measure in greater depth than the others. Although Feild (1982) evaluated the reliability of some of the performance indices, almost all the evidence on the psychometric quality of the measures has been devoted to the OAP. This evidence will be briefly reviewed and some suggestions will be made for future research.

The OAP survey contains 109 items that are divided into seven modules: (1) Background Information (16 items), (2) Job Inventory (34 items), (3) Job Desires (7 items), (4) Supervision (19 items), (5) Work Group Effectiveness (5 items), (6) Organization Climate (19 items), and (7) Job Related Satisfaction (9 items). Except for the items on the first module, respondents give their attitudes by responding on a seven-point scale, most frequently anchored at the extremes by "strongly disagree" and "strongly agree." Computer analyses of the OAP give mean scores on the 109 items, as well as mean scores on 20 factors based on a factor analysis of the items and 7 factors based on a non-statistical, rational clustering of items.

Hendrix and Halverson (1979) report the results of the factor analysis which determined how the survey should be

scored. The large number of factors that they found suggested that organizations within the Air Force are very multifaceted; however, later factor analyses (Hightower and Short, 1982; Hightower, 1982; Conlon, 1982) indicated that fewer factors were necessary. One explanation for the differences in results is that the nature of the organization has changed within the Air Force. Alternatively, they may reflect the different methods employed by the factor analysts. The latter explanation seems more plausible because (1) the analyses were all conducted within such a short period of time (three or four years) and (2) at least some of the analyses were conducted on overlapping samples. Given that statisticians have differing opinions as to which method is best for determining the number of factors, it is not surprising that applied researchers might disagree in their approaches to this problem. Nor is the choice purely an esoteric argument among academicians: different methods do produce different solutions when applied to real data.

Some methods for determining the number of factors include the following: eigenvalue-greater-than-one criterion (Kaiser, 1960), parallel analysis technique (Montanelli and Humphreys, 1976), scree test (Cattell, 1966), minimum average partial method (Velicer, 1975), the likelihood ratio criterion employed with maximum likelihood factor analysis, and interpretability of the factor solution. The

one method that has been applied most frequently by LADC researchers is the eigenvalue-greater-than-one criterion. This choice is a logical one in that it is by far the most popular method in application, the default option for most factor analytic computer programs, straightforward in application, and supported by some research with simulated data (Hakstian, Rogers, and Cattell, 1982). On the other hand, it is the approach that has been most frequently attacked in the factor analytic literature (e.g., Linn, 1968; Browne, 1968; Gorsuch, 1974). Zwick and Velicer (1982) recently made the following recommendation about this approach (referring to it as the K1 rule in their article): "Based upon the results of this study, which included samples drawn from simple, well defined population correlation matrices, we can conclude there is no evidence supporting the continued use of the K1 rule."

When Conlon (1982) applied the eigenvalue-greater-than-one criterion in his factoring of the OAP, he found 14 factors, but chose to interpret only nine of the factors. He dropped the remaining five factors on the basis of their lack of interpretability. This judgment is in agreement with that portion of the literature that suggests that the eigenvalue-greater-than-one criterion yields an overestimate of the true number of factors (e.g., Zwick and Velicer, 1982). It seems justified to conclude that the OAP may be represented with significantly fewer than the twenty

or more factors currently scored and available to consultants on computer printouts. In future analyses, LMDC should not use the eigenvalue-greater-than-one criterion, but rather use two or more of the other methods that were previously listed to determine the number of factors to retain.

The above, rather theoretical discussion probably overdramatized the importance of the judgment involving the number of factors underlying the OAP. Factors from analyses with a large number of factors do not differ interpretationally all that much from those with a fewer number of factors. What typically has occurred when too many factors were extracted was that a well-defined factor splits into a number of small, poorly defined factors. For example, Hightower (1982) reported a factor called Task Characteristics in her analyses which yielded 13 factors. This factor actually subsumes (with the exception of one item) four of the factors (OAP factors numbered 800, 801, 812, and 804) from the original factor analyses (Hendrix and Halverson, 1979).

A few additional comments can be made about the factoring of the OAP. The reported factor analyses were based on correlation matrices computed from individuals' responses on the OAP items. The choice of the individual as the unit of analysis needs to be reconsidered. Sociologists have

been aware of the difficulties associated with selecting the inappropriate unit of analysis since Robinson's classic article in 1950. He showed how correlations computed on aggregated data yield biased estimates of correlations that would have resulted if they had been calculated on individuals' results. Within the psychological and educational literature, rigorous discussions on the unit-of-analysis problem have occurred only recently (e.g., Burstein, 1975; Sockloff, 1975; Knapp, 1977; Sirotnik, 1980; Markham, Dansereau, Alutto, and Dumas, 1983). Perhaps some of the issues from this literature can be most easily understood with a few illustrations.

A researcher is interested in computing correlations between items on the Supervision module on the OAP. These items are completed by 50 groups of subordinates, with 10 subordinates per group. Each group is led by a different supervisor. The researcher could compute three types of correlations between items:

- (1) A total correlation could be computed by ignoring the grouping factor and basing any correlation on the 500 pairs of raw scores.

- (2) A between-group correlation could be calculated by obtaining a mean score for each group and computing a correlation on the 50 pairs of mean scores. Alternatively, a between-group correlation could be computed on weighted pairs of mean scores, with the weights

being the number of paired observations used to calculate the means for each group. These two versions of the between-group correlation yield the same value if the group sizes are equal.

(3) A within-group correlation is calculated by subtracting from each individual's score the mean score for his/her group and computing a correlation on the 500 pairs of deviation scores.

The three types of correlations have quite different interpretations:

(1) In general, researchers have argued that the total correlation is difficult to interpret because it confounds the between-group and within-group correlations. Specifically, the total correlation can be shown to be a function of the within-group correlation, the between-group correlation (if group sizes are unequal, the between-group correlation on weighted means), and the magnitude of the mean differences between groups (Robinson, 1950).

(2) The between-group correlation does address a relevant research question: Is there a tendency for supervisors who are perceived by their subordinates to show more of one characteristic also likely to be perceived as displaying more (or less) of another characteristic and are other supervisors who are rated as showing less of the first characteristic likely to be rated as

displaying less (or more) of the second characteristic?

(3) Because the within-group correlation is computed on deviation scores, the distinctions among supervisors have been eliminated (or, more accurately, mean differences among supervisors have been eliminated). Consequently, variation on an item indicates that subordinates have different perceptions of the same supervisors. They may reflect (a) different interpretations of the same behavior by different subordinates, (b) different behavior displayed by the same supervisor toward different subordinates, and/or (c) stylistic differences among subordinates in responding to survey items. The within-group correlation addresses the following research question: Is there a tendency for subordinates who perceive their supervisor to show more of one characteristic also likely to perceive their supervisor as displaying more (or less) of another characteristic and is there a tendency for other subordinates who rate the same supervisor as showing less of the first characteristic likely to rate him/her as displaying less (or more) of the second characteristic?

The logic used in the above example generalizes to more complex organizations. If the 50 work groups had come from 10 different squadrons, the total correlation could be

shown to be a function of three rather than two types of correlations, as well as the magnitude of mean differences among the organizational units (Knapp, 1977). The three types include within-work-group, between-work-group-within-squadron, and between-squadron correlations. The within-work-group correlation is identical to the within-group correlation with the simpler organization. A between-work-group-within-squadron correlation is calculated as follows: (1) compute item mean scores for the fifty work groups, (2) compute item mean scores for the 10 squadrons, (3) subtract the mean score for a squadron from the mean scores for work groups contained within that squadron, and (4) compute a correlation coefficient between the deviated mean scores for work groups (optionally weighted by the number of individuals in a work group). This index is an alternative to the between-group correlation, which ignores squadron effects. Finally, the between-squadron correlation is computed in the same way that the between-group correlation was for the simpler organization, but means and sample sizes for squadrons replace those for groups.

The number of coefficients that could be computed increases for organizations with a greater number of hierarchical levels. A different correlation can be computed for each level within the hierarchy. Also at any level within the hierarchy, a researcher has to decide whether to

control for or to leave uncontrolled the differences among divisions at the next highest level within the organization. To control for differences, the degree of covariation is computed on scores which are derived by subtracting the higher level means from the mean scores of interest.

Which type of correlation should LMDC choose when computing a correlation matrix among OAP items as input to a factor analysis? Although many of the different types of correlations might be interpretable, the choice is dependent on the focus of the items, the use of the test, and the objective of the analysis. For example, if items pertaining to supervisors are factored in order to develop scales, the choice is relatively clear. LMDC consultants primarily focus on differences among work groups with different supervisors, not individual group members' perceptions of the same supervisor. Although consultants are instructed to evaluate the distribution of scores for any work group, most consultants undoubtedly pay closest attention to the mean score for a group in comparison with other work groups. Also, LMDC has a limited number of consultants who visit a base and they must decide who should receive more intensive consulting efforts by evaluating OAP scores of any supervisor with those of other supervisors within that base. Consequently, because between-work-group-within-base correlations are based on covariability among work group means within a base, they would appear to be the statistic

of choice. Even though the factoring of these correlations may be most appropriate for the development of scales, the factoring of other types of correlations, such as the within-work-group correlation, may yield further insights into the Supervision items.

On the other hand, a different type of correlation should probably be computed between items on the Job Desires module. These items ask whether the respondent would like, for example, an easy job or work that can be done independently. Because they inquire about a respondent's personal needs, individual rather than organizational variables would seemingly underlie the responses to these items. If this hypothesis is correct, organizational level does not need to be controlled and a factor analysis of the total correlations would appear to be satisfactory. However, it would be possible to evaluate empirically this hypothesis (Markham, Dansereau, Alluto, and Dumas, 1983). If organizational variables did affect the relationships among these items, a coefficient could be chosen that was congruent with these empirical results.

A final example might be useful to illustrate further how a correlational index should be chosen. The Organization Climate module on the OAP contains items which describe characteristics of the respondent's organization. The General Information section differentiates between a

work group, "all persons who report to the same supervisor that you do," and the organization, "your squadron." In this case, the choice among types of correlations is a difficult one because the respondents are making judgments about one level of the organization, the squadron, while the consultants are frequently evaluating at another level of the organization, e.g., the work group. If LMDC were particularly interested in differences among squadrons, a factor analysis of the between-squadron correlations could be performed. Since the consultants do evaluate most often mean scores of work groups, a factor analysis of the between-work-group-within-squadron correlations might be more appropriate for scale development. The between-work-group correlation, which does not control for differences among squadrons, would probably be less appropriate than either of the other two correlations because of the anticipated impact of squadron variables on the correlations among the items from this module.

Based on the above example, it would seem preferable to evaluate the factor structure of the OAP items within each module rather than perform a factor analysis across all the attitudinal items on the survey. Using this alternative approach, a factor analyst may choose a correlation coefficient that is most appropriate for each module, instead of assuming the applicability of a single type of correlation for all pairs of items. Also, factoring within modules is

in line with the theory underlying the development of the OAP, in which the modules were conceived of as self-contained sets of items (Hendrix and Halverson, 1979).

A final note can be made about the factor analyses of the OAP items. Researchers in the social sciences have frequently had difficulty replicating factor analytic results. Like all multivariate statistical techniques, large sample sizes are required for stable results across replicates. Also, because small changes in the correlational pattern may produce alterations in the factor solution, consistency in results has to be evaluated across different segments within a population and/or across time periods with the same population of subjects. LADC researchers have recognized the importance of investigating the replicability and generalizability of the OAP factor structure. Hightower and Short (1982) have evaluated consistency of results across different population segments. Similarity in structure was found for different groups of individuals who perform different functions and across various demographic groups. In addition, they showed consistency in the OAP structure between results obtained before an intervention and those obtained after the intervention. Hightower (1982) investigated even more extensively factor stability over time. If the OAP is revised, LADC should evaluate the factor stability with as much care.

The consistency of the OAP factor structure across time is conceptually independent of the consistency or reliability of the survey's scale scores. In particular, the pattern of correlations among items may be similar for two testings, but a correlation between responses on an item for the two testings may be high or low. Short and Hamilton (1981) investigated the test-retest reliability of the OAP scale scores. The test-retest correlations, with a five-week interval between testings, were moderate in magnitude. Seventy-five percent of the values ranged from .70 to .90. However, the sample size for this study was very small (18 or 19 subjects). They also reported results with a much larger sample. These test-retest reliability coefficients were obtained with a six-month interval between testings. They were uniformly low; none of the coefficients exceeded .60. Short and Hamilton computed on the same samples a second type of reliability coefficient, Cronbach's alpha. As expected, the alphas were relatively high for factors with a large number of items and substantially lower (below .60) for factors with a few number of items. These results supported the commonsensical notion that two or three item scale cannot adequately represent an item domain.

Of the coefficients reported by Short and Hamilton, the test-retest reliability coefficients with a six-month interval between testings were the most meaningful. These

results indicated that OAP scores may markedly differ on two occasions, even when no intervention occurs between testings. The implication of this conclusion is quite disturbing: changes in scores on OAP factors between the first visit and the last visit by the LMDC consultants may be due to instability rather than due to an intervention effect. However, these reliability estimates indicate a lack of stability in the OAP factor scores for individuals, while LMDC makes judgments on the basis of mean scores. A mean score is typically more reliable than a score for an individual. Consequently, the low test-retest estimates do not accurately reflect the reliability of mean OAP scores. Reliabilities of means might be computed on the data analyzed by Short and Wilkerson. With the larger sample, mean scores for a work group could be calculated on both the test and the retest. These means would be correlated to yield test-retest coefficients on mean scores. Alternatively, reliabilities of mean scores could be estimated using generalizability theory (Brennan, 1983). An advantage to using generalizability theory is the ability to estimate the reliability of mean scores for different size work groups.

Additional research also needs to be conducted to establish the validity of the OAP. To LMDC's credit, the construct validity of the OAP has been investigated quite extensively using factor analysis. This body of research

has already been discussed. In addition, Short and Wilkerson (1981) have investigated the OAF's construct validity by evaluating mean differences among different functional groupings within the Air Force. They found that the pattern of mean difference across groups generally were in agreement with expected differences. More validity studies need to be done. The OAF scores need to be related to measures external to this survey and, particularly, measures of observed behaviors.

CHOICE OF RESPONDENTS

LMDC selects respondents to complete the OAF using a well-devised, stratified random sampling plan. A sufficient number of respondents are selected to estimate population means accurately. Of course, not every individual who is selected to participate can complete the survey and, therefore, the resulting sample is non-random. No change in method is recommended. However, the percent of individuals who were selected, but did not participate should be routinely calculated and recorded so that potential degree of bias in the sample can be assessed.

DATA ANALYSIS OF PRE-MEASURES

A major part of the consulting effort by LMDC is the feedback of subordinates' attitudes and opinions about work

to supervisors. The feedback is based almost exclusively on computer analyses of the OAP results. The analyses include descriptive statistics, which summarize the responses of individuals within the same work group, and inferential statistics, which evaluate if a work group's mean differs significantly from the mean of individuals who have previously taken the OAP and who belong to work groups which perform a similar function. The descriptive statistics are not stressed in the feedback of the OAP data because the raw scores from the OAP are not very interpretable. On the other hand, the results of the significance test are invariably included as part of the feedback package. A supervisor is told whether his/her group scored significantly lower (or higher) than individuals from comparable work groups. Although referring raw scores to a normative data base does give them greater meaning, the significance test currently in use has some drawbacks and alternative methods should be considered.

The significance test is an independent t-test. It is designed to evaluate if two population means are different from each other. The assumptions underlying the test are (1) that the population distributions are normally distributed, (2) that these distributions have equal variances, and (3) that a sample be obtained from these populations through random and independent sampling.

The data from the OAP do not appear to meet the requirements of the test. The third assumption appears to be untenable: the observations within the two samples drawn from their respective populations are not independent. One sample consists of the individuals from the work group of interest, while the other sample includes individuals from similar work groups. Although the observations between work groups are independent, the observations of individuals within a group are dependent. In other words, the probability of a supervisor receiving a high rating by a subordinate would increase given that another individual in the same group rated the supervisor as excellent. This dependency makes the statistical test inappropriate. (The test is also inappropriate for a number of other reasons. Some of these will become apparent in later discussion.) A slightly different perspective of this problem is that t-test currently in use is inappropriate because the unit of analysis, the individual, is incorrect and the more appropriate unit is the group.

A possible remedy to this problem might be to derive a group factor, with each group (including the one of interest) representing a different level of a factor. A planned comparison could then be performed to evaluate if the group of interest differs significantly from the average of the other groups. The problem with this analysis is that it is unclear what null hypothesis is being tested. The group

factor must be defined as a fixed-effects factor in order to perform a planned comparison. If it is a fixed-effects factor, LMDC evaluators must restrict their conclusions from the significance test to those groups included in the analyses. However, given such a restriction, no inferential statistics are necessary because the groups in the analysis are the population.

Alternatively, the levels of the factor could represent a sample of all possible groups; that is, the group factor might be defined as a random-effects factor. An f-test may be calculated to evaluate if the variance across all possible levels of the group factor is equal to zero. Post-hoc comparison tests are inappropriate following a significant omnibus f-test of a random-effects factor. However, it clearly is meaningful to determine where the group of interest stands in comparison with all possible groups, not only those that have been assessed up to this point in time. Any method that would obtain an accurate estimate of a group's percentile rank would be appropriate. These procedures are reviewed in standard texts on psychometric methods (e.g., Thorndike, 1971).

It should be noted that the f-test on the random-effects factor is not really necessary. Everybody would probably agree that the hypothesis of no differences among work groups has to be incorrect. Also, the f-test tells us

almost nothing about the work group of interest. Consequently, only the percentile ranks really need to be computed.

DATA ANALYSIS OF POST-MEASURES

After the third visit, LMDC analyzes the changes on the OAP factors from pre-intervention to post-intervention. A computer program calculates independent t-tests between the two sets of observations for each work group. As discussed in the previous section, an assumption underlying this test is independence among scores. In particular with the change score analysis, the pretest and the posttest data must be independent of each other. It is not likely that this condition is met with the LMDC data and, consequently, alternative statistical methods should be considered.

Before exploring the difficulties encountered with the LMDC data, it might be preferable initially to discuss the implications of performing the inappropriate analysis on data from a less complex example. A researcher was performed an independent t-test on data in which the same individuals took both the pretest and the posttest. It seems highly unlikely that the independence assumption would hold for these data. The standard test that is used to evaluate pretest-posttest change scores is the dependent t-test. By inspecting the equations for the independent

and dependent t-tests, one may conclude that no constant bias exists. In some cases, misapplication of independent t-test might yield an overly conservative test and with others, an overly liberal test (Green and Feild, 1970).

With the LMDC data, only some respondents complete both administrations of the OAP; other respondents complete either the pretest or the posttest. Although the degree of dependency between the pretest and the posttest for these data is not likely to be as strong as in the previous example, the partial overlap in subjects for the two testing sessions is still likely to distort the results of an independent t-test. Unfortunately, no alternative test is available with this type of data and, therefore, a significance test is not recommended.

LMDC researchers could refer a work group's mean to normative data to obtain a percentile rank. This alternative is similar to the one suggested for the analysis of the pre-measures. However, with the posttest data, a percentile rank for a mean score could be obtained in reference to the pretest data or the posttest data. Difficulties exist for both statistics.

With the first statistic, a report could include the following description:

At the time of our first visit to your base, the

supervisor received a mean rating by his/her subordinates which placed the supervisor at the nineteenth percentile in comparison with all other mean ratings received by supervisors who have not received consultation. At the time of our last visit, the supervisor had raised his/her mean rating so that he/she stood at the seventy-fifth percentile in respect to supervisors who have not received consultation.

The implication is clear: consultation by LMDC raised the score sixty percentile points.

Unfortunately, LMDC consultants did not necessarily cause this increase. It might have been due to a number of influences such as a change in personnel within the work group, reactivity to the pre-measure, measurement error, and a change in regulations. Of course, this criticism is not peculiar to changes in percentile ranks, but is true of change scores in general. A second difficulty is that transforming raw scores to percentile ranks produces scores which are not on an interval scale. If the underlying distribution is approximately normal, a change of 'k' percentile points in the middle of the distribution is easier to achieve than a 'k' change at the extremes.

If a posttest score were referred to a posttest distribution, the resulting percentile rank is not very meaningful. A report could include the following statement:

"The supervisor's ratings placed him/her at the forty-fifth percentile in respect to all other supervisors who have received consultation from LMDC." The statement is not very informative.

All the options considered for the analysis of post-measures appear to have difficulties. One option is to report both types of percentile ranks and to indicate that any change in percentile rank could be due to a multitude of causes, only one of which is the LMDC consultative effort. The problem with this statement is that it has the same connotation as the health warning that is attached to cigarette advertisements: "I do not want you to believe this cautionary note."

One final possibility is worth considering. Properly designed, experimental studies could be conducted to evaluate the effectiveness of consultation. An improvement score could then be developed on the basis of these studies that corrected for at least some of the irrelevant influences that affect change scores.

IV. ASSESSMENT OF EFFECTIVENESS

In a recent report, Conlon (1982) described a series of analyses that attempted to determine the effectiveness of the consultative effort by LMDC. He reached the following

conclusion:

The use of the current procedure for evaluating the ongoing impact of consultation on a workshop basis for feedback purposes is questionable given the regression effects and the dangers of trying to estimate and control for such effects based on limited control group data. I encourage LMDC to seriously weigh the benefits of continuing the present evaluation effort against its cost and to consider alternative evaluation strategies. As an alternative to the present evaluation methodology, it is suggested that LMDC consider a series of controlled studies which would investigate, by plan, the effectiveness of particular types of consultation efforts across various functional specialties in the Air Force.

In actuality, the analyses by Conlon represent the only effort by LMDC to formally evaluate the consultative services. Because I generally agree with Conlon's assessment of the evaluation system, I will restrict my discussion on this topic to a few comments about his analyses and to a consideration of another alternative approach to supplement his suggestion for a series of controlled studies.

On the basis of anecdotal evidence, he assumed that "the intensity of consultative efforts would systematically vary as a function of the pre-score." He suggested that assuming this relationship between the pretest scores and the

intervention effort did exist, a curvilinear relationship would exist between the pretest and posttest scores. The changes in the posttest scores with a unit change in the pretest scores would be greater with larger pretest scores. In addition, for individuals with low pretest scores, the amount of change in the posttest per unit change in the pretest should be greater for data that are collected at bases in which feedback to supervisors is given as opposed to bases in which no feedback is given (control bases). The results did not confirm his expectations and, on the basis of the data, he revised his ideas about the consulting process.

I found this analysis to be creative and informative. Conlon formulated a model of the consultative service, empirically tested the model, and formulated a revised model which awaits testing. In contrast to Conlon, I believe that continued analyses of this type would be fruitful; however, the approach could be improved upon by formulating a more precise, mathematical model. Also, such a model should include a sufficient number of variables to explain the consultative process. The model could then be evaluated following guidelines contained within the causal analysis literature. (For an initial exposure to this literature, see Fedhazur, 1982.)

Currently, LMDC could not easily perform such analyses.

A data base must be created that is designed for such analyses. It would contain pre-intervention and post-intervention means for certain types of work groups. A group would be included if it represented a type of work group that was on a number of bases. This data base could also serve as the normative data for the calculation of the percentile ranks necessary to interpret the OAP scores.

V. RECOMMENDATIONS

1. A well-defined, organizational theory is needed to produce better integration among measures, to insure that all important facets are being assessed, and to have greater continuity between the consulting and the evaluation efforts within LMDC.

2. An attempt should be made to enter data other than the OAP on the computer. For example, certain performance indices could be routinely collected and stored on the computer. Also, a change in format on the EIQ and SIQ would permit the resulting data to be stored on the computer.

3. The OAP may be represented with significantly fewer than the twenty or more factors currently scored and available to consultants on computer printouts.

4. In future analyses, LMDC should not use the eigenva-

lue-greater-than-one criterion, but rather use two or more of the other methods that were previously listed to determine the number of factors to retain.

5. Rather than perform a factor analysis across all attitudinal items on the OAP, it might be preferable to evaluate the factor structure of items within each module. If this approach is taken, the factor analyst may choose the correlation coefficient that is most appropriate for that module.

6. LMDC should continue to evaluate the factor stability of the OAP with as much care as they have shown in the past.

7. Reliability of mean OAP scores should be estimated in addition to the reliability of individual scores.

8. The OAP scores need to be related to measures external to this survey and, particularly, measures of observed behaviors.

9. LMDC selects respondents to fill-out the OAP using a well-devised, stratified random sampling plan. The percent of individuals who were selected, but did not participate should be routinely calculated and recorded so that potential degree of bias in the sample can be assessed.

10. In the feedback of subordinates' attitudes and opinions about work to supervisors, the computer analyses of the OAP results includes a t-test which evaluates if a work group's mean differs significantly from the mean of individuals who have previously taken the OAP and who belong to a work group which performs a similar function. This significance test is inappropriate for a number of reasons and should not be performed. In its place, a percentile rank for a work group should be calculated and given as feedback to supervisors.

11. LMDC should discontinue using a significance test to evaluate change scores from pre-intervention to post-intervention. Other analyses of the post-intervention measures were considered, but none were endorsed.

12. A data base should be created that would allow LMDC researchers to test empirically consultative models. It would contain pre-intervention and post-intervention means for specific types of work groups.

REFERENCES

- Brennan, R. L. Elements of Generalizability Theory. Iowa City, Iowa: ACT Publications, 1983.
- Brown, N. W. A comparison of factor analytic techniques. Psychometrika, 1968, 33, 267-334.
- Burstain, L. The use of data from groups for inferences about individuals in educational research. Unpublished Ph.D. dissertation, Stanford University, 1975.
- Catell, R. B. The scree test for the number of factors. Multivariate Behavioral Research, 1966, 1, 245-276.
- Conlon, E. J. Determining behavioral consultation effectiveness in The United States Air Force (AFOSR-81-013b). Bolling AFB, D.C., 1982.
- Feild, H. S. Using hard criteria to evaluate Leadership and Management Development Center consultations (USAF-SCEEE Summer Faculty Research Program). Maxwell AFB, Al., 1982.
- Gorsuch, R. L. Factor Analysis. Philadelphia: Saunders, 1974.

Green, S. B., & Feild, H.B. Assessing group change under conditions of anonymity: A problem in personnel research. Journal of Occupational Psychology, 1976, 49, 155-159.

Hakstian, A. R., Rogers, W. T., & Catell, R. B. The behavior of number-of-factors rules with simulated data. Multivariate Behavioral Research, 1982, 17, 193-219.

Hendrix, W. H., & Halverson, V. B. Organizational Survey Assessment Package for Air Force organizations (AFHRL-TR-78-93). Brooks AFB, Tx: Air Force Human Resources Laboratory, 1979.

Hightower, J. M., & Short, L. O. Factor stability of the Organizational Assessment Package (LMDC-TR-82-1). Maxwell AFB, Al: Air Force Leadership and Management Development Center, 1982.

Hightower, J. M. Temporal stability of the factor structure of the Organizational Assessment Package (LMDC-TR-82-2). Maxwell AFB, Al: Air Force Leadership and Management Development Center, 1982.

Kaiser, H. F. The application of electronic computers to factor analysis. Educational and Psychological Measurement, 1960, 20, 141-151.

Knapp, T. F. The unit-of-analysis problem in applications of simple correlation analysis to educational research. Journal of Educational Statistics, 1977, 2, 171-180.

Linn, R. L. A Monte Carlo approach to the number of factors problem. Psychometrika, 1968, 33, 37-71.

Markham, S. E., Dansereau, F., Alutto, J. A., & Dumas, M. Leadership convergence: An application of within and between analysis to validity. Applied Psychological Measurement, 1983, 7, 63-72.

Montanelli R. G., & Humphreys, L. G. Latent roots of random data correlation matrices with squared multiple correlations on the diagonal: A Monte Carlo study. Psychometrika, 1976, 41, 341-348.

Robinson, W. S. Ecological correlations and the behavior of individuals. American Sociological Review, 1950, 15, 351-357.

Short, L. O., & Hamilton, K. L. An examination of the reliability of the Organizational Assessment Package (LMDC-TP-81-2). Maxwell AFB, AL: Air Force Leadership and Management Development Center, 1981.

Short, L. O., & Wilkerson, D. A. An examination of the

construct validity of the Organizational Assessment Package Proceedings of the 23rd Annual Conference of the Military Testing Association (Vol. II). Arlington, Va.: U.S. Army Research Institute, 1981.

Sirotnik, K. A. Psychometric implications of the unit-or-analysis problem (with examples from the measurement of organizational climate). Journal of Educational Measurement, 1980, 17, 245-282.

Sockloff, A. L. Behavior of the product-moment correlation coefficient when two heterogeneous subgroups are pooled. Educational and Psychological Measurement, 1975, 33, 267-276.

Thorndike, R. L. (Ed.) Educational Measurement (2nd ed.). Washington, D.C.: American Council on Education, 1971.

Velicer, W. F. Determining the number of components from the matrix of partial correlations. Psychometrika, 1976, 41, 321-327.

Zwick, W. R., & Velicer, W. F. Factors influencing four rules for determining the number of components to retain. Multivariate Behavioral Research, 1982, 17, 253-269.

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

REINFORCEMENT INDUCED STEREOTYPY OF SEQUENTIAL

BEHAVIOR

Prepared by:	Dr. Arthur Gutman
Academic Rank:	Associate Professor
Department and University:	Psychology Florida Institute of Technology
Research Location:	Air Force Human Resources Laboratory Lowry Air Force Base, Colorado
USAF Research Contact:	Dr. Robert H. Summers
Date:	August 26, 1983
Contract No:	F49620-82-C-0035

REINFORCEMENT INDUCED STEREOTYPY OF SEQUENTIAL

BEHAVIOR

by

ARTHUR GUTMAN

ABSTRACT

The purpose of this report is to interface a major applied issue in Air Force technical training with basic research issues in the area of Learning and Memory. The final product is a proposal to conduct four basic research experiments relating to the deleterious effects of positive reinforcement on incidental learning and transfer of learning on a sequential learning task. The long-range purpose of the proposed research is to create a more general program of research to investigate the more global conditions that produce and, more importantly, prevent these deleterious effects.

Acknowledgements

This report is the result of an invaluable summer experience I had at the Human Resources Laboratory, Lowry Air Force Base, Denver, CO. I would like to thank the Air Force Systems Command, the Office of Scientific Research, and the Southeastern Center for Electrical Engineering Education for providing me with this opportunity.

There are so many individuals at Lowry who helped me with this project. In particular, I would like to thank Lt. Col. Allen J. Partin and his staff for their technical assistance. Also, there are a host of researchers with whom I shared very rewarding interactions. I would like to acknowledge Dr. James R. Burkett, Maj. Hugh L. Burns, Dr. Gerald M. Deignan, Capt. Richard T. Dineen, Capt. William E. Griffith, Dr. Karen P. Lane, Mr. Stephen V. Offutt, Dr. Roger J. Pennell, Maj. David L. Pohlman, Mr. Gerald S. Walker, and Dr. Joseph Y. Yasutake.

I am particularly indebted to Dr. Robert H. Summers, my focal Point Contact, for the knowledge and guidance he provided throughout the project. Also, I was quite fortunate to share many rewarding interactions with Dr. Phillip Langer, of the University of Colorado. His unanticipated presence greatly eased my task. Finally, I would like to acknowledge the collaboration and cooperation of my Graduate Assistant, Mr. Mark Blodgett. His effort and creativity make me extremely proud to be his teacher.

I. INTRODUCTION

The purpose of this proposal is to interface problems in Air Force technical training with current research findings in the area of Learning and Memory. My interest in technical training stems from my tenure as a Summer Faculty Research Fellow at Lowry Air Force Base (from 20 Jul 83 to 26 Aug 83). During this interval I interacted with a team of research psychologists and training officers who dealt with issues of training on a daily basis. I attended their planning sessions and, in addition, read technical reports and training manuals. On the basis of these experiences, the major theme that emerged was the question of need (or lack thereof) of classroom Residence Training (RT) programs for Air Force technicians.

It was interesting to note that a number of the training officers (and research psychologists) felt that RT provides little, if any advantage for ultimate job performance. In other words, the feeling was that most jobs can be adequately learned on the job site. Entering as an impartial observer, I had no stake in the truth or falsity of this belief. I did note, however, that most of the evidence favoring the belief is anecdotal. Clearly, the belief is widely held by the job site supervisors.

My own literature search (by computer and by hand) turned up only a single study that bore directly on the issue. The study (Black and Bottenberg, 1970) dealt with Category-B skills, and the results of it showed a myriad of interactions. In essence,

the results showed that some skills benefit from prior RT training, some benefit from on-the-job (OJT) training, and some benefit from neither.

It seems clear that the belief that RT is ineffective is not based as much on evidence as it is on hearsay and anecdote. Moreover, it would appear that the Air Training Command (ATC) is committed to the development of curriculum to be taught in the classroom setting. Thus far, the major efforts have been on software for computer assisted instruction (CAI), and on delivery systems such as the Unit Mastery approach. Given this, the relevant question becomes how to maximize transfer from the classroom to the job site.

The issue of transfer of training is itself complicated by two additional factors. First, with technological advances, the requisites of a job are often changed. Thus, what a technician is required to do on the job may be different from what he/she was trained to do in the classroom. Secondly, at the point of training, the trainee's ultimate job site is as yet unknown. Thus, even in the absence of technological changes, it is unlikely that the technician will end up doing (exactly) what he/she was trained to do. Given either of these factors, the important point is that if classroom education is to justify itself, it must produce flexible students.

II. OBJECTIVES OF THE RESEARCH EFFORT

From the point of view of the Learning/Memory literature, there are a number of topics which interface with the training issue discussed above. These include curriculum development, method of

instruction, issues of transfer (e.g., organization of material, state dependency, overlearning, etc.), and a host of others. But to this investigator, the single most relevant issue would appear to be that of positive reinforcement, per se.

Positive reinforcement (or reward) is used in any CAI program, particularly when the emphasis is on Unit Mastery. Although a large body of evidence suggests that rewards do strengthen the specific behaviors to which they are applied, there is a more current body of literature suggesting that this may be achieved at an undesirable cost (see Lepper & Greene, 1978 for a complete review). Given this, the objectives of this proposal are twofold:

- (1). to review relevant literature on the deleterious effects of positive reinforcement on initial learning and subsequent transfer; and
- (2). to outline a proposal for initial research on the effects of positive reinforcement on initial learning and subsequent transfer in sequential learning tasks.

III. DEFINITION OF TERMS

(A). Reinforcement: A reinforcer is any stimulus which when presented or removed increases the probability or frequency of the behavior it is contingent upon. The positive reinforcer (or reward) is the presented stimulus (e.g., a pellet of food for the rat's lever press response), and the negative reinforcer is the removed stimulus (e.g., termination or prevention of painful electric shock contingent upon the rat's running response). In the present research, the emphasis will be on positive reinforcement.

(B). Transfer: Transfer occurs when either present learning alters future learning (i.e., proactive transfer), or when present learning alters past learning (i.e., retroactive transfer). In either case, transfer is positive when the effect is to facilitate learning, and negative when the effect is to retard learning. In the present research, the emphasis will be on negative proactive transfer.

(C). CAI and Unit Mastery: The Unit Mastery system was formally proposed by Keller (1968). The most important feature of this system is that large units of material (e.g., a textbook) are broken down into much smaller, and presumably more manageable units. The student then proceeds unit by unit (usually at his/her own pace) until a mastery criterion has been reached on each unit. As the student is proceeding, there may be proctors (e.g., students who have previously taken the course) who administer tests and who offer help. Often, the role of the instructor is changed from the traditional lecture role to more of a motivating role. In more recent years, the tendency has been to create units for use on mainframes and microcomputers (hence the label "CAI").

IV. STATEMENT OF THE PROBLEM

As noted earlier, the ATC is committed to classroom education, and much time and resources have already been spent on CAI software. On face value, this is not a bad idea. Inspection of the various Job Description manuals reveals that most jobs are depicted in steps. For example, a report by Elwood, Warm, Thocher and Hritz (1983)

breaks down the job of refueling the B-52 into 36 distinct steps (i.e., from the review of safety precautions (#1) to disconnection of the LOX Cart Grounding Cable (#36)). From the Unit Mastery perspective, this job, and jobs like it, can be viewed as a sequence of units, each of which is mastered separately, and at one's own pace.

From the Learning/Memory perspective, there has been, since Skinner's (1938) initial proclamation, well over 40 years of evidence to support the contention that positive reinforcement strengthens specific behaviors. Skinner has always been a strong proponent of rewarding desirable and extinguishing undesirable behaviors (he argues in several places that punishment is not necessary). It is this belief that led Skinner (1968) to develop the teaching machine, and influenced Keller to develop the Unit Mastery system. But despite numerous documented successes (see the most recent issues of the Journal of Personalized Instruction), there is a growing body of literature suggesting that extrinsic rewards ^{can (many)} produce undesirable side effects. In particular, the evidence suggests three things:

- (1). Extrinsic rewards reduce intrinsic motivation to perform.
- (2). Extrinsic rewards for specific behaviors retard learning of nonreinforced, but related behaviors.
- (3). Extrinsic rewards for specific sequences of behaviors retards learning of new and/or different sequences of behaviors.

The implications of these possibilities are truly profound. From the point of view of CAI and Unit Mastery, extrinsic rewards (usually in terms of grades and/or points) are the cornerstone of

the teaching system. From the point of view of the job descriptions, it is sequential behavior that is the goal of technical training (i.e., from step #1 to step #N). Thus, if the aforementioned possibilities are true, then students may be learning very specific jobs and skills at the expense of more generalized knowledge needed in the field.

V. REVIEW OF THE RELEVANT LITERATURE

(A). Extrinsic Rewards: According to Skinner (1968), successful learning generates intrinsic rewards. That is, properly arranged programs of instruction (i.e., designed for student success) have built-in incentives for continued effort to achieve. Unfortunately, in most educational settings, there are always extrinsic rewards (e.g., grades or points), and it is the explicitness of these rewards which seems to create the undesirable side effects. The most basic finding in the literature is that when extrinsic rewards are provided for behaviors that are already preferred, the rewarded behaviors become less preferred.

For example, in a study by Lepper, Greene, and Nisbett (1973), children were first given the opportunity to choose among a list of attractive activities at playtime. Later, they were explicitly rewarded for engaging in the behaviors they most preferred. Relative to a nonreward control group, the reward group ultimately showed a marked reduction in preference for the initially preferred task.

This is a common finding that is not limited to children or type of task. But the important question relates to the mechanism by

which reduced incentive interferes with behavior. Is it simply a reduction in motivation to perform which affects any and all behaviors? Or is there an interaction between extrinsic reward and information processing?

Although early theorizing favored the simple notion of generalized decline in motivation, more recent evidence suggests that extrinsic rewards do, indeed, affect the manner in which information is processed. This evidence suggests that extrinsic rewards produce subjects who tend to choose easier problems to solve (Maher & Stallings, 1972). Also, they tend to focus on only those aspects of the problem they perceive as being related to the solution (Condry & Chambers, 1976).

The implications of such findings for Unit Mastery learning are obvious. Given a choice (which students often are), challenging problems will be avoided so that points may be built up. And, even if the system allows for no choices, students will limit their attention to what they believe is relevant to point production. This latter point relates to phenomenon called "Incidental Learning", which will be discussed next.

(B). Incidental Learning: The incidental learning paradigm is important because it reveals a distinction between what is known vs. what is done (i.e., learning vs. performance). Performance is directly observable, but learning must be inferred from performance. We tend to assume that good performance reflects good learning, and that poor performance reflects poor learning. However, given what is known about incidental learning, the implication of poor performance is not clear at all.

The demonstrative study in this area is by Tolman and Honzik (1930). Rats that were rewarded in the goalbox of a complex maze took ten sessions of training to reach asymptotic performance levels (as measured by time to complete the maze and by number of errors made). In comparison, nonrewarded rats showed random performance across the same ten-session interval. However, when given reward after 10 sessions, the previously nonrewarded rats took only a single session to reach the same asymptotic levels as the rewarded rats. Thus, despite nonreward, and despite random performance during nonreward, these rats had to have learned pathways to the goalbox during nonreward in order to do in one session, what rewarded rats did in ten sessions.

The implications of incidental learning are therefore twofold. First, learning is a cognitive process separable from performance; and secondly, reward, though seemingly necessary for good performance, is not necessary for learning.

In a surprising extension of these early findings with rats, more recent findings with humans suggest that nonrewarded subjects not only learn in the absence of reward, but also, may learn more than their rewarded counterparts.

In the demonstrative human study, Bahrick, Fitts, and Rankin (1952) asked adult subjects to track spatial events under reward vs. nonreward conditions. Peripheral stimuli unrelated to reward were occasionally turned off. Rewarded subjects not only performed more poorly on the target task, they also missed the incidental event with

greater frequency than the nonrewarded subjects.

This is clearly not an isolated finding (see also Bahrck, 1954; Davis & Lovelace, 1963; Johnson & Thomson, 1962). Moreover, it relates to the general issue of intrinsic vs. extrinsic motivation discussed earlier. If reward focuses attention on only those elements of problem solving that are directly related to reward, it makes sense that incidental tasks should be overlooked in the pursuit of the target behavior.

Returning to the classroom, the implication is that **students** will not attend to information unless it is an implicit part of the learning objectives. The danger here is that course objectives may not cover all the material that needs to be learned. If so, what is not covered explicitly, needs to be learned incidentally.

(C). SEQUENCES OF BEHAVIOR: The final relevant point in the literature is that sequential behavior, when explicitly rewarded, tends to fixate, or stereotype the response sequence. The origins are once again in the animal literature.

In a study by Schwartz (1980), pigeons were required to peck at either of two response keys. Rewards (access to grain) were presented after eight total responses, but only if an equal number of pecks were made to both keys. It is important to realize that in this procedure, specific order is not important - any sequence which leads to four pecks at each key is rewarded (and there are 70 such sequences in this particular task). Despite this, pigeons invariably pecked a given sequence 60-90% of the time.

More recently, Schwartz (1982) obtained the same effect with college students pressing either of two keys on a modified terminal keyboard (for money rewards). More importantly, once the subjects showed this stereotypy, **they** were retarded in learning new rules relative to control subjects who had never experienced any original training. It must be stressed that the experimental subjects showed negative transfer even though many of the component parts of the transfer tasks were present in the original learning task.

Once again, there are important implications for classroom education, particularly for the Unit Mastery system. If the goal of training is to obtain performance of one and only one task, and the performance requirements of this task never change, there is no major problem. However, if the nature of the task changes (e.g., due to technological changes), or if the student is asked to learn related, but different tasks, reinforcement induced stereotypy stands as a strong potential contributor to negative transfer.

It should be noted that there are also potential negative implications for the one-task/no-change situation. Suppose the task involves 30 or more steps, and suppose the performer loses his/her place? The literature on stereotypy suggests that the performer must start from the beginning (see Herrnstein, 1961).

There are other important implications, as well. For example, many applied researchers (particularly in the military) believe that overtraining induces greater long term retention of originally learned tasks. The evidence in the basic research literature on this point is

almost irrefutable. Nevertheless, incidental (but related) tasks should be equally as retarded as the originally trained sequence of behavior is strengthened.

VI. PROPOSED EXPERIMENTS

As noted earlier, the purpose of this proposal is to initiate a program of research. The point of initiation is Schwartz's (1982) procedure.

In Schwartz's procedure, subjects were instructed to accumulate as many points as possible by emitting sequences of left (L) and right (R) on two response keys. Points were accumulated on a counter as they were earned, and they were later converted to money rewards. Trials were administered in blocks of 50, with a 3-sec intertrial interval, and a 2-min rest period between successive blocks. A white light signalled the start of a given trial, and a red light signalled that points were earned. During a trial, subjects faced a matrix of light positions. As a left response was made, a light in the left column was turned off, and one under it was illuminated. As a right response was made, a light in the right column was turned off, and one under it was illuminated. Thus, a total of four presses to each key moved the illumination from top left to bottom right, at which point the red reinforcer light was illuminated, and the counter was incremented. In most of Schwartz's experiments, the reinforcer was delivered on only 50% of the correct response sequences (i.e., any of 70 possible sequences of four responses to each of the two keys).

A major concern with this procedure is that it limits response output. In a true learning situation, students must choose responses from a much wider set. In terms of the theoretical issues involved, factors such as incidental learning and cognitive vs. motor interference are difficult to study under such a restriction. Thus, the first task is to broaden the procedure.

(A). Experiment 1

All subjects will face a 2x2 matrix of keys. Each cell in this matrix will contain a L and R key. Thus, there will be a total of eight possible responses. In one condition, 50% reinforcement will be contingent on a sequence of Left/right responses (4 of each in any order) within the top left cell (i.e., Schwartz's condition). In the second condition, 50% reinforcement will be contingent on a sequence of left/right responses (4 of each in any order) across the top two cells. In other words, in this latter condition, a L response is defined as either of the two responses in the upper left cell, and a R response is as either of the two responses in the upper right cell.

The purpose of this procedure is twofold. First, it will allow systematic replication of Schwartz's procedure, and secondly, it will allow the creation of a Schwartz-like procedure that will allow testing of a number of theoretical issues. Given Schwartz's findings, there is every reason to expect equal amounts of stereotypy (i.e., a given sequence on more than 50% of the trials) in these two conditions.

(B). Experiment 2

One of the important broader issues is incidental learning. The second condition of experiment 1 permits a simple test of the

effects of reward vs. nonreward on incidental learning. In this cross-matrix procedure, a number of different response outputs may be defined, including:

1. Left/right within cell
2. Left/right across cells
3. Top/down across cells
4. Diagonal across cells
5. None of the above

There are others, of course, but these five will serve as basic "concepts". Concept #1 will be reinforced on the 50% contingency, but with a slight modification. If a correct sequence is made on a scheduled nonreinforced trial, another nonreinforced trial will be programmed. In other words, in at least 50% of the trials, nonreinforced sequences will be forced to occur. Each of the aforementioned concepts will produce a different colored light on the monitor. In the reward group, points will be accumulated on reinforced trials only. In the nonreward group, points will be accumulated randomly. The dependent measure here (other than degree of stereotypy) will be ability to recognize sequences other than the reinforced one (i.e., incidental learning).

(C). Experiment 3

A second broader issue concerns the degree to which Schwartz's interference effect is motor and/or cognitive. Schwartz's major finding is that once a subject shows stereotypy, he is retarded in a transfer task relative to a naive subject (who never experienced the initial learning condition). For example, the initial condition is four of each of L and R in any sequence, and the new condition may be a more specific requirement such as a double alternation (i.e., two of one

two of the other, two of the first, two of the second). The compelling factor here is that the experimental subject is retarded relative to a naive control even though the transfer condition involves the same motor elements as the original learning condition. Given the simplicity of Schwartz's task, it is entirely possible that the source of negative transfer is primarily motor. In experiment 3, both of the experimental groups will receive the left/right within cell condition as original learning, and one of them will be transferred to a double alternation within the same cell, whereas the other will be transferred to a double alternation across the top two cells of the matrix (i.e., two responses in one cell, two in the other, two in the first, two in the second).

Cognitively, these tasks have the same requirements - an initial general left/right sequence followed by a more restricted left/right sequence. However, motorically, the peripheral interference should be much greater in the within-cell condition. The two experimental groups will be compared to two naive control groups which will learn the transfer tasks without prior training.

(D). Experiment 4

Finally, in experiment 4, there will be a test of a third broader issue, namely overlearning. As noted earlier, overlearning should strengthen retention an initially learned task. But at the same time, there is every reason to expect that this increased retention may occur at the expense of incidental learning, and also, new learning. Thus, with stronger initial learning, there should be weaker incidental learning, as well as weaker transfer.

Operationally, criterion on original learning will be defined as two consecutive 50-trial blocks in which a stereotyped sequence occurred in 50% or more of the trials. Thus, 100% overlearning will be defined as two additional 50-trial blocks to the same criterion. Using these criteria, there will be two experiments.

Experiment 4a will replicate the experiment 2 incidental conditions with criterion vs. 100% overlearning, and experiment 4b will replicate the experiment 3 conditions, again with criterion vs. 100% overlearning.

VII. CONCLUSIONS AND IMPLICATIONS

To reiterate, the main purpose of the proposed experiments is to initiate a program of research. The questions asked in the proposed experiments are important in their own right, and hopefully, will lead to a large body of publishable data. But, in the longer run, if successful, they open the door to a series of important studies on the conditions which produce and prevent negative transfer in more generalized sequential learning procedures. In other words, the longer range purpose of the proposed experiments is more than just to demonstrate that reward ^{may} impacts learning and transfer. Rather, it is to investigate the conditions under which negative transfer may be prevented. Schwartz attempted to determine some of these conditions, but failed to find an "innoculation". The author thinks that this failure is due to the simplicity of the procedure, particularly as it confounds cognitive vs. motoric inputs to negative transfer. It is only after these two factors are separated that a true search for inoculation can begin.

REFERENCES

- Bahrick, H.P. Incidental learning under two incentive conditions. Journal of Experimental Psychology, 1954, 47, 170-172.
- Bahrick, H.P., Fitts, P.M., & Rankin, R.E. Effects of incentives upon reactions to peripheral stimuli. Journal of Experimental Psychology, 1952, 44, 400-406.
- Black, D. & Bottenberg, R.A. Comparison of technical school and on-the-job training as methods of skill upgrading. 1970, AFHRL-TR70-43.
- Condry, J.C. & Chambers, J.C. How rewards change the problem solving process. Manuscript in preparation, 1976.
- Davis, R.T. & Lovelace, W.E. Variable rewards and peripheral cues in discriminations by irradiated and nonirradiated monkeys. The Journal of Genetic Psychology, 1963, 103, 201-205.
- Elwood, W.F., Warm, R.E., Thocher, L.I., & Hritz, R.L. Validity and reliability assessment of performance evaluation instrument development procedures. Interim Report to AFHRL/ID, Lowry, AFB, 1983.
- Herrnstein, R.J. Stereotypy and intermittent reinforcement. Science, 1961, 133, 2067-2069.
- Johnson, R. & Thomson, C. Incidental and intentional learning under three conditions of motivation. American Journal of Psychology, 1962, 75, 284-288.
- Keller, F.S. Good-bye teacher. Journal of Applied Behavioral Analysis, 1968, 1, 78-89.
- Lepper, M.R. & Greene, D. The Hidden Costs of Reward. John Wiley & Sons, 1978, New York.
- Lepper, M.R., Greene, D., & Nisbett, R.E. Undermining children's intrinsic interest with extrinsic rewards: A test of the overjustification hypothesis. Journal of Personality and Social Psychology, 1973, 23, 129-137.
- Maehr, M.L. & Stallings, W.M. Freedom from external evaluation. Child Development, 1972, 43, 177-185.

- Schwartz, B. Development of complex, stereotyped behavior in pigeons. Journal of the Experimental Analysis of Behavior, 1980, 33, 153-166.
- Schwartz, B. Reinforcement-induced behavioral stereotypy: how not to teach people to discover rules. Journal of Experimental Psychology: General, 1982, 111(1), 23-59.
- Skinner, B.F. The Behavior of Organisms. New York: Appleton-Century-Crofts, 1938.
- Skinner, B.F. The Technology of Teaching. New York: Appleton-Century-Crofts, 1963.
- Tolman, E.C., & Honzik, C.H. Introduction and removal of reward, and maze performance in rats. University of California Publications in Psychology, 1930, 4, 257-275

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 TWO-DIMENSIONAL AEROELASTIC SYSTEM

Prepared by: Dr. Terry L. Herdman
Academic Rank: Associate Professor
Department and University: Department of Mathematics
Virginia Polytechnic Institute
and State University
Research Location: Air Force Flight Dynamics Laboratory, Flight
Control Division, Control Dynamics Branch,
Flying Qualities Group
USAF Research: Mr. Charles Suchomel
Date: September 14, 1983
Contract No.: F49620-82-C-0035

A TWO-DIMENSIONAL AEROELASTIC SYSTEM

by

Terry L. Herdman

ABSTRACT

In the classical formulation the aerodynamic loads depend on the instantaneous motion of the vehicle. That is, for example the lift at time t is modeled as a function of the velocity components at time t but not on any previous values. However, it is well known that the aerodynamic forces depend to some extent on the flow in the wake behind a lifting surface so that some "memory" effects are expected. The inclusion of these effects in flight-mechanics problems leads to a model including functional differential equations. A complete dynamic model is formulated for a two-dimensional system in which the elastic motions of a structure are coupled with the motion of the surrounding fluid. Suggestions for further development of a well-posed state space formulation 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 the opportunity to spend ten stimulating weeks at the Flight Dynamics Laboratory, Wright-Patterson Air Force Base, Ohio. In addition, he would like to acknowledge the hospitality of the members of Control Dynamics Branch.

I. INTRODUCTION:

The development of state space models for aeroelastic systems in unsteady aerodynamics is an important problem. For the Air Force in particular, the formulation of control and identification methods and the evaluation of potential designs establish the need for mathematical models that faithfully predict the dynamic behavior of the aeroelastic system.

The present project concerns the development of a state space model for motion of an airfoil in two-dimensional, unsteady flow of an inviscid, incompressible fluid and the examination of the well-posedness of the model. Although the actual model involves three-dimensional, unsteady, viscous, compressible flow we choose, as a first step in the development of a realistic mathematical model, to consider the simpler two-dimensional problem.

It is well known that aerodynamic forces depend to some extent on the flow in the wake behind a lifting surface so that some "memory" effects are expected (See [1,2,3,5,8]). The inclusion of these effects in flight-mechanics problems leads to a model which includes functional differential equations. Recent advances in the area of functional differential equations make it reasonable to include hysteresis in the model.

II. OBJECTIVES OF THE RESEARCH EFFORT:

The main objective of this project was to develop a useful mathematical model for the two-dimensional problem discussed above. We made no attempt to address the three-dimensional problem. Our specific objectives were:

- (1) Develop a mathematical model that faithfully predicts the dynamic behavior of an airfoil in two-dimensional, unstead flow of an inviscid, incompressible fluid.
- (2) Determine an appropriate state space so that the functional differential equation and the associated initial condition obtained in (1) is well-posed.

It is to be remarked that (2) above is essential for the development of efficient numerical approximating schemes.

III. Model

The physical situation is pictured in Figure 1. The functions $h(t)$, $\alpha(t)$, $\beta(t)$ denote the plunge, pitch angle and flap angle respectively.

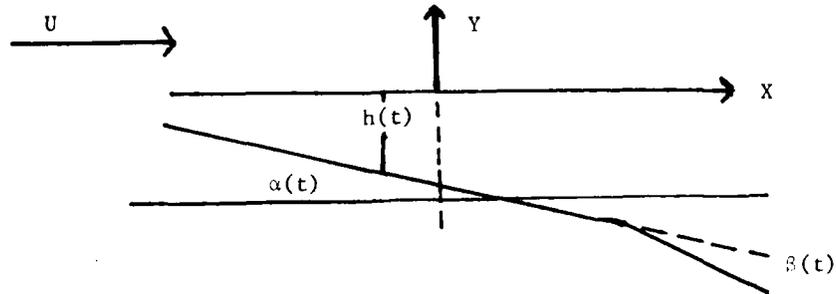


FIGURE 1 - TYPICAL SECTION GEOMETRY

Let $x(t) = [h(t), \alpha(t), \beta(t), \dot{h}(t), \dot{\alpha}(t), \dot{\beta}(t), \Gamma(t), \dot{\Gamma}(t)]$ where $\Gamma(t)$ is the circulation function. The mathematical model is given by the functional differential equation

$$\frac{d}{dt} \{Ax(t) + \int_{-\infty}^0 A(s)x(t+s)ds\} = Bx(t) + \int_{-\infty}^0 B(s)x(t+s)ds,$$

and the initial condition

$$x(s) = [h(0), \alpha(0), \beta(0), \dot{h}(0), \dot{\alpha}(0), \dot{\beta}(0), \Gamma(0), \dot{\Gamma}(s)]^T.$$

Note that the initial condition depends on the past history of $\dot{\Gamma}$ on $(-\infty, 0]$. For a more detailed development the reader is referred to references [1,2,3,8] and especially [4]. For a function $x: \mathbb{R} \rightarrow \mathbb{R}^8$ we define the function x_t by $x_t(s) = x(t+s)$ and define operators D and L on $L_1 \wedge L_\infty$ by

$$D\phi = A\phi(0) + \int_{-\infty}^0 A(s)\phi(s)ds$$

and

$$L\phi = B\phi(0) + \int_{-\infty}^0 B(s)\phi(s)ds.$$

Our model now takes on the form of a neutral functional differential equation

$$\frac{d}{dt} Dx_t = Lx_t, x_0 = \phi$$

where ϕ is a predetermined function. It is to be noted that the matrix A is singular (see [4]) and thus the operator D is non-atomic at $s = 0$. Consequently the recent results [7] concerning atomic operators can not be applied. At this time very little is known about the well-posedness of such problems (see [7] and [9]). In addition to the matrix A being singular (last row zero), the matrix function $A(s)$ is weakly singular, $A_{gg}(s) = \sqrt{\frac{Us-2}{s}}$, where U denotes the stream velocity. We point out that the singularity in A is at precisely the right position to be "cancelled" by the weak singularity of $A_{gg}(s)$ at $s=0$.

It is to be remarked that our model differs from those of Edwards and Balakrishnan [1,2,8] in that the states in our model are directly related to the physical states (i.e. pitch, plunge, circulation, etc.).

IV. WELL-POSEDNESS

The intent of our research effort is now to determine an appropriate state space so that the neutral functional differential equation of Section III is well-posed. The well-posedness is to be examined via semigroup theory for neutral functional differential equations with infinite delays. In particular, we wish to determine a state space Y of the form $R^8 \times L_p(-\infty, 0; R^8; g)$ so that the system of Section II induces a dynamical system on Y . The abstract dynamical system corresponding to $\frac{d}{dt} Dx_t = Lx_t$, $x_0 = \phi$ is conjectured to be

$$\frac{d}{dt} y(t) = A^*y(t)$$

with

$$y(0) = (\eta, \phi) \in D(A^*)$$

where the operator A^* is defined on

$$D(A^*) = \{(\eta, \phi) \in Y \mid \phi \in W_g^{1,p}, D\phi = \eta\}$$

by

$$A^*(\eta, \phi) = (L\phi, \dot{\phi})$$

where $W_g^{1,p}$ is the usual Sobolev space with measure g . The problem now becomes the problem of showing existence, uniqueness and continuous dependence on initial data of the solutions for the corresponding dynamical system given above. Based on earlier work [7, Section IV] for finite

delays it appears that the value of p (appearing in the state space Y) will have to be strictly less than two.

V. RECOMMENDATIONS

We have developed a mathematical model (Section III) for the motion of an airfoil in two-dimensional, unsteady flow of an inviscid incompressible fluid. Although the analysis is not complete at this time (Section IV) it appears that a state space can be determined so that the problem will be well-posed.

I propose to extend the results of Burns, Herdman and Stech [6,7] concerning the well-posedness of neutral functional differential equations with finite delay to include infinite delay equations and functionals that are not necessarily atomic at zero. In particular, the well-posedness of the model of Section III should be established. A second recommendation for follow-on research is the study of efficient numerical approximating schemes. It is to be noted that the well-posedness must be established before numerical schemes can be developed.

REFERENCES

1. A. V. Balakrishnam, "Active control of airfoils in unsteady aerodynamics", Applied Math. Optimization, 4 (1978), 171-195.
2. A. V. Balakrishnan and J. W. Edwards, "Calculations of the Transient Motion of Elastic Airfoils Forced by Control Surface Motion and Gusts", NASA-TM-81351, (1980).
3. R.L. Bisplinghoff, H. Ashley and R. L. Halfman, Aeroelasticity Addison-Wesley, (1955).
4. J. A. Burns, E. M. Cliff and T. L. Herdman, "State-Space Model for an Aeroelastic System," Proceedings of the 1983 CDC Conference.
5. J. A. Burns and E. M. Cliff, "Hereditary model for airfoils in unsteady aerodynamics, numerical approximations and parameter estimation," AFWAL-TR-81-3173. pp. 729-745.
6. J. A. Burns, T. L. Herdman and H. W. Stech, The Cauchy problem for linear functional differential equations, Integral and Functional Differential Equations, Lecture notes in Pure and Applied Mathematics Vol. 67, T. L. Herdman, S. M. Rankin III and H. W. Stech, Eds. Marcel Dekker, New York, (1981), 139-149.
7. J. A. Burns, T. L. Herdman and H. W. Stech, "Linear functional differential equations as semigroups on product spaces", SIAM J. Math. Anal., Vol. 14, No. 1, (1983), 98-116.
8. J. W. Edwards, "Unsteady aerodynamic modelling and active aeroelastic control", SUDARR 504, 1977, Stanford University.
9. Z.K. Pei, "On a Neutral Equation with Non-Atomic D-Operator", Ph.D. Dissertation, Univ. of Grätz, Grätz, Austria, 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

AN INTEGRATED APPROACH TO INTERFACE DESIGN

Prepared by: Stuart H. Hirshfield

Academic Rank: Assistant Professor

Department and Mathematics Department
University: Hamilton College

Research Location: Rome Air Development Center, Command
 and Control Division, Software Technology
 Branch, Artificial Intelligence Group

USAF Research: Ms. Sharon M. Walter

Date: August 16, 1983

Contract No: F49620-82-0035

AN INTEGRATED APPROACH TO INTERFACE DESIGN

by

Stuart H. Hirschfield

ABSTRACT

The question of how to design software systems requiring sophisticated user interfaces is investigated. Current systems have been evaluated and have been found wanting in terms of a variety of performance criteria. It is postulated that these shortcomings are attributable to (1) the lack of a formal design methodology which addresses properties idiosyncratic of such systems, and (2) the fact that two of the most sought-after performance criteria ("friendliness" and "domain independence") are to a certain extent incompatible. A semi-formal design methodology based on data flow techniques is proposed which represents an integrated approach to design wherein the extended interface and the tool are attended to equally. Furthermore, the methodology provides a framework for addressing the aforementioned incompatibility. Finally, suggestions for continued research in this area are offered.

ACKNOWLEDGMENTS

This research was sponsored by the Air Force Office of Scientific Research/AFSC, United States Air Force and was conducted under the auspices of the Southeastern Center for Electrical Engineering Education's Summer Faculty Research Program at the Rome Air Development Center. The author is indebted to these organizations and, in particular, to his Effort Focal Point, Ms. Sharon M. Walter, for their support. Through their efforts he has been afforded a thoroughly pleasant and rewarding research opportunity. Finally, the author would also acknowledge Dr. Northrup W. Fowler III at RAUC for his thoughtful guidance and sustained interest throughout this project.

1. INTRODUCTION

The computing community as a whole is investing a growing percentage of its resources in studies of what has come to be called "man-machine interaction." Many of these are research efforts which investigate the nature of such interactions and ultimately, when formalized and implemented as machines or programs, affect hardware and/or software design. Nowhere is the relationship between investigation and implementation more obvious than in the study of "user interface" - i.e., that aspect of man-machine interaction which focuses on the interface presented by an application program to its users.

In particular, much attention is being directed toward developing "friendly" interfaces. An interface is friendly to the extent that it facilitates and enhances communication between the users and application it mates. As recently as two years ago, an interface providing a high-level, 2-way communication channel was considered friendly. Studies have since extended the notion of friendliness to describe systems that are "graceful" [8], "robust" [4,7], "cooperative" [12], "adaptive" [15] and "efficient" [18] - almost human! It is, thus, no coincidence that researchers are now bringing many artificial intelligence theories and programming techniques to bear on the design and implementation of such interfaces.

Neither is it a coincidence that such research should be of interest to the Air Force. The more sophisticated a tool is, the more it stands to benefit from a sophisticated interface. As the number of computer controlled systems increases, so increases the demand for interfaces which can be used effectively in real-time situations with a minimum of specialized training.

The modest success of current friendly systems has spawned two independent, but related research thrusts. The

first concentrates on formalizing the notion of an interface. A number of abstract models have been proposed [11, 14, 20] which attempt to spell out all of the desirable properties of interfaces and organize them into a coherent whole. The common theme in all such models is a recognition that to completely specify an interface, one must take into account a model of the user - i.e., a description of a potential user in terms of his intentions, goals and training with respect to the application. Such a description is referred to as a "conceptual model". There is now a consensus that the degree to which an interface is friendly is reflective of the adequacy of its conceptual model.

Concurrently, other researchers have addressed the interface problem from a software management point of view. Friendly interfaces, particularly those endowed with natural language capabilities, tend to require large, complex programs which incorporate a great deal of application dependent information. In the interest of designing systems that lend themselves to effective implementation, studies have been conducted [5,6] which focus their attention on the notion of domain independence. That is, they have attempted to decompose their systems into components which distinguish information that is generic in nature from that which is specific to the application program. The common goal of these efforts is the development of an application independent natural language processing component.

The notions of domain independence and friendliness are clearly related in that both are desirable properties of an interface. Yet, most studies have pursued one at the expense of the other. Systems which have exhibited the greatest degree of friendliness (typically, natural language interfaces to database systems) have been the most difficult to decompose. There is little consensus as to what constitutes generic information, what constitutes domain dependent information, or the relationship between

these components. On the other hand, certain of the abstract interface models have clearly isolated the communicational aspects of a system from the application itself, but have mentioned natural language processing only in passing. In such systems, friendliness is achieved by means of complex conceptual models. Conceptual models have yet to be explicitly addressed in the context of natural language database systems.

II. OBJECTIVES

The main objective of this project was to investigate these research trends in an attempt to specify more precisely both the relationship between friendliness and domain independence, and the implications of this relationship to the design of user interfaces. While the limited time prevented us from implementing a prototype system, it was our intention to define a composite model which incorporated the critical aspects of each property.

More specifically, a representative collection of interface systems was assembled some of which concentrated on natural language capabilities, others on domain independence, still others on representing conceptual models. Each was evaluated in terms of (1) its method for defining conceptual models, and (2) the degree to which it achieved domain independence.

In the process of these evaluations, two additional objectives emerged which relate to all of these issues and, ultimately, became the focal points of this effort. One of the more striking common features of the evaluated systems was the arbitrariness with which they were designed, as indicated by their decomposition criteria. There seemed to be no method underlying their organization. One feature that was markedly different for every system was the notation used for representing it. Some were expressed in pseudo-English, others in frame-like formalisms and some in

code. I thus turned my attention to not only identifying the salient features of a composite interface model, but also to identifying a methodology and an associated notation for representing its features.

III. INTERFACES AND FRIENDLINESS

As the application domain of computers grows, so too grows the number of people who (often unwittingly) become computer operators. It is no longer the case that operators are highly-trained computer specialists. More often, they have some familiarity with the task they are putting the machine to, but are relatively unfamiliar and unconcerned with the workings of the computer itself. This generic class of user has been dubbed "naive" by the computing community whose task it is to provide friendly systems for this new clientele. Ideally, such a system would make transparent to its users that they are communicating with a computer at all and thus allow them to concentrate on the task at hand. Furthermore, the users should be able to communicate their needs and intentions to the system in as natural a way as possible - e.g., in natural language.

The value of providing natural language interfaces to a variety of real-world applications has been demonstrated beyond a doubt [3, 9, 18]. These successes, all of which restrict their domains so as to establish a "shared context" between the user and the tool, have shown that it is not enough merely to provide high-level access to a tool. A growing group of interface designers who share a psychological orientation now believe that this shared context must include not only information about the task domain, but also a model of the user - i.e., a conceptual model. This belief is based on the recognition that users take it upon themselves to construct models of systems they use. Such a model is thought to reflect the user's interpretation of a system's organization and functional

capabilities. The more closely the user's model corresponds to the system's underlying organization and capabilities, the more effective the communication between them.

Moran [14] defines a conceptual model as being composed of task level and semantic level information. The task level information represents the user's needs and demands on the system in terms of a hierarchy of user-conceived operations. These operations are redefined at the semantic level so as to coincide with the high-level operations actually provided by the system. The task level information is the part of the conceptual model brought to the interface by the user. The semantic level relates the user's view of the system to the system's functional capabilities. The other components of Moran's interface model, a communication component and a physical component, are seen to refine successively the conceptual model until a user's view of a particular task is mapped into a sequence of actions for carrying that task out. (See Moran for a complete description of Command Language Grammar.)

What is of practical interest about Moran's model is that it establishes a new "top" level of an interface - the conceptual model. Theorists such as Moran have recognized that, as the top level of a system, the conceptual model should be useful in guiding the design process. It should be defined first and should be allowed - even forced - to influence the design of the rest of the system. As Young [20] states, "One of the motivations for exploring the question of the user's conceptual model ... is that it provide a heuristic tool to the designer."

In practice, this has yet to come about. In fact, today's natural language interfaces make no explicit mention of a conceptual model, yet seem to perform reasonably well. This is attributable, in large part, to their common domain of application - database query. Within the realm of database theory, the importance of a

user model has long been recognized as critical to the interface. Date [2] defines an "external model" as "the view of the database presented by the transactions that may be invoked." For the user, this view is defined operationally - i.e., it is embedded in the logic of his transactions. The kinds of transactions that can occur are severely restricted and the contents of the database is well-known prior to implementation. The tasks that a user can ask of a DBMS are expressed in terms of the information in the database and the DBMS is designed to implement those tasks. Thus, the DBMS itself embodies an implicit conceptual model in the Moran sense.

Things are not so straightforward if we speak of designing natural language interfaces for other domains wherein we are faced with having to define a conceptual model explicitly. A user may approach an interface to a programming environment, for example, with an algorithmic description of his goals - i.e., they may be defined in terms of a structured sequence of operations. It is difficult at best to elicit from the user a conceptual model in such a relatively unconstrained case. Only recently have people come to recognize this as a more or less standard problem of systems analysis. Templeton and Burger [17] go as far as to say, "It is necessary to have extensive interaction with cooperative users in order to characterize the kinds of dialogues they will have with the system ... The importance of this interaction between potential users and system developers should not be underestimated as it is the basis for defining much of the knowledge base needed by the system."

Still, in a majority of cases, the developers of operational natural language interfaces go through no such interaction. Instead, they choose to design interfaces that interact with DBMSs (so as to take advantage of their implicit conceptual models) and then fortify the interface with an assortment of advanced reasoning and expert capabilities which reflect further assumptions about the

user's view of the application domain. It is precisely this tendency that clouds the issue of domain independence.

IV. INTERFACES AND DOMAIN INDEPENDENCE

A great deal of work has been invested in attempting to design natural language interfaces that are domain independent. Attempts have been made to identify a generic component that is distinct from those that describe domain specific information. The hope is that such a generic component could then be combined with descriptions of other domains to produce a variety of interfaces - i.e., it would be transportable. Current efforts toward this end, while commendable in practical intent, are tantamount to putting the cart before the horse.

Domain independence as espoused in the software engineering literature is not a guiding principle for design. It is but one byproduct of a rigorous, top-down design process based on the principle of modularity. If a system is designed to be highly modular, it tends to be relatively easy to code, debug, test, maintain and adapt to other domains. It is difficult, though, to render a system domain independent unless it was designed according to the rules of modularity in the first place.

This approach to attaining a transportable interface can be explained on two historical grounds: (1) most natural language interfaces are in fact "add-ons" - i.e., they have been appended to the front end of a pre-existing piece of code so as to afford a more facile means for communicating with it, and (2) software systems have traditionally been decomposed according to the input-process-output paradigm; in this case, natural language capabilities (input) being distinct from database query (processing). The early projects to address domain independence invariably distinguished the domain independent from the domain dependent information, or the

linguistic from the non-linguistic. More often than not, domain independent information was considered synonymous with linguistic (syntactic) information, while domain dependent information was regarded as primarily non-linguistic, or semantic.

A number of more recent studies have shown the distinctions to be quite arbitrary and, worse still, misleading in their impact on the design process. The work that ultimately led to the development of semantic grammars has demonstrated the relationship between semantic and syntactic linguistic information. More recently, Kaplan [12] and others have recognized that certain lexical information is in fact domain dependent. Hayes and Reddy [8] include both lexical and syntactic information in their tool description component. Ginsparg's "data base connection" [4] and Grishman's, et. al., "domain information schema" [5] both are intended to capture domain specific semantic information considered useful in processing natural language requests.

In addition to making evident the complex relationships between various types of information, these systems point to the prospect of an increasingly dominant domain dependent component. Grosz' work [6] is seen to further reduce the proportion of domain independent content by hypothesizing that much of the information needed to adapt an interface to a new domain can be obtained by using systems analysis techniques on users. That is to say, if the user is now considered part of the interface, much of his conceptual model is certainly domain dependent and thus must be adapted to each new application, as must all of the domain dependent linguistic information. Given the implicit nature of the conceptual model in most natural language interfaces and the complex relationship between linguistic information and the application domain, it is no wonder that attempts at domain independence (and thus portability) have been only moderately successful. In fact, this phenomenon is quite consistent with the trend to

make natural language interfaces "expert" in a particular domain.

V. AN INTERFACE METHODOLOGY

What, then, would we expect from a methodology for defining interfaces in order that it address the issues associated with friendliness and domain independence? Most generally, the methodology should embody an integrated approach to design which fosters modularity. Specifically, the methodology should: (1) be based on the assumptions that it is the responsibility of an interface to project a model of itself and the responsibility of the user to assimilate that model - i.e., it should recognize the user as part of the interface, (2) facilitate the attainment and description of a conceptual model describing the user's view of the system, (3) allow this conceptual model to influence the design of the interface and, when possible, the tool itself, (4) base its decomposition on formal criteria for modularity, and (5) be capable of specifying the complex interrelations between components of current "expert", or "knowledge-based" systems.

Such a methodology would address the concerns of both the psychologists and the software scientists. To the extent that these issues are addressed today, they are addressed only after the fact. What I am proposing is a methodology that not only addresses them, but allows them to influence the resulting design. Data flow analysis as described by Yourdon [21] is such a methodology.

Data flow analysis techniques were founded on the belief that most users can best describe a system in terms of its data and the transactions that affect this data. Such a data-oriented description often corresponds most closely with a naive users view of his job and, thus, enhances communication between the analyst and the user. Guidelines exist for transforming a data-oriented

description of a system into a hierarchical collection of processes. In the context of interface design, these techniques afford us a means for eliciting and representing a user view of the system, in addition to the design tools for systematically converting this information into a highly modular design.

Equally appealing is that it provides what I call an "informalism" - i.e., a less than formal, but still structured approach to designing a complex interface system. As an informalism, data flow techniques fall somewhere between the attempts of Woods [19], Reisner [16] and others to apply BNF-like formalisms to interface languages and the efforts of human factors researchers who apply behavioral criteria to the more physical components of an interface. Despite their informality, these techniques have been demonstrated to provide many of the advantages normally associated with the more formal methods as regards their impact on the design process itself.

Most interesting to me is the apparent correspondence between the data flow methodology and certain longer range issues that are due to impact the state of the art in natural language interfaces. The first issue is the as yet unspecified relationship between an interface and its tool. Recently, researchers have begun to investigate the possibility that both the interface and the tool can be enhanced if they are designed in concert - i.e., an integrated approach to design. The second, certainly broader in scope, is the relationship between data flow techniques and the so-called "fifth generation" computers. Both are based on data flow descriptions and make use of multiple, parallel, specialized (or, "expert") processors.

VI. RECOMMENDATIONS

The salient features of a composite interface model have been identified and a method for defining these features has been proposed. The value of incorporating a conceptual model of the user into the interface model and allowing this picture of the user to influence the design of the application itself is seen to outweigh the proposed advantages of domain independence. This result is seen to be consistent with both psychological and software-related treatments of interfaces. A strong case is made in support of data flow analysis techniques as a means for identifying and representing conceptual models. In that this approach supports an integrated, top-down design, it is anticipated that many of the benefits associated with domain independence may in fact be realized.

Still, the value of a model or a design methodology lies in their applicability to real problems. Also, value in this context is relative - i.e., it is not enough that such descriptive techniques apply to Air Force problems, they must compare favorably with other techniques according to some prescribed criteria. The criteria of most interest when discussing the design of interfaces include the performance of the interface and the ease with which the system was designed.

I propose as follow-on research to evaluate these methods by applying them to the design of an interface that has been developed by more traditional means. It is only by redesigning a system that we will be able to compare the approaches and the resultant designs and, thus, to evaluate the methodology. When comparable designs have been generated, they can be evaluated based on their friendliness and domain independence, as well as according to other formal design criteria.

The implications of this work are profound. As computer controlled systems become more sophisticated, the

need for equally sophisticated interfaces - as well as a methodology for their development - will continue to grow. Such work would overlap and further relate the areas of Artificial Intelligence and Software Engineering.

REFERENCES

1. Boguraev, B. K. and K. Sparck Jones, "A natural language analyzer for database access," Information Technology: Research and Development, 1, 1982, pp.23-29.
2. Date, C. J., An Introduction to Database Systems, Second Edition, Addison-Wesley, 1977.
3. Engleman, C., C. H. Berg and M. Bischoff, "KNOBS: An Experimental Knowledge Based Tactical Air Mission Planning System and a Rule Based Aircraft Identification Simulation Facility," Proc. Sixth Int. Jnt. Conf. on Artificial Intelligence, Tokyo, Japan, 1979, pp.247-249.
4. Ginsparg, J. M., "A Robust Portable Natural Language Data Base Interface," Proc. Conf. on Applied Natural Language Processing, Santa Monica, CA, Feb, 1983, pp. 25-30.
5. Grishman, R., L. Hirschman and C. Freidman, "Isolating Domain Dependencies in Natural Language Interfaces," Proc. Conf. on Applied Natural Language Processing, Santa Monica, CA, Feb. 1983, pp. 46-53.
6. Grosz, B. J., "TEAM: A Transportable Natural Language Interface System," Proc. Conf. on Applied Natural Language Processing, Santa Monica, CA, Feb. 1983, pp. 39-45.
7. Guida, G. and C. Tasso, "NLI: a robust interface for natural language person-machine communication", Int. J. Man-Machine Studies, (1982) 17, pp. 417-433.
8. Hayes, P. and R. Reddy, "Graceful Interaction in Man-Machine Communications," Proc. Sixth Int. Jnt Conf. on Artificial Intelligence, Tokyo, Japan, 1979.
9. Hendrix, G. G., E. D. Sacerdoti, D. Sagalowicz and J. Slocum, "Developing a Natural Language Interface to Complex Data," ACM Transactions on Database Systems, 3, 2 (June 1978).
10. Jacob, R. J. K., "Using Formal Specifications in the Design of Human Computer Interface," Comm. of the ACM, 26, 4 (Apr. 1983), pp.259-264.

11. Jagodzinski, A. P., "A theoretical basis for the representation of on-line computer systems to naive users," Int. J. Man-Machine Studies, (1983) 18, pp. 215-252.
12. Kaplan, S. J., "Cooperative Responses from a Portable Natural Language Query System," Artificial Intelligence, 19 (1982), pp. 165-187.
13. Ledgard, H., J. A. Whiteside, A. Singer and W. Seymour, "The Natural Language of Interactive Systems," Comm. of the ACM, 23, 10 (Oct. 1980), pp. 556-563.
14. Moran, T. P., "The Command Language Grammar: a representation for the user interface of interactive computer systems," Int. J. Man-Machine Studies, (1981) 15, pp. 3-50.
15. Mozeico, H. "A Human-Computer Interface to Accomodate User Learning Stages," Comm. of the ACM, 25, 2 (Feb. 1982), pp. 100-104.
16. Reisner, P., "Formal Grammar and Human Factors Design of an Interactive Graphics System," IEEE Trans. on Software Engineering, Vol. SE-7, No. 2, March 1981, pp. 229-240.
17. Templeton, M. and J. Burger, "Problems in Natural Language Interface to DBMS With Examples from EUFID," Proc. Conf. on Applied Natural Language Processing, Santa Monica, CA, Feb. 1983, pp. 3-16.
18. Warren, D. H. D. and F. C. N. Pereira, "An Efficient Easily Adaptable System for Interpreting Natural Language Queries," Amer. J. of Computational Linguistics, 8, 3-4 (July-Dec. 1982), pp. 110-127.
19. Woods, W. A., "Transition network grammars for natural language analysis," Comm. of the ACM, 13, 10 (Oct. 1970), pp. 591-606.
20. Young, R. M., "The machine inside the machine: user models of pocket calculators," Int. J. Man-Machine Studies, (1981) 15, pp. 51-85.
21. Yourdon, E. and L. L. Constantine, Structured Design, Prentice-Hall, 1979.

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

MESSAGE ROUTING METHODS FOR A TACTICAL AIR CONTROL SYSTEM

COMMUNICATIONS NETWORK

Prepared by:	Craig S. Holt
Academic Rank:	Assistant Professor
Department and University:	Department of Electrical Engineering Tufts University
Research Location:	Hanscom Air Force Base, Electronic Systems Division (MITRE location), Tactical Planning Branch
USAF Research:	Colonel Donald Miller
Date:	September 10, 1983
Contract No:	F49620-82-C-0035

MESSAGE ROUTING METHODS FOR A TACTICAL AIR
CONTROL SYSTEM COMMUNICATION NETWORK

by
Craig S. Holt

ABSTRACT

An investigation is made of possible routing methods for use in a communications network linking Tactical Air Control System Command and Control centers. Important qualities for the routing method are assumed to be ability to work in a distributed, amorphous network, ability to adapt reliably to frequent changes in the network, and ability to take advantage of one-way communication links. Previous routing methods are reviewed and a new method is proposed. In the proposed method each nodal unit reports the status of its inward links to all other units. Through these "updates" each unit is able to construct its own "model" of the system, and develop routing tables. A preliminary comparison of the methods is made, and areas for future research are recommended.

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 the opportunity for this summer project. He would also like to acknowledge the hospitality of the Electronic Systems Division/XR (MITRE location), especially the XRTE section. Numerous people in both the XR section and The MITRE Corporation were of great assistance in selecting and investigating a topic. Particular appreciation is due to Mr. Otto Wech and Mr. Murray Black for helpful discussions and key insights.

I. INTRODUCTION

In order for a deployable Tactical Air Control System (TACS) to perform its functions, reliable communications must be assured among the TACS command and control (C²) centers in the face of enemy counter action.

This report considers schemes for message routing within a TACS communication network. "Routing" refers to the process of directing each message through the proper communication links so that it ultimately arrives at its destination. The "routing algorithm" of a network is the collection of procedures used by elements of the network to accomplish the routing task.

In general, one desires a routing algorithm which adapts reliably to network changes, which provides the greatest possible throughput with the least possible delay, and which minimizes use of network resources (bandwidth, storage, and computation time) for routing "overhead".

Routing algorithms have been widely studied (see [1] for a survey). There are, however, some distinguishing characteristics of the distributed TACS network that make its routing requirements different from the requirements in most studies.

(1) The network is amorphous (has no specific structure) and is subject to rapid changes in its links and units. Links and units may be added, deleted or moved frequently. Multiple simultaneous failures are not unlikely.

(2) It is particularly important, in the TACS network, that when damage to a link or unit occurs there be as little disruption of service in the rest of the network as possible. The times when damage occurs are likely to be the times when proper operation of the network is most essential. (In some other networks, a temporary severe degradation of service, while the network recovers from failures, is acceptable.)

(3) Due to expected jamming effects, some directive communication links may work in one direction only. The requirement that the network maintain maximum service for as many units as possible in the face of jamming implies that the routing scheme should take advantage of one-way links as well as two-way links. (In most other types of networks, any communication path that works in only one direction is considered totally out of service.)

Though a number of routing algorithms satisfy requirements (1) and (2) to varying degrees, very few satisfy requirement (3).

II. OBJECTIVES OF THE RESEARCH EFFORT

The main objectives of this project were:

- (1) to investigate existing methods for routing in amorphous networks containing one-way links,
- (2) to propose a routing method which meets the requirement of a TACS network to the greatest possible degree, and
- (3) to compare the proposed method to other methods.

III. PROBLEM STATEMENT AND DEFINITIONS

A. The Routing Problem

We assume an abstract network which consists of a set of nodal units and a set of links connecting pairs of units. A unit represents any intersection of communication paths where the routing algorithm must determine which link to send a message out on. Links may be either one-way or two-way. The units at the ends of a link are referred to as the "link sender" and the "link receiver". Each unit refers to its own links by means of internal "port numbers" which are unknown to other units. It is assumed that units initially have no information about the network topology.

The routing algorithm must provide every unit with a "routing table" that tells the unit what port it should use to send a message destined for any other unit in the system.

B. Other Assumptions and Terms

The following notation will be used:

y is $O(x)$ -- (read "y is of order x") standard notation to indicate that the growth of quantity y is proportional to the growth of function x .

n -- the number of units in the network.

e -- the number of links in the network.

d -- the maximum number of links attached to one unit.

Due to hardware limitations in the TACS network, it is expected that d will be a small number, independent of n . Thus in assessing the complexity and communication load of each algorithm, it will be assumed that d is $O(1)$, and e is $O(n)$.

"Flooding" refers to a special procedure which provides a fast and reliable way to get data to every reachable unit in a network. When a unit wants to flood a network, it begins by sending copies of its message to all its neighbors. A unit receiving a flooded message first checks to see if it has already received another copy of the message. If no other copy has been received, the unit accepts the message, and sends copies out to all its neighbors (except possibly the one from which it was received). If a previous copy has been received, the new copy is discarded. Thus copies of the message spread rapidly outward from the originating unit, but the spreading stops wherever the message meets other copies of itself. Since each link is used at least once and no more than twice, the total network communication load involved in the flood is $O(e \times \text{message length})$.

Finally, a "directed message" is a message which has a particular destination, as contrasted with a flooded message. A directed message moves from unit to unit, each unit using its routing table to determine the next move.

IV. APPROACHES FOR DISTRIBUTED ROUTING

All the distributed routing algorithms of which the author is aware are based upon one or more of the following three methods:

Delay-Table-Passing: Each unit computes its delays to every other unit based on the delay tables passed to it by each of its neighbors. The neighbor showing the shortest delay to a given destination is used as a route to that destination. Changes in network conditions ripple through the network from neighbor to neighbor. Techniques have been developed to prevent certain "looping" instabilities that can occur with this type of algorithm [2, 3].

Flooded-Path-Searching: Flooded messages are used to discover the "best" paths from unit x to another unit y . When a message is flooded from x , the first copy to arrive at y is assumed to have taken the best path from x to y . If links are assumed to have the same delay in both directions, then y can record the port on which the message from x arrived and use the corresponding output port for sending messages back to x . (This is sometimes called "backwards learning".) If one-way links are involved, then y must somehow send the path information back

to x so that x will know the best route to y. In effect, a round trip from x to y to x is used to inform x of its path to y.

Model-Building: Each unit collects information about the operating links and units of the network, and uses this information to build its own "model" of the system, i.e. its own (possibly incomplete) view of the system topology. Once the system model has been constructed, an analysis is done to generate the routing table. There are a variety of methods for gathering information about the network and for analyzing the model to produce the routing table.

V. PAST WORK

In this section, several algorithms for routing in systems with one-way links are briefly described. For fully detailed descriptions of the Warner-Spragins and Brayer methods, see references [4] and [5]. The Wech algorithm has not been published.

A. Warner-Spragins Methods

Warner and Spragins described several routing algorithms, all of which are extensions of the Flooded-Path-Search approach, using Backwards Learning. In these methods, flooded search messages are sent periodically from every unit. An extra bit, called the Non Reciprocal Path Bit (NRPB), is added to each flooded message. The NRPB is initially zero, but is set to one if the message traverses any one-way link. By performing Backwards Learning using only those messages that arrive over two-way paths, units are able to determine the best two-way path to every destination (if one exists).

In the "Rapid Reciprocal Path Search" method, if a directed message must be sent to some destination to which there is no two-way path, then the message is flooded to all neighbors. The flood spreads only over those units that have no two-way path to the destination. With this method, a directed message will always arrive at its destination if any path exists, but if one-way paths are involved then multiple copies of the message may arrive.

In another Warner and Spragins method, the "Reciprocal Path Search Algorithm With Delay-Table-Passing", an extra Delay-Table-Passing mechanism is added to find one-way links. The unit at the receiving end

of each one-way link periodically floods out a special message containing:

- the location of the one-way link and
- a table of its own delays to all other units, each augmented by the one-way link delay.

This special message is called a "delay vector".

A unit receiving a delay vector can determine its delays to every destination through the one-way link by combining its own delay to the sending end of the one-way link with each delay vector value. Units keep track of not only the best two-way paths but also the best path through one-way links. The shortest path is used for routing messages. Entries in the "one-way path" table expire if they are not renewed within a certain time limit.

B. Wech Method

O. Wech has proposed another algorithm of the Flooded-Path-Search type. The algorithm will be presented first as two distinct steps (although we will see that the two steps are really performed by iterations of a single procedure).

First, each unit floods out an exploratory message to the network. The flooding process used differs from normal flooding in the following important detail: As the flood origin issues a message, it labels each copy of the message with the port on which it is being sent. Only the flood origin does this.

When the first copy of such a flooded message is received by any unit, the receiver records both the originating unit of the flood and the port used by that unit. After one round of floods, each unit *y* has a record of what port every other unit should use to send to *y*. This data is called the "experience vector" of unit *y*.

In the second step, each unit floods its experience vector back to all the other units. When a unit *x* receives the experience vector of unit *y*, it finds its own entry, which indicates the port *x* should use to send to *y*. Unit *x* enters this data in its routing table. By the end of the second round of floods, all routing tables are complete.

Though the above explanation treats the initial "exploratory" flood and the second "experience vector" flood as different steps, a single combination procedure actually combines both steps. Every set of flooded messages contains both the first step data (sender and port number) and the second step data (sender's experience vector). Thus each flood from x both explores paths away from x , and at the same time distributes the current experience vector of x .

Each unit continues periodic flooding so that any changes in the network will be sensed. To promote faster adaptation to network changes an immediate flood is sent whenever a unit detects a change in its experience vector. Routing table entries expire if they are not renewed within a certain time limit.

C. Brayer Method

The Brayer method is a Model-Building type method. Each unit's information about the network links is contained in its $n \times n$ "connectivity matrix". An entry of 1 in position c_{ij} of the matrix indicates that unit i has a link to unit j . An entry of 0 indicates no such link is known.

The connectivity matrix is constructed by means of path records appended to all messages. The path record is a list of all the units traversed by the message so far. Every time a unit handles a message, it examines the message's path record and adds any new links to its connectivity matrix. Then it adds itself to the message path record before sending the message onward.

Each message is routed along the best path that can be calculated from the connectivity table. If the connectivity table indicates no path, the message is randomly sent to any other unit which it has not already visited. Such "wandering" messages help explore new paths. If a message acknowledgement is not returned to the sender within a preset timeout, the sender will try again. After a certain number of failed attempts to send a message, the sender will conclude that the destination is unreachable.

So that failed links will be detected, each link is kept active by its sending end. If a unit stops receiving traffic on a link, that unit issues a "bad link" message. The "bad link" message is randomly

routed to all other units, informing them that they should remove a certain link from their connectivity tables.

The routing table is computed from the connectivity matrix by an algorithm of complexity $O(D.e.n)$, where D is the network diameter (i.e. the maximal number of units in a shortest path between two units).

VI. PROPOSED ROUTING METHOD

In this section the proposed routing algorithm is informally presented. A more detailed description of the algorithm is given in reference [6].

The method proposed here is a Model-Building type algorithm, based on the idea that the unit which first notices the failure or establishment of a link is best suited to inform the rest of the network of the change. In the proposed scheme, each link is kept active by its sending end. If there are no messages to send, the link is kept active with dummy traffic. Each unit constantly monitors the traffic on its input ports, and if a change is noted, floods an "update" message to the rest of the system. Update messages contain the identity of the source unit and information about the set of links on which that unit is receiving messages. Link information includes the identity of the link sender and the port number used by that sender.

From the updates received, each unit is able to construct a model of the part of the communication network accessible to that unit. This model is stored in graph form (unit records and link records, attached by pointers) so that the memory requirements are $O(n+e) = O(n)$. An algorithm of complexity $O(n+e) = O(n)$ is used to generate the routing table from the graph model. (See Section VI C.)

The proposed routing algorithm is quite similar to the one currently used in the ARPANET [7]. Both are Model-Building algorithms in which each unit distributes information about its own links. The proposed algorithm, however, is designed to make use of one-way links, while the ARPANET algorithm is not. Some major differences between the algorithms are:

- In the ARPANET each unit reports the status of its outward links, whereas in the proposed method, units can only report the status of inward links; outward link status may not be known.

- In the proposed algorithm, special methods are provided for adding and deleting nodes from the system model in the presence of one-way links.

- In the proposed algorithm, special measures are taken to avoid dependence on periodic updates from all units.

- The proposed algorithm for generating the routing table from the system model is different from the one used in ARPANET.

A. Adding Units to the Network

A problem arises when a new network (or a new single unit) is added to an operating network. The new connections will trigger updates informing the entire network about the new links along the border of the two networks. However, units on one side of the border will not find out about units on the other side unless those units experience some change that causes them to produce updates. This difficulty can be handled through use of the following four techniques:

1) A straightforward approach used by ARPANET is to require that all units issue periodic updates even if they experience no changes in their links. Such reports could be at relatively long intervals (ARPANET uses one minute) so that these extra flooded messages would not impose a large load on the network.

2) In the case of a single unit making its first connection to the network, the new unit could request a copy of its new neighbor's entire system model. Once the model was obtained, the new unit could keep up with future system changes like any other unit. Copying the model from a neighbor is an attractive method in that the generation of update floods from every unit is avoided.

3) It is often possible for units to detect the presence of new subnetworks and trigger their own updates using the following method: When a unit "x" receives an update from a unit "y" that was previously unknown to x, x issues a new update. The presence of new units thus stimulates the reporting of the entire system. These are referred to as "stimulated" updates, as contrasted with the normal "voluntary" updates.

In the case of two subnetworks joining, each unit would receive a rapid sequence of updates from previously unknown units. Rather than responding each time with unnecessary copies of the same update, the

unit should only respond to the first, then inhibit further simulated updates for a certain timeout period.

4) When two networks join by first creating a link in one direction, and then later creating a link in the other direction, then method 3 is not sufficient. This situation can be handled in the following manner: If a unit "x" receives an update reporting a link to a previously unknown unit "y", and if no update arrives from y within a short interval, then unit x floods the network with an "all report" message, asking all units to generate an update. Other units receiving the "all report" request are inhibited from generating their own "all report" requests for a short time.

The four methods just described can be combined as follows: When a single node enters a network it would use method 2 (i.e. inform the network of its presence and get a network model from a neighbor). The first update from the single new unit would have to include a special "no stimulate bit" so that all the other units would not respond with stimulated updates. When two networks each containing more than one unit combine, methods 3 and/or 4 would cause all units on both sides to produce updates at the proper times. Finally, a unit which senses no network changes would, after a relatively long delay, generate an update on its own. This is to assure that any routing information which is somehow lost or garbled will eventually be corrected.

B. Deleting Units From the Network

A second problem arises from the fact that only inward links are reported. Suppose some unit y becomes disconnected from the network. Unit y will flood an update indicating that its inward links are not working, but this update cannot reach the rest of the network. Hence the main network will not find out that unit y can no longer receive; it will only find out that unit y can no longer send. If no special action is taken, units could continue to try to send messages to unit y through broken links. The same reasoning holds if an entire subnetwork becomes disconnected.

This difficulty can be solved through use of some extra checking whenever an arc is removed from a system model. Whenever unit x receives an update which causes it to delete some arc from its model, the node representing the arc sender is checked to see if it still has a path to x. If not, the node is removed from x's model, along with any

arcs attached to that node. Removal of these arcs may, in turn, result in the removal of still more nodes. The entire process is accomplished by an algorithm of complexity $O(n+e) = O(n)$.

C. Algorithm for Computing Routing from a System Model

A simple algorithm is used to compute routes. This algorithm mimics the action of a flooded path search. The unit in which this algorithm takes place is referred to as u . Let M be the set of all nodes in u 's current system model. For each node x in M , the set of nodes that x has arcs to is referred to as $S(x)$ (successors of x).

For each node x in M there is also an entry in the route table, referred to as $ROUTE(x)$. It is assumed that whenever u receives an update listing a link from u to some node x , $ROUTE(x)$ is filled in with the port which u used to send to x . These entries provide the initial data for the REROUTE algorithm.

The algorithm is given in a psuedo-code form (similar to Pascal). In the algorithm, Q is a queue (first-in-first-out list) of nodes.

Procedure REROUTE

```
for each node  $x$  in  $M - S(u)$  do  $ROUTE(x) = \text{null}$ ;  
 $Q = S(u)$ ;  
while  $Q$  is not empty do begin  
  remove a node  $x$  from  $Q$ ;  
  for each node  $y$  in  $S(x)$  do begin  
    if  $ROUTE(y) = \text{null}$ ;  
    then begin  
       $ROUTE(y) = ROUTE(x)$ ;  
      add  $y$  to  $Q$ ;  
    end then  
  end for  
end while  
end procedure;
```

VII. COMPARISON OF ROUTING METHODS

In this section some observations are made about the expected overhead and performance of the various routing methods. These comments are intended to assist in future research by pointing out possible

strengths and weaknesses of the methods. Before any conclusions should be drawn about the "best" method for a particular application, empirical data must be obtained comparing the performance of the various algorithms under realistic conditions. Without such data it is difficult to assess the relative significance or insignificance of particular factors.

The comments in this section about the expected behavior of the proposed routing method are completely unconfirmed, since no simulation tests have yet been done with the proposed algorithm.

A. Communications Requirements

The Warner-Spragins methods and the Wech method require each unit to flood messages to all other units periodically. In both of the Warner-Spragins methods these flooded messages are of length $O(1)$, resulting in a total communication load of $O(n^2)$. In one of the Warner-Spragins methods, directed data messages are flooded out when no two-way path to a destination can be found. This flooding can represent an extremely high load when it occurs. In the other Warner-Spragins method, additional "delay vector" messages of length $O(n)$ are flooded out by every one-way link receiver, increasing the order of the load to some value between n^2 and n^3 .

In the Wech method the periodic messages are of length $O(n)$, resulting in a communications load of $O(n^3)$.

In the Brayer method the communication load is mainly in the form of path records which are attached to all messages. Whether a given path record represents a large or small overhead in a given system depends, of course, on the comparative size of the rest of the message.

The average size of path data is proportional to the average path length of a message, which grows very slowly as the number of units increases. (For a network with uniform connectivity and uniform traffic, the average path length increases roughly as $\log n$ [7].) When changes occur in the network, loading becomes larger due to the longer path lengths of messages randomly searching out new paths or taking nonoptimal paths. These effects, however, are difficult to quantify.

In the proposed routing method, initial model building requires flooding an "update" message of d items from every unit. This is a total communication load of $O(n^2)$. After this initialization, however, repeated frequent flooding by all units is not necessary. When links

or units fail, only the direct neighbors of the failure need to generate updates. Flooded updates from all units are essential only during initial system start-up, or when two networks are joined. (Periodic updates may be used as an extra reliability measure, but intervals between these updates can be very long.)

B. Storage Requirements

The Warner-Spragins and Wech methods require tables in each unit of size $O(n)$. The proposed method requires more space (since a system model must be stored) but the size order is still $O(n+e) = O(n)$. The matrices used in the Brayer method are of size $O(n^2)$.

C. Computation Requirements

The algorithms used in the Warner-Spragins methods are of complexity $O(d) = O(1)$, except for the algorithm for accepting delay vectors, which is of complexity $O(n)$. The complexity of the algorithms for the Wech method is $O(1)$, for the Brayer method approximately $O(n^2 \log n)$, and for the proposed method $O(n)$.

D. Performance in an Unchanging Network

Warner-Spragins Rapid Reciprocal Path Search method always uses the shortest two-way path if one exists. If no two-way path exists, the message is flooded out until it reaches units which do have a two-way path (if any). These policies can cause unnecessary congestion.

Warner-Spragins' "Delay Vector" method, the Wech method, and the proposed method all use the shortest available path, be it one-way or two-way.

The Brayer method has not been tested in systems with one-way links, and there is reason to suspect that the performance of the algorithm will be significantly degraded in such networks. Since each unit gathers network information from paths of arriving messages, each unit mainly finds out about the paths leading towards itself. If these cannot be assumed to be two-way paths, the information is of little value. The only way a unit x can find out about paths leading away from itself is through randomly routed messages which happen to traverse a path away from x , and then wander back through x . It may take considerable time for randomly routed messages to loop back to x , and it seems unlikely that such messages would ever give x a complete map of the system.

E. Response to Network Changes

In both the Warner-Spragins and Wech methods, the time required to adapt to network changes is governed mainly by the rate of the periodic updates. In the Wech method the time to determine unit x's route to unit y includes an additional factor related to the round-trip message time from x to y and back to x.

The ability of the Brayer method to adapt to changes is difficult to predict. Large delays are possible if it becomes necessary to send out a randomly routed message looking for a path. It is also possible that, due to incomplete random exploration, units will not always take advantage of new paths when they appear.

The proposed method is expected to adapt to network changes very quickly. The units first detecting network changes flood the new information to all other units. Thus the time before unit x adapts to a network change is limited only by the message transmission time from the damaged unit or link to unit x. This is a fundamental limit regardless of routing method.

The proposed method is expected to produce minimal overhead traffic when network failures occur. After a failure, units never send out flooded or wandering messages to search out new paths, since each unit already has all the current information about alternative paths in its model. Upon being informed of a failure, a unit will either derive a new path from its existing system model, or conclude that no path exists. The system-wide information gathered by each unit before failures effectively reduces the unit's need for new information immediately after a failure. This may be an important advantage in TACS applications.

Another possible advantage of the proposed method is its minimal dependence upon time-outs. The methods of purging old information and adding new information to the system model do not depend upon messages arriving within critical time-out intervals.

VIII. RECOMMENDATIONS

A number of aspects of routing need further work. Most of the routing algorithms studied here include, or can be extended to include, methods for adapting routing to local traffic loads. The effectiveness of these mechanisms needs further study. Policies for flow

control and task scheduling should be studied along with routing, since these mechanisms interact [8]. For routing in very large networks, some sort of "regional" routing scheme might be considered, wherein messages are delivered first to the proper region, then to the proper unit within the region.

As previously mentioned, empirical data is needed in order to make meaningful comparisons between routing algorithms. Such data could be obtained through careful computer simulations. It is important in such simulations that representative size, topology, data rate, traffic patterns, and damage patterns be used. A simulation program partially constructed for this project could serve as a starting point.

Many larger issues involved in the design of a distributed TACS network also need to be addressed. The following are only a few: -

Routing methods deal mainly with the task of delivering a message to a specified nodal unit. The problems of determining the unit in which a particular destination process resides also needs attention. (Brayer has considered this in his routing scheme.)

- The quantities, types, sources, and destinations of information to be sent are of critical importance in designing a communication system. Better characterization of future TACS communication loads are urgently needed.

- Possible network control methods (store and forward, packet switching, broadcast, etc.) for the TACS network need further study. The different types of information involved in TACS command and control impose conflicting demands on the network, making these design decisions very complex. Hence, appropriately focussed research in all of these areas seems indicated.

REFERENCES

1. M. Schwartz and T.E. Stern, "Routing Techniques Used in Computer Communication Networks," IEEE Transactions on Communications, Vol. COM-28, No. 4, April 1980, pp. 539-552.
2. P.M. Merlin and A. Segall, "A Failsafe Distributed Routing Protocol," IEEE Transactions on Communications, Vol. COM-27, No. 9, Sept. 1979, pp. 1280-1287.
3. J.M. Jaffe and F. H. Moss, "A Responsive Distributed Routing Algorithm for Computer Networks," IEEE Transactions on Communications, Vol. COM-30, No. 7, July 1982, pp. 1758-1762.
4. C.J. Warner and J. D. Spragins, "Simulation of Communications Protocols for a Tactical Radar Network," Final Report for AFOSR Grant 78-36/9, Directorate of Mathematical and Information Sciences, Bolling AFB, Aug. 1979.
5. K. Brayer, "Implementation and Performance of Survivable Computer Communication with Autonomous Decentralized Control," IEEE Communications Magazine, Vol. 21, No. 4, July 1983, pp. 34-41.
6. C.S. Holt, "Distributed Routing Methods for a TACS Communications Network," ESD Technical Report (in preparation), Hanscom AFB,
7. J.M. McQuillan and I. Richer, and E.C. Rosen, "The New Routing Algorithm for the ARPANET," IEEE Transactions on Communications, Vol. COM-28, No. 5, May 1980, pp. 711-719.
8. H. Rudin and H. Mueller, "Dynamic Routing and Flow Control," IEEE Transactions on Communications, Vol. COM-28, No. 7, July 1980, pp. 1030-1039.

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

ATTITUDE CONTROL ISSUES FOR LARGE FLEXIBLE SPACE SYSTEMS

Prepared by:	K.C. Howell
Academic Rank:	Assistant Professor
Department and University:	School of Aeronautics and Astronautics Purdue University
Research Location:	Air Force Rocket Propulsion Laboratory, Edwards Air Force Base, California
USAF Research Colleague:	Captain David Perkins
Date:	July 22, 1983
Contract No.:	F49620-82-C-0035

ATTITUDE CONTROL ISSUES FOR
LARGE FLEXIBLE SPACE SYSTEMS

by
K.C. Howell

ABSTRACT

Attitude control for large, flexible systems involves the on-orbit functions of disturbance cancellation, pointing, maneuvers, stationkeeping and shape control. The auxiliary propulsion system, although not solely responsible for these activities, will certainly interact with all of these functions. It is necessary to understand these interfaces in order to effectively design a control system. This work considered some of the issues involved in control/propulsion interactions. In particular, three methods, at various stages of development, were examined which attempt to determine the optimal number and distribution of actuators for effective attitude control.

ACKNOWLEDGMENTS

I would like to thank all those at the Air Force Rocket Propulsion Laboratory who offered their support and assistance during my stay. A special note of thanks to Dr. Richard Weiss. I was also aided a great deal by members of the Applications Analysis Branch including Mr. Richard Matlock and Captain David Perkins. Finally, I acknowledge the Air Force Systems Command, the Air Force Office of Scientific Research, and the Air Force Rocket Propulsion Laboratory for sponsorship of the program.

I. INTRODUCTION

Large space systems are expected to be deployed sometime within the next two decades. They will significantly change the way in which the space environment is utilized. For analysis of these large, lightweight structures, some "old" forces will have increased importance, such as aerodynamic drag and gravity-gradient forces in lower orbits and solar pressure forces at higher altitudes. Combine these with the flexible nature of a large space structure and it is apparent that the on-board propulsion systems may also be vastly different from those on the relatively small, dense spacecraft of today.

With work proceeding on the design of primary propulsion systems to move large, flexible vehicles, interest is growing in the area of auxiliary propulsion. The interaction of the secondary propulsion system and the structure will be of major importance in performing the on-orbit functions of attitude control, which includes cancellation of disturbances, pointing, and maneuvers, as well as shape control and stationkeeping.

To design auxiliary propulsion for the use of the attitude control system, it is necessary to understand the control problems faced in large, flexible systems and areas of control theory still under development. Actual implementation implies matching interfaces and determination of specific requirements. It is appropriate then, at this time, to examine control law design for these systems and to consider the impact upon propulsion system development. Preliminary work indicates that for electric systems, in particular, this is an attractive application. Ultimate understanding of exactly how to make that happen is a key.

II. OBJECTIVES OF THE RESEARCH EFFORT

This work was concerned with examining the attitude control system for a large, flexible space structure, in particular the questions involving auxiliary propulsion. Effort was to be directed at one aspect of the problem. The chosen area involved the number and distribution of actuators, which will include mass expulsion devices to some extent. Initially, some of the work done to date in this area was to be collected and summarized to indicate not only the state of progress in this subject but also to bring forth areas that have yet to be examined. The goal

was to find a method by which the selection process could be done. Because optimal choices were desired, the control law design was the determining factor.

III. LARGE SYSTEM CONTROL/PROPULSION

A. CLASSES OF STRUCTURES

The control of large, flexible space systems is a huge area of research that has been getting a significant amount of attention only in recent years. Contributing to the complexity is the fact that there is an infinite variety of vehicles in this category. Although the ultimate goal, of course, is to handle any configuration, it is common at this stage of analysis to consider generic classes usually with subdivisions. The following three classes show a popular grouping:

1. Planar Structures
 - a. large flat arrays (plates)
 - b. cross structures
2. Single Antenna Systems
 - a. box structures
 - b. modular antenna systems
 - c. maypole or hoop/column
3. Multiple Antenna Systems
 - a. Modular antenna systems
 - b. multiple antenna systems (series of antennas)

These are the types of structures which will require on-orbit control. Structural models can be obtained to aid in the control design. None of the models includes a particular secondary propulsion system.

B. CONTROL PROBLEM

In order to integrate the auxiliary propulsion system in the control design, it is necessary to understand the scope of the control problem. When on-orbit control activities are discussed under the title of "attitude control", it can be very misleading. Even for the small, dense spacecraft flown to date, the control problem included not only attitude but stationkeeping functions as well. Much useful information was obtained from analysis considering simple rigid bodies sometimes with flexible appendages and the control functions themselves could frequently be decoupled.

Moving to large, lightweight, flexible systems implies meeting stiffer performance requirements in the face of increasingly complicated vehicle characteristics. The control functions now include not only attitude control and stationkeeping, but also the complex problem of shape control. The complications combine a totally flexible vehicle with subsystem interactions, imperfect knowledge of the structural characteristics, disturbances related to large vehicles, inadequate passive shape control approaches, and drastic changes in the dynamic characteristics during the life of the vehicle. With the control system of a large, flexible structure asked to do a significantly better job than those in the past, in spite of the added complexities, we have added tremendously to the demands on the system.

C. CONTROL/PROPULSION INTERACTIONS

Some of the interfaces between the secondary propulsion system and on-orbit control functions can be understood by constructing a sensitivity matrix as seen in Fig. 1. It has been truncated for simplicity. Attitude control has been broken down into the three primary functions of disturbance cancellation, pointing, and reorientation maneuvers. Also listed are five significant propulsion system characteristics, not necessarily in order of importance, which will interact with the control functions. The strongest interactions are indicated with an X. Actually, at various levels of design, a good case could probably be made to mark all possible spaces. Also, of course, there are additional considerations which will be coupled with the items already listed such as structure size, lifetime, stiffness and strength, thermal effects, coordination with other forms of actuation, sensing, etc. It is obvious that the development work in this subject has barely begun.

Given the problem and the huge number of areas needing development, I chose to examine a particular piece, knowing that in reality it cannot be isolated. Implementation prompts three fundamental questions:

1. What types of sensing and actuation are required?
2. What is the least number required?
3. Where should they be located?

CONTROL FUNCTION APS CHARACTERISTICS	ATTITUDE CONTROL				SHAPE CONTROL	STATION-KEEPING
	DISTURBANCE CANCELLATION	POINTING	MANEUVERS			
THRUST LEVEL	X	X	X		X	X
MODULATION	X	X	X		X	X
RISE & DECAY		X			X	
NO. & DISTRIBUTION OF THRUSTERS	X				X	
MASS	X	X	X		X	X

Fig. 1. Control/Propulsion Interactions

Many people in a number of different groups have asked the same questions. How the problem is approached depends upon the goal of the study group.

IV. TYPE OF ACTUATION

In efforts to design a secondary propulsion system, studies, most notably by Boeing and Martin-Marietta, concentrate heavily on type of actuation. They are particularly interested in comparing chemical and electric systems in an effort to guide future development efforts. Assumptions made in such studies include perfect knowledge of the state, perfect knowledge of disturbances and distributed propulsion. In one case, it was assumed that mass expulsion devices would do the whole control job which is unlikely. Although the optimal number and location of propulsion units is acknowledged as a key characteristic, they appeared to be selected more from an intuitive feel for rigid body motion and symmetry than from any optimizing procedure.

Such studies have been very beneficial in detailing some requirements for an auxiliary propulsion system design. They have shown this to be an extremely promising application for electric systems especially at geosynchronous orbit and electric propulsion is being pursued on the strength of that promise. There are, however, additional issues which will need to be addressed in more detail, most of which, while not necessarily being unique to large, flexible structures, are certainly much more significant for such systems. The following should be included, although there are certainly others. (i) Contamination may affect type and placement of a unit and/or any auxiliary equipment. (ii) Mass and volume penalties for any extra power requirements will certainly be a tradeoff, particularly with electric systems. (iii) Reliability and life considerations will determine the need for "back up" capability. (iv) Ultimately the number of propulsion units may be limited, perhaps due to volume limits during storage and deployment. Gimbaling ability, not the current norm for attitude control, may help to offset such a constraint. (v) There will most certainly also be limitations on locations. Propulsion equipment can only be located, for example, on structural members which are stout enough to support it. But also

affecting location will be deployment methods. Not only must it be possible to store the units before deployment, there must be no interference during the exercise, which also means no cable interference. (vi) The subject of location overlaps with inertia considerations both during and after the spacecraft is deployed. Although large moment arms are desired for fuel conservation, the dynamic characteristics are affected. This is one of the issues in centralized versus distributed propulsion and may affect shape control. (vii) Not only location of the propulsion unit, but also its associated equipment for long term operation, is to be selected. There naturally are tradeoffs such as distance to the power source versus electric losses; kind and amount of cabling for least interference; number and distribution of fuel tanks. Clustering of units and/or equipment has been discussed as a way to maximize the system. On the unit itself there are characteristics which are currently ill-defined, as far as their effect on shape control is concerned. Even if a mass expulsion device is not actually a part of the control loop for shape control, the shape interactions of a propulsion unit may overcome the abilities of the chosen actuator. Modulation would affect all control functions, but transients involved would particularly impact shape. For example, the minimum impulse bit capability with no extra impulse added. Thrust level has an obvious effect on the structure, but even start up activities may require a damper between the structure and propulsion units. It follows that start up delays and "clean" shutdowns may also be important.

V. NUMBER AND DISTRIBUTION OF SENSORS AND ACTUATORS

Of the research efforts of those most interested in the control design, a few have addressed the questions listed in Section III. But concentration has been on number and distribution of "generic" sensors and actuators. (Sensing and actuation are intimately related and completeness is lost without consideration of both.) The effort is directed to development of a method which will produce the optimal number and locations. The different methods will frequently produce different results. There are some common assumptions used frequently. The desired sensors may be assumed to be available although much work is currently being done to

develop the sensing information which is required. Perfect actuation which instantaneously produces the desired control history may be assumed, but it is certainly not a reality. This translates into no time delays and no mass included for the control system. With no discrimination between types of actuation, it is unclear exactly what roles are to be played by traditional momentum exchange and mass expulsion devices as well as new types of actuation. Electric systems can be considered "new" as well as some non-mass expulsion ideas for which some initial designs are being tested in a few places. Of course, without differentiating type, no actuator dynamics are being included in system models. It is quite possible that this is a significant source of error. The most immediately available way to determine the sensitivity of control design to type of actuator, is to include the dynamic model of, for example, an electric propulsion system, in the total system model. Then determine optimal locations.

In preparing to do such a test, let me first review some of the technology areas upon which the methods for selection of sensors and actuators are based. First, there are different ways of modelling for distributed control and distributed estimation. For writing differential equations of motion, design options include the following. (i) Partial differential equations produce control and estimation procedures developed on continuous models. It is useful particularly for early control concept design. (ii) Modal analysis retains physical insight by considering the vehicle dynamics in terms of the natural modes. (iii) Finite-element models produce local controllers where each actuator command depends only on the nearby sensor measurements; currently the most widely used method. Any disturbance models should at least include radiation pressure, gravity gradient, aerodynamic forces and orbital perturbations. Magnetic and thermal effects do not appear significant at this stage.

Second, some amount of model order reduction is necessary for a realistically usable model, usually done in steps. It is certainly not arbitrary, the significant characteristics must be recognized and retained. Seen in Fig. 2, from the actual large structure is produced a ground evaluation model which still may contain as many as 10,000 modes.

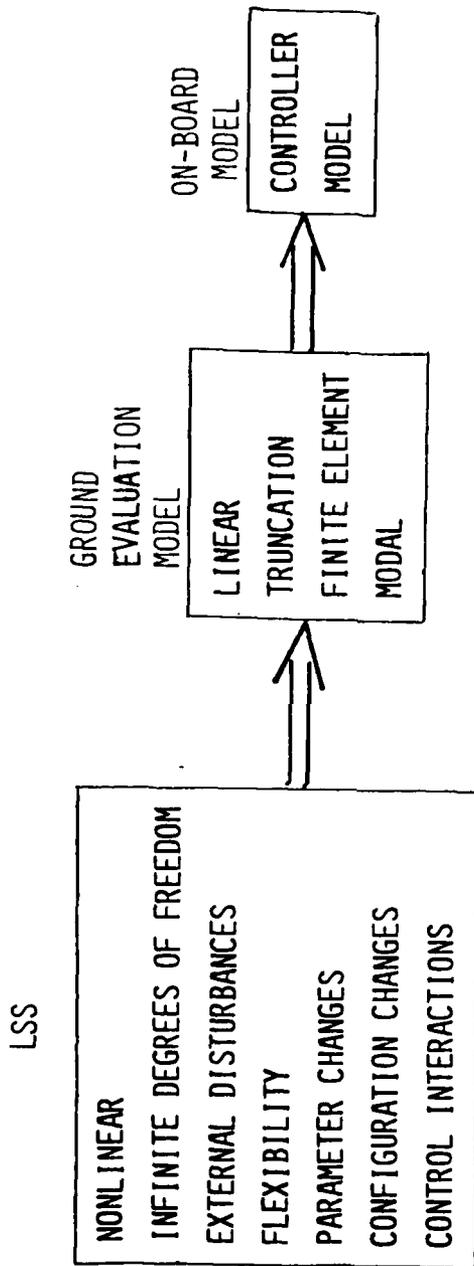


Fig. 2. Model Order Reduction

A major step in model reduction one which is included in all the distribution methods discussed ahead is that of linearization. On-board computing capability will limit the model still further. It is desirable to delay model reduction as long as possible in the design process.

Third, the last area to be mentioned is model error estimation. It is very important, particularly for sensing, to be able to detect deficiencies in our models. They may originate from unmodelled external disturbances, truncated dynamics, neglected nonlinearities, parameter uncertainties, etc. The currently popular studies of spillover are in this area. It is also a basis for the development of adaptive control techniques.

Using models developed from advances in the above areas, the three most popular methods of determining optimal number and location of sensors and actuators are

1. Degree of Controllability/Observability
2. Independent Modal Space Control
3. Modal Cost Analysis

I will briefly discuss each one.

V. DEGREE OF CONTROLLABILITY/OBSERVABILITY

Separately calculating degree of controllability and degree of observability is based on steady-state solutions of the optimal control and optimal estimation problems. In particular, it involves minimizing the steady-state optimal cost function and state estimation error. Both the control and estimation problems use the same reduced order model, represented by the vector equations

$$\dot{x} = Ax + Bu$$

$$z = Mx + v$$

where x is the state, the elements of z are the measurement (sensor) information available, and v represents errors in the system model. In the control problem, u indicates inputs to the control system and it represents dynamic model errors in the estimation problem.

LSS

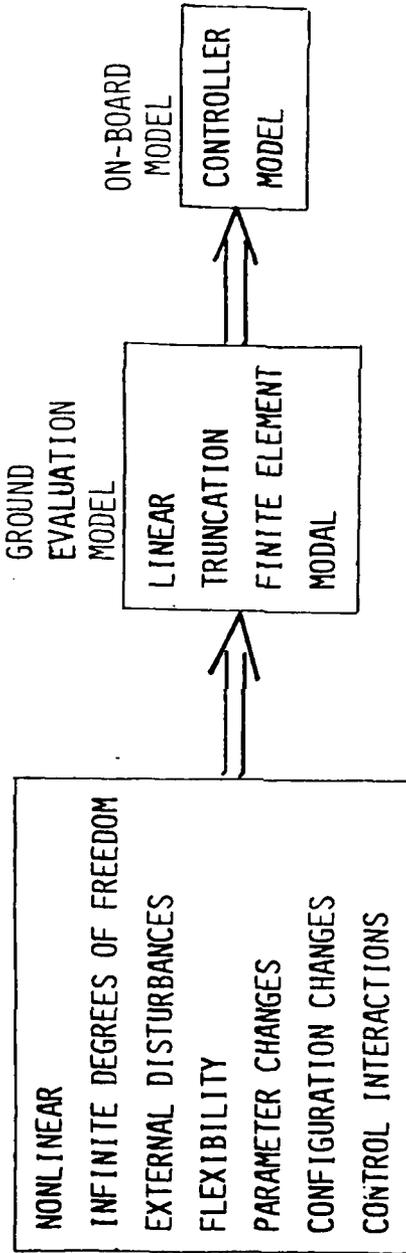


Fig. 2. Model Order Reduction

A major step in model reduction one which is included in all the distribution methods discussed ahead is that of linearization. On-board computing capability will limit the model still further. It is desirable to delay model reduction as long as possible in the design process.

Third, the last area to be mentioned is model error estimation. It is very important, particularly for sensing, to be able to detect deficiencies in our models. They may originate from unmodelled external disturbances, truncated dynamics, neglected nonlinearities, parameter uncertainties, etc. The currently popular studies of spillover are in this area. It is also a basis for the development of adaptive control techniques.

Using models developed from advances in the above areas, the three most popular methods of determining optimal number and location of sensors and actuators are

1. Degree of Controllability/Observability
2. Independent Modal Space Control
3. Modal Cost Analysis

I will briefly discuss each one.

V. DEGREE OF CONTROLLABILITY/OBSERVABILITY

Separately calculating degree of controllability and degree of observability is based on steady-state solutions of the optimal control and optimal estimation problems. In particular, it involves minimizing the steady-state optimal cost function and state estimation error. Both the control and estimation problems use the same reduced order model, represented by the vector equations

$$\dot{x} = Ax + Bu$$

$$z = Mx + v$$

where x is the state, the elements of z are the measurement (sensor) information available, and v represents errors in the system model. In the control problem, u indicates inputs to the control system and it represents dynamic model errors in the estimation problem.

To examine the procedure, consider first degree of controllability. The control problem can be defined as bringing the system from an initial state $x(t_0)$ to a terminal state $x(t_f)$ using acceptable levels of control $u(t)$ and not exceeding acceptable levels of the state along the way. There are slight variations in how this is achieved, but the following is a representative example from Rodriguez. We choose to minimize the quadratic performance index

$$J = \frac{1}{2} \int_{t_0}^{t_f} (x^T Q x + u^T R u) dt$$

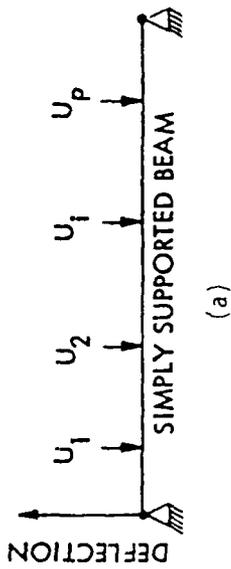
Q and R are weighting matrices which indicate relative importance of vector elements. Their values must be assumed. The optimal steady-state cost function is found as the solution of the Hamilton-Jacobi-Bellman equation, and is

$$J^\circ = \frac{1}{2} x^T(t_0) P^\circ x(t_0)$$

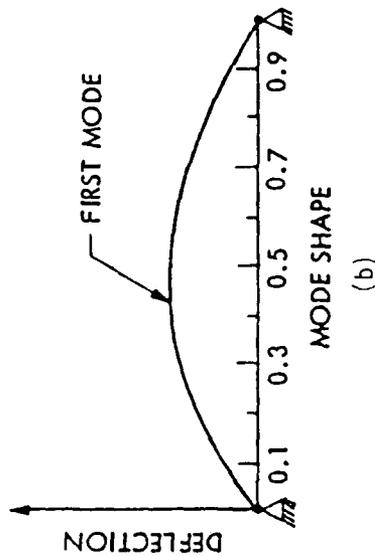
where the initial state, $x(t_0)$, is a random variable and P° is the solution of the steady-state Riccati equation. If we define ξ as the actuator locations, then the expectation of J° is a function of ξ through P° . Let ξ_b indicate the optimal locations. Then

$$E[J^\circ(\xi_b)] \leq E[J^\circ(\xi)]$$

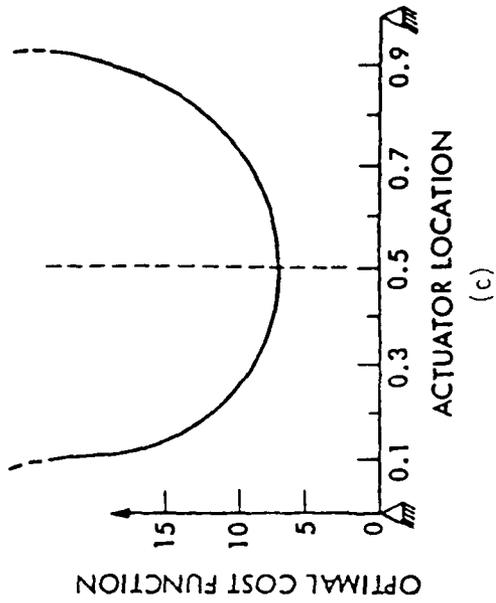
or the optimal location of actuators is where the absolute minimum of the expectation of the optimal cost function occurs. As an example, use a simply supported beam as seen in Fig. 3. One actuator, located anywhere along the beam (Fig. 3a), is to be used to control the first mode (Fig. 3b). If the degree of controllability is calculated and plotted versus location, Fig. 3c results. The minimum value, or optimal location of the single actuator, is at the midpoint. This coincides with our intuition for this simple example. Skelton and Likins use a similar, although much more complicated method to produce degree of controllability on a



(a)



(b)



(c)

Fig. 3. Degree of Controllability

two dimensional surface. Their results is shown in Fig. 4. It should already be clear that all possible combinations must be tested to produce a complete map. This would be a severe complication for a large complex structure.

The degree of observability is calculated in the estimation problem in a similar manner. The estimator for the reduced order model is

$$\dot{\hat{x}} = (A - KM)\hat{x} + Kz$$

where \hat{x} is the estimated state and K is the Kalman gain. The estimation error inherent in the process is $e(t)$ or

$$e(t) = x(t) - \hat{x}(t)$$

The error is, of course, dependent on the model errors which can be seen through the differential equation for e ,

$$\dot{e} = (A - KM)e + Bu - Kv$$

The solution for the state error from the above equation requires knowledge of the model errors. Because specific knowledge of u and v is not available, they can be approximated as zero mean processes with covariances. Then the optimal location of the sensors will be where the absolute minimum of the state estimation error occurs. Plots can be produced which are similar to those shown in Figs. 3 and 4 for degree of controllability.

There are pros and cons to using the concepts of degree of controllability/observability to determine optimal locations of sensors and actuators. This method does give some intuitive feel for "how controllable" or "how observable" a system is. The plots, at least for simple systems, do give physical insights. The results are consistent with the generally accepted trend that increasing the number of actuators will always reduce the overall deflections of the structure. However, the disadvantages are many and include the following points. This method treats actuators

**ACTUATOR LOCATIONS
CONTROLLABILITY SURFACE**

SEVENTH MODE

[TWO TORQUE ACTUATORS COINCIDENT AT (a_x, a_y)]

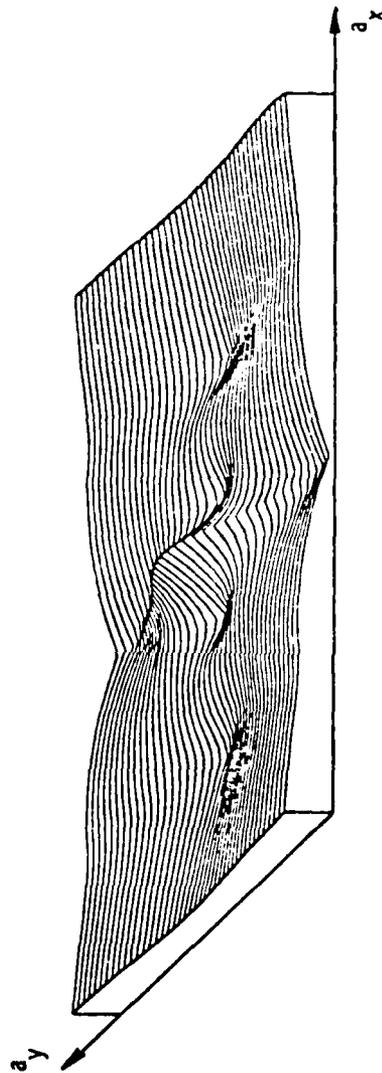


Fig. 4. Degree of Observability

and sensors as separate issues. In reality, determination of actuator positions requires use of sensor information. One way of including uncertainty in the control problem is by adding process noise which does not appear here. Of course, no actuator dynamics are included which have previously been mentioned as possibly significant. This means there is no distinction between methods of actuation. Even with all the simplifications, the method is complex in practice. All combinations of locations must be checked individually which is unweildy for a large structure with many actuating units. Finally, the result produced is for a possible arbitrary selection of the weighting matrices Q and R. The best choices may not be obvious for a large, complex structure.

VI. INDEPENDENT MODAL SPACE CONTROL

A second method of control design which is less developed but getting more attention recently is independent modal space control. The state space representation of the reduced order system is similar to that detailed above with one significant difference,

$$\dot{x}(t) = A x(t) + W(t)$$

The term previously containing information on actuator location, $B u(t)$, has been replaced by $W(t)$, the "generalized controls". This formulation is presented to emphasize that control laws can be determined without regard to actuator location. Of course, ultimate implementation means that some method still must be used to determine the locations which will produce the "generalized forces" that are the result of independent modal space control. A modified performance index is still being minimized,

$$J_{IMSC} = \int_0^t (x^T Q_{IMSC} x + W^T R_{IMSC} W) dt$$

where the new weighting matrices Q_{IMSC} and R_{IMSC} must again be chosen, with a less "physical" representation. It has been shown possible, however, to decompose this into n 2nd order decoupled optimal control problems which makes it much easier to solve.

So independent modal space control does introduce the advantages of smaller computational requirements. Also the control law and actuator distribution problems are decoupled. The major disadvantages again include a separation of control and estimation. In addition, assuring a physically realizable control means a careful selection of actuator locations by some other method. Also, in that selection, at least one actuator is required for each mode to be controlled. (Significant progress on removing that constraint is being made at Columbia by Longman.) To date, no actuator dynamics have been included in this process, so no differentiation between type of actuation is available.

VII. MODAL COST ANALYSIS

The third method which I considered was modal cost analysis. It also uses a linearized reduced order model, but the goal is to find sensor/actuator combinations to meet output specifications by minimizing the sum of normalized input variances. The model is written as

$$\dot{x} = Ax + Bu + Dw$$

$$D = [B \ D_0]$$

$$y = Cx$$

$$z = Mx + v$$

where y has been added as the vector of outputs to be regulated. $D_0 w$ represents unmodelled system behavior and Bw is the unmodelled actuator noise. v again indicates unmodelled or noisy sensors, and u the actuator signals. Recall that the noise term in the differential equation did not appear in the previous models. The unknowns $x(t_0)$, w , v are all assumed as random, zero mean processes. To keep the outputs and inputs (controls) within the specified ranges,

$$-\sigma_j \leq y_j(t) \leq \sigma_j$$

$$-\mu_j \leq u_j(t) \leq \mu_j$$

the following variance constraints are placed on the variables throughout the procedure,

$$E y_i^2(t) \leq \sigma_i^2$$

$$E u_i^2(t) \leq \mu_i^2$$

E is the expectation operator. The selection problem then becomes, in the simplest terms, one of using \bar{m} out of m possible actuators and \bar{l} out of l possible sensors, to find the combination which produces the smallest values for

$$\sum_{i=1}^k E_{\infty} y_i^2 / \sigma_i^2 \text{ subject to } E_{\infty} u_i^2 = \bar{\mu}_i^2$$

in the input constrained problem or

$$\sum_{i=1}^m E_{\infty} u_i^2 / \bar{\mu}_i^2 \text{ subject to } E_{\infty} y_i^2 = \sigma_i^2$$

in the output constrained problem with the following definition,

$$\lim_{t \rightarrow \infty} E y_i^2(t) \triangleq E_{\infty} y_i^2$$

$$\lim_{t \rightarrow \infty} E u_i^2(t) \triangleq E_{\infty} u_i^2$$

The values \bar{m} , m , \bar{l} and l must all be specified. One of the keys in using this method, however, is use of effectiveness values from component cost analysis (developed primarily by R. Skelton). Rather than calculations for each possible combination of sensors and actuators, each unit can be assigned a component cost, V_i^{sen} and V_i^{act} . It can be interpreted as the cost to the system of using the unit at that particular location. Total system cost, V , is again quadratic,

$$V = E_{\infty} (y^T Q y + u^T R u)$$

In the process of calculating total cost, the best values of the elements of the weighting matrices are selected as well. It is a very novel approach to the problem with some surprising results.

A large positive contribution of this idea is the treatment of control and estimation as one problem. They are indeed coupled and this confirms such thinking. Second in significance only to that general philosophy is the contribution of component cost analysis, that is, the elimination of locations without testing individually all possible combinations. A surprising result of work done to date with this method stems from the addition of process noise. When it is added, a "best" number of actuators emerges. No longer will an increase in number of units always result in a decrease of deflections. In other words, unmodelled behavior may have a significant effect on the system. Other advantages of this method include the fact of starting with "possible" locations so only physically realizable locations are even considered. Also, this technique has potential use in component failure analysis. Eliminating a unit from operation may be similar to eliminating it from consideration. Modal cost analysis does have disadvantages as well. It shows unmodelled behavior as important but currently includes no actuator dynamics and so does not distinguish between types of actuation. Inclusion of an actuator model appears to be the next logical step here. The relative importance of that information can be tested. A disappointing aspect of this analysis is that the final selection may or may not be optimal. Theoretical attempts at a proof continue but have yet to be successful. However, I believe that modal cost analysis is the most unique approach to sensor/actuator placement and has showed the most interesting results so far.

VIII. SUMMARY AND RECOMMENDATIONS

The information presented above summarizes development in some areas of research necessary for design of an effective attitude control system for a large, flexible space system. Its purpose was twofold. First, some areas of concern in the implementation of control system elements were highlighted,

which had previously required little attention in small, dense spacecraft. Their effects on the dynamics and stability of the vehicle is largely unknown, although a number of theories have been offered. Answers to such questions will certainly have a bearing on the selection of mass expulsion devices as part of the secondary propulsion system, as well as the use of non-thrusting units to perform a given task. In the second part of the report, the currently available theory in one particular area is explored, i.e., the placement of actuating and sensing units. The methods most prominent in the literature in recent years are summarized. The details have been examined and are ready for use in follow-on work in this topic area.

In the control theory used for determination of "how many" actuating units and "where" to place them, no mention is ever made of the particular device employed to perform the control function. The next phase of this work would involve incorporating a model of an actuator, such as the actual dynamics of a low thrust electric unit, in the dynamic model of a vehicle. Currently accepted practice is to use a generic actuator which performs perfectly in producing required forces at precisely the right instant of time. Untidy elements such as cabling are usually neglected as well. It would be very valuable to know if such approximations are always justified. A test case comparing the required actuator placement to control a given vehicle both with and without actuator dynamics in the model is suggested. Should the inclusion of this information prove to be important, it may lead to further investigation of the effects of different devices.

BIBLIOGRAPHY

Conference and journal publications:

1. Bush, H.G., M.M. Mikulas, Jr., and W.L. Heard, Jr., "Some Design Considerations for Large Space Structures," AIAA Journal, Vol. 16, No. 4, April 1978, pp. 352-359.
2. Gevarter, W.B., "Basic Relations for Control of a Flexible Vehicle," AIAA Journal, Vol. 8, No. 4, p. 666.
3. Gorland, S.H., "Drawing a Bead on Propulsion for Large Space Systems," Astronautics and Aeronautics, Jan. 1982, pp. 57-59.
4. Hughes, P.C., "Dynamic of Flexible Space Vehicles with Active Attitude Control," Celestial Mechanics, Vol. 9, 1974, p. 21.
5. Hughes, P.C., and R.E. Skelton, "Controllability and Observability for Flexible Spacecraft," Journal of Guidance and Control, Vol. 3, No. 5, Sept. - Oct. 1980, pp. 452-459.
6. Kline, "Construction of Large Structures in Space," Journal of the Astronautical Sciences, Vol. XXVII, No. 4, Oct. - Dec., 1979, pp. 401-418.
7. _____, "Large Space Systems / Low-Thrust Propulsion Technology," "Proceedings, NASA Conference Publication 2144, Lewis Research Center, Cleveland, Ohio, May 1980.
8. _____, "Large Space Systems / Propulsion Interactions Workshop," "Proceedings, NASA TM 82904, Lewis Research Center, October 1981.
9. _____, "Large Space Systems Technology - 1981," "Proceedings, NASA Conference Publication 2215, Langley Research Center, Hampton, VA., Nov., 1981.
10. Meirovitch, L. (ed.), "Dynamics and Control of Large Flexible Spacecraft," Proceedings of the First VPI and SU/AIAA Symposium, Blacksburg, VA., June 1977.
11. Meirovitch, L. (ed.), "Dynamics and Control of Large Flexible Spacecraft," Proceedings of the Second VPI and SU/AIAA Symposium, Blacksburg, Va., June 1979.
12. Meirovitch, L. (ed.), "Dynamics and Control of Large Flexible Spacecraft," Proceedings of the Third VPI and SU/AIAA Symposium, Blacksburg, VA., June 1981.

13. Meirovitch, L., H. Baruh and H. Öz, "A Comparison of Control Techniques for Large Flexible Systems," Journal of Guidance, Control, and Dynamics, Vol. 6, No. 4, Jul.-Aug. 1983, pp. 302-310.
14. Ogg, G., "Procedures to Integrate Electric Secondary Propulsion Systems to Large Deployable Space Systems," AIAA Paper 83-1392, AIAA/SAE/ASME 19th Joint Propulsion Conference, Seattle, Washington, June, 1983.
15. Santini, P., "Stability of Flexible Spacecraft," Acta Astronautica, Vol. 3, Sept.-Oct. 1976, pp. 685-713.
16. Schaechter, D.B., "Hardware Demonstration of Flexible Beam Control," Journal of Guidance and Control, Vol. 5, January 1982, p. 48.

Reports:

17. Bainum, Kumar, and James, "Dynamics and Control of Large Flexible Space Structures - Part B," Final Report NASA Grant NSG-1414, Department of Mechanical Engineering, Howard University, May 1978.
18. Carignan, C.R. and W.E. Vander Velde, "Number and Placement of Control System Components Considering Possible Failures," Report - NASA Grant #NAG1-126, Dept. of Aeronautics and Astronautics, Massachusetts Institute of Technology, March 1982.
19. Chiu, D., "Optimal Sensor/Actuator Selection, Number, and Placement for Linear Stochastic System," Ph.D. Dissertation, School of Aeronautics and Astronautics, Purdue University, West Lafayette, IN., May 1981.
20. DeLorenzo, M.L., "Selection of Noisy Sensors and Actuators for Regulation of Linear Systems," Ph.D. Dissertation, School of Aeronautics and Astronautics, Purdue University, West Lafayette, IN., June 1983.
21. Likins, P.W., "Dynamics and Control of Flexible Space Vehicles," JPL Technical Report 32-1329, January 1970.
22. Lindberg, Jr., R.E., "Actuator-Placement Considerations for the Control of Large Space Structures," Naval Research Laboratory Report 8675, May, 1983.
23. Maloy, J.E. and W.W. Smith, "An Insight into Auxiliary Propulsion Requirements of Large Space Systems," NASA TM-82827, May 1982.
24. _____, "MPD Thruster Definition Study - Phase II Review," Martin Marietta Denver Aerospace and Eagle Engineering, Inc., Air Force Rocket Propulsion Laboratory Contract F04611-82-C-0049, April, 1983.

25. Skelton, R.E., "Component Cost Analysis of Large Scale Systems," Final Report, Jet Propulsion Laboratory, Contract No. 955369, School of Aeronautics and Astronautics, Purdue University, January 1982.
26. Smith, W.W. and J.P. Clark, "Study of Electrical and Chemical Propulsion Systems for Auxiliary Propulsion of Large Space Systems," NASA CR-165502, Boeing Aerospace Company, November, 1981.
27. Vander Velde, W.E. and C.R. Carignar, "A Dynamic Measure of Controllability and Observability for the Placement of Actuators and Sensors on Large Space Structures," Report-NASA Grant #NAG1-126, Dept. of Aeronautics and Astronautics, Massachusetts Institute of Technology, Jan. 1982.

Books:

28. Wertz, J.R. (ed.), Spacecraft Attitude Determination and Control, D. Reidel Publishing Co., Dordrecht, The Netherlands, 1978.

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

Title: Scanning Electron Microscopical Study of
Periosteum from Rat and Rhesus Monkey

Prepared by: Dr. Gwendolyn B. Howze

Academic Rank: Associate Professor

Department and University: Department of Biology
Texas Southern University

Research Location: Aerospace Medical Research Laboratory
Biodynamics & Bioengineering Division
Biodynamic Effects Branch
Wright-Patterson Air Force Base

USAF Research: Dr. Leon E. Kazarian, Mr. Clarence M. Oloff

Date: August, 1983

Contract No.: FF49620-82-C-0035

A Scanning Electron Microscopical Study of Periosteum from Rat and Monkey

by

Gwendolyn B. Howze

Abstract

The normal or unstimulated structure of the periosteum has been studied by scanning electron microscopy. In the two species studied, rat and Rhesus monkey, five distinct morphological regions can be discerned. The top-most fibrous layer is a coherent layer of fibers in an interknitted array. Intermixed with the fiber bundles are many blood vessels. Below the top layer is a second fibrous layer in which the fibers are more loosely arranged and mixed with a few cells. A third region contains cells, fibers and blood vessels in an apparently random innominate distribution. Cells are plentiful in this region. The fourth layer is highly cellular. In the monkey, the cells are stratified in layers, with each layer separated by a network of fibrils. The fifth region, "juxta-matrix" is apparently one cell layer thick, in which the cells are strongly attached to the subjacent bone matrix.

In both species collagenous fibers are most plentiful and are composed of fibrils in parallel array. The fibers characteristically are covered with a meshwork of fibrils which usually contain particulate structures of unknown composition. The collagenous fibrils appear to differ in the two species. In both species the majority of the cells contain numerous cytoplasmic processes.

ACKNOWLEDGEMENTS

The author would like to thank the Air Force Office of Scientific Research and the South Eastern Center for Electrical Engineering Education for providing the opportunity to spend a profitable and interesting summer at the Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio. Particular thanks is given to the Biodynamic Effects Branch (Biodynamics & Bioengineering Division) for its hospitality and quite adequate working conditions.

Among the numerous people who contributed to making this a very satisfying experience. I would like to especially thank Dr. Leon E. Kazarian and Mr. Clarence M. Oloff for suggesting an area of mutual interest which could profitably be studied in the short time period available. They also provided continued excellent support during the research period. Miss K. C. Smith made essential contribution to the effort. Special acknowledgement is also given to Lt Col R. A. Bruner and Mr. David R. Mattie of the Pathology Branch (Toxic Hazards Division) for giving me full run of their Scanning Electron Microscope (SEM) facility. Finally, I would like to thank Mr. David Mattie for instruction and advice in scanning electron microscopy. Miss Lisa Wooley typed the manuscript.

I. Introduction

In a live human, a bone such as the femur, for example, is a living organ. It has three main compartments: the matrix, which contains a labyrinth of channels; the periosteum; and the periosteal appendages, the osteocytes, which derive from the periosteum but reside in matrix channel spaces called lacunae (1,2).

When most people speak of bones they have the matrix in mind. It is the most obvious and massive portion (3). The bone matrix, however, is a nonliving secretion product of the periosteum (1, 2). By contrast, the periosteum is a thin unobtrusive tissue which covers the outer surface of the matrix of a bone. In fact, except for the articular surfaces at the joints, the exterior bone surface consists in the periosteum. Though inconspicuous, the periosteum is composed of living cells and their secretion products, the most prominent of which are the collagenous fibers (1, 2). The living periosteum and its appendages, the osteocytes, manufacture and maintain the non-living bone matrix (1, 2).

Interestingly but distressingly, most of the research on bone has been and, even now, is devoted to the characterization of the bone matrix. See for example references 3, 4, and 5. Thus only a minor portion of the massive literature on bone is concerned with the periosteum. This point of view is changing, however, as more and more scientists realize the need for an intense effort to uncover the complete story of periosteal influence on bone function.

Many normal essential Air Force activities place human bones under considerable stress and risk of injury. For example, the three main bone

functions in humans; protection, support and calcium homeostasis; are placed under stress by such activities as: parachuting and ejection from aircraft, and space flight. Bone injury is usually diagnosed and understood in terms of changes in the matrix. Accordingly, the Air Force has a considerable involvement in the study of bone matrix. In light of the evidence that the periosteum and its appendages, the osteocytes, manufacture and maintain the matrix (1, 2), it is clear that a profound understanding of the periosteum is essential to a successful program of bone health, injury prevention, and repair.

This project is concerned with the structural organization of the periosteum as revealed by scanning electron microscopy. It adds to the information obtained by light microscopy transmission electron microscopy and other methods.

II. Objective

The original goal for this project was to investigate the possibility that the osteogenic region of the periosteum plays a significant role in bone demineralization which is induced by hypokinesia. A search of the literature, however, indicated that the available information about the periosteum was not sufficiently detailed to generate a description of normal periosteum which could serve as a baseline for our studies.

Therefore, the immediate purpose of the project was to produce a detailed scanning electron microscopical (SEM) description of the organization of the periosteum in the normal or unstimulated state. A secondary aim is to combine the SEM information with the published information in order to synthesize a description which can serve as a baseline picture of the normal periosteum.

III. Summary of Methods

In order to achieve the stated objective, standard scanning electron microscopical procedures (4, 8) have been slightly modified to suit the periosteum. The sequence of steps in summary are the following: 1) obtaining the specimen. 2) preservation of native state by fixation. 3) dehydration. 4) preparation for SEM. 5) observations and producing a permanent record.

Obtaining the specimen: The rats were sacrificed and the femurs were dissected out and washed in buffered saline. The rhesus monkey periosteum were obtained by biopsy. They were not washed before fixation and unfortunately were contaminated by adhering red blood cells.

Fixation: The specimens were fixed in 2% glutaraldehyde in pH 7.3, 0.2 M phosphate from 3 to 24 hours.

Dehydration: A graded series of ethyl alcohol was used in the dehydration, 35%, 70%, 95% and 100%.

Preparation for SEM: Normally after the dehydration, specimens are transferred to iso-amyl acetate in preparation for the critical point drying procedure. Carbon dioxide was used for the critical point drying. Next the specimens were attached to the stubs with a conductive cement and covered with a gold coat estimated to be 20 nm thick. The observations were made on an ETEC Autoscan Scanning Electron Microscope. Uniformly, the accelerating voltage was 20 Kv. A permanent record was obtained upon Polaroid 55 P/N film. The negatives were used to make the prints exhibited in this report. The printing process resulted in an additional enlargement of 2.21X. This factor should be used in estimating the total magnification in the prints. This procedure was modified slightly for rat specimens.

Modified procedure for rat femur: This procedure was developed for the rat femur, but it probably should be used uniformly since it eliminates certain artifacts, e.g. dissection injury and shrinkage artifacts due to fixation and dehydration. The sequence is as follows: 1) the whole femur is fixed. 2) the dehydration was the same as above, but a subroutine to separate the periosteum from the matrix was inserted before the critical point drying procedure. 3) isolation of the periosteum by embedding it in a collodion membrane. The details of this procedure are listed below. 4) critical point drying and continuation with the standard SEM procedure.

The details of the procedure for isolation of the periosteum are the following:

1. The bones were placed in a 1:1 mixture of absolute ethonol and omyl-acetate, two changes, 15 min each.

2. Next, the bones were placed in 1.5% collodion (dissolved in the ethonol/omyl acetate solvent) for 15 min.

3. They were next transferred to a small container of 5% collodion for 5 minutes.

4. The top of the containers were then removed as the solvent evaporated for 5 min.

5. The bones were removed and held in the air for a minute or more until a definite solid film could be seen, and the smell of solvent had decreased. Caution: do not allow it to dry completely, rotate to insure a uniform film.

6. The semi-dry sample was placed in a container of distilled water for 5 min.

7. A circular cut (through the film down to the matrix) was made at each end. One longitudinal cut extending between the two circular cuts was made.

8. The bone was split along the longitudinal line.

9. The membrane was peeled off the split bone, either as one or two pieces.

IV. Results

The exterior surfaces (of the periosteal) in the two species are similar, consisting of a coherent interlocking array of collagenous fibers. Figures 1, 2, 3 and 4 are from rat specimens. The rhesus specimens are very similar, except that the adhering red blood cells obscure the details at low power magnifications, Figures 5 and 6. The fibers are actually bundles of fibrils aligned in parallel array. Measurements of the fibrils range from 60 nm to 100 nm. (the thickness of the gold coating was subtracted.) Series A, Figures 7, 8, 9 show an example of the fibers as seen in rat specimens. In Figure 7 a bundle of fibers is shown. There are also numerous cells associated with the bundle. The cells are possibly fibroblast. Figure 8 shows a region from figure 7 in higher magnification. The magnification progression is continued in Figure 9. The expected collagen bending is not detected in the fibrils possibly due to the use of ether in the processing regimen. The fibers are often in bundles as shown in Figures 7 and 31.

An unexpected finding was the presence of a network of fibrils covering fibers, blood vessels and cells. The network can be seen very easily in Figures 7 and 8 and most clearly in Figure 10. Also in Figure 10 can be seen particulate structures which are often seen associated with the fibril network. Figure 11 also shows the relationship of the network to the particles, which seem to be inorganic in composition.

Series B, Figures 12 through 15, shows an increasingly magnified view of a fiber located on the exterior or fibrous surface of a periosteal sample from the monkey. This series is actually an enlarged view of a region in Figure 5. The red blood cells are an unfortunate contaminant. The fibrils within the fiber can easily be seen to be in parallel array. The covering

network is most obvious in Figure 12. The particulate structures are abundant. The fibrils show a periodic decoration which are either buldges or exogeneous accretions. These are belived to be projections of the collagen banding. Figures 16 and 17 shows a blood vessel and nerve entering the periosteum. The top or fibrous surface of the periosteum contains many blood vessels.

As one proceeds interiorly and downward toward the bone matrix, fibers are more loosely arranged and are accompanied by more cells, see Figures 18, 19, 20. These are from the rat. The fiber bundle seen in the center of Figure 20 is an enlargement of a region in Figure 7.

As one approaches the interior or "juxta-matrix" surface the tissue becomes more cellular, Figures 21 and 22 which are from the rat. The cells seen in Figure 21 are shown magnified in Figure 23. Figures 24 through 28 show the result of teasing fixed monkey femur periosteum to reveal the cells which are associated with the fibers. This series shows an increasingly magnified view of the cell type commonly associated with the fibers. They are approximately kidney shaped containing many cytoplasmic processes and they are enclosed in the ubiquitous fibril network. A similar cell type, but smaller in size, is shown in Figures 29 and 30, from rat periosteum. A portion of a red blood cell is shown for size comparison. Figure 30 is a magnification of the scene in Figure 29.

Figure 31 shows a high power view of a free-hand median section through the monkey femur periosteum. An admixture of cells and fiber bundles are visible. Figure 32 shows part of a bundle of fibers and the cell types which are characteristic from the mid region and downward toward the bone matrix. As shown in Figure 33, the deeper layers are composed of the sheets of flattened waxy looking cells. Figures 34 and 35 show increasingly

enlarged views of this same type of cell. This cell type is predominate in the region immediately above the matrix. It corresponds to the osteogenic layer visualized by light microscopy (1, 2).

In the rat the deep region, presumably the osteogenic region contains more fibers than the similar region in the monkey. Some of the fibers in the rat might be elastic fibers, they are non-banded and exhibit branching. See Figures 36 and 37. Figures 38 and 39 show examples of cells and fibers in the deep region (rat). The series figures 40 to 43 show a close up enlarged views of the cells shown in Figure 39. The cell type shown in this series, Figures 39 through 43, are shown in their presumably normal confluent configuration in Figure 44 which apparently shows the surface which lies directly upon the matrix or more likely upon a one cell thick layer of osteoblast which are firmly attached (6, 9) to the matrix. Figure 45 is another view of the inner surface at lower magnification.

The cells on the inner surface exhibit many cytoplasmic processes, which seem to anastomoses to form a network. There is the definite perception that the cells form the meshwork of fibrils.

The "juxta-matrix" surface of the monkey periosteum is seen in Figure 46. The lighter material at the lower border of the micrograph is seen with regularity. A higher magnification examination shows that it consists of cells. The cells are probably osteoblast from the "one cell thick layer of osteoblast which adhere firmly to the matrix. The putative osteoblast are shown in greater detail in Figure 47 and 48. These cells appear to be mineralized. Presumably the visualized cells were in the process of laying down bone matrix.

**Note to the reader: Because photographs do not copy well, they are not included in the report. A set of photographs are available for loan from the author.

V. Recommendations

The observations presented in the results section of this report will be used in conjunction with the published literature to produce a detailed description of the structural organization of the periosteum. Before this is entirely practical, however, two additional studies are necessary. It should be determined that the organic solvents used do not modify the cell surface. The current conclusions about periosteal stratification should be challenged by a freeze fracture study.

The proposed detailed description can serve as a baseline for studies of changes in periosteal structure coincident with various experimental conditions which cause change in the bone matrix. In particular, it can serve as a guide as to which components are likely candidates for observation when looking for periosteal changes coincident with bone matrix demineralization induced by hypokinesia. On the basis of the present study:

- 1) the possibility of changes in the number of contour of the cytoplasmic processes on the cells in the "osteogenic" layer should be considered.
- 2) the ubiquitous fibril network is a likely specimen for study.
- 3) changes in the cell population associated with blood vessels, should also be studied.

REFERENCES

1. Ham, A. W. and D. H. Cormack, Histophysiology of Cartilage, Bone, and Joints, Physiology, J. B. Lippincott Company, 1979.
2. Ham, A. W. and W. R. Harris, "Repair and Transplantation of Bone," In G. H. Bourne (ed.): The Biochemistry and Physiology of Bone, Vol. III, N. Y. Academic Press Inc., 1973.
3. Prtichard, J. J., "General Histology of Bone," In G. H. Bourne (ed.): The Biochemistry and Physiology of Bone, Vol. I., N. Y. Academic Press Inc., 1972.
4. Boyde, A., "Scanning Electron Microscope Studies of Bone" In G. H. Bourne (ed.): The Biochemistry and Physiology of Bone, Vol. I., N. Y. Academic Press Inc., 1972.
5. Cameron, D. A., "Ultrastructure of Bone", In G. H. Bourne (ed.): The Biochemistry and Physiology of Bone, Vol. I., N. Y. Academic Press Inc., 1972.
6. Jones, S. J., A. boyde and J. B. Pauley, "Osteoblast and Collagen Orientation," Cell Tiss. Res., Vol. 159, pp. 73-80, 1975.
7. Graham, M. D. and H. P. House, "Human Stapes Crura," Arch. Otolaryngol., Vol. 101, pp. 548-551, 1975.
8. Hayat, M. A., Introduction to Biological Scanning Electron Microscopy, Baltimore, University Park Press, 1978.
9. Peck, W. A., J. K. Burks, J. Wilkins, S. B. Rodan, and G. A. Rodan, "Evidence of Preferential Effects of Parathyroid Hormone, Calcitonin and Adenosine on Bone and Periosteum," Vol. 100, pp. 1357-1364, 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

CYCLOCONVERTER MODELING FOR VARIABLE SPEED DRIVES

Prepared by: Dr. Medhat A. H. Ibrahim
Academic Rank: Professor
Department and University: Department of Electrical Engineering
California State University, Fresno
Research Location: Air Force Wright Aeronautical Laboratories
Aero Propulsion Laboratory
USAF Research
Colleague: Dr. William U. Borger
Date: August 10, 1983
Contract No: F49620-83-C-0035

CYCLOCONVERTER MODELING FOR VARIABLE SPEED DRIVES

by

Medhat A. H. Ibrahim

ABSTRACT

A mathematical model of the cycloconverter was developed. The model would accept a variable voltage variable frequency supply. The output frequency and SCRs conduction angle can be controlled. A three-phase static load made of resistance in series with an inductance was used for testing the model and results of the simulation were given. A model of the brushless dc motor was suggested for the dynamic loading of the cycloconverter. Suggestions for further research in this area, are offered.

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 conduct research at the Air Force Wright Aeronautical Laboratories, Wright Patterson AFB, Ohio. A special thanks is offered to the Power Systems Branch of the Aerospace Power Division of the Aero Propulsion Laboratory for the excellent working conditions and hospitality.

Finally, the author would like to thank Dr. William U. Borger for suggesting the topic and providing support and guidance throughout the research period, and Ms Angela Pajak for helping with the graphics, and Ms Tammy Ayers for typing the report.

I. INTRODUCTION

The constant speed drive represents the lowest installed cost, but may not provide the flexibility or operating efficiency which can be attained from an adjustable speed drive. Adjustable speed drives are needed in aircraft systems as well as countless industrial applications.¹⁻⁶ An adjustable speed drive can be accomplished by prime movers such as: 1) gas turbine, 2) constant speed electric drive with slip coupling, 3) dc motor with adjustable voltage or adjustable field control, 4) ac motor with adjustable voltage or adjustable frequency control. The dc commutator machine is suitable for variable speed and variable torque operation; but due to the brush maintenance and other problems, the machine is not suitable for aircraft application. Fortunately, the brushless dc motor offers the same desirable speed-torque characteristics as the dc commutator machine without the disadvantages. The most common configuration of brushless dc motors is the polyphase permanent magnet synchronous motor with power supply conditioning circuit which is capable of bidirectional power flow. This circuit is a static frequency changer which converts a source frequency to another frequency corresponding to the motor rotation and is called cycloconverter.³⁻⁷ Through the use of the cycloconverter the speed of the motor can be changed by the triggering frequency of the SCR set, and the motor torque is controlled by controlling the length of on-time of the SCRs. The use of cycloconverter as the power conditioning circuit is preferred compared to the use of a dc link.⁵ From the above it is obvious that a good model for the cycloconverter is needed for the simulation of, steady state as well as transient behavior, a generator-motor set coupled through a cycloconverter circuit.

II. OBJECTIVES

The main objective of this project was to develop a model for representing the cycloconverter circuit which is linking a variable frequency variable voltage supply to a passive load made of a resistance

and inductance elements or a permanent magnet brushless dc motor. Due to the research period frame, the loading with the brushless dc motor and the variation of the conduction angle, in wider range, was left for further studies in the future.

The specific objectives that were pursued are given:

1. To choose a mathematical model to represent the cycloconverter.
2. To write a computer program to represent the mathematical model.
3. To load the cycloconverter by a passive load.
4. To suggest further studies to be conducted.

III. MATHEMATICAL MODEL USED FOR THE SIMULATION OF THE CYCLOCONVERTER

The cycloconverter is a nonlinear device made of switches, SCRs or transistors, which are either closed or open according to a prescribed function to shape the output wave from the input wave. If the SCRs are triggered and have forward bias, it conducts and continues to conduct after the trigger signal is removed until commutated. With this in mind, the cycloconverter works like a conditioning function which transforms the supply wave into a desired output; or in mathematical terms

$$\underline{v}_m = F(\underline{v}_s, \underline{i}_m) \quad (1)$$

where

v_m : the output wave from the cycloconverter

v_s : the input wave to the device

i_m : the output current from the cycloconverter, i.e.; currents in the different SCRs

F: the cycloconverter transformation function which depends on the triggering function and commutation in the SCRs.

Because of the large number of SCRs used in the cycloconverter, 36-SCRs as in Figure 1 or 18-SCRs as in figure 2 for the three-phase input three-phase output case, the use of matrices in the formulations is very useful. For the case 36-SCRs shown in figure 1, the transformation function F given in equation 1 can be represented by the matrix D as follows:

$$D = \begin{bmatrix} d_{11} & d_{12} & d_{13} & d_{14} & d_{15} & d_{16} \\ d_{21} & d_{22} & d_{23} & d_{24} & d_{25} & d_{26} \\ d_{31} & d_{32} & d_{33} & d_{34} & d_{35} & d_{36} \\ d_{41} & d_{42} & d_{43} & d_{44} & d_{45} & d_{46} \\ d_{51} & d_{52} & d_{53} & d_{54} & d_{55} & d_{56} \\ d_{61} & d_{62} & d_{63} & d_{64} & d_{65} & d_{66} \end{bmatrix} \quad (2)$$

The elements of this matrix are related on a one to one basis with the SCRs in the circuit of figure 1. Assuming that the SCRs are ideal switches, the elements of the matrix d_{ij} are either set to "1" when conducting or "0" when not conducting or open.

The state of each element in this matrix depends on three factors which are:

- 1) that the trigger signal is on or off
- 2) that the bias on th SCR is positive or negative
- 3) the commutation state of the SCR after the trigger signal is removed.

Thus it follows that the matrix D can be written as a function of three matrices, D_1 , D_2 , and D_3 , which are dependant on the three criteria listed above by the equation.

$$D = D_1 * D_2 + D_3 \quad (3)$$

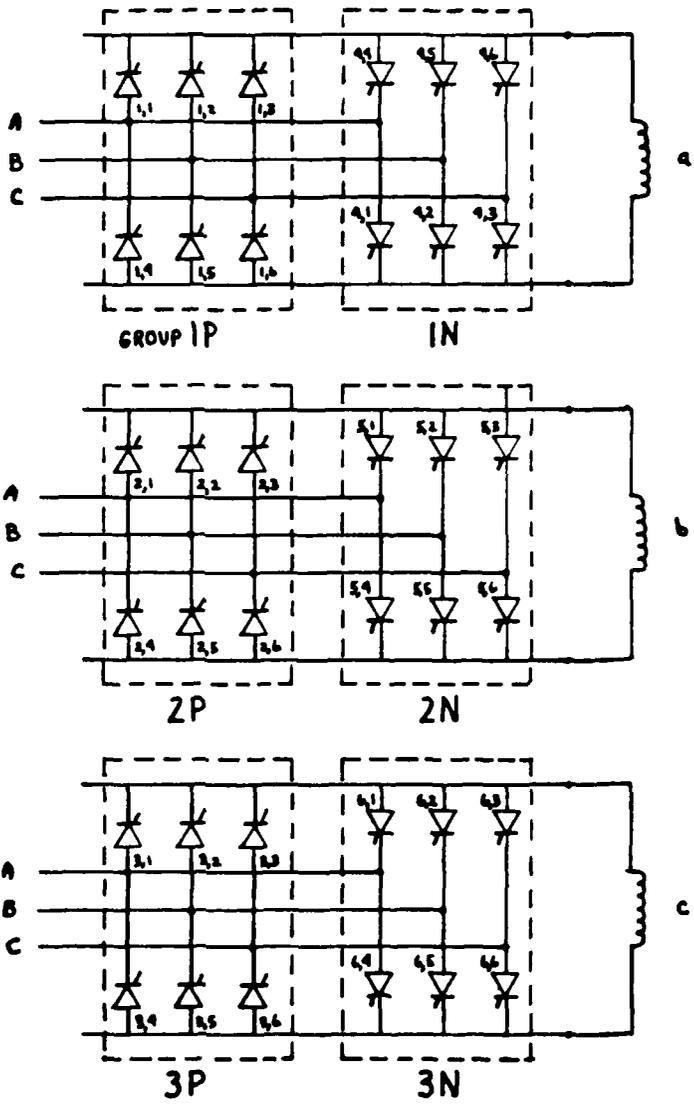


Figure 1 - 36 SCR CYCLOCONVERTER

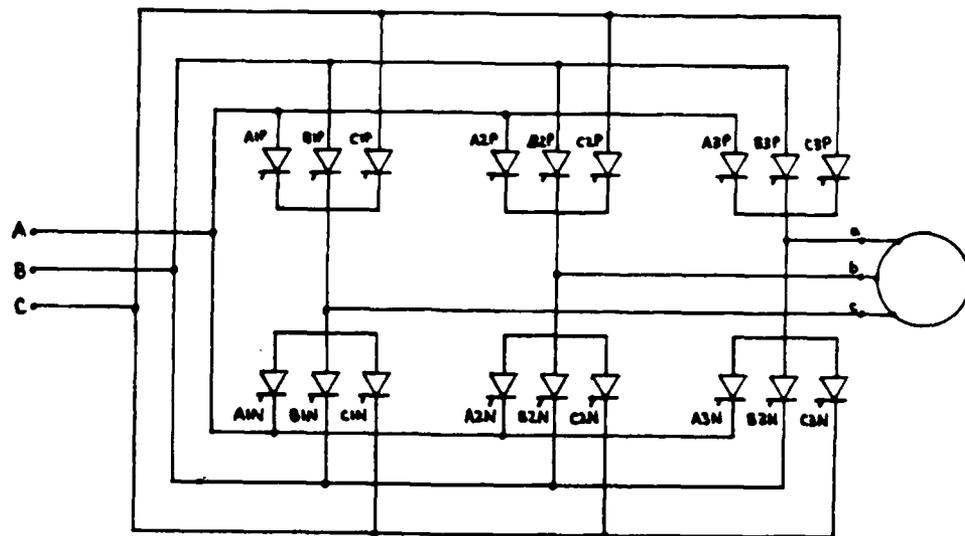


Figure 2 - 18 - SCR CYCLOCONVERTER

In equation 3 the (*) designates corresponding terms multiplications of the two matrices. The elements of these matrices are either "0" or "1" only and the values are set according to the following rules:

Matrix D_1 : The elements of this matrix are set to "1" if the corresponding SCR in the cycloconverter is triggered, and "0" if there is no trigger signal applied.

Matrix D_2 : The elements of this matrix are set to be "1" if biased in the forward direction and to be "0" otherwise.

Matrix D_3 : The elements of this matrix for which the corresponding SCRs are still conducting, after the trigger signal is removed, are set to "1". These elements are then set to "0" following commutation.

The values of the elements of the matrix D, at any given time, are determined by the prearranged control signals to give a specific output voltage wave. From the matrix D and the knowledge of the circuit, a matrix C could be formed:

$$\underline{I}_s = \underline{C} \underline{I}_m \quad (4)$$

where \underline{I}_s : represents the input current to the cycloconverter
 \underline{I}_m : represents the output current from the cycloconverter
 \underline{C} : the connection matrix which is constructed of the elements of the D matrix

From equation 4 and the power equation the output voltages can be obtained:

$$\underline{V}_m = \underline{C}^T \underline{V}_s \quad (5)$$

where \underline{V}_m : the output voltage
 \underline{V}_s : the supply voltage
 \underline{C}^T : the transpose of the matrix C.

The matrix C formation will be demonstrated in the next section for a particular cycloconverter circuit.

IV. CYCLOCONVERTER LOADING AND SIMULATION

A. System Description - A 36-SCRs cycloconverter shown in figure 1 was simulated by means of a Fortran program. In this simulation all the SCRs in one group are simultaneously triggered for a certain portion of the wave. By this method a three-phase output voltage waves, which are 120 electrical degrees apart, were produced. The output voltage frequency and magnitude are determined by the control signal and the input voltage.

B. Cycloconverter - For the circuit shown in figure 1 the matrix C could be written as follows:

$$\underline{C} = \begin{bmatrix} d_{11} & -d_{14} & d_{21} & -d_{24} & d_{31} & -d_{34} \\ -d_{41} & d_{44} & -d_{51} & d_{54} & -d_{61} & d_{64} \\ d_{12} & -d_{15} & d_{22} & -d_{25} & d_{32} & -d_{35} \\ -d_{42} & d_{45} & -d_{52} & d_{55} & -d_{62} & d_{65} \\ d_{13} & -d_{16} & d_{23} & -d_{26} & d_{33} & -d_{36} \\ -d_{43} & d_{46} & -d_{53} & d_{56} & -d_{63} & d_{66} \end{bmatrix} \quad (6)$$

To demonstrate the choice of values for the d_{ij} of the matrix C: for motor angle $0 < \theta_m < \alpha$ (conduction angle) the elements $d_{1j} = 1$ for $j = 1 - 6$, in the matrix D_1 and all the other elements are zero. Also for a generator or supply angle of $120^\circ \leq \theta < 180^\circ$, v_B is the most positive and v_C is the most negative, $d_{i2} = 1$ for $i = 1 - 6$, and $d_{i6} = 1$ for $i = 1 - 6$ in the matrix D_2 . From this information the matrix D can be obtained which indicates that $d_{12} = d_{16} = 1$ or $I_B = I_b$, $I_C = -I_b$ and $v_a = v_B - v_C$.

C. Static load - The load in this case was made of simple resistance and inductance in series connected to the three phase output as shown in figure 1. This system is a nonlinear system because of the switching operation by the cycloconverters SCRs. The load could be represented by three first order differential equations which are written in the vector form :

$$\frac{di}{dt} = -(Ri - v)/L \quad (7)$$

where i : the current in the load
 R : the resistance in ohms
 L : the inductance in henries
 v : the voltage applied to the load in volts

A fourth-order Runge-Kutta method, with fixed increment, was utilized for the numerical integration subroutine.

D. BRUSHLESS DC MOTOR LOAD MODEL - The model for the brushless dc motor per phase would be represented by induced emf in series with resistance and inductance as shown in figure 3. Also in this case similar to the static load case, a set of first order differential equations of the currents are given by the vector form in equation 8:

$$\frac{di}{dt} = [A(i)] \underline{i} + [B(\omega_1, \omega_2)] \underline{v} + [H] \underline{e} \quad (8)$$

where

- i: the current in the motor phases
- v: the motor terminal voltages
- e: the motor phase induced emf
- ω_1, ω_2 : the supply and load frequency respectively

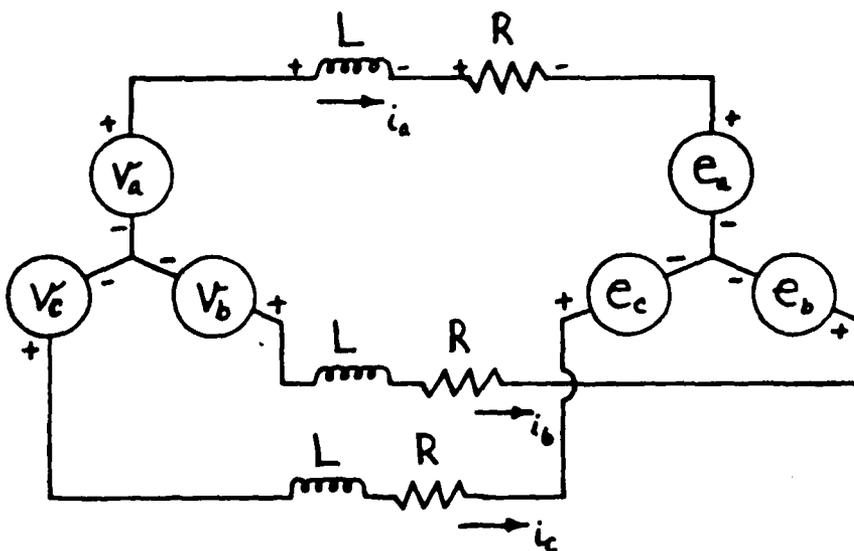


Figure 3 - BRUSHLESS DC MOTOR CIRCUIT

In figure 3, the flow of and the direction of the current in the different lines will depend on the cycloconverter's SCRs connecting the source to the motor lines at a particular time. For this reason the motor angle between 0 - 360 electrical degrees should be divided in six or more segments, depending on the conduction angle, for the purpose of the simulation.

E. Simulation Results - Due to the time frame of this work, only the static load simulation case was performed. The brushless dc motor case is left for future study. The resistance in the load was taken to be 2.6 ohms and the inductance to be 0.175 mH. The ratio of input to output frequencies was varied and also different values of conduction angles were considered. The simulation results are shown in figures 4 and 5. The output of the simulation proves that the model is correct and the program is working. Figures 4a, 4b, 4c are for the case of input to output frequency ratio of 4:1 and conduction angle of 59°. Figures 5a, 5b, 5c is for the case of input to output frequency of 2:1 and conduction angle of 30°.

V. RECOMMENDATIONS

In order to expand the application of the computer program, several augmentation and refinements are suggested.

1. Adaptation of the program for dynamic loading by brushless dc motor.
2. Modification of the program to simulate the 18-SCRs cycloconverter model.
3. Extention of the conduction angle range for angles above 60°.

Also additional research is recommended regarding the following topics:

- A. To study the effect of parameter variation on the steady state as well as the transient behavior in case of the

brushless dc motor loading.

- B. To fabricate a model and verify the test results against the computer simulation.
- C. To study in detail the control circuits needed for the cycloconverter, including the microprocessor selection or design to best suite the purpose.
- D. To investigate the conditions which might create a discontinuity in the output current of the cycloconverter feeding the dc brushless motor.

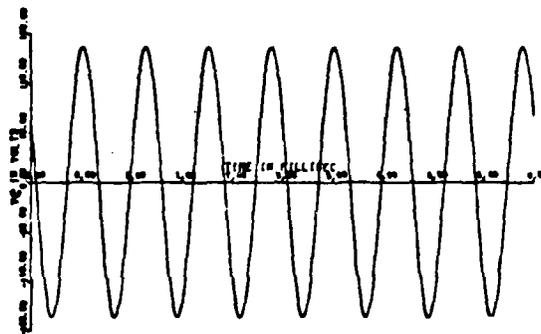
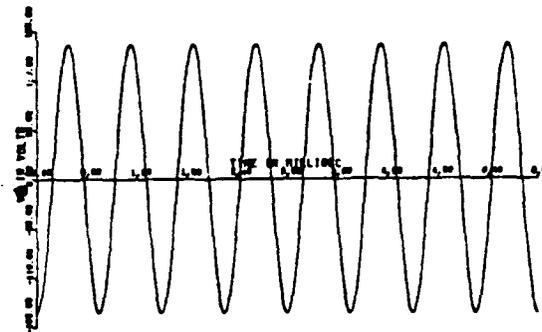
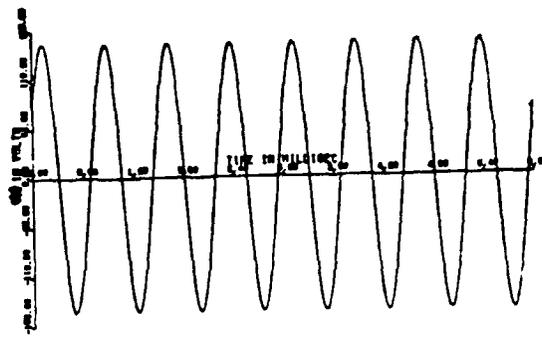


Figure 4a - THREE PHASE INPUT VOLTAGES v_A , v_B , and v_C $\alpha = 59.0^\circ$, $f_s/f_r = 4$

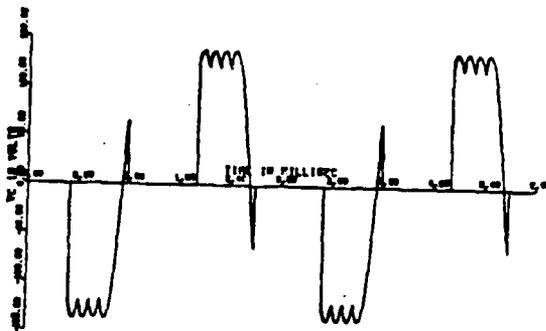
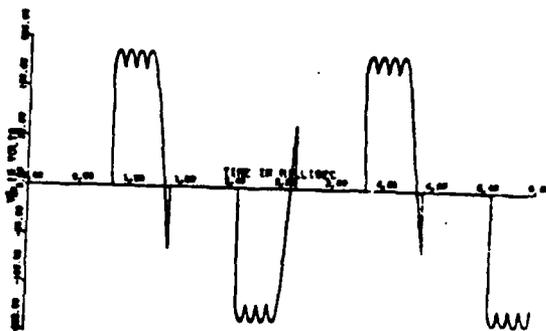
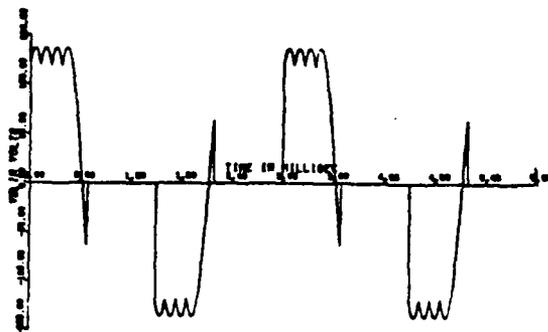


Figure 4b - THREE PHASE OUTPUT VOLTAGES v_a , v_b , and v_c $\alpha = 59.0^\circ$, $f_s/f_r = 4$

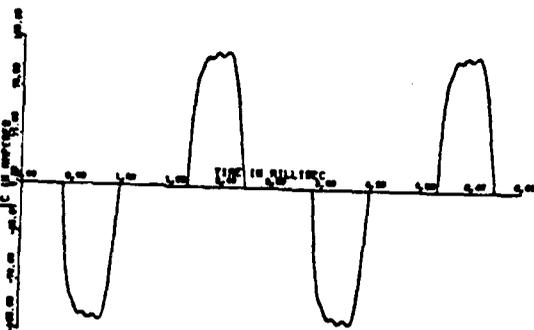
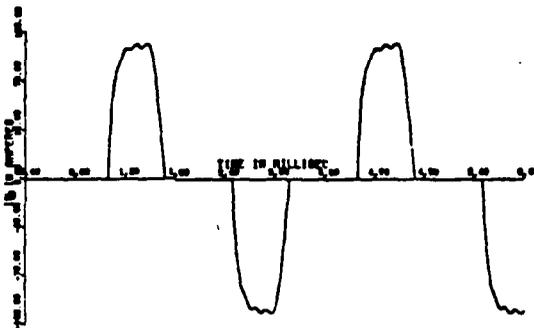
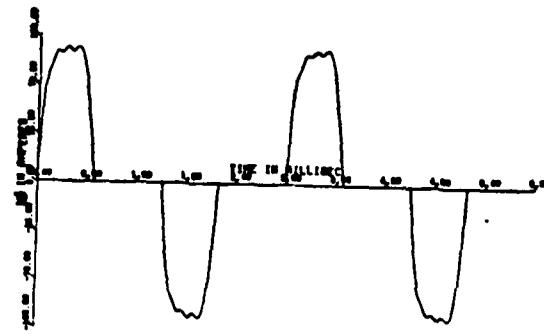


Figure 4c - THREE PHASE OUTPUT CURRENTS i_a , i_b , and i_c $\alpha = 59.0^\circ$, $f_s/f_r = 4$

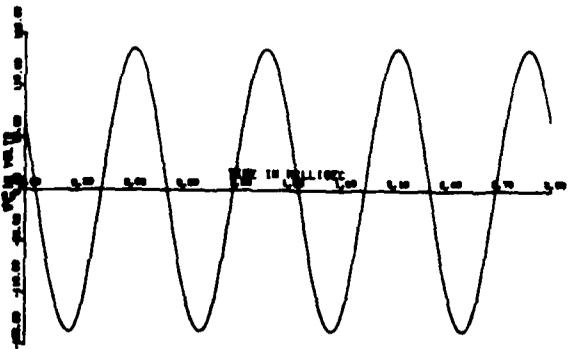
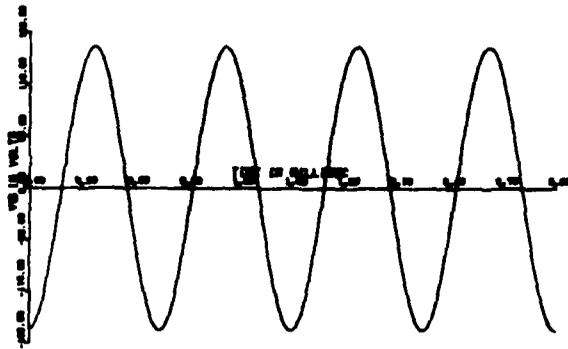
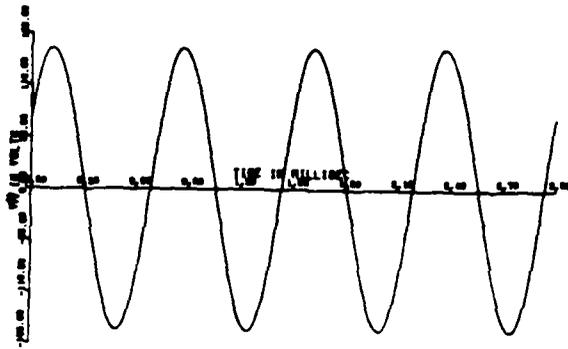


Figure 5a - THREE PHASE INPUT VOLTAGES v_A , v_B , and v_C $\alpha = 30^\circ$, $f_s/f_r = 2$

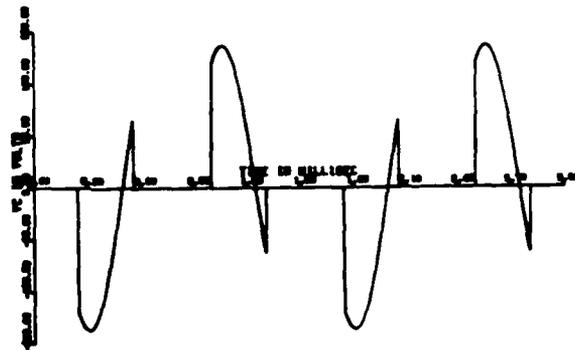
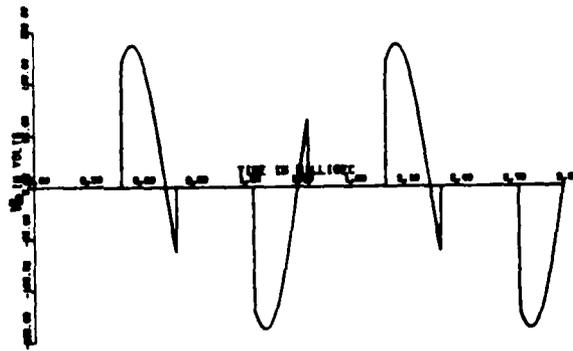
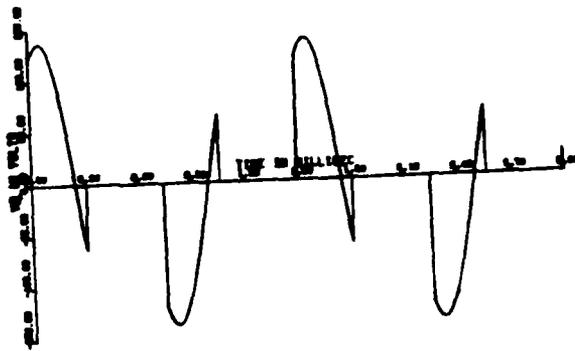


Figure 5b - THREE PHASE OUTPUT VOLTAGES v_a , v_b , and v_c $\alpha = 30^\circ$, $f_s/f_r = 2$

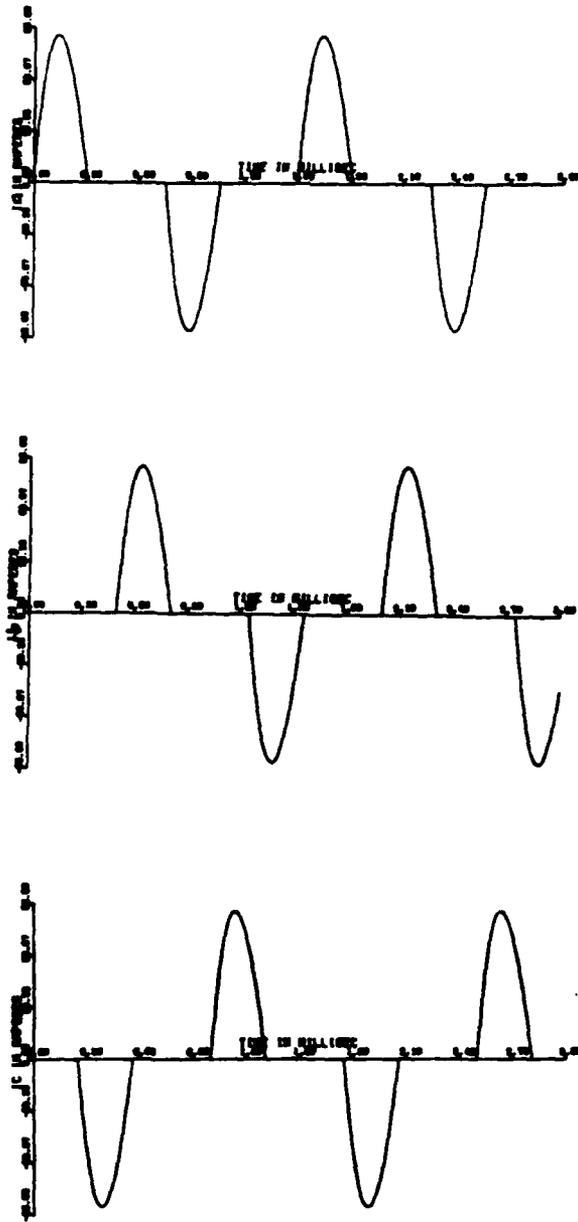


Figure 5c - THREE PHASE OUTPUT CURRENTS i_a , i_b , and i_c $\alpha = 30^\circ$, $f_s/f_r = 2$

REFERENCES

1. Weiss, Herbert W., "Adjustable Speed AC Drive Systems for Pump and Compressor Applications," IEEE Transaction on Industrial Applications, Vol IA-10, Jan/Feb 1974, pp 162 - 167.
2. Hamilton, R. A., and Lezan G. R. , "Thyristor Adjustable Frequency Power Supplies for Hot Strip Mill Run-Out Tables," IEEE Transaction on Industry and General Applications, IGA-3, Mar/Apr 1967, pp 168 - 175.
3. Tsuchiya, T., "Basic Characteristics of Cycloconverter-Type Commutatorless Motors," IEEE Transaction on Industry and General Applications, IGA-6, No. 4, July/August 1970, pp 349 - 360.
4. Jacovides, L. J., "Analysis of a Cycloconverter - Induction Motor Drive System Allowing for Stator Current Discontinuities," IEEE Transaction on Industry and General Applications, Vol. IA-9, No. 2, March/April 1973, pp 206 - 215.
5. Cathey, J. J., "Electrically Compensated Constant Speed Drive," SCEEE Summer Faculty Research Program Final Report 1982.
6. Bowler, P., "The Application of a Cycloconverter to the Control of Induction Motors," IEE Conference Publication 17 on Power Applications of Controllable Semi-Conductor Devices, Nov. 1965, pp 137 - 145.
7. Amato, C. J., "An AC Equivalent Circuit for a Cycloconverter," IEEE Trans on Industry and General Applications, Voi. IGA-2, Sep/Oct 1966, pp 358 - 362.

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 NOVEL MODULATION TECHNIQUE FOR FDM

FOR OPTICAL FIBER COMMUNICATION

(March 21 - May 28, 1983)

Prepared by:	Charles S. Ih
Academic Rank:	Professor
Department and University	Dept. of Electrical Engineering University of Delaware, Newark, De. 19711
Research location:	Rome Air Development Center, Hanscom AFB, Massachusetts, Electro-Optical Device Technology Branch
USAF Research:	Drs. Andrew C. Yang and Richard Payne
Date:	September 9, 1983
Contract No:	F49620-83-C-0035

A NOVEL MODULATION TECHNIQUE FOR FDM

FOR OPTICAL FIBER COMMUNICATION

by

Charles S. Ih

University of Delaware

Newark, De. 19711

ABSTRACT

A new laser beam modulation technique suitable for FDM (Frequency-Division-Multiplexing) for Optical Fiber Communication (OFC) is described. Two methods for implementing this new modulation technique have been studied. One uses a SWAOM (Standing-Wave-Acousto-Optical-modulator) and the other employs a Mach-Zehnder interferometer with an AOM and injection-locked lasers. The first is simpler in operation and can operate to several GHz. The second is more complicated but is capable of operating to several tens of GHz. A fairly detailed experimental investigation was made for the first. We also studied the characteristics of the second, including some experimental techniques. The FDM system using this modulation technique is suitable for multi-channel simultaneous analog and digital information transmissions.

Acknowledgement

The author wishes to express his sincere appreciation for 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 for spending a rewarding and worthwhile spring/summer at Electro-Optical Device Technology Branch, Rome Air Development Center, Hanscom AFB, Ma. He is particularly indebted to Drs. Andrew Yang and Richard Payne for many helpful suggestions and discussions during the course of this investigation and for providing him with excellent technical and moral support without which carrying out these experiments would not be possible. He would like to thank Charles Tsacoyeanes and Steve Spaziani for their patient and industrious technical assistance. He is very grateful to Steve Spaziani for his effort to the smallest details in making my stay at Hanscom pleasant. He would also like to acknowledge the assistance from Lionel Bouthillette, Charles Woods, and William Miceli including borrowing many optical components from them. The speedy service in preparing the many Brief Reports of Effort by Meg Dutra is appreciated. Last but not the least, he would like to thank Dr. Harold Roth for his encouragement.

I. Introduction

Since its inception in 1966 [1], Optical Fiber Communication (OFC) has been progressing very rapidly in recent years. Fibers with an attenuation of less than 0.2 db and transmission distance over 100 km without a single repeater at speed more than 400 mb/s have been demonstrated [2],[3]. In a coherent OFC system in which the laser diode operates essentially in a single frequency and the fiber in a single mode, the system bandwidth can exceed 100 GHz-km [4]. Fiber systems have been largely used for long haul communication using digital techniques. The advantages of light weight, security, immunity to EM interference and lower per channel cost make them also very attractive for military and other applications. For many of these applications, simultaneous transmission of digital and analog signal without elaborate signal processing is often very important. A flexible FDM (Frequency-Division-Multiplexing) system will satisfy this need. The current technique of direct modulating a laser diode with a frequency multiplexed electrical signal is less desirable for the following reasons. First the intermodulation distortion between channels is very severe due to the non-linearity of the laser diode. This can be minimized by using FM modulated channels. Secondly it is difficult to transmit digital and analog signals simultaneously. Thirdly the maximum frequency that can be used is limited by the laser whose frequency response is currently about a few GHz.

We investigated a rather novel modulation technique suitable for FDM for optical fiber communications. This technique was originally demonstrated by the author using a He-Ne laser. In this novel technique, the baseband signal is directly modulated on the laser beam and the carrier is later added on externally and separately using an AOM (Acousto-Optical-Modulator). In our investigation during the summer, we discovered that the system is potentially capable of operating at very high frequencies, to several tens of Ghz. Therefore this new technique can extend the fiber system's information bandwidth beyond the frequency response of the laser.

There are several variations of the basic system [5,6]. We investigated two configurations which offer a greater potential for practical use. The first uses a SWAOM (Standing-Wave AOM) to insert the (sub)carrier. The second incorporates a Mach-Zehnder interferometer with a TWAOM (Traveling-Wave AOM) and injection-locked lasers [7]. The first has the advantages of being relatively simple and can be used for multimode fiber systems. Current technologies limit its operation to several GHz (3 - 5 GHz) which is already equal or better than the direct modulation method. The frequency may be extended to several tens of GHz possibly using magneto-striction and optical interactions.

The second configuration requires extensive preparation for specially designed equipment. Intricate operations are necessary when discrete components are used. The operation would be greatly simplified when it is opto-electronically integrated. The

second configuration offers the advantages of very high operating frequency and the possibility for fiber dispersion compensating [6]. The second configuration is most effective in a coherent optical fiber system operating at very high frequencies.

We did an in-depth investigation for the first configuration using a SWAOM and studied the requirements and characteristics including some simple experiments for the second. Unlike the previous demonstrations [5,6], we used semiconductor lasers instead of a He-Ne laser for these experiments. This would extend this technique toward practical applications.

II Objectives

The objectives for my 1983 USAF-SCEE Summer Faculty Research Program are to investigate related issues of a new modulation technique suitable for FDM for optical fiber optical communications. Different implementation configurations will be studied. Specifically the objectives are:

1. Investigate the new modulation technique suitable for FDM for optical fiber communications using a SWAOM and GaAlAs lasers.
2. Study the feasibility of using a SWAOM to modulate an incoherent light source, such as an LED.
3. Study the requirements and characteristics for injection-locking of two single-mode GaAlAs lasers with a

frequency shift between the two. The required frequency-shift is accomplished by a TWAOM.

III. System Description and Experimental Results.

A. Laser Modulation Using an SWAOM

The optical layout is shown Fig. 1. The operation of the system is the following. The pulsed laser beam (digital modulation) is first collimated by the lens L1. A wavefront-correction-optics (WCO) may be used to correct the wavefront distortion of the laser diode and thus improving the coupling efficiency to the optical fiber. The collimated laser beam is then passing through a standing wave AO modulator (SWAOM) at the Bragg angle. The diffracted beam is focused by lens L2 and coupled to an optical fiber for transmission. If the SWAOM is driven (CW) at a frequency $(1/2)fc$, the diffracted beam then contains two frequency components, one at $f + (1/2)fc$ and another at $f - (1/2)fc$, where f is the frequency of the laser. A photo-detector which is a square-wave detector will detect the beat frequency which is fc . Therefore this system is effectively a DSB modulation with the carrier being suppressed. We shall call this technique the DBM (Double-Beam-Modulation). Therefore the (digitally) modulated laser beam (in the baseband) is automatically put on a (sub)carrier. The (sub)carrier frequency is equal to twice the frequency input to the SWAOM. Even the system has been

described for digitally modulation, it works equally well for analog modulations.

We first used an RCA laser diode, Model C30130. It was a multimode and CW laser with a maximum output of 15 mw and an unknown radiation pattern. The laser was operated at an output level of 4 mw. A 0.5-km multimode fiber with an attenuation of 5 - 6 db/km at the .83 um was used for the experiment. Because of the multimode of the laser, the coupling was estimated about only 20% or a 7 db loss. No attempts were made to improve the coupling efficiency, since our main purpose of these experiments was to investigate the characteristic of the modulation technique. The use of an anamorphic lens will improve the coupling efficiency. Using this modulation technique, we can simultaneously send many channels of digital and analog signals. Since these signals are separated by the different RF carriers, the exact wavelengths of the optical waves are unimportant and can be very close together.

We later used a single mode laser, Model number ML--3101, supplied by Applied Invention. Because the radiation pattern is nearly a gaussian, the coupling efficiency was improved to about 50%. The Intra-Action Model SWM-40 modulator (SWAOM) was used for all the experiments. The driving frequency was 40 mHz and therefore the carrier was 80 mHz. The modulated laser beam transmitted through the same 0.5-km multimode fiber and the detected signal had an excellent SNR. The results are shown in Fig. 2. The laser was modulated up to 20 Mb/s, therefore there

were 4 cycles in each pulse.

B. DBM of an LED

Since a SWAOM produces the frequency up- and down- shifted beam with exactly the same optical path length, it was suggested that the SWAOM can be used to modulate even an incoherent source as such an LED. We used an LED with a fiber pigtail in place of a laser diode. The modulated beam was indeed observed. However, because of the inherent large beam divergence of the LED, the modulation efficiency (or depth) was very low. The SNR was therefore also low and estimated to be only 5-to-1.

C. Investigation of Injection-Locked Laser Modulator

The modulator is schematically shown Fig. 3. Basically this modulator is a Mach-Zehnder interferometer incorporating an TWAOM (Traveling-Wave-Acousto-Optical-modulator) and two injection-locked lasers. The laser LR1 is directly injection-locked to laser LR3 but LR2 is injection-locked to the frequency-shifted beam of LR3. The frequency is shifted by F_c by the TWAOM. Therefore the recombining of the two beams produces the same effect as DBM. The laser LR1 and LR2 are modulated directly and simultaneously by the modulating signal. A complete implementation of this modulator requiring many specially designed equipment was therefore difficult to do for the short time. We tempted only to study the characteristics and requirements for injection-locking of two lasers at different frequencies. The experimental set-up for this study is shown in Fig. 4.

The output of laser LR1 was frequency-shifted by the TWAOM. The frequency-shifted beam was then used to injection-lock the laser LR2. In order to have an injection-locking to take place, the frequencies of the two lasers must be within a few GHz of each other. The frequency of a laser diode can be controlled by changing its temperature and the injection current. Two lasers (Model ML - 3101) of nearly identical frequencies are individually mounted on two thermal-electrical coolers. The wavelengths of the lasers were monitored by a high resolution holographic grating and a TV camera. This simple monitor was capable of resolving 0.05 nm and more than adequate to reveal the mode structure of these lasers. The two laser beams were displaced slightly so that they can be seen simultaneously on the TV screen. The monitor was first calibrated (relatively) using a beam splitting arrangement to simulate the two beams from a single laser. After calibration, the two laser beam were then displayed on the TV screen. Their positions which represent their frequencies (or wavelengths) can then be compared on the TV screen. By adjusting the laser temperatures and the injection currents, the wavelengths of the two lasers could be brought together. Injection locking of the two lasers were attempted. However, we could not positively confirm even a short term locking because we did not have at hand the necessary monitor equipment. A longer term locking, say only a few minutes, was not expected because the current and temperature were not automatically controlled. However, during the experiments, the changing of the laser mode due to the temperature and injection change was observed. The laser

mode shifted slightly before jumping rather quickly to another mode when the current was changed. The temperature could change the the position of the lasr mode gradually over a much larger range before jumping to the next mode. Therefore by changing the temperature and the injection current, the wavelengths of the two lasers were brought together. The laser mode changing due to optical feedback and interaction between the two lasers were also studied. Two optical isolators were assembled so that the optical feedback from the second laser to the first was eliminated. This is an important step toward injection locking of two lasers.

IV RECOMENDATIONS

The modulation techniques demonstrated during my 1983 USAF-SCEEE summer research program offer many possible practical applications particularly for multi-channel and two-way short distance communications. Since the communication channels can operate essentially independently, analog and digital information initiated and received at different locations can simultaneously be transmitted through an optical fiber using an FDM system employing this modulation technique. Since the baseband information is modulated on the laser beam directly and the (sub)carrier is added separately, very high carrier frequency is possible and not limited by the laser. Also separated lasers are used for each channel, the intermodulation due to non-linearity of the laser modulation characteristics is less important. Even though the injection-locking of two or more lasers offers the potential of operating the carrier frequency to several tens of GHz, the

simpler modulator using a SWAOM with a carrier to several GHz is sufficient for most applications. The operating characteristics of an FDM system using the new modulation technique, such as, cross-talk, intermodulation, SNR, etc. can all be learned at lower frequencies.

I proposed to study the FDM system for optical fiber communication using the new modulation technique operating at carrier frequencies near 1 GHz. Several analog and digital channels will be operating simultaneously so that their transmission characteristics can be studied. I intend to submit a proposal for the RISE program which replaces the Mini-grant this year. If time and funds permit, I intend to continue some of the experiments for injection-locking of two lasers.

References

1. C. K. Kao and G. A. Hockman, Dielectric fiber surface waveguides for optical frequencies, Proc. IEE, 133, 1151-1158, July 1966.
2. K. Iwashita and K. Nakagawa, 400 Mbits/s Transmission Test using A 1.53 μm DFB Laser Diode and 104 km Single-Mode Fibre, Electronics Letters, 18, 937 (28th October, 1982).
3. M. M. Boenke, R. E. Wagner and D. J. Will, Transmission experiments Through 101 km and 84 km of Single-Mode Fibre at 274 Mbits/s and 420 Mbits/s, Electronics Letters, 18, 897 (14th October, 1982).
4. J. Yamada, A. Kawana, H. Nagai and T. Kimura, 1.55 μm Transmission Experiments at 2 Gbits/s Using 51.5 km Dispersion-Free Fibre, Electronics Letters, 18, 98 (21th January, 1982).
5. U.S. Patent No. 4, 210,803, July 1, 1980. Inventor: C.S. Ih assignee: the University of Delaware.
6. C. S. Ih, Feasibility and Requirements for Dispersion Compensation in Coherent FOC, Optical Waveguide Sciences - Proceedings of the International Symposium, Kweilin, China, June 20 - 23, 1983 Martinus nijhoff Publishers, 1983.
7. S. Kobayashi and T. Kimura, Injection Locking in GaAlAs Semiconductor Laser, IEEE J. Quantum Electron., QE-17, 681, May (1981).

Fig. 1.

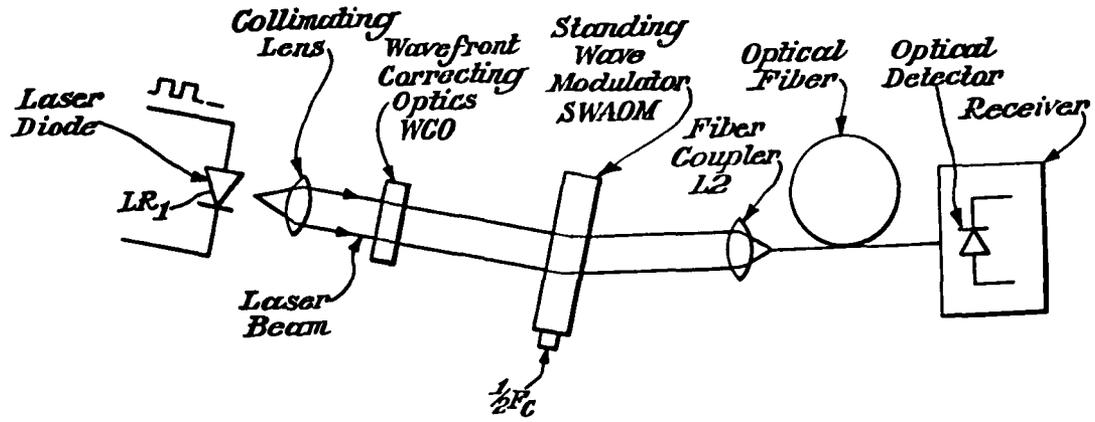


Fig. 1 Double-Beam Modulator (DBM)

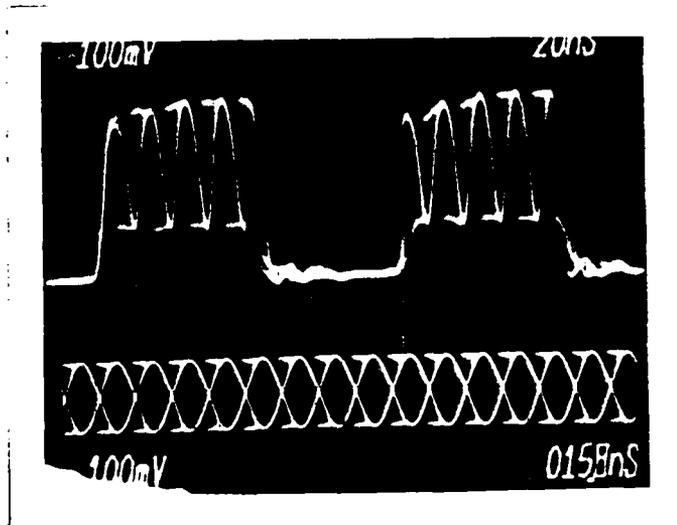


Fig. 2 Output of DBM

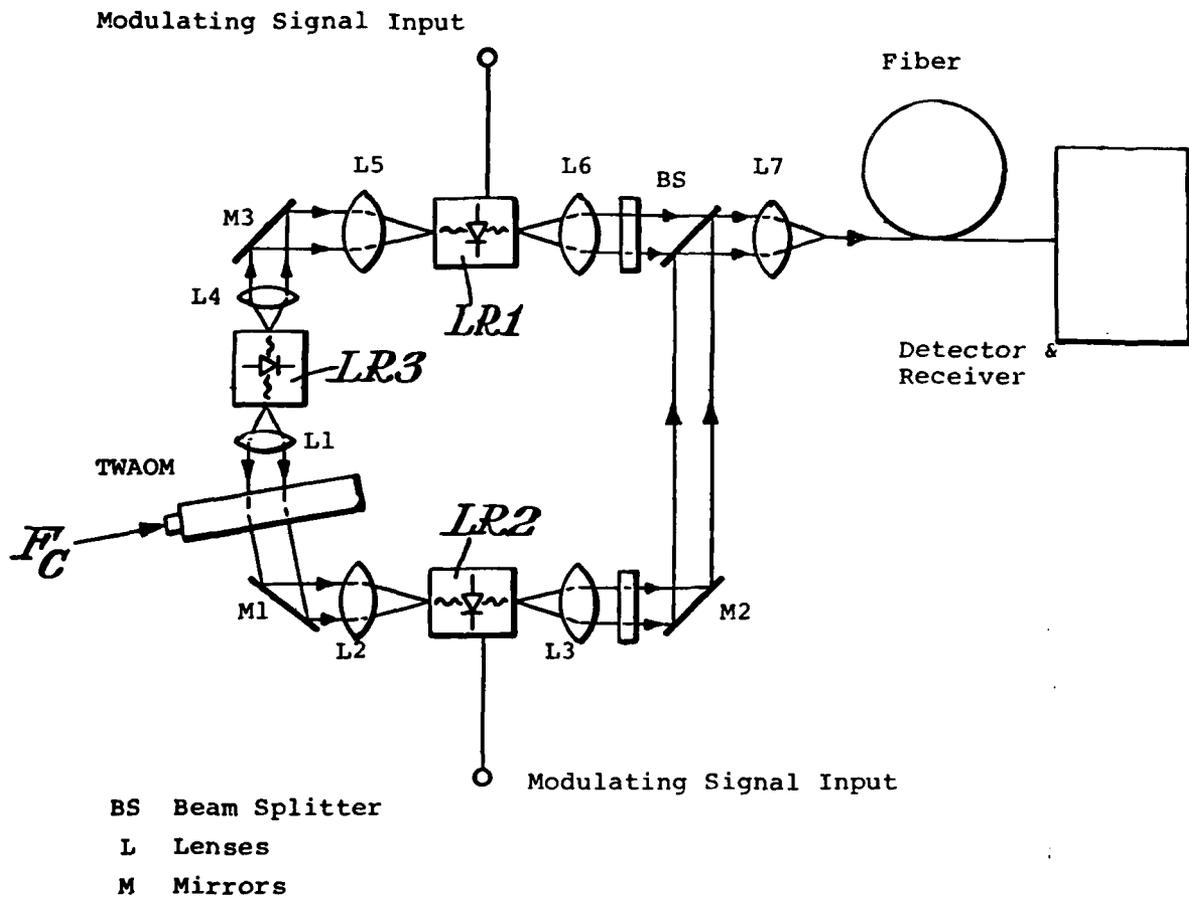


Fig. 3 Injection-Locked Laser Modulator

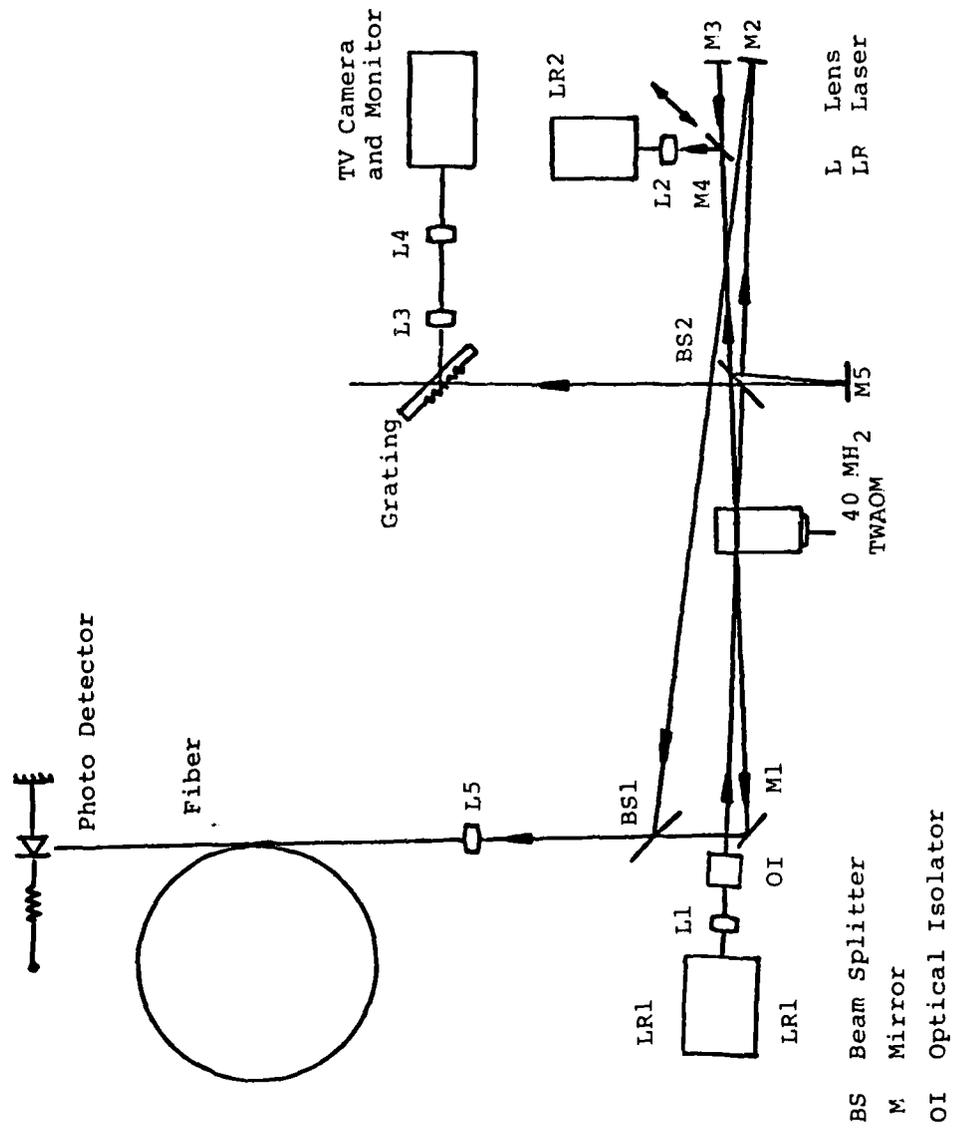


Fig. 4 Optical setup for Injection Locking Experiment

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

SOFTWARE FAULT-TOLERANCE/DIAGNOSTICS FOR SINGLE-USER SYSTEMS

Prepared by: Dr. Gregory W. Jones
Academic Rank: Associate Professor
Department and University: Department of Computer Science
Utah State University
Research Location: Air Force Human Resources Laboratory,
Training Systems Division,
Technology Development Branch
USAF Research: Maj. Richard E. Bolz
Date: September 19, 1983
Contract No: F49620-82-C-0035

SOFTWARE FAULT-TOLERANCE/DIAGNOSTICS FOR SINGLE-USER SYSTEMS

by

Gregory W. Jones

ABSTRACT

The rapidly increasing use of microcomputer systems by isolated, relatively inexperienced Air Force personnel puts a severe strain on software. Programs must be designed to handle routine errors without user intervention and to give maximum support to the user when errors occur that can't be handled automatically. Failure to do this causes frustration, resentment, error and even rejection of the software.

This report covers the basic concepts underlying fault-tolerance in software and diagnostics for users, as well as describing the general techniques for implementing them. The report then focuses on detailed, specific techniques for supporting the goals of error -avoidance, -handling and -investigation. Hardware considerations are basically ignored.

The Air Force has contracted for an Instructional Support System, modeled on the current Advanced Instructional System. This software will be widely used by non-experts in remote locations, and the report describes current and projected diagnostic needs for the AIS and ISS. Suggestions are made for specific changes and additions to ISS, as well as for additional research in the general area of fault-tolerance and diagnostics.

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 him with the opportunity to spend a worthwhile and informative summer at the Air Force Human Resources Laboratory, Lowry Air Force Base, Colorado.

The Laboratory has been cooperative in providing research opportunities. Lt. John Prentice, Lt. Brad Poulliot, Alan Marshall and Maj. Dick Bolz have shown interest in the project. Bob Summers has helped considerably as a local contact to SCEEE. Finally, Dave Pflasterer of McDonnell Douglas has been most helpful in providing information on the ISS and AIS systems.

I. INTRODUCTION

The failure of a computer system to perform as expected can be caused by hardware or software faults or by mistakes of the user. These failures are manifested in lost or corrupted data, wasted time and effort and incorrect results. The secondary effects are frustration, anger and distrust of the system. Since the user's feelings about a computer system play an unexpectedly large role in the decision to utilize it after its installation, the perception that it is unreliable can cause an expensive and useful system to be an unused ornament and a waste of money.

These problems are compounded when a relatively inexperienced user in a location remote from any expert help is given use of a modern microcomputer system. These systems are generally capable of performing an astounding variety of complex operations. However, their complexity leaves them open to serious mistakes, which possibility hangs like a threatening cloud over the head of the user. Having little experience, he or she reacts to any failure as if it were a disaster, even when it can be easily corrected. There is no one to help and they fear that any effort to fix things up could lead to worse damage.

In this situation, the user could be helped by a number of common resources. They are, in order of usefulness: expert advice, good operating manuals, descriptions by the system itself of what has gone wrong, and subsidiary software to allow investigation of the current state of affairs. Given today's shortage of trained personnel, expert advice is often unavailable; it may be automated, though it rarely is, in expensive software "expert" systems. Manuals are often automated in the form of "help" packages; in fact, these are one of the hallmarks of well-designed software. Error messages, dumps, debuggers, etc. represent the last two resources, but it is difficult for an inexperienced person to use them effectively.

The Instructional Support System being developed by McDonnell Douglas and AFHRL/ID is a very complex piece of software which will eventually be available to the Air Force, Navy and Army. It can be used for training purposes wherever needed by personnel with a minimum

of experience in computing. ISS is based on an existing system, AIS, which has been used for years, but has always had available experts to handle problems. It is foreseen that ISS will suffer from decreased utility and acceptance unless additional efforts are made to provide inexperienced users with sufficient diagnostic aids to support them.

II. OBJECTIVES

The objective of this research was to provide strategies for supporting user management of failures of ISS.

The specific goals were to:

- 1) Analyze the operational history of AIS and the design of ISS with regard to failure and failure management;
- 2) Investigate the general principles of fault-tolerance and diagnostics in the light of the needs of ISS; and
- 3) Develop a set of suggested procedures for diagnostics for ISS.

III. FUNDAMENTAL CONCEPTS OF FAULT-TOLERANCE AND DIAGNOSTICS

There are three levels of response to the possibility of failure in computing systems. The first is termed "fault-intolerance" and represents the well-known effort to ensure that hardware be so dependable and software so well-designed that failure simply does not occur. It is appropriate that this should be the first line of defense, and that perfecting design should be always in the mind of the designer. But perfection requires a great deal of talent and a huge expenditure of resources; realistically, designers must be aware of the possibility of failure in their systems, and must provide for that possibility.

The second level of response is "fault-tolerance". It represents design which attempts to deal automatically with faults that may originate in the hardware, software or operation of the system. Examples of this are: retry when an I/O operation is in error; alternate routes to avoid "divide by zero"; and "end of file" conditions.

The third level is that of diagnostics. It provides the same kind of failure management that fault-tolerance does, but under user control. These three techniques form a spectrum in terms of user-

friendliness, cost, value and necessity. It is up to the system designer to determine where on the spectrum his software should sit. The decision usually balances reliability and ease of use against cost. Occasionally a designer, through thoughtlessness or laziness, decides not to provide even the bare minimum of diagnostics.

In terms of software, accepted terminology distinguishes¹ between a fault, which represents incorrect design (a bug), an error, which is the manifestation of the fault within the program, and a failure, which is the irrevocable action taken by the software which affects the user. Fault-tolerance comprises the set of techniques used to automatically correct errors and avoid failures, while diagnostics aid the user in managing failure after the fact. The goals of fault-tolerance/diagnostics^{4,12,14} are:

- to detect errors quickly and perhaps categorize them;
- to confine the effect of errors so as to minimize damage;
- to correct the error in a way invisible to the user;
- to restore the system to a consistent state when correction is impossible;
- to make a permanent record of errors, their effects and correction.

Detection of errors is vital to localization and elimination of faults, but the latter is an activity of the designer, not the user.

IV. DESIGN CONSIDERATIONS FOR FAULT-TOLERANCE/DIAGNOSTICS

Specific techniques promoting the above goals will be covered in section V. It is hard to find any overall structure in the details, perhaps inevitably so, since the problem is serious, forcing designers to use any tool that comes to hand. On the other hand, the principles of software engineering demand a coordinated, methodical approach to the design of programs. So it is desirable to impose some general organization on the specific techniques. Two are possible.

The first is obtained by dividing the methods into backward and forward procedures. Backward procedures rely on the detection of an

error and the ability to back up to a consistent past state and either retry the same path or choose an alternate path which will avoid the fault. Backward procedures have the advantage of being general; that is, they can be implemented in the form of software tools, applicable in a wide variety of situations with little customizing. This allows for great savings in time and effort on the part of the software designer. The disadvantage of backward techniques is their lack of intelligence. They fail to solve problems that could be solved by thoughtful consideration of specific details^{11,14}.

In the other category, forward procedures involve steps specific to the particular circumstances. They cost more, but are more efficient in the long run. However, the additional design complexity required by forward techniques may outweigh the long-term savings unless the software will be very widely used. Forward procedures should generally be the second choice of the designer.

The second organization of specific techniques is based on their different general approaches. Of course these are rather crude umbrellas, but they do point out similarities in techniques which are used in widely varying circumstances. In the following section, these groupings will be referred to by the single-letter codes used here:

- R -- Redundancy
- D -- Double-checking
- L -- Localization
- S -- Restoration
- A -- Analysis
- C -- Reconfiguration

There are other, very complicated techniques used for multiprogrammed or timeshare systems, where backward procedures are much more difficult to implement. Since most microcomputers are single-user systems, this question will be ignored here.

Redundancy techniques refer to those which provide spare parts: duplicate hardware components, alternate software modules, multiple

copies of data, etc. Double-checking procedures require extreme caution in making decisions. Some examples of double-checking are handshaking routines, special privilege requirements and parity.

Localization is the design method which groups processes and data and isolates the groups from each other. This has the purpose of avoiding unintentional interference and confining damage. Restoration is a combination of backup and procedures to correct damaged data. Analysis techniques attempt to examine data, software and hardware logically and locate errors. Finally, reconfiguration is used to return the system to a known (and perhaps restricted) state.

Software designers are typically practical and optimistic. This creates a willingness to accept "anything that works" and a belief that there are no faults in the product. Consequently, formal design techniques are not usually palatable to software designers, particularly those which have fault-tolerance/diagnostics as their goal. The lack of methodology in software design has long been recognized as a significant part of the "software crisis", but failure management techniques, when they appear at all, are usually ad hoc and not adequately tested. Anybody who has tried to cope with error messages in a typical program has a feeling for the quality of work that designers put into diagnostics.^{3,11,12}

Structured, modular programming promotes localization, and all standard design techniques (modular & top-down, data flow, Jackson, etc.) promote fault-intolerance. Consequently, some such technique should always be used. . If it is justice that the punishment fit the crime, then designers who fail to use a design methodology or to consider fault-tolerance/diagnostics should be imprisoned together and condemned to 10 years of using each other's software.

V. DETAILED TECHNIQUES FOR FAULT-TOLERANCE/DIAGNOSTICS

The methods outlined here are organized into categories, essential and recommended and sophisticated. Since the introduction of fault-tolerance and diagnostic capabilities requires considerable expenditure of time and effort, the designer must balance cost against value.

There is, however, a minimum capability which should be provided in all cases. This is represented by the essential category here. Throughout this section specific techniques are tagged with the letter code of the "umbrella" under which they fit (see Section IV).

ESSENTIAL

Documentation¹² (A) - It is necessary to provide planned, well-written information which includes all operating procedures, examples of input and output for actual use of the procedures, structure of all data, defaults, assumptions and restrictions. A famous computer scientist has suggested one way to improve reliability is to decrease user optimism, so that performance is in line with expectations. While this seems somewhat cynical, documentation should give reasonable estimates of system performance. Documentation, in common with the other analytic techniques, is forward.

Consultation (A) - This of course amplifies the function filled by documentation and provides expert guidance based on experience in the search for the cause of failure.

Duplication (D) - There should be a provision for creating duplicates of data (backup files) and of software. Hardware duplicates are more expensive, but might sometimes be worthwhile.

File Access (A) - Users should be able to determine the presence, creation date and time of update of files, as well as to transfer files from one medium to another. Believe it or not, there are packages provided by respectable companies that do not permit this.

Restart^{1,4,8,11,12} (C) - The user should be able to suspend or abort execution of a package, as well as restart it with a new disk image loaded into main memory or reboot the entire computing system. All these gradations of action should be available. While restart usually involves significant losses, it does return the system to a predictable state. Separate reset should also be available for peripheral devices like printers and terminals. Restart techniques do not generally reset files, but other techniques for this purpose will be discussed later. This is obviously a backward and general procedure.

Startup diagnostics⁸ (A) - Programs, often ROM resident to protect them from failure, can be used to test the integrity hardware, software and data. A well designed program will depend at any point only on those capabilities which have already be verified by the program.

Screen feedback (A) - Software should never be designed to allow the user to spend more than 5 seconds looking at a static screen after specifying some action. During long procedures a blinking figure will reassure the user, while a line visibly growing toward a termination point will give greater information. It is advisable that the I/O filtering routine be set to respond to a control character (like the ^T in VAX VMS) to determine asynchronously to the executing procedure whether the CPU is still responding to the terminal and whether the executing procedure is halted.

RECOMMENDED

Low coupling and high cohesion^{3,11} (L) - Accepted standards of structured, modular programming promote the grouping of code and data which is closely related into a module (high cohesion) with a minimum of interaction between modules (low coupling). This has the effect of containing errors and making detection of errors and location of the underlying fault easier. If communication between modules is restricted to carefully designed interfaces, there is decreased likelihood of unwanted side effects (the kind produced by using global storage like FORTRAN Common), thus containing errors.

Generalized error procedures^{5,7,8,12,14} (L) - It is desirable to separate error detection, which is carried out in applications modules, from error handling, which should be carried out by general purpose handling routines. Communication of the detected error should be by means of a set of integer error codes that has a well-designed structure. The handler can decide whether to fix the error and return to the signaling module or exit to a higher level in an attempt to backup. It may be that the applications module itself will transmit this decision to the handler, rather than allowing it to make the decision. One advantage is that the handling routines can be table-driven software

tools, thus lowering cost. Another is that testing of handling procedures, always a great challenge, can be facilitated by this uniform treatment.

Verification^{4,10,12,14} (A,S) - Values should be checked upon receipt or creation and when returning to them after an absence. This means that all input data (whether from a peripheral device or as parameters in a procedure invocation) and all calculated data should be checked for reasonableness before being acted upon. The error handler can either request retransmission or recalculation of the data or, where this is impossible, replace the data by some default value if this will not compromise the correct working of the system. Modules should also verify the correctness of internal data structures which were created earlier as the first step when the module is invoked again.

Retry and voting^{1,4} (R) - It is sometimes profitable in the case of bad input to attempt the same action again. It is possible in the case of multiple attempts to utilize the data which resulted most often. Vital systems may sometimes benefit from parallel execution of identical processes, proceeding only when they agree.

Backup and checkpoints^{12,14} (S) - In the case that no correction of an error is possible, the error handler may decide to back out of the current situation till a point is reached from which the program can safely proceed. Backing out may require the undoing of any changes in the values of variables that occurred after the safe point. This can be facilitated by creating checkpoints at which the values of variables are stored, so that they can be easily restored when backing up.

Handshaking¹² (D) - This is a technique which requires that all communication of information be a two-way process. In other words, it is not sufficient to tell somebody something, you must be sure he or she heard what was said. Handshaking tends to insure that decisions made on the basis of incomplete information and mistaken identification are avoided. In computing systems, this may involve the reverse transmission of data ("Did you really say to delete all files?"), the comparison of the key stored in a record which has been found via a table with the original key sought or the avoidance of "ELSE" cases in CASE structures.

Error messages and logs^{8,10} (A) - The system must provide clear, accurate error messages in the case of faults which can't be handled automatically. These should provide sufficient information to allow the user to correct whatever error may have occurred. High quality error messages are another hallmark of a well-designed system. Still, it would be a mistake to inform the user of the errors that were automatically dealt with, thus causing an interruption in the flow of information on the screen. It is important to have a record of these events, so as to be able to isolate and eliminate faults. For this reason, a disk file is used to log the time, place, results and, if possible, cause of each error. This log can be examined periodically for diagnostic purposes.

Help facilities (A) - These represent the automation of documentation. A well designed help package will aid the user in determining the correct procedure to follow at any point. Different help menus should be available at different points in the program, so that the user is only burdened with the information which is currently applicable. This technique can be expanded by addition of intelligence to the package, thus turning it into an expert system.

Standard file methods^{13,14} (S,D) - One of the standard diagnostic techniques for files is the use of checksums, cyclic redundancy checks and parity to aid in the detection of errors. Most hardware implements one of these methods in its circuitry. Fault-tolerance is often obtained by the use of grandparent, parent, child file versions together with a separate audit file which documents the changes made since the time when the backup file was current. The backup files can be created by the user at times that seem appropriate, or can be created automatically by the system. The latter method has historically been detrimental to system throughput, since backups are created more often.

Differential files and careful replacement¹³ (L,D) - These are somewhat similar techniques involving the avoidance of inplace replacement until it is deemed safe. Careful replacement is often used in database management systems, where it delays the recording of changes in a record until all changes are present and verified. Differential

files involve marking of records in a file as "not current", but writing the updated record to another file. These are merged into a single updated file when it is safe to do so.

Salvation programs ¹³ (S) - These programs are used when a file has been corrupted and needs to be restored to a consistent state suitable for further use. They generally allow the user to automatically or manually follow the structure of the file, determining the data stored and possibly altering it or reinitializing it. It may be necessary to flag all records which have been altered so as to broadcast to future processors that they are suspect. Such "data editors" are generally designed so as to be able to trace the structure of the file in terms of logical records and linked lists, where these are used. A well-designed salvation program can be used as a software tool.

Cross-connection (A,R) - It is desirable, when economically feasible, to provide the possibility of cross-connecting different components so as to diagnose failures in hardware or software. This may be done by, say, replacing a suspect terminal by another one located at the same place, or by use of remote components via telephone or other data communication lines. Ideally, it is possible to use a remote computer to test each individual hardware component of the failing system, as well as the software and data stored on it. These diagnostic tests should involve heavy use of the various components and comparison of results with the known behavior of a correct system.

SOPHISTICATED

Layered design ^{4,5,6,8,11,14} (L) - The use of layered design, in which each level sees the next one down as a virtual machine, is common in large software packages, including ISS. As regards fault-tolerance and diagnostics, the designer must be careful to remember that the interfaces between modules at the same level are standard communication interfaces, but between modules in different levels they are abstraction interfaces. Error-handling at a lower level should be

completely transparent to modules at higher levels. When lower-level modules are ignorant of the implications of an error, they must pass it to the upper-level invoking routine as an exception. The invoker in turn may cancel the operation (possibly including the transmission of a reset message to the lower level) or call on a standard error-handling routine for help. The errors most correctable at any level are those which involve parameters passed to or from a lower level and the corruption of data structures local to this level.

Standby software^{6,10,11} (R,S,C) - When an error is detected, one way of handling it is to obtain the desired results using a different method. The new method should be orthogonal (that is, totally distinct from) the previous method. In the case of further error, another way of obtaining results may be tried. Good design techniques should be used to avoid increasing the complexity of the system. The general trend should be to use the sophisticated but efficient method first, then proceed toward the simple-minded and brutal. These alternate modules, the main and standby software, can be grouped into a "recovery block", a kind of "IF it works THEN send results to invoker ELSE IF the first alternate works THEN send out those results ELSE IF...ELSE punt" structure. Checkpoints must be established so as to reprovide the initial data each time a retry is made.

Soft failure¹ (C) - It may sometimes be appropriate to handle an error by shrinking the system to an essential core. This has the effect of avoiding those processes which have been found to be faulty, but still carrying out the most important duties of the system. Future expert analysis may be necessary to restore the system to full function, or it may be that when operating on new data, the system can return to full capacity.

Statistics (A) - Information about the efficiency of a running system, CPU time and frequency of calls to routines, can locate faults which do not result in outright failure, but still impair function because of repeated but tolerated faults. While statistical analysis capabilities usually accompany larger systems at the time of purchase, these functions have to be specifically designed for micro systems.

Tracing and debugging^{2,3,9} (A) - Tracing is the process of allowing the user to know the order in which elements of the software are executed. For diagnostic purposes, it is good to allow the user to specify the granularity of the trace: first, module by module, then when the target module is identified, structure by structure, and finally source line by source line. Tracing can involve a great deal of overhead cost, so it should be possible to cut it out until it is requested. The use of trace traps triggered by the call interrupt handler or for each instruction may cut down the overhead. At the module level, trace overhead may amount to no more than 4%, and in this case it may be worthwhile to constantly trace the software, putting the output into a disk file in case of a failure. It is important to provide actual module names, source line text, etc. in order to make the trace meaningful. Debuggers are general tools which provide the trace capability and additionally give the user access to the values of variables at each moment. They also allow the user to control execution in real time. Debuggers should be chosen so as to provide data in an appropriate format, which may even include the ability to define special formats to represent complex data structures.

Language capabilities^{5,7,9,14} (A) - In choosing the source language for software, it is possible to make the implementation of some of the above techniques easier. Where possible, a source language should include: ESCAPE (go immediately to error handler and don't come back), NOTIFY (go to error handler but return here afterwards), SIGNAL (let error handler decide what to do), ENDED (what to do in end-of-file type situations) and the ability to pass the identity of the error handler to be used by an invoked procedure as a parameter to that procedure. The compiler for the language should provide the hooks and symbol table information necessary to implement traces at the module, structure and statement levels, names of modules or structures, text of source code for the traces and variable names and types for the debugger.

Time-out protection^{6,8,12} (D) - When entering a procedure that has the capability of looping indefinitely, it may be wise to schedule an

asynchronous interrupt after a given amount of time unless the procedure has been exited.

Self-identifying structures^{4,7,14} (D) - This technique associates with each internal data structure a key or name, much like the primary key or ID found in file records. It is useful for procedures wandering lost in core, enabling them to avoid making updates in the wrong place. This is particularly important in the case of pointers and descriptors. In this form, self-identifying structures represent an extension of handshaking techniques. It is possible to store, along with the key, a set of codes to indicate which procedures are authorized to alter the structure. This makes it easier for the procedure to determine whether it is in the right place. Self-identifying structures are just as useful in the case of a user who is using a debugger to run through main memory in address order.

Recovery stacks**** (S) - These data structures are used to store the information required by recovery blocks for retry by standby software modules. Certain specialty languages have recovery blocks as a syntactic element, while the compilers provide for the automatic creation of recovery stacks.

VI. THE AIS AND ISS ENVIRONMENTS

AIS has been used for 10 years in one form or another. It has undergone continuous debugging, so that it is relatively fault-intolerant. However, it has never really escaped from the development environment; that is, it has not achieved true status as a distributed product. Instead, it remains a one-of-a-kind package subject to constant tinkering. In its design it benefits from good verification checks, a good help package and good modular structure. The overall design structure has little logic, making study of the system somewhat overwhelming. The greatest number of failures seem to occur because of alteration of the operating system on the CYBER 73-16 on which it runs. These modifications have done an end run around the failure management resources of the operating system in many cases. The major failures of the system involve structure corruption (traceable to the

operating system modifications), terminal hanging and bad reaction to control characters (traceable to lack of tailoring of the software-hardware interface), buffer overflow (a design fault) and the ability of naive users to bring the system down (again, the operating system modifications).

The vast majority of ISS is the result of a mechanical translation of AIS, which is written in CAMIL, into Ada. For this reason, ISS suffers from the same lack of overall design as AIS, and suffers further from the fact that the Ada code, having been created automatically, is not always comprehensible. Add to this the fact that internal documentation of AIS is sparse and that the translation effort does not include documentation, and the result is a package that will be very difficult to work with if it is to be modified.

If ISS is distributed for extensive use throughout the armed services, as is currently anticipated, then it will be used by inexperienced people in remote locations. They will certainly have great need of help in failure management. There will also be pressure to provide versions of ISS to run on equipment which has already been acquired for some other purpose and may not be very well adapted to ISS. The current paucity of documentation and apparent insufficiency of error messages will be frustrating to users. The problems AIS has had due to operating system modification, however, should not propagate into these small ISS installations so long as the current decision is kept, not to install ISS except where there is a full Ada support environment.

Current plans are to continue to involve McDonnell Douglas personnel as expert consultants in order to provide the bulk of failure management. This in turn will tend to keep the software shackled to the development environment, so that it will still be in a constant state of flux. Installation and control of the software is supposed to be in the hands of a separate entity, but this will be difficult if the users are in constant communication with the developers.

VII. SUGGESTIONS

The Air Force should pay to have McDonnell Douglas provide a considerably expanded documentation for the ISS system, one that will provide some structure to the system as a whole. Then ISS should be taken out of the development environment so far as distribution, installation, failure management and update are concerned. Careful control of the systems on which ISS is installed must be used to insure that there is no tendency to try to run it on hardware that can't support it adequately.

ISS distributed materials should include:

- a trouble-shooting manual
- instructions on how to use the Ada debugger in the ISS context
- a good salvation program for files
- a file and disk directory structure verifier
- an automated file backup procedure
- an automatic error logging system.

The hardware should provide:

- automatic I/O retry
- trace traps
- time-out checking
- freedom from unwanted control-code interaction
- separate access to components for remote diagnostic programs
- reset and screen feedback capabilities
- ROM-based startup diagnostics
- Checksum, parity or CRC capabilities.

The software should be augmented to include:

- an expanded help package
- better error messages
- screen feedback
- trace capabilities.

There should be a separate ISS facility to handle consultation, remote diagnostics, distribution, installation and update.

Under no circumstances should ISS be installed in systems that require operating system modification to make the package run, or whose operating system is inadequate.

In the case that a major modification of ISS is undertaken, this should additionally provide for the use of a layered design with general error-handling routines, recovery block structure, careful replacement and statistics.

There should also be continued research for the purpose of further clarifying the general principles of fault-tolerance/diagnostics, providing a sound method for weighing cost versus worth in failure management, development of a design methodology for failure management, creation of templates for promotion of failure management in modular programming and evaluation of the prerequisites in the host environment for failure management.

REFERENCES

1. Avizienis, Algirdas, "Fault-Tolerant Systems", IEEE Transactions on Computers, vol. C-25, pp. 1304-1312, 1976.
2. Bowie, W. S. and J. G. Linders, "A Software Trace Facility", Software Practice and Experience, vol. 9, pp. 535-545, 1979.
3. Compton, Michael T., "Easing Fault Location in Large Systems", Communications of the ACM, vol. 23, pp. 440-442, 1980.
4. Denning, Peter J., "Fault-Tolerant Operating Systems", ACM Computing Surveys, vol. 8, pp. 359-390, 1976.
5. Goodenough, John B., "Exception Handling: Issues and a Proposed Notation", Communications of the ACM, vol. 18, pp. 683-696, 1975.
6. Horning, J. J., H. C. Lauer, P. M. Melliar Smith and B. Randell, "A Program Structure for Error Detection and Recovery", in Operating Systems (ed. Gelenbe & Kaiser), Lecture Notes in Computer Science, vol. 16, pp. 171-189, 1974.
7. Johnson, Mark S., "Translator Design to Support Run-Time Debugging", Software Practice and Experience, vol. 9, pp. 1035-1041, 1979.
8. Kopetz, H., Software Reliability, (Springer Verlag, New York), 1979.
9. Perrott, R. H. and A. K. Raja, "Quasi-Parallel Tracing", Software Practice and Experience, vol. 7, pp. 483-492, 1977.
10. Randell, Brian, "System Structure for Software Fault Tolerance", IEEE Transactions on Software Engineering, vol SE-1, pp. 220-232, 1975.
11. Randell, Brian, P. A. Lee and P. C. Treleaven, "Reliability Issues in Computing System Design", ACM Computing Surveys, vol. 10, pp. 123-166, 1978.
12. Rees, R. L. D. (ed.), Software Reliability, (Infotech, Maidenhead), 1977.
13. Verhofstad, Joost S. M., "Recovery Techniques for Database Systems", ACM Computing Surveys, vol. 10, pp. 167-196, 1978.
14. Wulf, William A., "Reliable Hardware/Software Architecture", IEEE Transactions on Software Engineering, vol. SE-1, pp. 223-240, 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

A THERMAL EVALUATION OF THE "LSSI" LIQUID-COOLED SYSTEM:

AN ENGINEERING PERSPECTIVE

Prepared by:	Dr. Amir Karimi
Academic Rank:	Assistant Professor
Department and University:	Division of Engineering University of Texas at San Antonio
Research Location:	School of Aerospace Medicine, Brooks Air Force Base
USAF Research:	Captain Charles A. Flick III
Date:	August 22, 1983
Contract No.	F49620-82-C-0035

A THERMAL EVALUATION OF THE "LSSI" LIQUID-COOLED
SYSTEM: AN ENGINEERING PERSPECTIVE

by
Amir Karimi

ABSTRACT

The thermal performance of each component of the "LSSI" (Life Support Systems, Inc.) liquid-cooled system has been investigated. The cooling system consists of a vest-hood arrangement, a heat exchanger unit, and a control display unit (housing a pump and a temperature control valve). This study covers the evaluation of the freezing time of the ice cartridge, determination of the thermal efficiencies of the vest and heat exchanger units, and evaluation of thermal gains of the working fluid from the environment. Comments are also made on the reliability of the control display unit components.

This study suggests that the vest, in its present configuration, is not satisfactory to meet demands of the groundcrew personnel. Several recommendations are offered to improve the present system. Additionally, it is suggested that further investigation be conducted to improve the vest and the heat exchange unit.

Acknowledgments

I 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 the opportunity for me to conduct summer research under a SCEE Faculty Research Fellowship at the School of Aerospace Medicine, Brooks Air Force Base.

I would also like to thank the staff of the School of Aerospace Medicine for creating an enjoyable atmosphere for carrying out this research. In particular, I would like to thank Dr. Wes Baumgardner for his support and encouragement; Captain Charles Flick for his valuable suggestions and discussions during the course of this research; and, last but not least, Mr. Chen for his invaluable help and assistance in conducting experiments.

This summer research opportunity was a rewarding experience for me and has served to stimulate ideas for additional studies.

I. INTRODUCTION:

To provide the USAF groundcrew personnel protection against chemical warfare agents, they must be placed in a completely impermeable environment. It is more effective and less costly to encapsulate one man (micro-environment) than to encapsulate the entire operational environment (macro-environment). For this purpose, in the event of chemical warfare, groundcrew personnel are expected to perform their combat duty wearing chemical defense attire.

While the recent design improvements have increased the chemical resistance of the chemical defense (CD) clothing, it has also placed additional thermal stress on the groundcrew personnel. Results show that the insulation and impermeable material used in the construction of the CD ensemble severely limit the body's evaporation/heat dissipation.

To alleviate the thermal stress, three (3) methods of heat removal have been suggested:¹

- a) evaporative cooling by circulating a flow of ambient air close to the body;
- b) convective cooling by circulating a flow of conditioned air close to the body; and
- c) conductive cooling by circulating a cold liquid through a vest placed in contact with the body.

The studies^{1,2} have shown that of the three (3) methods of cooling, the liquid cooling (conductive) is superior. In this method, the heat absorbed from the body must be stored in a "heat sink".

There are a number of heat sinks available which can be used with a liquid-cooled system. For example, a small refrigeration unit, operating from a vehicle power or an aircraft power source, can be used to cool the working fluid circulating in the system. The use of ice packs as heat sinks provides a novel "portable" cooling system

that can be used by groundcrew personnel who must move freely to perform their combat duties.

In general, the portable-cooling systems consist of a vest, a pump and a heat exchanger unit (HEU). The working fluid is pumped through the vest and removes heat from the body. The warm liquid then enters the heat exchanger unit where it comes into indirect contact with the ice (The warm liquid may be circulated through a bladder in contact with an ice cartridge or it may be passed through a tubing coil that is in contact with melting ice.).

II. OBJECTIVES:

In recent years the efforts of investigators in the U.S. Air Force and the Royal Air Force have been concentrated on evaluating the effect of cooling of humans by three (3) available portable cooling systems. These included the cooling system designed and manufactured by the Engineering Physics Department of the Royal Aircraft Establishment (RAE) in Farnborough, Hampshire; ILC Dover; and Life Support Systems, Inc. (LSSI), Mountain View, California. The majority of the studies conducted on these systems have dealt with the physiological effects of cooling on humans.

Scientists at the Defense and Civil Institute of Environmental Medicine (DCIEM) in Canada have rejected the use of the ILC cooling system as a viable source of cooling³ since "the unit could not be configured to work beneath the outer layers of clothing necessary for [man's] protection, next to the skin." In addition, the ILC ice cartridges cannot be recharged without removing the CD overgarment.⁴

Currently the USAF is in the process of reevaluating the LSSI liquid-cooled system, for design improvements. The objective of the present study is to conduct an engineering thermal evaluation of the LSSI

liquid-cooled system. In the present effort, no attempt was made to redesign the entire system, nor have we completed the thermal evaluation of the system in depth. Because of time limitations and the experimental equipment available to us at the time, we tried to keep our methods of evaluation as simple as possible. We left the detailed and complicated experiments and the task of evaluating the effectiveness of the alternate designs for future studies. In order to accomplish our objectives, the study was divided into the following tasks:

- (1) evaluation of the cooling capacity of the LSSI system;
- (2) evaluation of the freezing time of the ice cartridges;
- (3) determination of the thermal efficiencies of the vest and the heat exchanger unit (HEU);
- (4) comment on the reliability of the components (pump, temperature control valve) in the control display unit (CDU);
- (5) evaluation of the thermal gains of the working fluid from the environment; and
- (6) recommendations and suggestions for further studies.

III. A BRIEF DESCRIPTION OF THE LSSI LIQUID-COOLED SYSTEM:

This is a closed-loop system consisting of three (3) main units: heat source (the vest and hood); heat sink (heat exchanger unit); and the control display unit. The cooling system is shown in detail in Fig. 1⁵.

The vest consists of two layers of "flexitherm" heat exchanger panels which have been heat-sealed to form a 36 m channel network through which the working fluid circulates. The outer layer of the vest is covered by 0.5 cm of insulated quilted nylon^{5,6}. The vest covers approximately 0.209 m² of the body. A hood, made of the heat exchanger channels, is also attached to the vest. The hood covers approximately 0.056 m² of the heat surface.

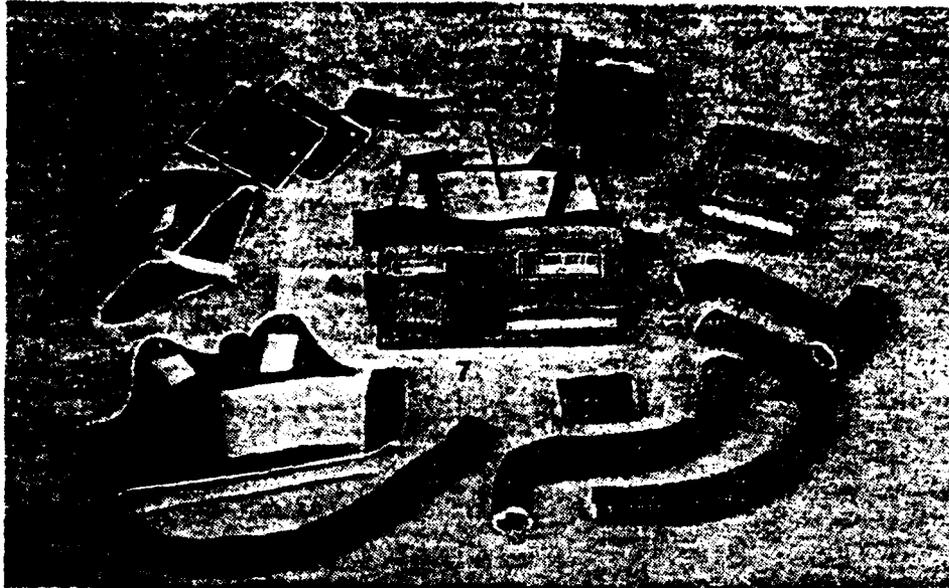


Fig. 1. The basic components of the complete system are the Control Display Unit (4), Heat Exchanger Unit (6), Power Supply (rechargeable) (7), The Garment (Headliner, vest, or heat-vest) (1 & 2), the Suspension System (belt or suspension vest) (3), and Cooling Cartridge (5).

Each heat exchanger unit contains a one-quart cartridge filled with a solution of 5% (by weight) Isopropyl alcohol and 95% (by weight) water. The frozen cartridge provides a heat sink which removes the heat (from the working fluid) stored in the vest. The cartridge's outer surface is in contact with a flexible heat exchanger bladder through which the working fluid flows.

The control display unit houses a pump which circulates the coolant fluid from a distensible reservoir through the vest. The coolant operating pressure (through the vest) is in the range 0.07-0.1 MPa

(10-15 psig). A magnetic clutch in the CDU disengages the pump when the operation pressure across the vest exceeds the 15 psig limit. This ceases the flow of coolant fluid into the vest.

A temperature-control valve is incorporated into the CDU to control the supply temperature of the vest. The valve, in theory, should regulate the amount of working fluid that passes through the HEU. The remaining portion of the flow bypasses the HEU. Therefore, when the valve is in the full open position, it provides the lowest possible temperature into the vest.

The system is powered by a 6-volt rechargeable battery. It provides two (2) hours of continuous operational power. Twelve (12) hours are required to fully recharge the battery.

The coolant fluid recommended by the manufacturer is a solution of 20% propylene glycol/80% water. A schematic description of the system's operation is sketched in Fig. 2.

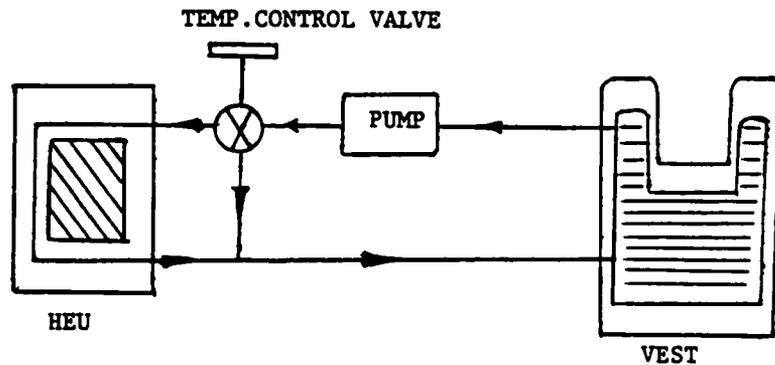


Fig. 2. A SCHEMATIC DESCRIPTION OF THE LSSI LIQUID-COOLED SYSTEM

IV. COOLING CAPACITY OF THE SYSTEM:

The cooling capacity of the system is determined by the amount of ice present in the heat exchanger unit. Each cartridge in the HEU contains approximately 1.05 kg of an isopropyl/water solution. When frozen and subcooled to a temperature (for illustration purposes) of -15°C , it has a potential of absorbing 484 kJ of heat if the temperature of the melt is to reach 25°C .

Heat can be added to the cartridge in three (3) modes (See Fig. 3).

- (1) To increase the temperature of the ice from -15°C to the melting point (-2°C), 29kJ of heat are needed.
- (2) The ice can absorb 341 kJ of heat (heat of fusion) to complete the melting process.
- (3) Finally, 115 kJ can be added to the melt to increase the temperature from the melting point to 30°C .

In practice, however, the cartridge will probably absorb 30 kJ of heat from the environment during the charging of the heat exchanger unit. Furthermore, when the entire ice in the cartridge is melted, the cooling system will not operate efficiently. Therefore, for all practical purposes each cartridge will provide 340 kJ of cooling.

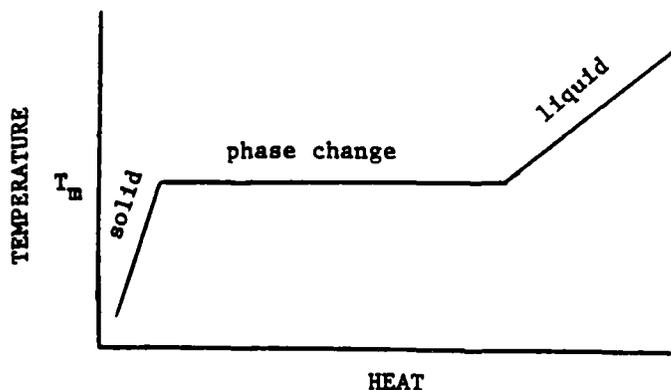


Fig. 3. A TYPICAL PHASE DIAGRAM

A C T I V I T Y	RANGE OF HEAT OUTPUT, W
1. Sedentary - Asleep, lying, relaxed, sitting at rest	82 - 117
2. Light Activity - writing, drafting, rapid typing	126 - 161
3. Light Work - slow walking, ironing	205 - 352
4. Medium to Heavy Work - swimming, cycling (10-12 mph), digging climbing stairs	440 - 703
5. Very Heavy Work - wrestling, running (8 mph)	879 - 1407

TABLE 1. TYPICAL HEAT OUTPUT OF HUMAN ADULTS⁶ - Typical heat output for young, healthy males of 150-lb. weight, according to their activities.

Considering that the groundcrew personnel perform duties that fall in the range of medium to heavy work, the system (with 2 HEU) will provide between 16 and 26 minutes of cooling. This determination is based on the assumption that the heat gained from the environment is minimal and, furthermore, that the system is highly efficient, allowing removal of all metabolically-generated heat. In practice, we find that the present system cannot remove heat at this rate.

V. TIME REQUIREMENT FOR FREEZING THE ICE CARTRIDGE:

The time required to fully freeze the ice cartridge (dimensions: 17.2 x 11.4 x 5.7 cm or 6-3/4 x 4-1/2 x 2-1/4 in.) creates a logistic problem in support and maintenance of the liquid-cooled system. Terrian and Nunneley³ have suggested that nine hours, at a temperature of -23° C, is required to freeze an ice cartridge.

When a liquid-filled cartridge is placed in a sub-freezing environment, a layer of ice will form on the inner surface of the ice cartridge. The temperature gradient between the liquid and the environment results in heat flow from the liquid towards the sub-freezing environment.

The removal of the latent heat of solidification, at the liquid-ice interface, is equal to the heat transferred to the environment if we consider the heat capacities of the container wall and the ice to be negligible and further assume the liquid to be at freezing temperature. The removal of latent heat results in the formation of continuous new layers of ice. This process proceeds inward until the entire cartridge is solidified.

A simple one-dimensional steady state heat flow model through composite walls can be employed to approximate the freezing time of an ice cartridge. A layer of ice formed against the wall of a container is shown in Fig. 4. The temperature distribution in the wall and through the ice layer is also illustrated in Fig. 4.

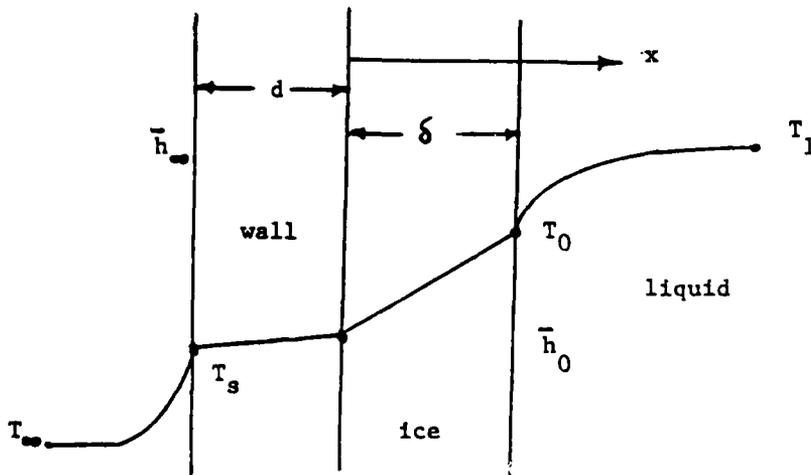


Fig. 4. TEMPERATURE DISTRIBUTION IN AN ICE-FORMATION PROCESS

If only the two (2) largest surface areas of the cartridge are introduced to subfreezing environment (insulating the other four surfaces) the freezing problem reduces to a one-dimensional heat flow problem.

To simplify the analysis further, let's assume that the outer surface of the cartridge is maintained at a constant temperature, T_s , and the liquid is at freezing temperature.

It can be shown that the time, t , required to form a layer of ice of thickness, δ , is represented by*:

$$t = \frac{\rho h_{if} k_i}{(T_o - T_s)} \left[\left(\frac{\delta}{k_i} \right) \left(\frac{d}{k_c} \right) + \frac{1}{2} \left(\frac{\delta}{k_i} \right)^2 \right] \quad (1)$$

where:

- ρ = the density of ice
- h_{if} = the latent heat of fusion
- k_i = the thermal conductivity of the ice
- k_c = the thermal conductivity of the container wall
- T_o = the freezing point temperature
- T_s = the surface temperature of the cartridge
- d = the wall thickness of the container

The characteristics of the ice cartridge (wall thickness, thermal conductivity), with $(T_o - T_s)$ as the parameter, was used to illustrate the time/ice thickness relationship of Equation (1). The results are presented in Fig. 5.

* The physical properties of the ice are considered to be uniform throughout the ice layer.

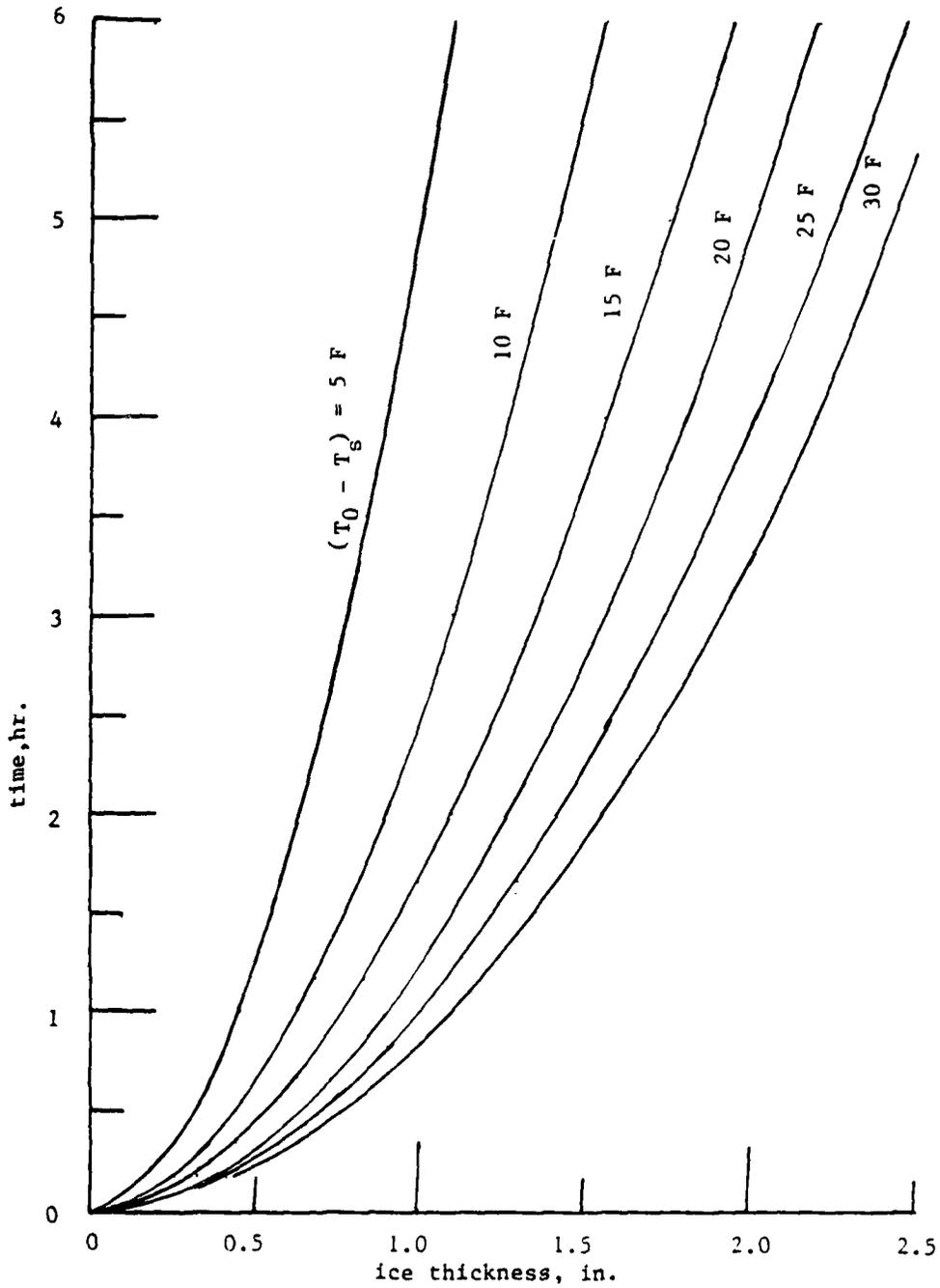


Fig. 5. TIME - ICE THICKNESS RELATIONSHIP IN A ONE-DIMENSIONAL FREEZING PROCESS [Equation (1)]

For a given temperature gradient ($T_o - T_s$), Fig. 5 displays a sharp increase in freezing time with an increase in ice thickness. Therefore, the use of thinner ice cartridges in the liquid-cooled system will reduce the logistic problems of freezing time. To achieve this result, either the HEU must be redesigned or a number of thin cartridges can be combined to accommodate the present HEU configurations.

VI. THERMAL EVALUATION OF THE VEST

Procedure: A large cylindrical tank was used as a heat source for thermal evaluation of the vest. The vest was placed around the tank and both the tank and the suit were thermally insulated from the environment. The surface of the tank was maintained at a uniform temperature by circulating warm water ($27 - 40^\circ \text{C}$) at a relatively high flow rate (6 l/min). Mean surface temperature was measured by means of three thermisters placed on various locations on the tank. To ensure adequate mixing, water was supplied through the bottom and outleted at the top of the tank.

A closed-loop system was used to supply the vest with liquid coolant. The system included the vest, a pump and the heat exchanger unit. A valve and a flow meter were placed in the loop to regulate the flow. Two thermisters were used to measure the inlet and the outlet temperatures of the coolant.

From the first law of thermodynamics for open systems, the rate of heat removal by the vest was calculated, using the relation:

$$\dot{q} = \dot{m} c_p (T_{\text{out}} - T_{\text{in}}) + m c_p \frac{dT}{dt} \quad (2)$$

where

- \dot{q} = the rate of heat removal
- \dot{m} = mass flow rate

c_p = specific heat of the coolant
 T_{in}, T_{out} = inlet and outlet temperatures, respectively
 m = mass of coolant within the vest
 dT/dt = variation of average temperature with time

Results: For various coolant flow rates, the observed values for the rate of heat removal are plotted in Fig. 6 against the temperature gradient, $(T_s - \bar{T})$, where \bar{T} and T_s are the average coolant and surface temperatures, respectively. The results indicate that the rate of heat removal on these coordinates is virtually independent of the coolant flow rate. They also suggest a linear dependence of the heat removal rate, \dot{q} , on the temperature gradient $(T_s - \bar{T})$. The maximum rate of heat removal (116 W) by the vest corresponds to an average coolant temperature of 22° C and a surface temperature of 40.4° C.

Fig. 7 exhibits the behavior of the observed data on $\dot{q} - (T_s - T_{in})$ coordinates. Again it indicates that the rate of heat removal is independent of the coolant flow rate. In order to improve the rate of heat removal of the vest, Fig. 7 suggests either increasing T_s or decreasing T_{in} . There are, however, limitations imposed on these two temperatures in practical situations.

The minimum possible operational temperature for T_{in} is the melting temperature of ice in the HEU. (This corresponds to a temperature of -2° C if direct mixing of the coolant fluid and the ice occurs in the HEU.) In the present configuration, the inlet temperature for the "LSSI" vest is always much greater than this value. Harrison and Belyavin⁷ have argued that in human subjects when the skin temperature drops below 30° C, the blood vessels contract, reducing the blood circulation, and hence reducing the rate of heat removal from the body's core. Thus, to maximize the coolant effect, the skin temperature should be maintained in the range of 30-37° C. In

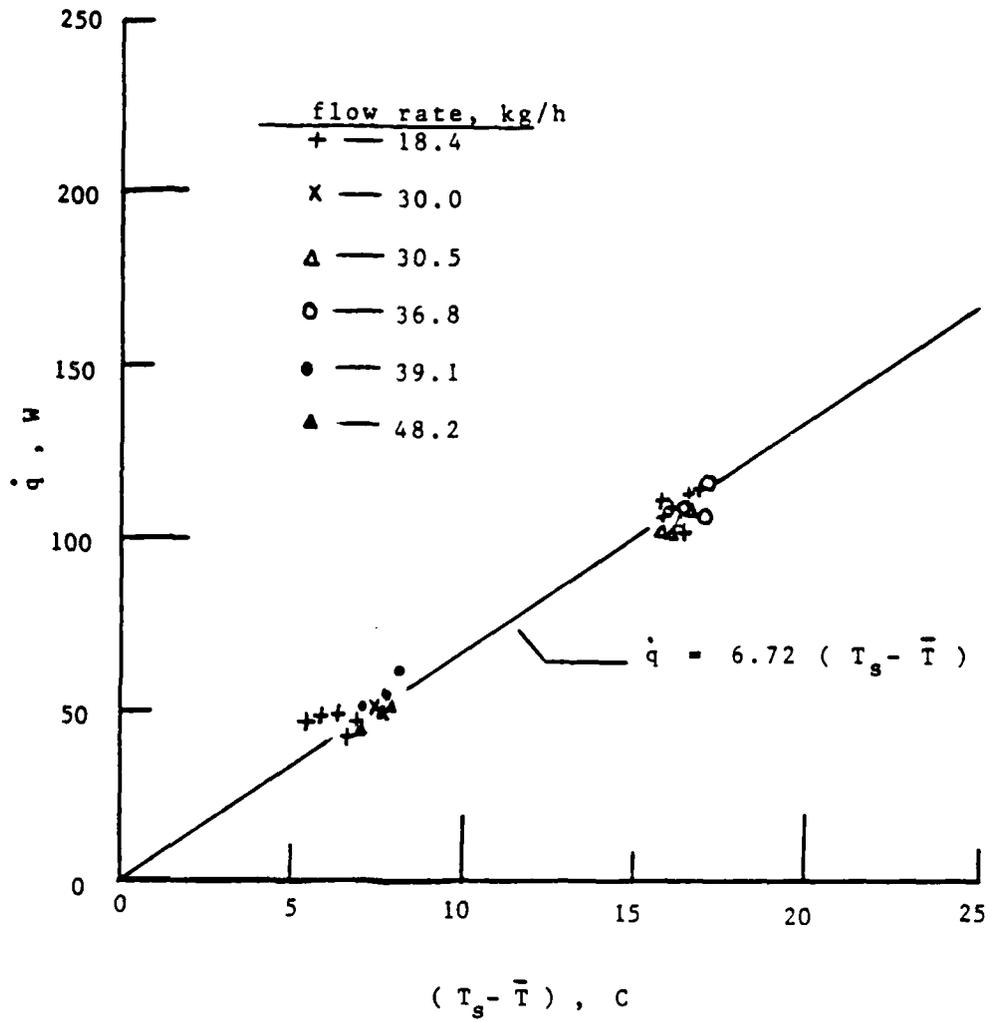


Fig. 6. RATE OF HEAT REMOVAL BY THE VEST VERSUS TEMPERATURE GRADIENT $(T_s - \bar{T})$

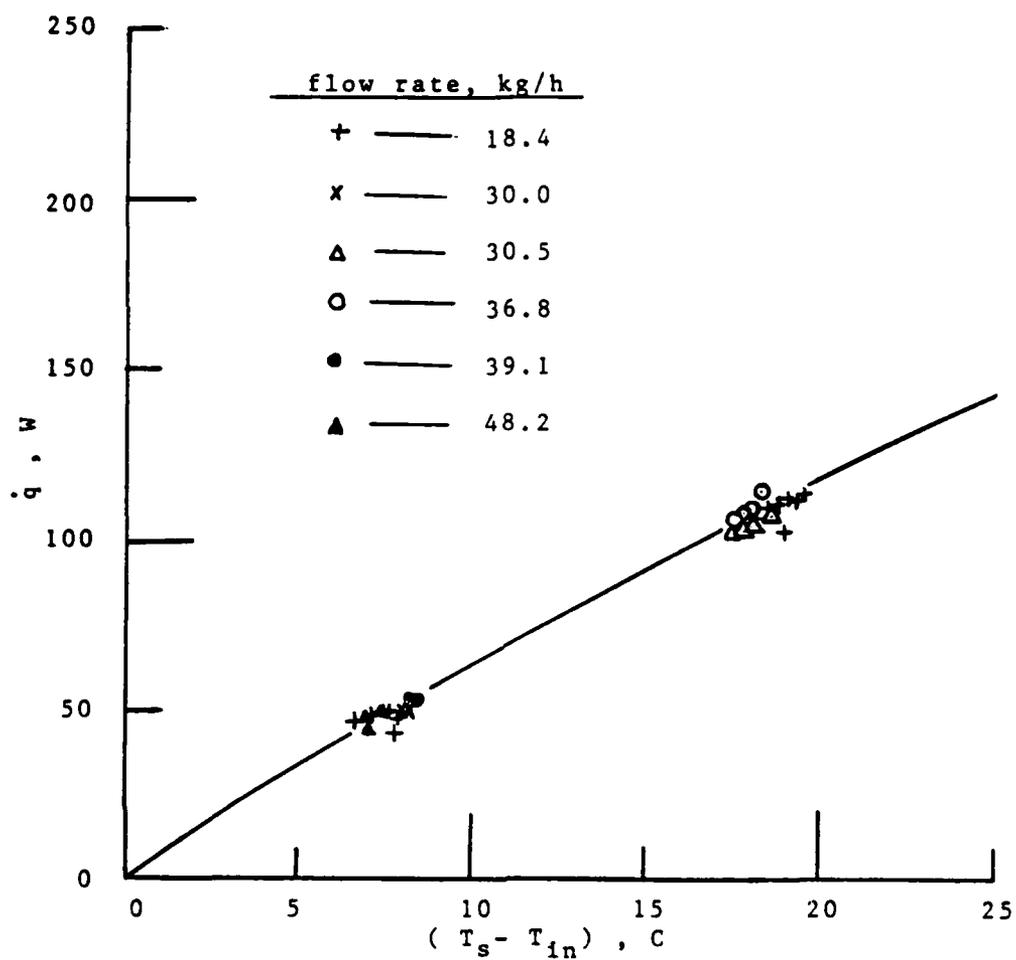


Fig. 7. RATE OF HEAT REMOVAL BY THE VEST VERSUS TEMPERATURE GRADIENT ($T_s - T_{in}$)

general, the skin temperature is influenced by both the coolant temperature and the rate of metabolic heat generated inside the body. In this work, we did not attempt to determine the relationship which might exist among the above factors. This task has been left to future studies.

It may be concluded from the above discussion that it is unlikely that the present system operates at a temperature gradient ($T_s - \bar{T}$), over 25° C. Hence, Fig. 6 suggests that the vest (without the hood) is capable of removing only up to 170 Watts of generated body heat. To improve the rate of heat removal, the area of contact between the tubing in the vest and the skin, must be increased.

VII. THERMAL EVALUATION OF THE HEU:

Procedure: A closed-loop system, shown in Fig. 8, was designed to evaluate the efficiency of the heat exchanger unit.

The system includes one heat exchanger unit, the control display,

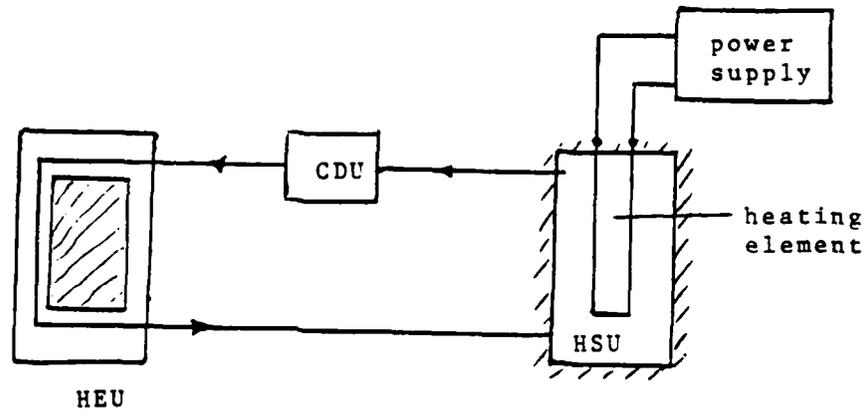


Fig. 8. SCHEMATIC DESCRIPTION OF EXPERIMENTAL SYSTEM USED IN THE THERMAL EVALUATION OF THE HEU

and a heat source unit.* The heat source unit consisted of a small, cylindrical container, into which an electrically-heated element was inserted. The entire heat source unit was thermally insulated from the environment. Two thermisters, located near the inlet and outlet of the HEU, provided the measure for the average temperature of the working fluid. The fluid was circulated through the system by the pump, located in the CDU (control display unit).

Our experimental system limited the HSU heat input to a maximum of 200 W. This input is insufficient to effectively evaluate the thermal efficiency of two heat exchanger units. Therefore, our evaluation of the system was based on a single HEU.

An energy balance on the closed system determined the rate of heat removal in the HEU. The following relation was used to evaluate the rate of heat removal:

$$\dot{q}_{HEU} = \dot{q}_{HSU} - m c_p \frac{dT}{dt} \quad (3)$$

where

- \dot{q}_{HEU} = the rate of heat removal by HEU
- \dot{q}_{HSU} = the rate of heat input in the HSU (heat source unit)
- m = the mass within the system
- c_p = specific heat
- dT/dt = variation of the average temperature with time

Results: The observed rates of heat removal, \dot{q}_{HEU} , for a single heat exchanger unit, as a function of time, with heat input \dot{q}_{HSU} as parameters are plotted in Fig. 9. This figure indicates that a cooling system with a single heat exchanger unit can completely remove the heat gain in the vest if this gain is below 100 Watts.

* The heat source unit was designed by Mr. Chen and constructed at The School of Aerospace Medicine/Brooks Air Force Base.

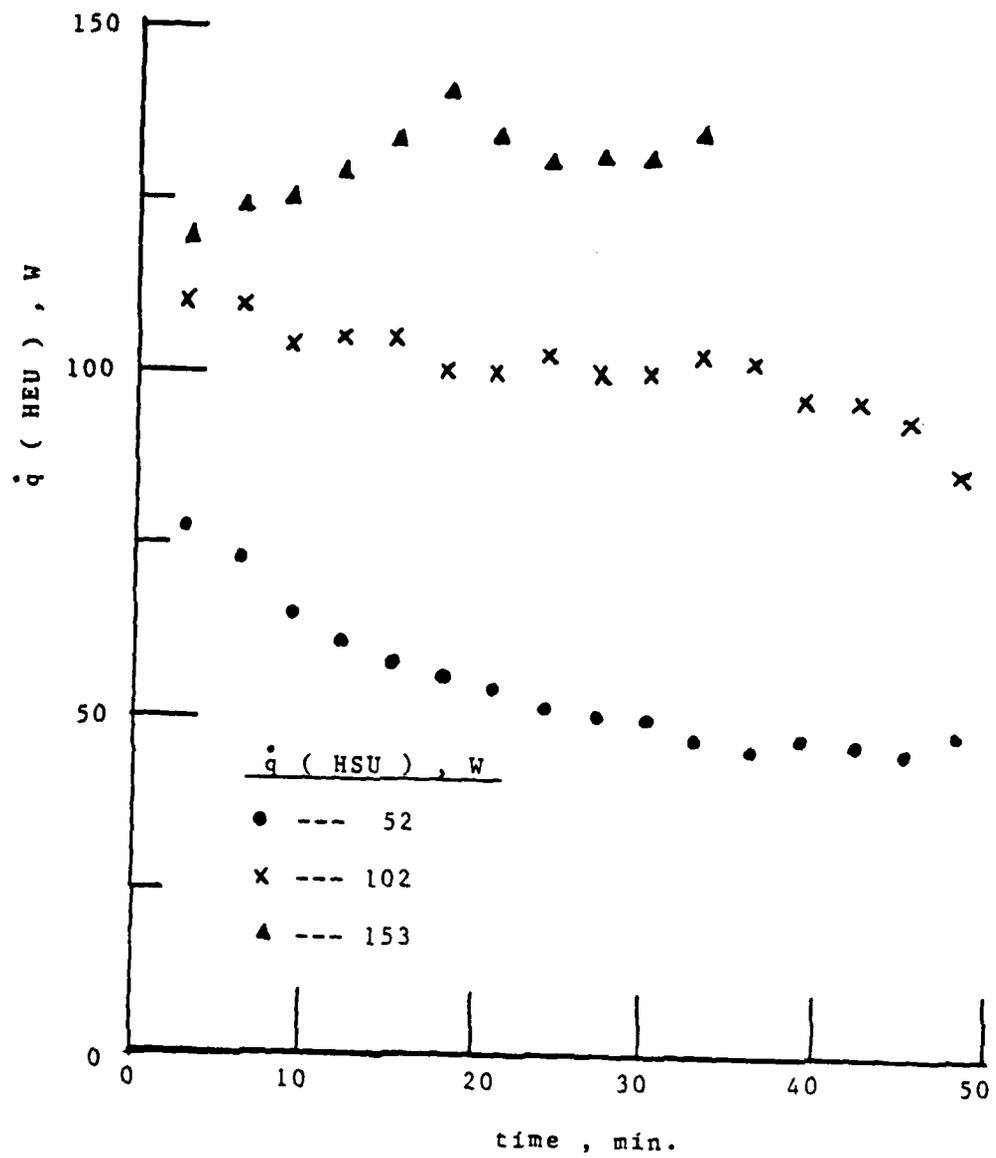


Fig. 9. RATE OF HEAT REMOVAL BY THE HEU

For values above 100 W, only partial removal is possible. This results in an increase in the average temperature of the coolant fluid and thus reduces the heat absorbed in the vest.

One problem that hinders the performance of the heat exchanger unit is the melting process in the ice cartridge. As the melting process proceeds, a liquid forms between the ice and the cartridge wall. Because of the low thermal conductivity, the melted ice acts as thermal resistance to heat transfer. Therefore, for a constant inlet temperature, the rate of heat removal in the HEU decreases with time.

A simple heat conduction solution, such as that obtained for the freezing process (see Section V), does not provide an accurate means for evaluating the heat transfer rate. In this case, such factors as natural convection between the liquid-solid interface and the degree of agitation become important considerations. Therefore, further studies must address themselves to such problems.

VIII. EVALUATION OF THE THERMAL GAINS OF THE WORKING FLUID FROM THE ENVIRONMENT

Assuming that chemical defense attire are perfect insulators, the coolant within the vest should have no thermal gains from the environment. We therefore attempted to evaluate the thermal gains through those components of the system exposed to the environment. The components tested included the HEU (excluding the ice cartridge), the CDU, and tubing and connectors exposed to the atmosphere. These items were arranged to form a closed-loop system. The pump in the HEU circulated the working fluid throughout the system.

Two models were tested. In one model the components were exposed to the atmosphere; in the other, they were enclosed by a special covering that prevents penetration of chemical agents into the system. Both models were placed in thermal heat chambers.

The thermal gains of the working fluid in both models were calculated using the relation:

$$\dot{q} = m c_p \frac{dT}{dt} \quad (4)$$

where

m = the mass of liquid within the system

c_p = specific heat of the working fluid

dT/dt = the variation of liquid temperature with time

The results are presented in Fig. 10 on $\dot{q} - (T_{amb} - T_{liq})$ coordinates, where T_{amb} and T_{liq} denote the environmental and the liquid temperatures respectively.

The data for the older model is presented with open symbols and that for the new system is shown with solid symbols. Fig. 10 exhibits no evidence to indicate that one model is superior to the other. It does indicate, however, that the thermal gains are influenced by the relative humidity of the environment. In fact,

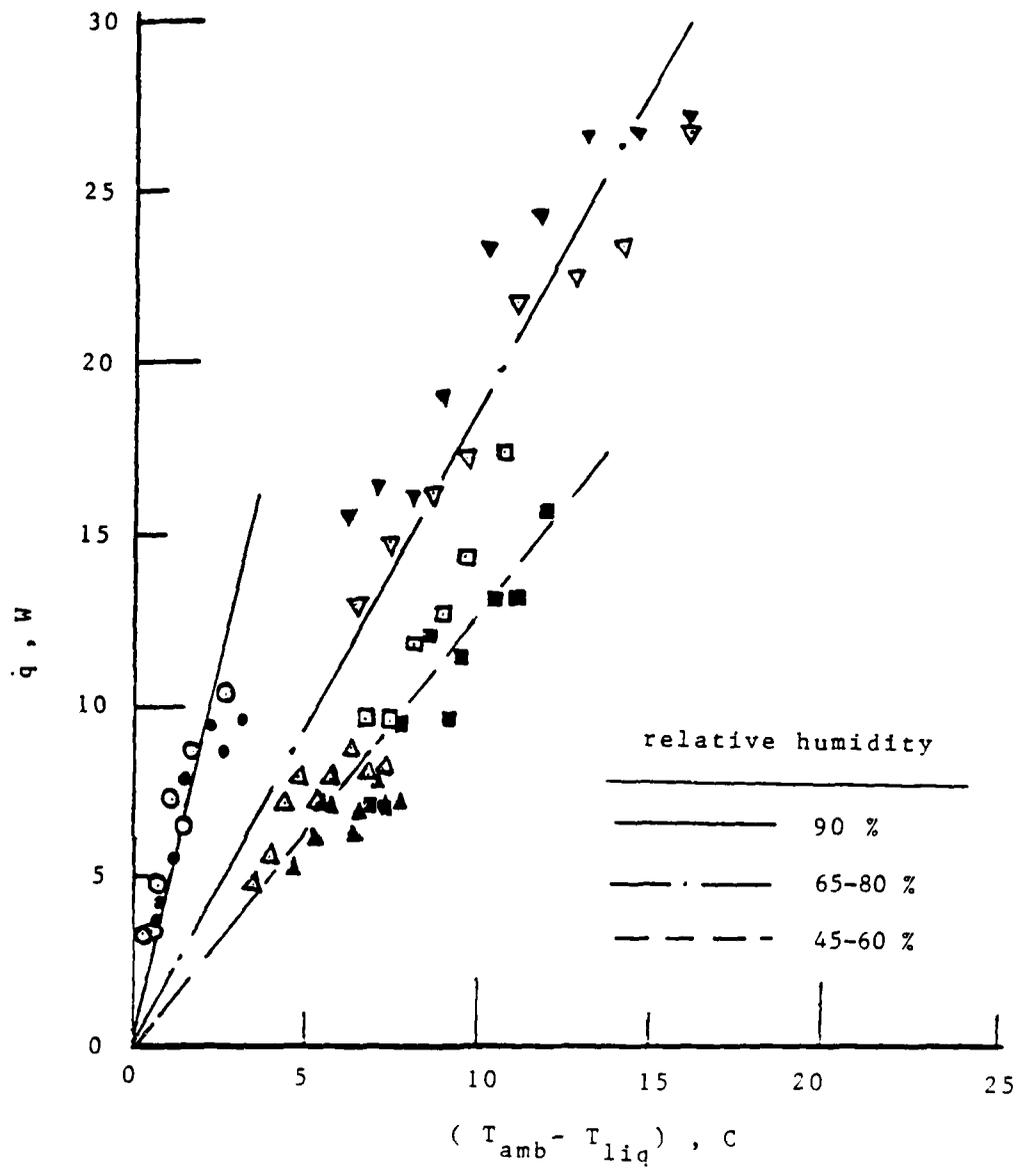


Fig. 10. COOLANT'S ENVIRONMENTAL THERMAL GAINS

during the experiments condensation on the tubing and connectors was observed when the relative humidity in the thermal chamber exceeded 90%. The heat released by condensation would be absorbed by the working fluid. Although the thermal gains from the environment decrease the thermal capacity of the cooling system by less than 5%, the insulation of the exposed components (particularly the connectors) would be an improvement.

IX. RELIABILITY OF THE CDU COMPONENTS

During the course of our experiments, the following observations were made:

Pump: When the operating pressure within the vest exceeds the limit of 15 psig (due to kinks or bends in the tubing), the magnetic clutch provided in the CDU disengages and causes the pump to terminate the circulation of the coolant. This causes a logistics problem because there is no warning to the user that coolant circulation has ceased, other than the sensation of increased heat.

Temperature Control Valve: Two models were tested by placing a flow meter between the CDU and HEU. In the older model the flow indicator dial which regulates the flow into the HEU had a 120° range between the fully closed and fully open position. (0° referring to fully closed and 120° referring to fully open position) Between 0° and 65°, we did not observe any flow into the HEU. At 95°, the maximum flow position had already been reached. In short, the range of control was actually about 30° of rotation on the valve.

The indicator on the newer model, which was a modification of the above, displayed a range of 65° (rather than 120°) between the fully closed and fully open positions. By turning the flow indicator dial, the flow meter indicated either no flow or maximum flow, with no intermediate flow. Therefore, the valve on these units should be modified to allow the user to regulate the temperature.

Battery: After two hours of operation, the coolant flow rate dropped

by 14%, indicating a decrease in power output by the battery.

X. RECOMMENDATIONS

The results of this study suggest two sets of recommendations be pursued. The first deals with design improvements of the existing LSSI system and the second requires further investigation prior to implementation of the system.

1. Design improvements of the present LSSI unit -
 - a. Thinner ice cartridges to reduce the freezing time. This requires design modification of the HEU.
 - b. A warning system to be incorporated into the CDU to alert the user to failure of the system supplying coolant to the vest.
 - c. The temperature control valve to be replaced with one that provides the user a wider range of control.
 - d. The exposed tubing and connectors to be insulated, thereby reducing environmental thermal gains by the working fluid.
 - e. A single strap securing the vest, to be replaced by multiple straps, thus providing a snugger fit of the vest.
2. Further investigation is suggested in the following areas -
 - a. We have shown that the vest in the present configuration cannot provide adequate cooling in the range of medium to heavy work (Table 1). Therefore, further investigation in this area is necessary.

It has been established that the liquid condition cooling is more effective in the proximity of large blood vessels (neck and head).^{4,7} To identify such areas, it is suggested that cooling liquid units be placed at various locations on the body. It could then be determined which areas provide the greatest rate of heat removal.

Further studies should be made to determine the effect of the following on the cooling rates: inlet temperature, skin

temperature and surface areas of the vest in direct contact with the skin.

- b. We suggested that the efficiency of the HEU for heat removal decreases with time because of the melting process occurring in the ice cartridge. Further studies should be conducted to establish the relationship that might exist among the following parameters: the rate of heat removal, inlet temperature, degree of agitation, time and amount of ice remaining in the ice cartridge.

By following the above recommendations it should be possible to implement compatible improvements in the vest and the HEU, thereby creating an optimal working system. It should be reiterated that the heat absorbed in the vest must be completely dissipated in the HEU to ensure an ideal system.

R E F E R E N C E S

1. Vincent D. Iacono et. al. "Development of a Microclimate Cooling System for Combat Vehicles," United States Army Natick Research & Development Laboratories, Natick, Massachusetts, February, 1982.
2. George C. Rannenberg, "Study to Determine the CB Threat and Define Alternative Crew Protection Systems for the Advance Attack Helicopter"(AAH) (Vol. 1), United States Army Natick Research and Development Laboratories, Natick, Massachusetts, June, 1981.
3. S. Livingston et. al. "The Effect of Body Cooling on Aircrew in Hot Humid Environments," (Abstract), ASCC Working Party 61 Symposium, November, 1979.
4. D.M. Terrian and S. A. Nunneley, "A Laboratory Comparison of Portable Cooling Systems for Workers Exposed to Two Levels of Heat Stress," USAFSAM Technical Report, 1982.
5. "Cool Head TM Operation and Maintenance Instructions" (Mobile Configuration), Document Number 900109E, August, 1982.
6. P. Webb, Editor. "Biastronautics Data Book", NASA SP 3006, National Aeronautics and Space Administration, 1964.
7. M.H. Harrison and A.J. Belyavin, "Operational Characteristics of Liquid-Conditioned Suits", Aviation, Space, and Environmental Medicine, August, 1978, pp. 994-1003.

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 STATISTICAL METHOD FOR THE SERIAL COMPARISON OF

VECTOR CARIOGRAMS

Prepared by:	Dr. Jerome P. Keating
Academic Rank:	Assistant Professor
Department and University:	Division of Mathematics University of Texas at San Antonio
Research Location:	USAF School of Aerospace Medicine Biomathematics Modeling Branch Brooks AFB, Texas
USAF Researcher:	Dr. James T. Riley
Date:	August 19, 1983
Contract No:	F49620-82-C-0035

A STATISTICAL METHOD FOR THE SERIAL
COMPARISON OF VECTOR CARDIOGRAMS

by

Jerome P. Keating

ABSTRACT

In the absence of clinical standards to detect serial changes in cardiograms, statistical procedures are proposed as an alternative. These procedures are all preceded by a dimension reducing orthonormal transformation of the original digitized cardiogram into a lower dimensional feature space. In feature space, multivariate test criteria are given for the detection of changes in covariance matrices or mean vectors of the cardiograms. The flexibility is provided to compare the cardiograms of the same individual pairwise or simultaneously. In addition, a peripheral test of independence of beats within a cardiogram is given. Some pertinent remarks are also made about controlling the overall level of significance and its impact on the application of these techniques to cardiograms of USAF pilots.

ACKNOWLEDGEMENT

The author thanks the United States Air Force School of Aerospace Medicine, the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for the opportunity to participate in the Summer Faculty Research Program. He acknowledges Dr. James T. Riley for suggesting this area of research.

I. Introduction

For the last decade the School of Aerospace Medicine (SAM) at Brooks AFB has been involved with programs for the automated processing of vector cardiograms (VCG's). Recently the School has taken an interest in the serial comparison of VCG's. In this effort, cardiologists hope to develop an automated process which would detect serial changes in cardiograms (i.e., is there a significant clinical difference between the initial and current cardiogram of an individual pilot?). From Tolan (1982) it is clear that cardiologists have great difficulty in quantifying clinical standards for serial changes in cardiograms. Thus, the use of statistical methods for detection of serial changes is an alternative approach in the absence of clinical standards.

In the statistical sense, a serial change in two cardiograms will be present if there is a statistically significant difference in the individual covariance matrices or subsequently the mean vectors of the respective cardiograms. The statistical tests of hypotheses for equality of covariance matrices and mean vectors presented in this paper are based on certain intrinsic assumptions, which should be listed for those wishing to use this approach.

- 1) Rate compensation of cardiograms to a common dimension does not distort or lose pertinent statistical information.
- 2) The distribution of a single beat from a cardiogram has an approximate multivariate normal distribution.
- 3) Consecutive beats within a cardiogram may be effectively treated as a random sample.

Violation of assumption 2 can be tested statistically by methods given in Malkovich and Afifi (1973). Since the population of interest in this study consists of 40,000 almost exclusively very healthy USAF pilots, the practical observations of Day and Kerridge (1967) seem appropriate. They note that in medical applications variables are often approximately normally distributed in control patients but definitely skew in the abnormal group. Although this note does not imply that the detection of nonnormality in a variable is cause for clinical concern it does support some belief that the population of cardiograms of healthy pilots will reasonably approximate a multivariate normal distribution. A test for multivariate normality is not conducted on each sample but the robustness of all subsequent tests to nonnormality is addressed for completeness.

Violation of assumption 3 can also be tested statistically and is given in section IV. There is no statistical significance test for the violation of assumption 1. However, there is medical evidence that both the duration of the QRS complex and the locus of the QRS complex are heart rate invariant. These results are established by Simoons and Hugenholtz (1975) and Rataharju et al (1973). Thus the resolution of the digitizations of all cardiograms to the same dimension is crucial and reasonable in this multivariate approach.

In section III some practical issues are given for any method which uses the Karhunen-Loève expansion (or principal components) as a method of dimension reduction. The test for homogeneity of covariance matrices is given in section V and the conditional test of equality of

mean vectors, given that homogeneity is present, is presented in section VI. Finally the computation of exact critical values of Wilks' Λ distribution is given in section VII due to unusual levels of significance, small sample size and large dimension. Recommendations for further studies in the analysis of cardiograms are given in section VIII.

II. Objectives

The objectives of this paper are the presentations of exact statistical tests for three multivariate hypotheses. The three tests are:

- 1) a test of independence of rate compensated heartbeats within a record;
- 2) a test of equality of covariance matrices from two different cardiograms;
- 3) a conditional test of equality of mean vectors from two different cardiograms given that homogeneity of covariance matrices is present.

The aspects of power and robustness to nonnormality of the tests in 2) and 3) are investigated.

III. The Karhunen-Loève Expansion as a Technique in Dimension Reduction of VCG's

The process of applying the Karhunen-Loève expansion as a method of dimension reduction in VCG's/ECG's has been a recurring topic in the literature in recent years. Young and Huggins (1963), Womble et al (1976, 1977), Tolan et al (1981), Hsu et al (1981) and Tolan (1982) have all applied the Karhunen-Loève expansion in the presence of an

estimated training set covariance matrix, S . Since S is the basis from which the eigenvalues and eigenvectors of Σ are estimated, it is crucial for n , the sample size to be sufficiently large so that the asymptotic properties of S are applicable. Thus the following question naturally arises. How large must n be for the asymptotics of S to hold? The answer to this question is predicated on which asymptotic property is being met.

Suppose $\underline{X}_i \sim N_p(\underline{\mu}, \Sigma)$ are independent for $i = 1, \dots, n$. $N_p(\underline{\mu}, \Sigma)$ denotes a p -dimensional random vector having a multivariate normal distribution with mean vector, $\underline{\mu}$, and covariance matrix, Σ . The maximum likelihood estimator (mle) of Σ is S where

$$nS = \sum_{i=1}^n (\underline{X}_i - \bar{\underline{X}})(\underline{X}_i - \bar{\underline{X}})' \quad \text{and} \quad \bar{\underline{X}} = \sum_{i=1}^n \underline{X}_i / n. \quad (1)$$

It is well-known (see Anderson (1958)) that $nS \sim W_p(\Sigma, n-1)$, where $W_p(\Sigma, n-1)$ denotes the distribution of a random $p \times p$ matrix having a p -dimensional Wishart distribution with covariance matrix, Σ , and degrees of freedom, $n-1$.

In most advanced textbooks on Multivariate Analysis (e.g., Srivastava and Khatri 1979) one can find the following distributional result that

$$|nS| \sim |\Sigma| \prod_{i=1}^p u_i \quad (2)$$

where the u_i 's are independently distributed χ^2 random variables with degrees of freedom $n-i$. Since the u_i 's are independent then it follows that

$$E(|nS|) = |\Sigma| \prod_{i=1}^p E(u_i) = |\Sigma| \prod_{i=1}^p (n-i).$$

Taking logarithms and rearranging terms yields

$$\ln[E(|S|)/|\Sigma|] = \sum_{i=1}^p \ln((n-i)/n) \leq - \sum_{i=1}^p i/n = -p(p+1)/2n.$$

Equivalently,

$$E(|S|) \leq |\Sigma| \exp(-p(p+1)/2n). \quad (3)$$

Using this inequality, the $|S|$ (in expectation) begins to approach $|\Sigma|$ when $n \gg p(p+1)/2$. This result is sensible because $p(p+1)/2$ is precisely the number of unknown parameters in the covariance matrix itself.

For the purposes of this VCG study, the dimension p is on the order of 750. Thus from (3) it follows that

$$n \gg 750(751)/2 = 281,625. \quad (4)$$

For the $|S|$ to be approximately 95% of $|\Sigma|$ requires that

$$n \doteq 20(281,625) = 5,632,000. \quad (5)$$

From the size of the numbers in (4) and (5) it seems absurd to assume that this asymptotic property will hold for the covariance matrix of the training set based on 10,000 VCG's. Notice that in this case

$$E(|S|) \leq |\Sigma| \exp(-28.1625) = 5.9 \times 10^{-13} |\Sigma|. \quad (6)$$

These results are no doubt related to the entropy maximizing property (see Keating et al (1983)) of the KL and they reveal that the estimated KL may be in tremendous violation of this property. However, the KL transformation matrix, Φ , derived from an estimated covariance matrix, S , is still orthonormal (i.e., $\Phi\Phi' = I_r$ but $\Phi' \neq I_p$).

Some preliminary simulation studies conducted by Dr. James T. Riley indicate that the mean squared error minimization property of the KL

may in fact be valid for significantly smaller values of n than those given in (4) and (5). Thus all asymptotic properties do not require numerical values as large as (4) and (5).

The previous remarks highlight the importance of accurate estimation of Σ when n is not sufficiently large. In a recent lecture (see Lin and Perlman 1983), Perlman presented results on comparison of estimators of Σ . In this lecture he heavily referenced the Seattle Lectures (1982) and Reitz Lectures (1975) for which we have been unable to find published proceedings.

It is well-known that the sample covariance matrix, S , can be extremely sensitive to atypical observations (outliers). Atypical observations from multivariate data sets can be treated through robust M-estimators of the mean vectors and covariance matrices. From Campbell (1980) the robust estimators of the mean vectors and covariance matrices are as follows:

$$\tilde{X} = \frac{\sum_{i=1}^n w_i X_i}{\sum_{i=1}^n w_i} \quad \text{and} \quad \tilde{S} = \frac{\sum_{i=1}^n w_i^2 (X_i - \tilde{X})(X_i - \tilde{X})'}{\left(\sum_{i=1}^n w_i^2 - 1\right)}, \quad (7)$$

where $w_i = \phi(d_i)/d_i$ and $d_i^2 = (X_i - \tilde{X})' \tilde{S}^{-1} (X_i - \tilde{X})$. ϕ can be any one of a set of bounded influence functions (see, Hampel 1974) whose argument is based on the Mahalanobis distance of the i th observation. \tilde{X} and \tilde{S} in (7) are found iteratively.

Huber (1964) and Hampel (1973) have given very popular and practically useful forms for the influence function. Maronna (1976) has established the results on the existence and uniqueness of M-estimators of the covariance matrix in addition to the properties of consistency and asymptotic normality. Huber (1977) gives a thorough discussion of

robust covariance estimation. The M-estimators can also be used in the sense of robustness to nonnormality by first transforming the multivariate data to achieve approximate symmetry (see Gnanadesikan 1977) and then applying the robust M-estimator procedure.

The estimated KL transformation matrix, based on S or S , still satisfies the property of orthonormality. In dimension reduction it seems reasonable to expect more of a dimension reducing transformation than just the property of orthonormality even in the estimated case. Recent works on the topic of dimension reduction are given in Krishnaiah and Kanai (1982).

In the subsequent tests for serial changes, under an orthonormal transformation, there are three statistical tests to perform, one each for independence, homogeneity of covariances and equality of mean vectors given that homogeneity is present. Let α_1 , α_2 and α_3 be the levels of significance of the respective tests. The overall level of significance, α , can be bounded by Bonferroni's inequality (see e.g. Johnson and Wichern (1982)).

$$P[\text{all null hypotheses are true}] \geq 1 - (\alpha_1 + \alpha_2 + \alpha_3). \quad (8)$$

Actually in this case a tighter bound can be obtained for α than $\alpha \leq \alpha_1 + \alpha_2 + \alpha_3$ since the test for homogeneity of covariance matrices and the test for equality of mean vectors (given that covariance matrices are equal) are actually independent (see Anderson 1958). Thus $\alpha \leq \alpha_1 + \alpha_2^*$ where $\alpha_2^* = 1 - (1 - \alpha_2)(1 - \alpha_3)$.

Tolan (1982) warns that for the purposes of clinical practice one should maintain a relatively small value of α (i.e., $\alpha = .001$). He explains that the USAF has 40,000 pilots (a population which is extremely

healthy) and at an α level of .05, 2,000 aviators would be cited for false positive serial changes. Since catherization is a distinct possibility for an individual cited for serial cardiogram change, the cardiologist would become skeptical of repeatedly performing an unnecessary and somewhat dangerous test. Moreover, if the pilot is falsely diagnosed as having a serial change his career can be adversely effected until the discrepancy is resolved.

In the sections that follow a test for homogeneity of two covariance matrices and subsequently for k covariance matrices is given. The first test allows the clinician to compare the covariance matrices from two different cardiograms (e.g., the first and the last) whereas the second test of homogeneity allows one to compare the current VCG to all previous cardiograms. It is apparent from Tolan (1982) that the flexibility to perform either test is desirable from the clinician's viewpoint. Likewise, the capability is provided to test the equality of two mean vectors given that homogeneity of covariance matrices is present. Differences in mean vectors would also be indicative of serial change. This test is also extended to test the equality of k mean vectors in light of Tolan's (1982) remarks.

IV. Independence

Consider the consecutive beats X_1, \dots, X_{2n} to be the digitizations of the first $2n$ beats taken from a single VCG record. If the data pass a hypotheses test for independence then the recordings of consecutive beats can be treated as a random sample. However, it has been suggested that since the recorded beats are recorded consecutively a time dependence

will exist between consecutive beats and $\underline{X}_1, \dots, \underline{X}_{2n}$ will not form an independent set of random vectors.

Assume that $\underline{X}_i \sim N_p(\underline{\mu}, \Sigma)$, for $i = 1, \dots, 2n$. Under the null hypotheses, H_0 , $\underline{X}_1, \dots, \underline{X}_{2n}$ are independent whereas under the alternative, H_1 , they are dependent. Since the tests for independence involve an estimate S of Σ , which can be guaranteed to be nonsingular if $n > p$, the test of the null hypotheses in p dimensional pattern space would require approximately 13 minutes of cardiogram information just to be able to run the test. Consequently, an orthonormal dimension reduction technique is employed. Let T be an $r \times p$ transformation matrix such that $TT' = I_r$ but $T'T \neq I_p$. This generalization away from the Karhunen-Loève transformation is in keeping with the findings of section III. However, all the results presented here will be applicable to the KL expansion as a special case.

Consider the vectors $\underline{Y}_i = T\underline{X}_i$ for $i = 1, \dots, 2n$. Clearly, it follows $\underline{Y}_i \sim N_r(T\underline{\mu}, T\Sigma T')$. Under H_0 \underline{X}_i and \underline{X}_j are independent for $i \neq j$. Consequently

$$\text{Cov}(\underline{Y}_i, \underline{Y}_j) = T E[(\underline{X}_i - \underline{\mu})(\underline{X}_j - \underline{\mu})'] T' = T \text{Cov}(\underline{X}_i, \underline{X}_j) T' = 0. \quad (9)$$

Hence the independence of $\underline{X}_1, \dots, \underline{X}_{2n}$ implies the independence of $\underline{Y}_1, \dots, \underline{Y}_{2n}$. Since it is impractical to conduct the test of independence on the p -dimensional \underline{X}_i 's, we shall conduct the test on the r -dimensional \underline{Y}_i 's.

In the ordering of the beats of the cardiogram it is anticipated that the greatest dependency with \underline{Y}_i will come from \underline{Y}_{i-1} and \underline{Y}_{i+1} . Thus, separate $\underline{Y}_1, \dots, \underline{Y}_{2n}$ into two sets. $\underline{Z}_1^{(1)}, \dots, \underline{Z}_n^{(1)}$ and

healthy) and at an α level of .05, 2,000 aviators would be cited for false positive serial changes. Since catheterization is a distinct possibility for an individual cited for serial cardiogram change, the cardiologist would become skeptical of repeatedly performing an unnecessary and somewhat dangerous test. Moreover, if the pilot is falsely diagnosed as having a serial change his career can be adversely effected until the discrepancy is resolved.

In the sections that follow a test for homogeneity of two covariance matrices and subsequently for k covariance matrices is given. The first test allows the clinician to compare the covariance matrices from two different cardiograms (e.g., the first and the last) whereas the second test of homogeneity allows one to compare the current VCG to all previous cardiograms. It is apparent from Tolan (1982) that the flexibility to perform either test is desirable from the clinician's viewpoint. Likewise, the capability is provided to test the equality of two means vectors given that homogeneity of covariance matrices is present. Differences in mean vectors would also be indicative of serial change. This test is also extended to test the equality of k mean vectors in light of Tolan's (1982) remarks.

IV. Independence

Consider the consecutive beats X_1, \dots, X_{2n} to be the digitizations of the first $2n$ beats taken from a single VCG record. If the data pass a hypotheses test for independence then the recordings of consecutive beats can be treated as a random sample. However, it has been suggested that since the recorded beats are recorded consecutively a time dependence

will exist between consecutive beats and $\underline{X}_1, \dots, \underline{X}_{2n}$ will not form an independent set of random vectors.

Assume that $\underline{X}_i \sim N_p(\underline{\mu}, \Sigma)$, for $i = 1, \dots, 2n$. Under the null hypotheses, H_0 , $\underline{X}_1, \dots, \underline{X}_{2n}$ are independent whereas under the alternative, H_1 , they are dependent. Since the tests for independence involve an estimate S of Σ , which can be guaranteed to be nonsingular if $n > p$, the test of the null hypotheses in p dimensional pattern space would require approximately 13 minutes of cardiogram information just to be able to run the test. Consequently, an orthonormal dimension reduction technique is employed. Let T be an $r \times p$ transformation matrix such that $TT' = I_r$ but $T'T \neq I_p$. This generalization away from the Karhunen-Loève transformation is in keeping with the findings of section III. However, all the results presented here will be applicable to the KL expansion as a special case.

Consider the vectors $\underline{Y}_i = T\underline{X}_i$ for $i = 1, \dots, 2n$. Clearly, it follows $\underline{Y}_i \sim N_r(T\underline{\mu}, T\Sigma T')$. Under H_0 \underline{X}_i and \underline{X}_j are independent for $i \neq j$. Consequently

$$\text{Cov}(\underline{Y}_i, \underline{Y}_j) = T E\{(\underline{X}_i - \underline{\mu})(\underline{X}_j - \underline{\mu})'\}T' = T \text{Cov}(\underline{X}_i, \underline{X}_j)T' = 0. \quad (9)$$

Hence the independence of $\underline{X}_1, \dots, \underline{X}_{2n}$ implies the independence of $\underline{Y}_1, \dots, \underline{Y}_{2n}$. Since it is impractical to conduct the test of independence on the p -dimensional \underline{X}_i 's, we shall conduct the test on the r -dimensional \underline{Y}_i 's.

In the ordering of the beats of the cardiogram it is anticipated that the greatest dependency with \underline{Y}_i will come from \underline{Y}_{i-1} and \underline{Y}_{i+1} . Thus, separate $\underline{Y}_1, \dots, \underline{Y}_{2n}$ into two sets. $\underline{Z}_1^{(1)}, \dots, \underline{Z}_n^{(1)}$ and

$z_{-1}^{(2)}, \dots, z_{-n}^{(2)}$ where $z_{-i}^{(1)} = y_{2i-1}$ and $z_{-i}^{(2)} = y_{2i}$ for $i = 1, \dots, n$.

Once again independence of x_{-1}, \dots, x_{-2n} implies independence of the $z_{-j}^{(i)}$'s. Now, a test can be employed to see if $z_{-1}^{(1)}, \dots, z_{-n}^{(1)}$ are independent from $z_{-1}^{(2)}, \dots, z_{-n}^{(2)}$.

The likelihood ratio test (see Anderson 1958) for testing the independence of the $z_{-i}^{(1)}$'s and $z_{-j}^{(2)}$'s is such that

$$\lambda^{2/n} = |A| / (|A_{11}| \cdot |A_{22}|), \quad (10)$$

where

$$A = \sum_{i=1}^n (\underline{u}_i - \bar{\underline{u}})(\underline{u}_i - \bar{\underline{u}})' = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \quad \text{and} \quad \underline{u}_i = \begin{pmatrix} z_{-i}^{(1)} \\ z_{-i}^{(2)} \end{pmatrix}.$$

Under the truth of the null hypothesis

$$\lambda^{2/n} \sim \Lambda(r, n-1-r, r) \quad (11)$$

where $\Lambda(p, m, n)$ represents Wilks lambda distribution with dimension p and degrees of freedom m and n (see Mardia et al 1979). Thus for small values of $\lambda^{2/n}$, reject the null hypothesis that $z_{-1}^{(1)}, \dots, z_{-n}^{(1)}$ are independent from $z_{-1}^{(2)}, \dots, z_{-n}^{(2)}$ if

$$|A| / (|A_{11}| \cdot |A_{22}|) < \Lambda_{\alpha}(r, n-1-r, r) \quad (12)$$

where $\Lambda_{\alpha}(p, m, n)$ is the $100\alpha\%$ point of Wilks' lambda distribution. The distribution in (11) was originally derived by Hotelling (1936) and Girschick (1939). A discussion of computation of the percentage points in (12) is given in section VII. Unfortunately the well-known asymptotic approximations to the distribution in (11) are not applicable for our purposes because of small values of $n-1-r$, large values of r and extreme values of α .

Rejection in (12) does imply that X_1, \dots, X_{2n} are not independent. However, acceptance of the null does not necessarily imply the independence of X_1, \dots, X_{2n} . This is an unsavory aspect of this test but if the $Z_i^{(1)}$'s are independent of the $Z_j^{(2)}$'s this is a pretty good indication of independence.

Pillai and Hsu (1979) have compared four multivariate tests of independence on the basis of robustness to nonnormality. They conclude that slight nonnormality does not effect Wilks' criterion seriously. However, their numerical work is only complete for $r = 2$. Davis (1980) delineates the effect of deviations from normality on Wilks' criterion. The effects of nonnormality may be specified in terms of Mardia's (1970) measures of multivariate skewness and kurtosis and an additional skewness measure, whose effect seems to be negligible. He notes that increases in kurtosis tend to lower the true α -level, whereas increases in skewness tend to raise it. However, he only supplies numerical results for the case $r = 2$. Unfortunately, there is basically no literature on the case of large r .

A second aspect of the tests of independence deals with segmenting the signal in the P wave, QRS complex and T wave. Keating (1982) showed that forming the KL transformation matrix from the individual covariance matrices Σ_{11} , Σ_{22} and Σ_{33} of the P wave, QRS complex and T wave, respectively, was meaningful if and only if the overall covariance structure is of the form

$$\Sigma^* = \begin{bmatrix} \Sigma_{11} & 0 & 0 \\ 0 & \Sigma_{22} & 0 \\ 0 & 0 & \Sigma_{33} \end{bmatrix} \quad (13)$$

Fortunately there exists a test of such a structured covariance matrix. The test given here can be generalized from 3 to m blocks (see Anderson 1958). Suppose that $X_1, \dots, X_n \sim N_p(\underline{\mu}, \Sigma)$ are independent. Under the null hypothesis, $H_0: \Sigma = \Sigma^*$ and the likelihood ratio test statistic is given by

$$\lambda^{2/n} = |A| / \prod_{i=1}^3 |A_{ii}| \quad (14)$$

where $A = nS$ and $A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}$. Anderson (1958) shows

that $\lambda^{2/n} \sim \prod_{i=2}^3 \left\{ \frac{p_i}{\pi} X_{i\ell} \right\}$ where the $X_{i\ell}$'s are independent beta random variables with parameters $(n - \bar{p}_i - \ell)/2$ and $\bar{p}_i/2$ where $\bar{p}_i = p_1 + \dots + p_{i-1}$ and Σ_{ii} and A_{ii} are $p_i \times p_i$ matrices.

This distribution can be found directly from the work of Mathai (1971a) and Nandi (1980) and is explained in greater detail in section VII. Fortunately, in this case, n is quite large relative to p so that an intermediate approximation to the distribution of $w = \lambda^{2/n}$ by Box (1949) is suitable. Box's expansion yields

$$\Pr\{-a \ln W < w\} = \chi_f^2(w) + \frac{\gamma^2}{a^2} \left[\chi_{f+4}^2(w) - \chi_f^2(w) \right] + O(a^{-3}) \quad (15)$$

where $\chi_f^2(w)$ is the cdf of a chisquare random variable with f degrees of freedom where $f = [p(p+1) - \sum_{i=1}^3 p_i(p_i+1)]/2$, $a = n - 3/2 - (p^3 - \sum_{i=1}^3 p_i^3)/6f$ and $\gamma^2 = (p^4 - \sum_{i=1}^3 p_i^4)/48 - 5f/48 - (p^3 - \sum_{i=1}^3 p_i^3)^2/144f$. Narain (1950) has shown that $\lambda^{2/n}$ is strictly unbiased.

V. Homogeneity of Covariance Matrices

In keeping with the thoughts of Section I we shall detect a serial change in two cardiograms if $\Sigma_1 \neq \Sigma_2$. Hence, test the null hypothesis $H_0: \Sigma_1 = \Sigma_2$ versus the alternative $H_1: \Sigma_1 \neq \Sigma_2$. All tests will be carried out in the r -dimensional space (known in pattern recognition studies as feature space) as opposed to the p -dimensional space of the observed cardiogram (known as pattern space).

Let $\underline{Y}_{ij} \sim N_r(\underline{\mu}_i, \Sigma_i)$ be a random sample of size n_i for each $i = 1, 2$ and $j = 1, \dots, n_i$. Moreover, suppose that the two samples are independent. It then follows directly that $n_i S_i = A_i \sim W_r(\Sigma_i, n_i - 1)$ and are independent for $i = 1, 2$. S_i is the mle of Σ_i and $\min(n_1 - 1, n_2 - 1) \geq r$. Since $\Sigma_i^{-1/2} A_i \Sigma_i^{-1/2} \sim W_r(I, n_i - 1)$ are independent for $i = 1, 2$ then under the truth of H_0 it follows that

$$|A_1| / |A_1 + A_2| \sim \Lambda(r, n_1 - 1, n_2 - 1), \quad (16)$$

(see Mardia et al 1979). Thus if $|A_1| / |A_1 + A_2|$ is very small or very large we would reject H_0 . Hence, we accept H_0 if

$$\Lambda_{\alpha/2}(r, n_1 - 1, n_2 - 1) < |A_1| / |A_1 + A_2| < \Lambda_{1-\alpha/2}(r, n_1 - 1, n_2 - 1) \quad (17)$$

and we reject H_0 otherwise. This straightforward test of homogeneity of covariance matrices is known as Wilks' criterion. The issue of determining the exact critical values of the test of homogeneity of covariance matrices depends upon percentage points of Wilks' Λ -distribution.

The test in (17) was chosen over the modified likelihood ratio (mlr) test (see Anderson (1958)) on three basic reasons: robustness to nonnormality, simplicity of calculation of critical values and power.

The mlr test statistic is given as

$$B_2 = (w_1^{rw_1})(w_2^{rw_2}) |A_1|^{w_1} |A_2|^{w_2} / |A_1 + A_2| \quad (18)$$

where $w_i = v_i/v$, $v_i = n_i - 1$ and $v = v_1 + v_2$. Thus reject H_0 if $B_2 \leq C_\alpha(p, n_1, n_2)$. The critical values are usually found through Box's (1949) approximation.

It is well-known that (18) is extremely sensitive to changes in Mardia's measure of kurtosis (see Mardia et al 1979) whereas (16) is relatively robust to slight deviations from normality according to Davis (1980). The exact critical values (17) of Wilks' criterion can be determined from the methods described in section VII. However, the calculation of exact critical values of (18) are only alluded to in an article by Nandi (1980). The calculation of exact critical values in (18) is important since the necessary assumptions for good convergence of Box's approximation is that each $n_i > 20$ and $r \leq 5$. Sugiura and Nagao (1968) give the proof that the mlr test is unbiased. Srivastava et al (1978) also show that the power of the mlr test of $H_0: \Sigma_1 = \Sigma_2$ increases if λ_1 increases from 1 and λ_r decreases from 1, where $\lambda_1 > \dots > \lambda_r > 0$ are the distinct ordered nonzero eigenvalue of $\Sigma_1 \Sigma_2^{-1}$. If $p = 1$, (18) reduces to Bartlett's test (1937). For this simple case, Dyer and Keating (1980) have made available computer programs for the calculation of exact critical values in the unequal sample size case.

Chu and Pillai (1979) compare six tests of homogeneity of covariance matrices against two-sided alternatives with the finding that Roy's (1945) largest root test is the test of choice. The numerical work of Chu and Pillai is restricted to the case $r = 2$ (the bivariate normal) and is

therefore by no means conclusive for our situation. Further, in this power study (18) has greatest power if $\lambda_1 < 1$ and $\lambda_2 < 1$ or $\lambda_1 < 1$ and $\lambda_2 > 1$ or $\lambda_1 > 1$ and $\lambda_2 < 1$ subject to the constraint that $\lambda_1 + \lambda_2 \leq 2$. In the case when both $\lambda_1 > 1$ and $\lambda_2 > 1$, Wilks' criterion does not perform appreciably worse than Roy's largest root test. In the cases when the mlr test is preferred, Wilks' criterion does not perform unsatisfactorily in comparison with Roy's largest root test based on the summary of Chu and Pillai (1979). Moreover the exact nonnull distribution of Wilks' criterion can be calculated exactly as a mixture of beta distributions (see, e.g., Tretter and Walster (1975), Walster and Tretter (1980) and Nagarsenker (1979)). These results although exact in distribution do not readily lend themselves to simple calculation of the power function. In general all the eigenvalues of $\Sigma_1 \Sigma_2^{-1}$ must be specified in order to obtain a power calculation (i.e., the exact form of the alternative hypothesis must be stated). Secondly, the calculation of the coefficients σ_j in the beta mixtures are nontrivial as they involve the calculation of certain zonal polynomials. Unfortunately Wilks' criterion is not generalizable to test the hypothesis $H_0: \Sigma_1 = \dots = \Sigma_k$ versus the alternative $H_1: \Sigma_i \neq \Sigma_j$ for some $i \neq j$. The mlr test is generalizable to the following form:

$$B_k = \prod_{i=1}^k (w_i^{r w_i}) \prod_{i=1}^k |A_i|^{w_i} / |A|, \quad (19)$$

where $w_i = v_i/v$, $v_i = n_i - 1$, $v = \sum_{i=1}^k v_i$, $A_i = n_i S_i$ and S_i is the mle of the covariance matrix of the i th sample. Nandi (1980) alludes to a method for the calculation of the exact critical values of (19).

Chmielewski (1980) has shown that (19) is exactly robust to test

$H_0: \Sigma_1 = \dots = \Sigma_k$ within the class of elliptically symmetric distributions. Critical regions and powers developed under a multivariate normal assumption automatically extend to the larger class. The class of elliptically symmetric distributions contains the variance mixture of multivariate normal laws and the symmetric stable laws as special cases.

As an alternative to (19), the traditional approach is the use of multivariate nonparametric tests of hypotheses. Hawkins (1981) proposes a distribution free test of multivariate homogeneity through a modified measure of Mahalanobis distance. Under the assumption of multivariate normality and homogeneity of covariance matrices

$$T_{ij}^2 = v_i (\underline{y}_{ij} - \underline{y}_i^*)' S^{*-1} (\underline{y}_{ij} - \underline{y}_i^*) / n_i - vr F(r, v-r+1) / (v-r+1). \quad (20)$$

\underline{y}_i^* and S^* denote the mean vector of the i th group and the pooled covariance matrix resulting from the removal of \underline{y}_{ij} from the sample. Denote that the T_{ij}^2 's are definitely dependent. Furthermore n_i will not be substantially larger than r so that the asymptotics needed for the independence of the T_{ij}^2 's is not present. Hawkins relates that this phenomenon is the Achilles heel of his approach.

Puri and Sen (1971) derive a class of rank order tests for

$H_0: \Sigma_1 = \dots = \Sigma_k$ (not assuming that $\underline{\mu}_1 = \dots = \underline{\mu}_k$) versus $H_1: \Sigma_i \neq \Sigma_j$ for some $i \neq j$. The class of rank ordered tests is derived by centering the observations at the respective estimates of the location parameters. The centering of the observations once again violates the assumption of independence. The test advocated by Puri and Sen is valid when large sample properties are considered. Since this is known not to be

the case for our problem the nonparametric tests of hypotheses fail to fill the void.

In the nonparametric framework, the sample covariance matrix, S , is not considered a viable estimate due to its susceptibility to atypical observations. In Section III, some alternative estimators of S were suggested based on M -estimators which are less sensitive to atypical observations. Furthermore, the conditions of the test of homogeneity by Puri and Sen are quite similar to the conditions of elliptic symmetry. It seems worthwhile, therefore, to see if the two sets of conditions are equivalent.

A last point to note is that one can test the hypothesis that the transformation matrix dictated by the KL basis of Σ does indeed diagonalize the covariance matrix of \underline{Y}_i . Test $H_0: \Sigma = \text{diag}[\lambda_1, \dots, \lambda_r]$ versus $H_1: \sigma_{ij} \neq 0$ for some $i \neq j$. The likelihood ratio criterion for testing H_0 , that the covariance matrix is diagonal is given by

$$\lambda^{2/n} = |V| / \prod_{i=1}^r V_{ii} \quad (21)$$

where n is the sample size, $V = \sum_{i=1}^n (\underline{Y}_i - \bar{\underline{Y}})(\underline{Y}_i - \bar{\underline{Y}})'$,

$V_{ij} = \sum_{k=1}^n (Y_{ik} - \bar{Y}_i)(Y_{jk} - \bar{Y}_j)$ and Y_{ij} is the j th component of \underline{Y}_i ,

for $j = 1, \dots, r$. The exact null distribution of (21) is given by Mathai and Saxena (1973) and the exact percentage points of (21) under H_0 are given by Mathai and Katiyar (1979). Unfortunately, for the obtuse values of α recommended in Section III, the large values of r and small values of $n-r$, the exact percentage points of (21) are not tabulated. However, Professor Mathai of McGill University has graciously provided computer programs for the calculation of such critical values.

This test, although valid in the statistical sense only conditional on S , does give some measure of how close the estimated KL comes to diagonalizing Σ .

VI. Equality of Mean Vectors

If the test of homogeneity of covariance matrices is passed then a serial change will be detected in r -dimensional feature space if the mean vectors differ. Thus, test $H_0: \underline{\mu}_1 = \underline{\mu}_2$ given that $\Sigma_1 = \Sigma_2$ versus $H_1: \underline{\mu}_1 \neq \underline{\mu}_2$. This test is well-known in the literature as Hotelling's T^2 test. Let $\underline{Y}_{ij} \sim N_r(\underline{\mu}_i, \Sigma)$ be independent random samples for $i = 1, 2$ and $j = 1, \dots, n_i$. Johnson and Wichern (1982) (pages 467-8), Graybill (1976) (page 357) and Mardia et al (1979) (page 77) all give the critical region of this test as

$$W > F_{\alpha; r, n_1 + n_2 - r - 1} \quad (22)$$

where $F_{\alpha; m, n}$ is the 100(1- α)% point of an F distributed random variable with m and n degrees of freedom and W is given by

$$W = \frac{(n_1 + n_2 - r - 1)n_1 n_2}{r(n_1 + n_2 - 2)(n_1 + n_2)} (\bar{\underline{Y}}_1 - \bar{\underline{Y}}_2)' S^{-1} (\bar{\underline{Y}}_1 - \bar{\underline{Y}}_2) \quad (23)$$

where $\bar{\underline{Y}}_i = \sum_{j=1}^{n_i} \underline{Y}_{ij} / n_i$, $S = n_1 S_1 + n_2 S_2 / (n_1 + n_2 - 2)$ and S_i is the mle of Σ_i . $D^2 = (\bar{\underline{Y}}_1 - \bar{\underline{Y}}_2)' S^{-1} (\bar{\underline{Y}}_1 - \bar{\underline{Y}}_2)$ is the sample Mahalanobis distance between the populations in feature space. Likewise $\Delta^2 = (\underline{\mu}_1 - \underline{\mu}_2)' \Sigma^{-1} (\underline{\mu}_1 - \underline{\mu}_2)$ is the population Mahalanobis distance. The quantity D^2 plays an important role in classification problems, which Lachenbruch (1975) thoroughly discusses.

Mardia (1970, 1974) has shown that the size of Hotelling's T^2 test is overall robust to nonnormality. In fact if $n_1 = n_2$ the two sample test is little influenced by the skewness or kurtosis of the alternative

distribution. It is also worthwhile in this case to determine the power of the test in (22). Under the truth of H_1 it follows that the distribution becomes

$$\frac{(n_1 + n_2 - r - 1)n_1 n_2}{r(n_1 + n_2)(n_1 + n_2 - 2)} W^2 \sim F_{r, n_1 + n_2 - r - 1}(\delta) \quad (24)$$

where $F_{m,n}(\delta)$ represents the noncentral F-distribution with m and n degrees of freedom and in (24)

$$\delta = n_1 n_2 \Delta^2 / (n_1 + n_2). \quad (25)$$

Thus the power, which is the probability of rejecting the null given that it is false is

$$\begin{aligned} P(\Delta^2, n_1, n_2, \alpha) &= P(W > F_{\alpha; r, n_1 + n_2 - r - 1} | H_1) \\ &= 1 - F(F_{\alpha; r, n_1 + n_2 - r - 1}; r, n_1 + n_2 - r - 1, \delta) \end{aligned} \quad (26)$$

where $F(\cdot; m, n, \delta)$ is the cdf of the noncentral F distribution with non-centrality parameter δ and degrees of freedom m and n . Power is especially sensitive to the size of δ and the denominator degrees of freedom, $n_1 + n_2 - r - 1$. In fact Graybill (1976) gives a theorem that power is

- a) monotonic increasing in δ for fixed values of n_1 , n_2 , r and α ,
- b) monotonic increasing in $n_1 + n_2 - r - 1$ for fixed values of α , r and δ

and c) monotonic decreasing in r for fixed values of n_1 , n_2 , α and δ . Thus for the purposes of good power it is not only important for n_1 and n_2 to grow large but for them to grow large relative to r , the dimensionality of feature space.

Once again in light of Tolan's (1982) remarks let us extend the theory to test $H_0: \mu_1 = \dots = \mu_k$ given that $\Sigma_1 = \dots = \Sigma_k$ against $H_1: \mu_i \neq \mu_j$ for some $i \neq j, i, j = 1, \dots, k$. The likelihood ratio test of this hypothesis once again leads us to Wilks' Λ criterion. Suppose that $Y_{ij} \sim N_r(\mu_i, \Sigma)$ constitutes a random sample from the i th population of size n_i for each $i = 1, \dots, k$ and $j = 1, \dots, n_i$. Then the likelihood ratio test is derived (see Mardia et al 1979) as

$$\lambda^{2/n} = |W| / |nS| \quad (27)$$

where $W = \sum_{i=1}^k n_i S_i$ is the "within-groups" sum of squares and products (SSP) matrix, $n = \sum_{i=1}^k n_i$ and $T = nS$ is the total SSP matrix, derived by regarding all the observations as if they composed a single sample.

Define the matrix

$$B = T - W = \sum_{i=1}^k n_i (\bar{Y}_i - \bar{Y})(\bar{Y}_i - \bar{Y})' \quad (28)$$

which is the "between groups" SSP matrix. It follows that $W \sim W_r(\Sigma, n-k)$ and $B \sim W_r(\Sigma, k-1)$ are independent so that

$$\lambda^{2/n} \sim \Lambda(r, n-k, k-1). \quad (29)$$

Consequently we reject $H_0: \mu_1 = \dots = \mu_k$ given that $\Sigma_1 = \dots = \Sigma_k$ if

$$|W| / |nS| > \Lambda_{1-\alpha}(r, n-k, k-1). \quad (30)$$

Thus the determination of percentage points of Wilks' Λ distribution becomes ever more important.

The robustness of Wilks' criterion to slight deviations from normality has been documented by Davis (1980). The power of the test is however decidedly more complicated to settle than the issue of robustness. It is clear that under H_0 or H_1 $W \sim W_r(\Sigma, n-k)$. However,

under H_1 B has a noncentral Wishart distribution with $k-1$ degrees of freedom, associated matrix Σ and noncentrality matrix $\Sigma^{-1}\Delta$ where

$$\Delta = \sum_{i=1}^k n_i (\underline{u}_i - \bar{\underline{u}})(\underline{u}_i - \bar{\underline{u}})'. \quad (31)$$

For a derivation of these results see Kshirsagar (1972), page 349. When H_1 is true the distribution of B and hence that of $|W|/|W+B|$ depends upon the nonzero eigenvalues, $\delta_1, \dots, \delta_q$ of the noncentrality matrix $\Sigma^{-1}\Delta$ where $q < \min(r, k-1)$. Thus to specify the nonnull distribution of $|W|/|W+B|$ one must specify $\delta_1, \dots, \delta_q$. However in the special case where there is only one group which is different from the other $k-1$ groups then $\Sigma^{-1}B\Sigma^{-1/2} \sim W(I, k-1, \lambda^2)$ where λ^2 is the linear noncentrality parameter. Hart and Money (1975) derive the distribution of $|W|/|W+B|$ under these conditions and give some tables for power calculations.

For the purposes of conservatism, serial change was defined to be any statistically significant change in the covariance matrices or mean vectors of cardiograms. However, in practice this may be too conservative. For example suppose we define a serial change as a statistically significant difference in the mean vectors only. Anderson (1963) gives critical regions for testing the hypothesis $H_0: \underline{u}_1 = \dots = \underline{u}_k$ (regardless of equality of covariance matrices). This general problem is known as the Behren's-Fisher problem.

VII. Critical Values of Wilks' Λ Distribution

Kshirsagar (1964) has written an article on the many applications of Wilks' Λ criterion. The tests of hypotheses mentioned in this paper heavily depend upon the percentage points of the Wilks' Λ distribution. A representation of Wilks' Λ distribution that has been exploited

directly for the purposes of the calculation of percentage points is

$$\Lambda(p,m,n) = \prod_{i=1}^n u_i \quad (32)$$

where u_1, \dots, u_n are independent random variables and u_i has a beta distribution with parameters $(m+i-p)/2$ and $p/2$ for $i=1, \dots, n$. This result can be found in Anderson (1958), Kshirsagar (1972), Mardia et al (1979), etc.

Of course Wilks (1932) introduced the distribution in a slightly more complex form. He (1935) was able to state the exact null distribution as a $p-1$ fold integral, which he evaluated for $p=1,2$, for $p=3$ with $n=3,4$ and $p=4$ with $n=4$. Emphasis on the calculation of the percentage points of (32) then shifted to asymptotic expansions. Bartlett (1938) gave a chisquare approximation to $-n \ln(\Lambda(p,m,n))$ whereas Wald and Brookner (1941) gave an asymptotic expansion of the cdf of (32). Rao (1948) extended the asymptotic expansion of Wald and Brookner and determined a more rapidly convergent series for the cdf of $-[m - \frac{1}{2}(p-n+1)] \ln(\Lambda(p,m,n))$. Box's (1949) approximation, which was given in section IV, yielded approximations to the cdf of $\Lambda(p,m,n)$ that were not as bounded to the asymptotic situations. Rao (1952) gave a further approximation to the cdf of (32) as a series of beta functions. Rao (1951) also gave the following approximation through the F-distribution

$$\frac{[m+n - (p+n+1)/2]s + 2\lambda}{2r} \cdot \frac{1 - [\Lambda(p,m,n)]^{1/s}}{[\Lambda(p,m,n)]^{1/s}} \sim F_{2r, [m+n - (p+n+1)/2] + 2\lambda} \quad (33)$$

where $r = np/2$, $\lambda = -(np-2)/4$ and $s = (n^2 p^2 - 4)^{1/2} / (n^2 + p^2 - 5)^{1/2}$.

Mardia and Zemroch (1978) have also given an approximation to the cdf of $\Lambda(p,m,n)$ which involves the F-distribution with noninteger degrees of freedom.

With the advent of the computer in Statistics, several statisticians began working on techniques for the calculation of exact critical values of $\Lambda(p,m,n)$ through numerical methods. Kabe (1962) actually started off the new surge in finding simpler expressions for the exact distribution of $\Lambda(p,m,n)$ by inverting the characteristic function of $-2 \ln L(p,m,n)$ and obtaining a multiple infinite series representation of the cdf.

However, this result improved little on the original work of Wilks (1955) in terms of numerical computation. Consul (1966), using the concept of inverted Mellin transforms, obtained the distribution of $\Lambda(p,m,n)$ for a few special cases. Schatzoff (1966) and Pillai and Gupta (1969) found explicit expressions of the cdf of $\Lambda(p,m,n)$ by expressing

$\ln[\Lambda(p,m,n)] = \sum_{i=1}^n \ln(u_i)$ and determining the convolution of the independent variables $\ln(u_i)$. However, their calculations are limited to small values of p and even values of n . When both p and n are odd the convolution involves infinite series. Lee (1972) was able to circumvent these problems for odd values of $p, n \leq 9$ by reducing the problem to the numerical computation of certain univariate integrals.

Following Consul, several authors began pursuing the distribution of $\Lambda(p,m,n)$ through inverse Mellin transforms. Mellin transforms conveniently lend themselves to the determination of the distribution of a product of independent random variables. Mathai (1973) gives an excellent discussion of the various techniques, which are useful in determining the distributions of many multivariate decision rules. The

technique of Mellin transforms ultimately leads to the expression of the pdf of $\Lambda(p,m,n)$ in terms of Meijer's G-function. Mathai and Saxena (1973) and Springer (1979) give excellent discussions on the applications of Meijer's G-function in Statistics. Using the Mellin transform technique, Mathai (1971a) and Mathai and Rathie (1971) first gave the pdf of $\Lambda(p,m,n)$ as

$$h(u) = \prod_{i=1}^n \frac{\Gamma((m+i)/2)}{\Gamma((m+i-p)/2)} G_{n,0}^{n,0} \left[u \left| \begin{matrix} (m+1)/2-1, \dots, (m+n)/2-1 \\ (m+1-p)/2-1, \dots, (m+n-p)/2-1 \end{matrix} \right. \right], \quad (34)$$

where G of course represents Meijer's G-function. Nandi (1977) gave a simplification of (34) by obtaining a special identity for Meijer's G-function. This identity leads one to the result that the cdf of $\Lambda(p,m,n)$ can be expressed as an infinite sum of incomplete beta functions. Mathai (1971b) also obtained the distribution of $\Lambda(p,m,n)$ in terms of Psi and Riemann-Zeta functions, whereas Kabe (1971) gave a multiple infinite sum representation for the cdf of $\Lambda(p,m,n)$ through the notions of ordered random variables.

Tretter and Walster (1975) had previously recognized that the cdf of $\Lambda(p,m,n)$ can be expressed as a mixture of incomplete beta functions. However, the computation of the coefficients of each incomplete beta function is far from trivial. This difficulty led Davis (1979) to actually solve for the cdf through a fundamental solution of a differential equation related to Meijer's G-function. Richards and Gupta (1979) have also expressed the pdf of $\Lambda(p,m,n)$ through inverse Mellin transforms. Gupta and Richards (1979) obtained the distribution of $-\ln(\Lambda(p,m,n))$ by identifying the distribution as a sum of independent gamma random variables. They further express the cdf of $-\ln(\Lambda(p,m,n))$

as an infinite sum of incomplete gamma functions.

Mathai (1971a) and Nandi (1980) actually give a more general solution to the distribution of (32) by considering the product of any n independent beta variates. Using Nandi (1980) the cdf of $\Lambda(p,m,n)$ becomes

$$F(u;p,m,n) = C_0 \sum_{j=0}^{\infty} \psi_j B((m+n-p)/2, np/2+j) I_u((m+n-p)/2, np/2+j) \quad (35)$$

where $B(\alpha, \beta)$ is the beta function and $I_x(\alpha, \beta)$ is the cdf of a beta variate evaluated at x with parameters α and β . The values of C_0 and ψ_j greatly simplify in the case of Wilks' Λ to

$$C_0 = \left\{ \prod_{i=1}^n B((m-i+p)/2, p/2) \right\}^{-1} \quad \text{and}$$

$$\psi_j = \sum_J \prod_{i=1}^{n-1} \frac{B(p/2, ip/2 + \sum_{\ell=1}^i j(\ell))}{[(p+1)/2 + j(i)] B((p+1)/2, j(i)+1)} \quad (36)$$

where \sum_J denotes the sum over all sequences J , $J = j(1), \dots, j(n-1)$ for which $j(i)$ is a nonnegative integer and $j(1) + \dots + j(n-1) = j$.

Thus $\Lambda_{\alpha}(p,m,n)$ can be calculated through (35) by determining u such that $F(u;p,m,n) = \alpha$. Glaser (1980) outlines a procedure for the construction of an interval (A_q, B_q) containing the critical value $\Lambda_{\alpha}(p,m,n)$. The length of the interval decreases monotonically to 0 as $q \rightarrow \infty$ so that in practice one may choose the proper number q of coefficients that achieves a specified degree of accuracy. The q th partial sum in (35) is such that

$$S_q(u) < F(u;p,m,n) < S_q(u) + \epsilon_q h_q(u) \quad (37)$$

where $h_q(u) = I_u(np/2+q+1, (m+n-p)/2)$, $\epsilon_q = 1 - \sum_{j=0}^q r_j$ and

$$r_j = C_0 \psi_j B(np/2+j, (m+n-p)/2).$$

Thus it follows that (35) may be used to obtain values of $\Lambda_{\alpha}(p,m,n)$ to any specific degree of accuracy through (37). In addition, Professor Mathai of McGill University has graciously provided a computer program to calculate $\Lambda_{\alpha}(p,m,n)$ for integer values of m,n and p and levels of significance, α .

VIII. Recommendations

In Section III, doubts were raised about the use of the Karhunen-Loève transformation based on the sample covariance matrix when large sample properties are absent. However, the techniques described in Sections IV - VI depend only on an orthonormal transformation matrix, T . As a candidate for a more fitting transformation matrix, the eigenfunctions from the solution of the differential equations describing the "dipole" effect in the heart should be considered. In this approach the electrical dynamics of the heart actually assist in the selection of an appropriate set of eigenfunctions, which in turn can be digitized into eigenvectors. The subsequent eigenvalues and nuisance parameters could be estimated by statistical methods.

A second avenue for new research is the application of regression approaches to the digitized vector cardiogram in spherical coordinates. Time series techniques could then be employed to model the subsequent residuals.

References

1. Anderson, T. W. (1958), An Introduction to Multivariate Statistical Analysis, John Wiley and Sons, New York.
2. Anderson, T. W. (1963), "A Test for Equality of Means When Covariance Matrices are Unequal," Annals of Mathematical Statistics, 34, 671-2.
3. Bartlett, M. S. (1937), "Properties of Sufficiency and Statistical Tests," Proceedings of the Royal Society, A, 160, 268-282.
4. Bartlett, M. S. (1938), "Further Aspects of the Theory of Multiple Regression," Proceedings of the Cambridge Philosophical Society, 34, 33-40.
5. Box, G. E. P. (1949), "A General Distribution Theory for a Class of Likelihood Criteria," Biometrika, 36, 317-46.
6. Campbell, N. A. (1980), "Robust Procedures in Multivariate Analysis I: Robust Covariance Estimation," Applied Statistics, 29, 231-7.
7. Chmielewski, M. A. (1980), "Invariant Scale Matrix Hypothesis Tests Under Elliptical Symmetry," Journal of Multivariate Analysis, 10, 343-50.
8. Chu, S. S. and Pillai, K.C.S. (1979), "Power Comparisons of Two Sided Tests of Equality of Two Covariance Matrices Based on Six Criteria," Annals Institute of Statistical Mathematics, 31, 185-205.
9. Consul, P. C. (1966), "Exact Distribution of the Likelihood Ratio," Annals of Mathematical Statistics, 37, 1319-30.
10. Davis, A. W. (1979), "On the Differential Equation of Meijer's $G_{p,p}^{p,0}$ function, and Further Tables of Wilks's Likelihood Ratio Criterion," Biometrika, 66, 519-31.

11. Davis, A. W. (1980), "On the Effects of Moderate Multivariate Non-normality on Wilks's Likelihood Ratio Criterion," Biometrika, 67, 419-27.
12. Day, N. E. and Kerridge, D. F. (1967), "A General Maximum Likelihood Discriminant," Biometrics, 23, 313-23.
13. Dyer, D. and Keating, J. P. (1980), "On the Determination of Critical Values for Bartlett's Test," Journal of the American Statistical Association, 75, 313-9.
14. Girschick, M. A. (1939), "On the Sampling Theory of Roots of Determinantal Equations," Annals of Mathematical Statistics, 10, 203-224.
15. Glaser, R. E. (1980), "A Characterization of Bartlett's Statistic Involving Incomplete Beta Functions," Biometrika, 67, 53-8.
16. Gnanadesikan, R. (1977), Methods for Statistical Data Analysis of Multivariate Observations, John Wiley and Sons, New York.
17. Graybill, F. E. (1976), Theory and Application of the Linear Model, Duxbury Press, North Scituate.
18. Gupta, R. D. and Richards, D. (1979), "Exact Distributions of Wilks' Λ Under the Null and Non-Null (Linear) Hypotheses," Statistica Bologna, 39, 333-42.
19. Hampel, F. R. (1973), "Robust Estimation: A Condensed Partial Survey," Z. Wahr. verw. Geb., 27, 87-104.
20. Hampel, F. R. (1974), "The Influence Curve and Its Role in Robust Estimation," Journal of the American Statistical Association, 69, 383-93.

21. Hart, M. L. and Money, A. H. (1975), "Exact Powers for the Linear Case of Wilks' Likelihood Ratio Criterion," South African Statistics Journal, 9, 11-26.
22. Hawkins, D. (1981), "A New Test for Multivariate Normality and Homoscedasticity," Technometrics, 23, 105-10.
23. Hotelling, H. (1936), "Relations Between Two Sets of Variates," Biometrika, 28, 321-77.
24. Hsu, K., Womble, E., Tolun, G. D. and Zied, A. M., (1981), "Simultaneous Noise Filtering and Data Compression of ECG's," Biomedical and Scientific Instrumentation, 17, 47-52.
25. Huber, P. J. (1964), "Robust Estimation for a Location Parameter," Annals of Mathematical Statistics, 35, 75-101.
26. Huber, P. J. (1977), "Robust Covariances," In Statistical Decision Theory and Related Topics II, 165-91, Academic Press, New York.
27. Kabe, D. G. (1962), "On the Exact Distribution of a Class of Multivariate Test Criteria," Annals of Mathematical Statistics, 33, 1197-1200.
28. Kabe, D. G. (1971), "Some Distributions of Ordered Values from Eurr and Beta Distributions," Canadian Mathematics Bulletin, 14, 167-73.
29. Keating, J. P. (1982), "Theoretical Aspects of the Application of the Karhunen-Loève Expansion to EKG's," SCEEE Final Report, September 23.
30. Keating, J. P., Michalek, J. E. and Riley, J. T. (1983), "A Note on the Optimality of Karhunen-Loève Expansion," Letters in Pattern Recognition, 1, 203-4.

31. Krishnaiah, P. R. and Kanal, L. (1982), Classification, Pattern Recognition and Reduction of Dimensionality, North Holland, New York.
32. Kshirsagar, A. M. (1964), "Wilks' A Criterion," Journal of the Indian Statistical Association, 2, 1.
33. Kshirsagar, A. M. (1972), Multivariate Analysis, Marcel Dekker, New York.
34. Lachenbruch, P. A. (1975), Discriminant Analysis, Hafner Press, New York.
35. Lee, Y. S. (1972), "Some Results on the Distribution of Wilks's Likelihood Ratio Criterion," Biometrika, 59, 649-64.
36. Lin, S. P. and Perlman, M. D. (1983), "An Evaluation of Several Improved Estimators of a Covariance Matrix," Multivariate Analysis VI (ed. P. R. Krishnaiah, to appear).
37. Malkovich, J. F. and Afifi, A. A. (1973), "On Tests for Multivariate Normality," Journal of the American Statistical Association, 68, 176-9.
38. Mardia, K. V. (1970), "Measures of Multivariate Skewness and Kurtosis with Applications," Biometrika, 57, 519-30.
39. Mardia, K. V. (1974), "Applications of Some Measures of Multivariate Skewness and Kurtosis in Testing Normality and Robustness Studies," Sankhya, B, 36, 115-128.
40. Mardia, K. V., Kent, J. T. and Bibby, J. M. (1979), Multivariate Analysis, Academic Press, New York.
41. Mardia, K. V. and Zemroch, P. J. (1978), Tables of the F- and Related Distributions with Algorithms, Academic Press, London.

42. Maronna, R. A. (1976), "Robust M-estimators of Location and Scatter," Annals of Statistics, 4, 51-67.
43. Mathai, A. M. (1971a), "An Expansion of Meijer's G-Function and the Distribution of Products of Independent Beta Variates," South African Statistics Journal, 5, 71-90.
44. Mathai, A. M. (1971b), "On the Distribution of the Likelihood Ratio Criterion for Testing Linear Hypotheses on Regression Coefficients," Annals of the Institute of Statistical Mathematics, 23, 181-197.
45. Mathai, A. M. (1973), "A Review of the Various Methods of Obtaining the Exact Distributions of Multivariate Test Criteria," Sankhya, A, 35, 39-60.
46. Mathai, A. M. and Katiyar, R. S. (1979), "Exact Percentage Points for Testing Independence," Biometrika, 66, 353-6.
47. Mathai, A. M. and Rathie, P. N. (1971), "The Exact Distribution of Wilks' Criterion," Annals of Mathematical Statistics, 42, 1010-19.
48. Mathai, A. M. and Saxena, R. K. (1973), Generalized Hypergeometric Functions with Applications in Statistics and Physical Sciences, Lecture Note Series No. 348, Springer-Verlag, New York.
49. Nagarsenker, B. N. (1979), "Noncentral Distributions of Wilks' Statistic for Three Tests of Hypotheses," Sankhya, A, 41, 67-81.
50. Nandi, S. B. (1977), "The Exact Null Distribution of Wilks's Criterion," Sankhya, B, 39, 307-15.
51. Nandi, S. B. (1980), "On the Exact Distribution of a Normalized Ratio of the Weighted Geometric Mean to the Unweighted Arithmetic Mean in Samples from Gamma Distributions," Journal of the American Statistical Association, 75, 217-220.

52. Narain, R. D. (1950), "On the Completely Unbiased Character of Tests of Independence in Multivariate Normal Systems," Annals of Mathematical Statistics, 21, 293-8.
53. Pillai, K.C.S. and Gupta, A. K. (1969), "On the Exact Distribution of Wilks's Criterion," Biometrika, 56, 109-118.
54. Pillai, K.C.S. and Hsu, Y. S. (1979), "Exact Robustness Studies of the Test of Independence Based on Four Multivariate Criteria and Their Distribution Problems Under Violations," Annals Institute of Statistical Mathematics, 31, 85-101.
55. Puri, M. L. and Sen, P. K. (1971), Non-parametric Methods in Multivariate Analysis, John Wiley and Sons, New York.
56. Rao, C. R. (1948), "Tests of Significance in Multivariate Analysis," Biometrika, 35, 58-79.
57. Rao, C. R. (1951), "An Asymptotic Expansion of the Distribution of Wilks's Criterion," Bulletin of the Institute of International Statistics, 33, 177-180.
58. Rao, C. R. (1952), Advanced Statistical Methods in Biometric Research, John Wiley and Sons, New York.
59. Rataharju, P. M., Punsar, S., Blackburn, H., Warren, J. and Menotti, A. (1973), "Waveform Patterns in Frank-lead Rest and Exercise Electrocardiograms of Healthy Elderly Men," Circulation, 48, 541-
60. Roy, S. N. (1945), "The individual sampling distribution of the maximum, the minimum and any intermediate of the p-statistics on the null hypotheses," Sankhya, 7, 133-158.
61. Richards, D. and Gupta, R. D. (1979), "The Exact Distribution of Wilks' Criterion," South African Statistics Journal, 13, 7-14.

62. Schatzoff, M. (1966), "Exact Distributions of Wilks's Likelihood Ratio Criterion," Biometrika, 53, 347-58.
63. Simoons, M. L. and Hugenholtz, M. D. (1975), "Gradual Changes of ECG Waveform During and After Exercise in Normal Subjects," Circulation, 52, 570-7.
64. Springer, M. D. (1979), The Algebra of Random Variables, John Wiley and Sons, New York.
65. Srivastava, M. S. and Khatri, C. G. (1979), An Introduction to Multivariate Statistics, North Holland, New York.
66. Srivastava, M. S., Khatri, C. G. and Carter, E. M. (1978), "On Some Monotonicity Property of the Modified Likelihood Ratio Test for the Equality of Two Covariances," Journal of Multivariate Analysis, 8, 262-7.
67. Sugiura, N. and Nagao, H. (1968), "Unbiasedness of Some Test Criteria for the Equality of Two Covariance Matrices," Annals of Mathematical Statistics, 39, 1686-1692.
68. Tolan, G. D. (1982), "Vectorcardiogram Serial Comparison Using Modern Communication and Control Theory Techniques," IEEE Computers in Cardiology, 493-8.
69. Tolan, G. D., Schick, I. C. and Gustafson, D. E., (1981), "Space Technology Applied to VCG Analysis," IEEE Computers in Cardiology, 395-8.
70. Tretter, M. J. and Walster, G. W. (1975), "Central and Noncentral Distributions of Wilks' Statistic in MANOVA as Mixtures of Incomplete Beta Functions," Annals of Statistics, 3, 467-72.

71. Wald, A. and Brookner, R. J. (1941), "On the Distribution of Wilks's Statistic for Testing Independence of Several Groups of Variates," Annals of Mathematical Statistics, 12, 137-52.
72. Walster, G. W. and Tretter, M. J. (1980), "Exact Noncentral Distributions of Wilks' A and Wilks Lawley U Criteria as Mixtures of Incomplete Beta Functions: For Three Tests," Annals of Statistics, 8, 1388-90.
73. Wilks, S. S. (1932), "Certain Generalizations in the Analysis of Variance," Biometrika, 24, 471-94.
74. Wilks, S. S. (1935), "On the Independence of k Sets of Normally Distributed Variables," Econometrica, 3, 309-26.
75. Womble, M. E., Halliday, J. S., Mitter, K. S., Lancaster, M.C. and Triebwasser, J. H., (1977), "Data Compression for Storing and Transmitting ECG's/VCG's," Proceedings of the IEEE, 65, 702-6.
76. Womble, M. E., Keiser, H. N., Meyer, C. R., Triebwasser, J. H. and Lancaster, M. C., (1976), "USAFSAM ECG/VCG Digitizing and Averaging System," Report SAM-TR-76-31, Aerospace Medical Division (AFSC), USAF School of Aerospace Medicine, September.
77. Young, T. Y. and Huggins, W. H. (1963), "The Intrinsic Component Theory of Electrocardiography," IRE Transactions on Bio-Medical Electronics, BME-10, 86-95.

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

SHORT CRACK BEHAVIOR FOR FLAWS EMANATING FROM FASTENER HOLES

Prepared by: George C. Kirby III
Academic Rank: Assistant Professor
Department and University: Mechanical and Aerospace Engineering
West Virginia University
Research Location: Air Force Wright Aeronautical Laboratories
Flight Dynamics Laboratory
Structures and Dynamics Division
Structural Integrity Branch
Fatigue, Fracture, and Reliability Group
USAF Research: John M. Potter
Date: July 26, 1983
Contract No: F49620-82-C-0039

SHORT CRACK BEHAVIOR FOR FLAWS

EMANATING FROM FASTENER HOLES

ABSTRACT

Apparent stress intensity factors are found for short fatigue cracks emanating from fastener holes by the Anderson-James backcalculation technique. Empirical formulas for the stress intensity factor as a function of crack length are also derived. For crack lengths of 0.01 inches (0.25 mm) to 0.2 inches (5 mm), the empirical formulas agree well with accepted solutions of cracks radiating from holes. For cracks lengths of less than 0.01 inches (0.25 mm) the accepted solutions diverge from those derived herein. It is found that the stress intensity factor is independent of the crack location and load transfer through the fastener. Also, the "short crack effect" reported by other investigators was not corroborated. Suggestions for follow up research to this project are offered.

Acknowledgement

The author wishes to express his gratitude to the Air Force Systems Command, the Air Force Office of Scientific Research, the Southeastern Center for Electrical Engineering Education and the Air Force Wright Aeronautical Laboratory, Wright Patterson AFB, Ohio for the opportunity to perform meaningful and interesting research for the summer.

In particular he would like to express his debt to John Potter who provided helpful guidance throughout this project. Finally, the author would like to thank the personnel of the Fatigue, Fracture, and Reliability Group for their friendship and suggestions.

I. INTRODUCTION:

The principal cause of failure of aircraft structural components is the fatigue growth of a microscopic crack to a critical crack size. Most fatigue cracks originate in regions of high stress concentrations that are caused by sudden changes in geometry or notches. The most common stress raiser in aircraft vehicles is the fastener hole, 250,000 of which exist in a General Dynamics F-16 fighter.

The fatigue growth of flaws is a function of several variables and is usually expressed by an empirical formula similar to:

$$\frac{da}{dN} = C(\Delta K)^n \quad (1)$$

where da/dN is the crack growth rate, C and n are material parameters, and ΔK is the difference between the maximum and minimum stress intensity factor (SIF). Because of the exponential nature of flaw growth, it is important to know how the SIF changes during flaw growth in order to predict the growth of flaws in aircraft structural components, and to set inspection schedules and maintenance criteria. The SIF is dependent upon several variables (crack length, loading conditions, crack location, etc.) and is expressed in a generic form as:

$$K = \sigma \sqrt{\pi a} \beta \quad (2)$$

where σ is the remote stress, a is the crack length and β is a correction factor for free surface effects, flaw shape and other variables.

Because a major portion of a component's fatigue life occurs during the formation and growth of short cracks, considerable attention has

been recently directed to the study of this kind of crack (1-4). These studies suggest an irregularity in the growth of short cracks. This anomaly is manifested as an apparent high rate of growth at the onset of crack propagation, which is then followed by a period of decreasing growth rate. The flaw growth rate is then seen to increase monotonically, and thereafter normal flaw growth is observed (figure 1).

The prevailing explanations for the short crack effect (SCE) are:

- 1) The crack length is smaller than the associated plastic zone around the crack tip. Thus, linear elastic fracture mechanics is not applicable (2,3,4).
- 2) Crack closure causes retardation in the growth rate (1,2,4).
- 3) The crack length is of the same order of magnitude as the grain's dimensions, and violates the continuum mechanics requirements (3).

Although the above studies (1-4) indicate a SCE, there is evidence that this may be a fictitious effect. Noronha et al (5) obtained data for cracks emanating from filled fastener holes under a 0 and 15% load transfer across the fastener. Because the flaw that caused the failure was fractographically traced to its origin, crack lengths of 0.001 inches (.025mm) were routinely observed. It should be noted that the fastener holes in these specimens were not preflawed. Thus, the flaw growth behavior in the above study represented the growth behavior of flaws found in service. In a subsequent study using the data from Noronha et al (5), Potter and Yee (6) showed that the flaw growth rates from the above data were smooth and monotonic; thus exhibiting no SCE.

II. OBJECTIVES:

The goal of this project was to investigate the behavior of short fatigue cracks emanating from fastener holes. The objective of the

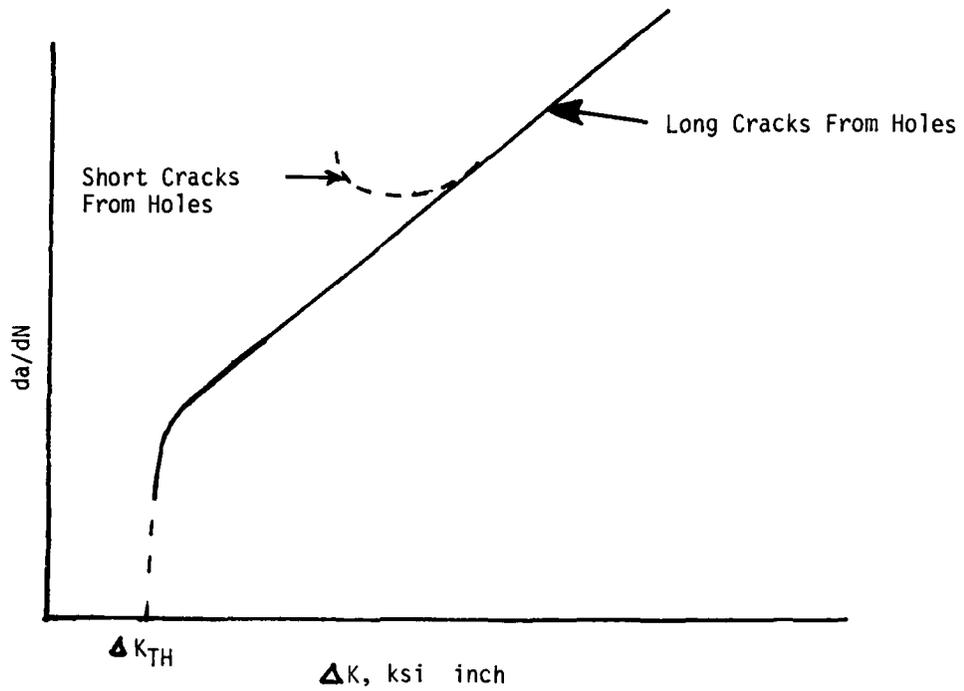


Figure 1. Crack Growth Rate Curve Showing Short Crack Effect

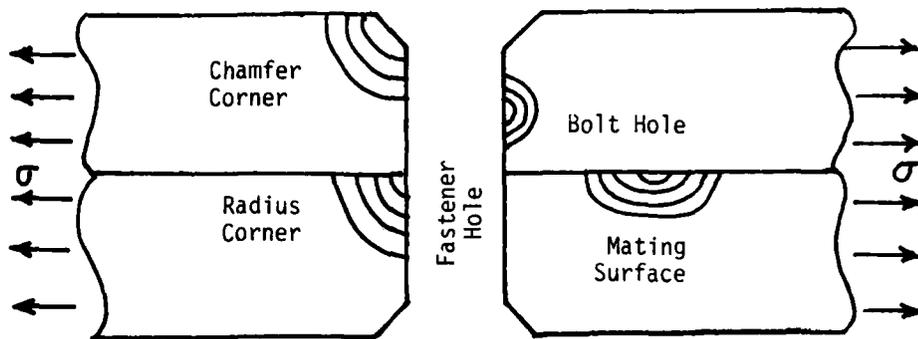


Figure 2. Crack Initiation Sites

study was attained by answering three questions:

- 1) How is an apparent SIF extracted from fatigue crack growth data?
- 2) How do the load transfer and crack location influence the apparent SIF for short cracks?
- 3) Do short cracks exhibit abnormal growth behavior and what is an explanation to the short crack effect seen by other investigators?

III. METHODOLOGY:

The basic crack growth data used in this study was documented by Noronha et al (5). In the Noronha study, fatigue tests were performed on a 7475-T7351 aluminum alloy using a random load spectrum based upon a fighter load history. At fracture, the flaw causing the failure was fractographically traced to its origin. Crack lengths were determined at intervals of 400 flight hours, and the origination of the failure flaw was recorded. From the above tests, flaws initiated from one of the following sites: the bore hole, the chamfer corner, the radius corner, or the mating surface (see figure 2). Potter and Yee (6) observed that the tests showed a high degree of reproducibility, and an exponential relationship existed between crack length and flight hours.

To answer the first question of the objective, the Anderson-James (7) backcalculation technique was used. The principal assumption of this technique is that the crack knows nothing of the gross specimen boundaries. Therefore, at any given instant the flaw growth rate is associated with an unique value of the SIF. Hence, it is necessary to obtain the material constants C and n under the same conditions as the fatigue tests performed on the specimens. These constants were obtained previously by Wilkinson and Potter for a similiar aluminum alloy under

the same fighter load spectrum (8).

Since the raw data is in terms of crack length and flight hours, the crack growth rate was found in terms of these quantities using the ASTM seven point incremental polynomial method (9) and is denoted by da/dF . Hereafter, the crack growth rate refers to da/dF .

With the growth rate constants C and n , and a crack growth rate da/dF associated with individual crack lengths, the quantity ΔK for each crack length was calculated by:

$$\Delta K = 10^n \left[\log \frac{da}{dF} - \log C \right] \quad (3)$$

The apparent SIF (K_{app}) was defined as

$$K_{app} = \frac{\Delta K}{1-R} \quad (4)$$

where R is the load ratio $\sigma_{min} / \sigma_{max}$ based upon a root mean square average of the load spectrum. For tests run under constant amplitude loading, K_{app} is equal to K_{max} .

The second problem of the objective was solved by determining a relationship between crack length and K_{app} for the various cases of load transfer and crack location. Upon examination of log-log graphs of K_{app} versus crack length (figures 3, 4, and 5), it was apparent that the following relationship existed:

$$K_{app} = A(a)^b \quad (5)$$

The constants A and b were determined by a least squares analysis. The above equation provided a means for comparing the effects of crack

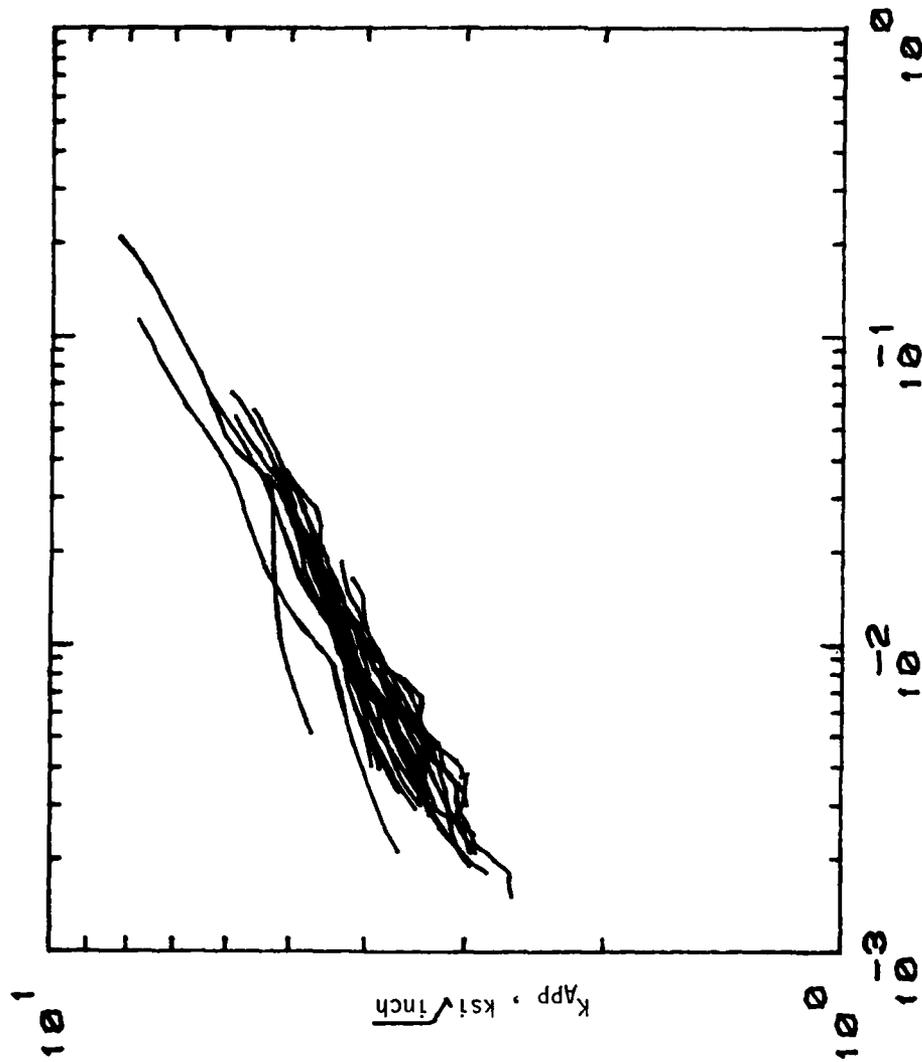


Figure 3. Apparent Stress Intensity Factor for WPF, 33 specimens

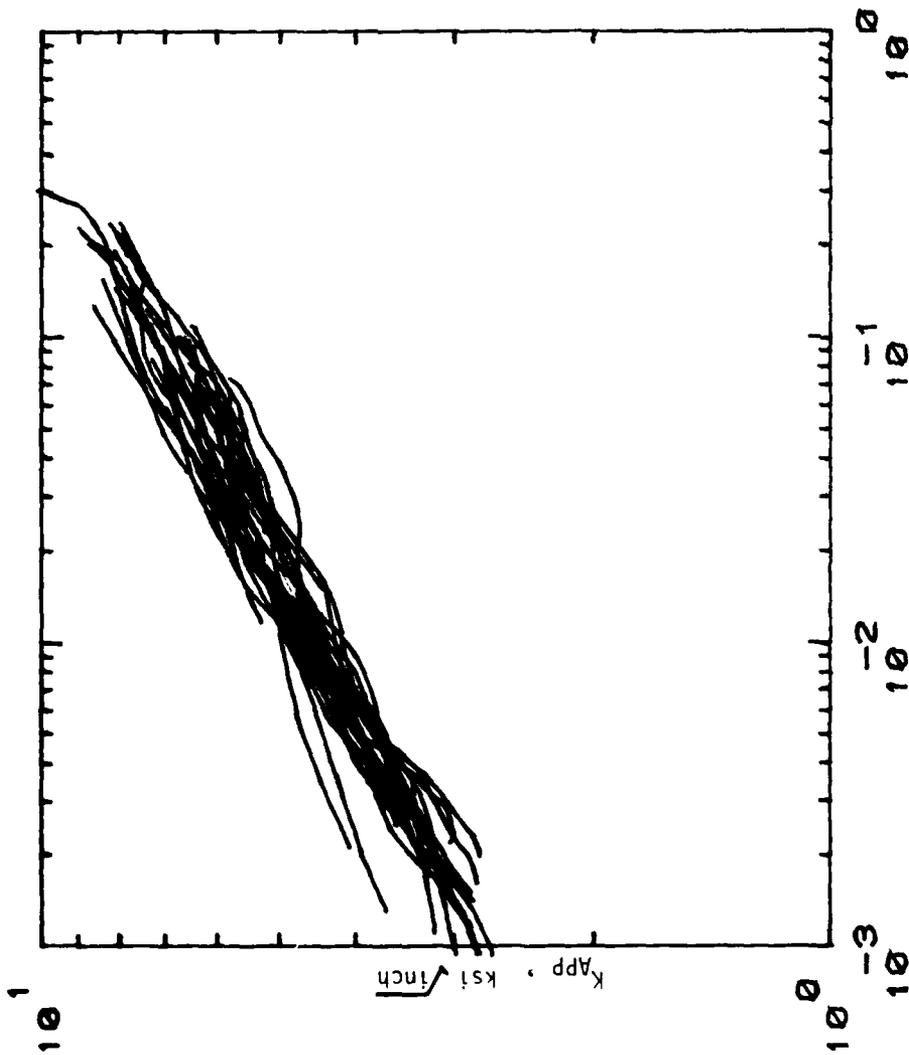


Figure 4. Apparent Stress Intensity Factor for XMPF, 34 specimens

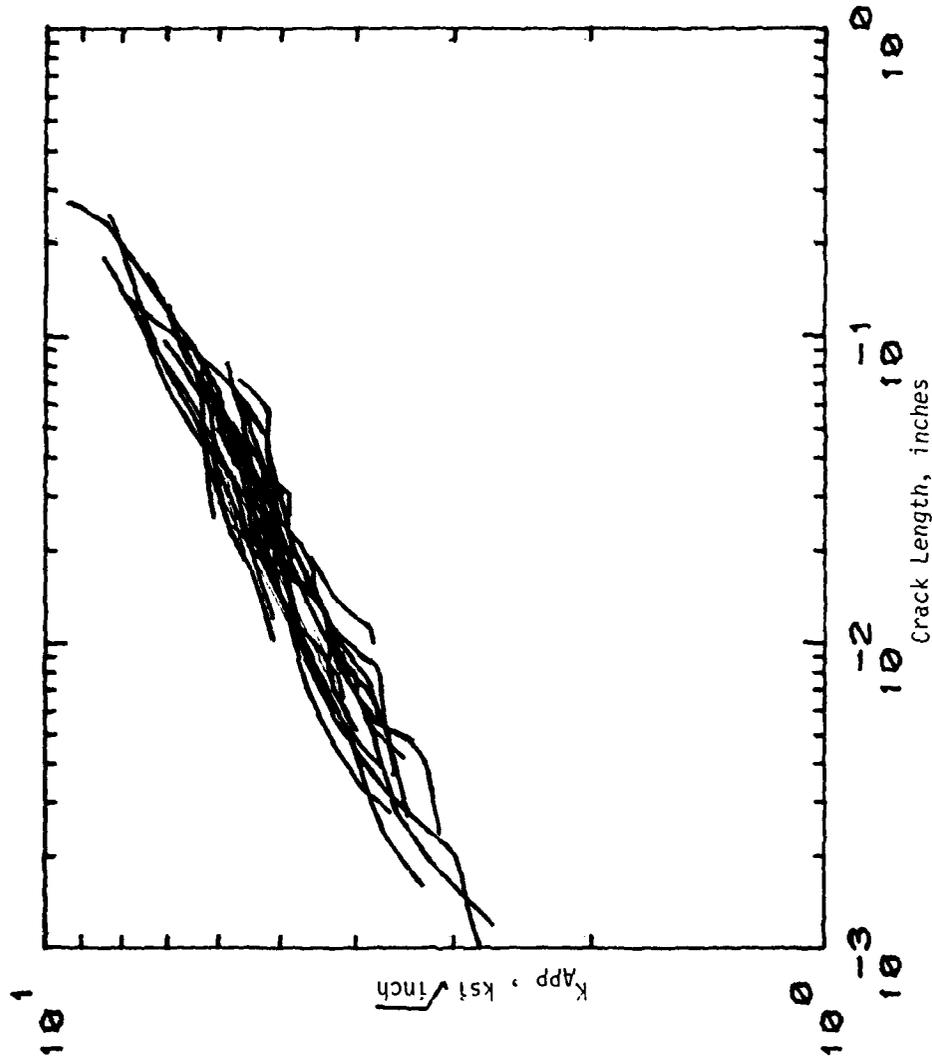


Figure 5. Apparent Stress Intensity Factor for YMPF, 25 specimens

initiation site and load transfer on the apparent SIF.

The final question of whether short cracks grow differently than long cracks was resolved by examining a plot of crack growth rate versus flaw length. Normal crack growth should appear as smooth, continuous and monotonically increasing in the above graph. The short crack phenomena was explained by examining how studies that reported the SCE would have handled the data by Noronha et al (5).

IV. RESULTS and DISCUSSION:

Figures 3, 4, and 5 show how K_{app} varies with crack depth for specimens from three test series, WPF, XWPF, and YWPF respectively. The WPF series were a 0% load transfer through the fastener. The "W" stands for a Winslow Spacematic drill, the "P" for proper drill speeds, coolant and feeds, and "F" for fighter load history. The XWPF and YWPF series were 15% load transfer; the YWPF specimens used an improved drilling assembly. With the exception of 9 out of 92 tests, the data for each series (WPF, XWPF, and YWPF) fall within 10% of the mean behavior. These figures indicate a high degree of repeatability. With the exception of 2 out of 11 cases of crack initiation sites, coefficients of correlations for the fitted lines are 0.95 or higher.

To determine the influence of crack location upon the apparent SIF solution, crack length versus apparent SIF was plotted for the various cases of initiation sites (figure 6). While these results are for the YWPF, tests they are typical of the results for the XWPF series. The solid lines are empirical relations for distinct cases of crack initiation site between crack length and K_{app} calculated by a least squares analysis. Coefficients of correlations are 0.90 or higher. The dashed lines indicate $\pm 5\%$ deviations in K_{app} from the empirical

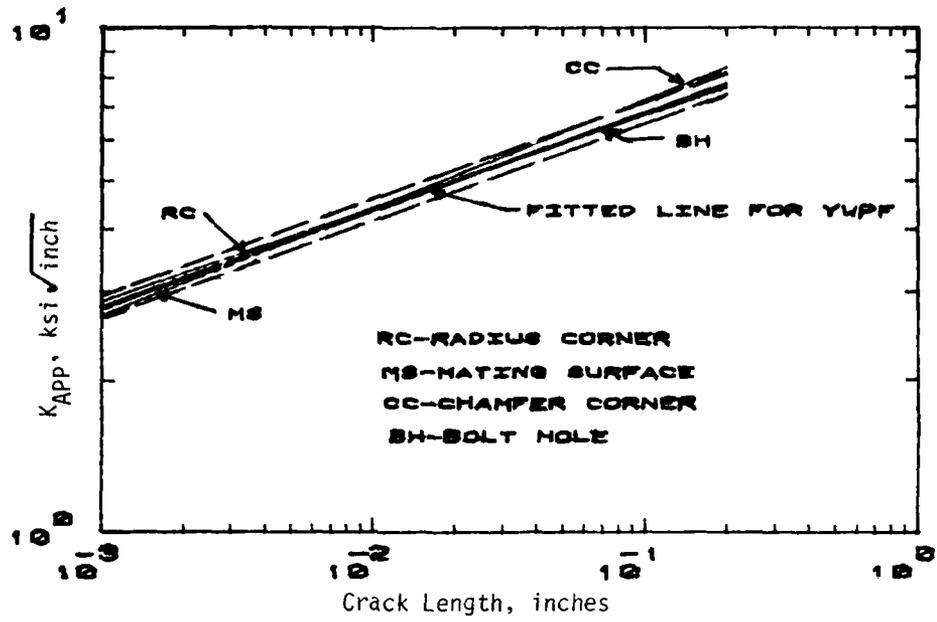


Figure 6. Influence of the crack initiation site upon the SIF

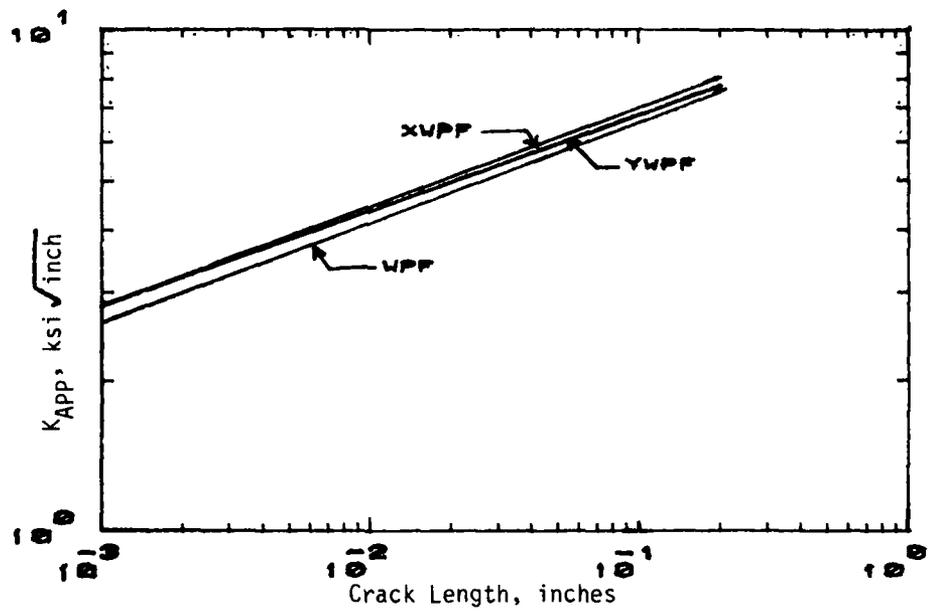


Figure 7. Comparison of the Apparent SIF for WPF, XWPF & YWPF Series

solutions for all the specimens of the YWPF series. A maximum difference of 9% in K_{app} occurs for the cases of chamfer and radius corner initiations. Based upon these observations one may conclude that for short cracks ($a \leq 0.200$) the initiation site does not affect K_{app} .

The effect of load transfer through the fastener and the influence of the drilling procedure is summarized in figure 7. The figure shows that the derived solutions of WPF, YWPF, and XWPF all lie within 8% of each other for the range investigated. Thus, one may say that neither the 15% load transfer through the fastener nor the drilling technique affect the apparent SIF solution.

Table 1 tabulates the apparent SIF for each series and case of crack initiation site as an equation of the line of best fit and the corresponding coefficient of correlation (r). The simple power relationship of equation 5 was used to fit a line through the data. Because coefficients of correlations greater than 0.90 were obtained, a strong exponential relationship between crack length and apparent SIF exists.

Three analytical fastener hole SIF solutions are compared with the results of this study (figure 8). One should note that the present solution was multiplied by 4.35 to achieve agreement of the apparent SIF in the range of crack lengths of 0.01 to 0.1 inches (.25 to 2.5 mm). The justification of this correction lies in the ambiguity of the effects of random amplitude loading on the load ratio R . However, this correction does not alter the general trend seen in this study. At crack lengths less than 0.01 inches (.25 mm), Grandt's solution (10), Tweed's and Rooke's solution (11), and Bowie's solution (12) for cracks at holes begin to diverge from the present solution and underpredict

TABLE 1

INITIATION SITE	WPF 33 specimens	XMPF 34 specimens	YMPF 25 specimens
All Sites	$K_{app} = 10.358a^{0.192}$ $r = 0.958$	$K_{app} = 11.031a^{0.197}$ $r = 0.971$	$K_{app} = 10.585a^{0.192}$ $r = 0.952$
Bolt Hole		$K_{app} = 11.327a^{0.200}$ $r = 0.975$	$K_{app} = 10.435a^{0.190}$ $r = 0.972$
Chamfer Corner		$K_{app} = 10.985a^{0.198}$ $r = 0.992$	$K_{app} = 11.754a^{0.213}$ $r = 0.980$
Radius Corner		$K_{app} = 9.376a^{0.171}$ $r = 0.965$	$K_{app} = 10.268a^{0.184}$ $r = 0.900$
Mating Surface		$K_{app} = 11.066a^{0.225}$ $r = 0.971$	$K_{app} = 10.568a^{0.194}$ $r = 0.940$

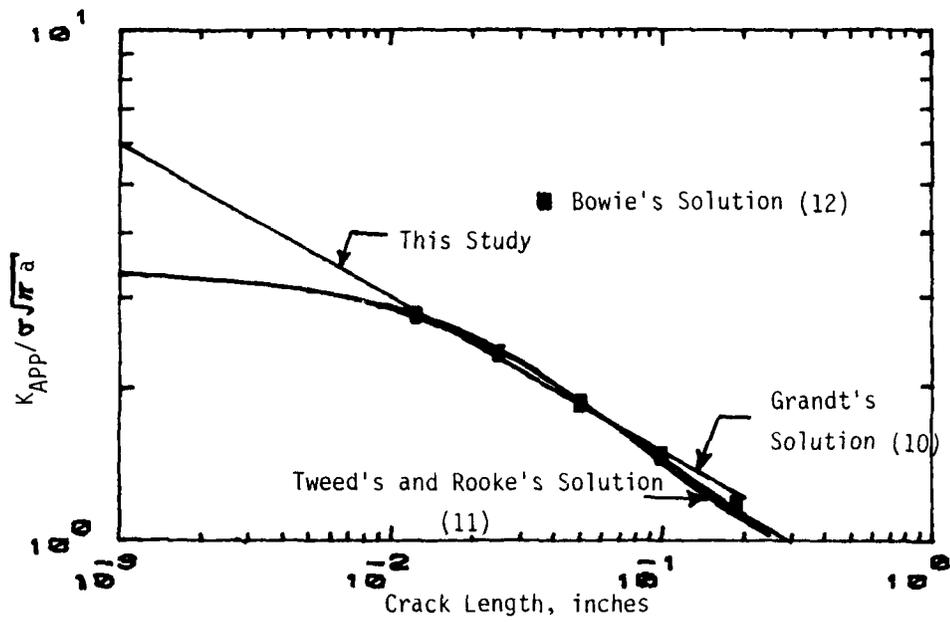


Figure 8. Comparison of Various Solutions for cracks emanating from a hole

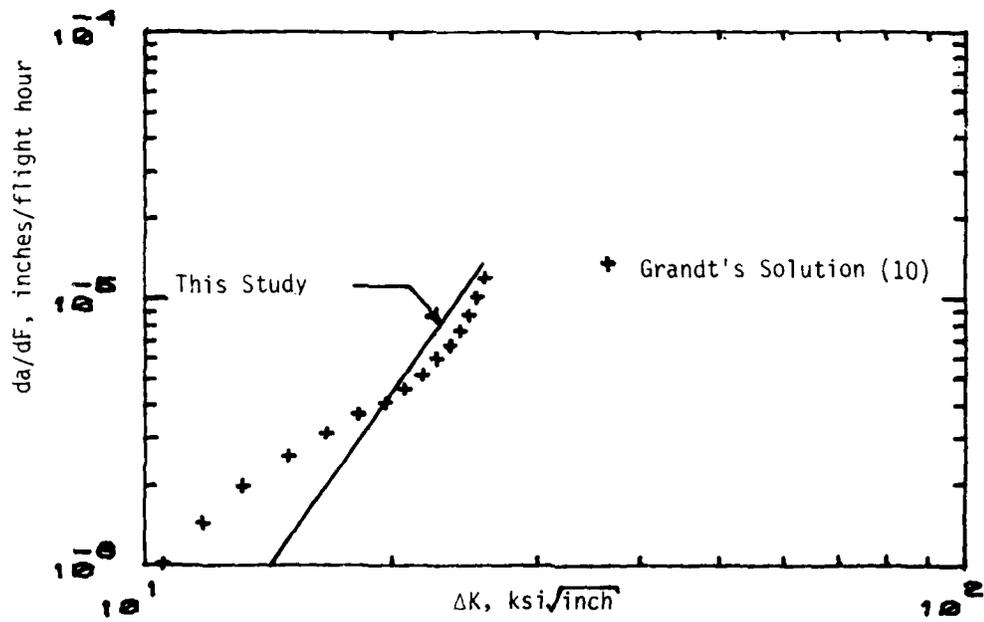


Figure 9. Short Crack Effect predicted by Grandt's Solution (9) for Noronha et al data (5).

K_{app} by approximately 45% for a crack length of 0.001 inches (0.025 mm). This underprediction of K_{app} for short crack lengths may be the cause of the SCE seen by many investigators. For example, using the data from a typical test from Noronha et al (5), Grandt's solution indicates a presence of a short crack effect (figure 9).

An unanticipated result from this study is that the apparent SIF is independent of crack initiation site and load transfer for short cracks. One possible explanation for this unexpected behavior is that short cracks cannot feel the presence of the boundaries of the specimen. Thus, all short cracks behave the same regardless of the crack location or load transfer through the fastener. Although this explanation is satisfactory for cracks having lengths of less than 0.01 inches (.25 mm), it does not explain this behavior for flaws with lengths of 0.20 inches (5 mm).

Since the cracks studied in this report indicated that the SIF was independent of crack location and load transfer, the comparison of two dimensional solutions with the derived solutions of this study is justified even though the cracks in this study were part-through. Also, to compare the results of this study accurately with three dimensional solutions, it is necessary to know how the flaw shape continuously changes throughout the flaw growth. For the cracks studied in this report only the final flaw shape was known.

V. CONCLUSIONS AND RECOMMENDATIONS:

The following conclusions were drawn from this study:

- 1) The SIF varies as an exponential function for short fatigue cracks.
- 2) For short fatigue cracks emanating from fastener holes the SIF is independent of load transfer through the fastener hole and the

location of the crack initiation.

- 3) For specimens examined in this study, no short crack effect was seen. However, it was found that fastener hole SIF solutions similar to Bowie's solution (12) underpredict the SIF for crack lengths less than 0.01 inches (0.25 mm). This underprediction of the apparent SIF causes a short crack effect to be seen.

Two questions are proposed for follow up research. First, is the apparent SIF dependent upon higher or lower stress levels? All the tests examined in this study were performed with a maximum stress level of 34 ksi and these results may be biased in this respect. Second, what is the effect of pin clearance on the apparent SIF? Since the load transfer through the fastener is dependent upon the clearance between the fastener and the hole, pin clearance may indeed influence the apparent SIF. Thus, to fully answer the question of the effect of load transfer upon the apparent SIF, the effect of fastener fit must be examined.

REFERENCES

1. S. Pearson, "Initiation of Fatigue Cracks in Commercial Aluminum Alloys and the Subsequent Propagation of Very Short Cracks," *Engineering Fracture Mechanics*, Vol. 7, pp. 235-247, 1975.
2. B. C. Fisher and F. Sherratt, "A Fracture Mechanics Analysis of Fatigue Crack Growth Data for Short Cracks," *Fracture Mechanics in Engineering Practice*, Proceedings of the Annual Conference, Sheffield, England, Sept. 1976.
3. S. J. Hudak, "Small Crack Behavior and the Prediction of Fatigue Life," *Journal of Engineering Materials and Technology*, Transactions of ASME, Vol. 103, p. 26-35, Jan. 1981.
4. B. N. Leis and R. D. Galliher, "Growth of Physically Short Corner Cracks at Circular Notches," Proceedings International Symposium on Low Cycle Fatigue and Life Prediction, ASTM STP (In Press).
5. P. J. Noronha, S. P. Henslee, Z. Wolanski, D. E. Gordon, and B. G. W. Yee, "Fastener Hole Quality," AFFDL-TR-78-206, Vol. I & II, USAF Flight Dynamics Lab Dayton, Oh, 1979.
6. J. M. Potter and B. G. W. Yee, "Use of Small Crack Data to Bring About and Quantify Improvements to Aircraft Structural Integrity," AGARD Specialists Meeting on Behavior of Short Cracks in Aircraft Structures, Toronto, Canada, Sept. 1982.
7. W. E. Anderson and L. A. James, "A Simple Experimental Procedure for Stress Intensity Calibration," *Engineering Journal of Fracture Mechanics*, Vol. 1 pp. 565-568, Apr. 1969.
8. R. L. Wilkinson and J. M. Potter, "Effects of Bonding Temperatures on Fatigue Crack Growth," AFWAL-TR-82-3058, Air Force Wright Aeronautical Laboratories, Dayton, Oh, 1982.
9. ASTM, "Standard Test Method for Constant-Load-Amplitude Fatigue Crack Growth Rates Above 10^{-8} m/cycle". E647-81.
10. A. F. Grandt, "Stress Intensity Factors for Some Thru-Cracked Fastener Holes," AFML-TR-74-71, USAF Materials Lab, Dayton, Oh, 1971.
11. J. Tweed and D. P. Rooke, "The Distribution of Stress Near the Tip of a Radial Crack at the Edge of a Circular Hole," *International Journal of Engineering Science*, Vol. 11, pp. 1185-1195, 1973.
12. O. L. Bowie, "Analysis of an Infinite Plate Containing Radial Cracks Originating at the Boundary of an Internal Circular Hole," *Journal of Mathematics and Physics*, Vol. 35, pp. 60-71, 1956.

1983 USAF-SCEEE 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

AN EVALUATION OF THE MATHEMATICAL PROCESS AND FORMULATION FOR
CASE MOUNTED DISPLACEMENT SENSORS

Prepared by:	James Kirpatrick Susie A. Hobbs (Graduate Student)
Academic Department:	Department of Mathematics Department of Physics (Graduate Student)
University:	Alabama Agricultural & Mechanical University
Research Location:	Arnold Engineering and Development Center Arnold Afs, TN.
USAF Research Contact:	Mr. Marshall K. Kingery
Date:	August 13, 1983
Contract No.:	F99620-82-C-0035

AN EVALUATION OF THE MATHEMATICAL PROCESS AND FORMULATION FOR
CASE MOUNTED DISPLACEMENT SENSORS

by

James Kirkpatrick

Susie A. Hobbs (Graduate Student)

ABSTRACT

A review of the data Analysis Principles of case mounted displacement sensors are validated relative to their algebraic significance and derivations from the standard equation of motion.

Some basic properties of fourier series and transform are discussed. We are primarily concerned with the recapture of f from f if additional information concerning f is known. The proof relies basically on Cesaro Summability, Planchereal Theorem, and the Lebesgue Dominated Convergence Theorem. Meanwhile, we can relate the fourier transform to the prominent Paley-Weiner Theorem in the theory of entire functions, that is an entire function is analytic throughout the finite complex plane. This gives the assurances for symmetric inversions formulations for the recapture of f from f .

Acknowledgement

We would like to express our appreciation to the organizations and individuals involved in the research activity. Namely, the Air Force Command, the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for providing an excellent opportunity this summer to do research at the Arnold Engineering Development Research Center, Arnold Afs, TN. We appreciate the hospitality and the magnificent working conditions exemplified by the Dynamic Systems and Analysis Unit.

We wish to thank Mr. Paul E. McCarty, Mr. Henry T. Jones, and Mr. Gordon W. Konyndyk, whose helpful discussions, suggestions, and encouragements were invaluable.

Finally, we would like to acknowledge our USAF contact, Mr. Marshall K. Kingery for his cordiality and guidance through the research period.

I. INTRODUCTION

Vibrations occur frequently in the physical world. These deformations caused by forces acting on mechanical systems, electromotive forces in electrical circuits, etc. are not simple periodic functions but represent discontinuous functions. Under investigation at the Arnold Center is the vibratory stress of rotary blade deflections in turbomachines which represents these kind of functions. The program is designed to use case mounted optic sensors as the source of data collection and eliminate the excessive maintenance cost of the present Strain Gage Method. The problem with these functions is to find a mathematical system that will provide the analysis necessary to represent the periodicity accurately. The Fourier Series function and Transform named after the French Physicist Jean Baptise Fourier is the mathematical duality selected to give the desired results. The mathematical theory clearly indicates a mirror image recapture of the system if suitable Symmetric Inversion formulas are applied.

II. OBJECTIVES

The main objective of this report is to analyze the mathematical principles for blade tip deformation, in turbomachines through the use of displacements sensor. Secondly, verify the algebraic summations, approximation, derivations, and substitutions based on the classical equation of motion to detect the tangential and axial deflections of rotary blade for nonintegral and integral order vibrations. Finally, to justify the existence of Fourier Transform and verify the analysis of the spectra that will provide the recapture of the original function. Our specific objectives consist of:

- (1) Analysis of basic vibratory formulas to determine the validity

of data analysis principle for case mounted stationary optic sensor measurements for turbo machine engines.

- (2) To determine if the equation for tangential deflections gives solutions to the problem, suitable interpolation for rules are sought to establish the uniqueness of the solutions.
- (3) To determine if the equation for axial deflections are derivable from standard equation of motion and are algebraically sound.
- (4) To show that the existence of Fourier Transform is the tool to recapture the original function.

FOURIER TRANSFORM AND APPLICATIONS

We will now discuss the fourier transform, first for \underline{L}^1 and then for \underline{L}^2 by the Plancherel Theorem.

Given a periodic and integrable function, it is possible to define fourier coefficients and then try to recapture the function from its fourier series. We shall do a similar thing here except that the starting point is not a periodic function but simply a function integrable on the whole real line. The motivation comes from formally considering fourier series for functions of period 2ℓ and letting ℓ tend to infinity. The fourier coefficients for such a function f is given by

$$\frac{1}{\ell} \int_{-\ell}^{\ell} f(t) \cos \frac{n\pi t}{\ell} dt, \quad \frac{1}{\ell} \int_{-\ell}^{\ell} f(t) \sin \frac{n\pi t}{\ell} dt \quad 1.0$$

Assuming the validity of all operations, we have for the fourier series of f at x

$$\begin{aligned} & \frac{1}{\ell} \sum_{n=1}^{\infty} \int_{-\ell}^{\ell} f(t) \cos \frac{n\pi}{\ell} (t-x) dt \\ &= \frac{1}{\pi} \left[\frac{\pi}{\ell} \int_{-\ell}^{\ell} f(t) \cos \frac{n\pi}{\ell} (t-x) dt \right. \\ & \left. + \frac{\pi}{\ell} \int_{-\ell}^{\ell} f(t) \cos \frac{2\pi}{\ell} (t-x) dt + \dots \right] \end{aligned}$$

Taking $\Delta y = \frac{\pi}{\ell}$ and letting $\ell \rightarrow \infty$, the sum approaches formally the integral

$$\begin{aligned} & \frac{1}{\pi} \int_0^{\infty} \int_{-\infty}^{\infty} f(t) \cos y(t-x) dt dy \\ &= \int_0^{\infty} \left\{ \left[\frac{1}{\pi} \int_{-\infty}^{\infty} f(x) \cos yt dt \right] \cos yx \right. \\ & \left. + \left[\frac{1}{\pi} \int_{-\infty}^{\infty} f(x) \sin yt dt \right] \sin yx \right\} dy \end{aligned}$$

$$\begin{aligned}
&= \int_{-\infty}^{\infty} \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} f(t) \cos yt \, dt - \frac{i}{2\pi} \int_{-\infty}^{\infty} f(t) \sin yt \, dt \right] \\
&\quad \times (\cos yx + i \sin yx) \, dy \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-iyt} \, dt \right] e^{-iyx} \, dy \quad 1.1
\end{aligned}$$

We thus get a formal analogy with fourier series, and so define the fourier transforms of $f \in L^1(-\infty, \infty)$ by

$$\hat{f}(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-iyt} \, dt \quad 1.2$$

other definitions of $f(y)$ are also in use, primarily

$$\int_{-\infty}^{\infty} f(t) e^{-2\pi i y t} \, dt \quad \text{and} \quad \frac{1}{2\pi} \int_{-\infty}^{\infty} f(t) e^{-iyt} \, dt$$

We will now show that the form in 1.2 leads to a systematic inversion formula between f and \hat{f} when certain conditions are satisfied. If f is even, then

$$\hat{f}(y) = \sqrt{\frac{2}{\pi}} \int_0^{\infty} f(t) \cos yt \, dt \quad 1.3$$

if f is odd we get

$$\hat{f}(y) = -i \sqrt{\frac{2}{\pi}} \int_0^{\infty} f(t) \sin yt \, dt \quad 1.4$$

for functions defined on $(0, \infty)$, (1.3) and (1.4) without the $-i$ and called the cosine and sine transforms of f . Theorems from advanced calculus can be used to assert the existence of \hat{f} from certain f .

If there are intervals $(-\infty, -a)$ and (a, ∞) on which f is continuous and $f(x)$ tend to 0 as $x \rightarrow \pm \infty$, then the cosine integrals

$$\int_a^{\infty} f(x) \cos yx dx \quad \text{and} \quad \int_{-\infty}^{-a} f(x) \cos yx dx$$

1.5

exist, and similarly for the sine integrals.

If $f(x) = e^{-|x|}$, then $\hat{f}(y)$ exist by (1.5)

Integration by parts $\hat{f}(y) = \sqrt{\frac{2}{\pi}} / (1 + y^2)$

Let $f(x) = a^{-ax}$ for $x > 0$, where a is some positive constant. Then the cosine and sine transform are respectively,

$$\sqrt{\frac{2}{\pi}} \frac{a}{a^2 + y^2} \quad \text{and} \quad \sqrt{\frac{2}{\pi}} \frac{y}{a^2 + y^2}$$

If $f(x) = e^{-(x^2/2a^2)}$, then $\hat{f}(y) = ae^{-(a^2 y^2/2)}$

Note that if $a = 1$, then $f = \hat{f}$

Consider the functions measurable in $L^1 [a, b]$.

If $f \in L^1(-\infty, \infty)$, then (FT means Fourier Transforms)

- a. FT of $f(x + h)$ is $e^{-ihy} \hat{f}(y)$
 - b. FT of $e^{-ihx} f(x)$ is $\hat{f}(y + h)$
 - c. FT of $f(ax)$ is $\frac{1}{|a|} \hat{f}\left(\frac{y}{a}\right)$, $a \neq 0$
 - d. FT of $\frac{1}{f(-x)}$ is $\frac{1}{\hat{f}(y)}$
 - e. $\text{Sup}_y |\hat{f}(y)| \leq \|f\|_1$
- 1.6

If $f \in L^1$, then \hat{f} is uniformly continuous on the real number line

by part a in (1.6) we have for all h

$$\left| \hat{f}(y + h) - \hat{f}(y) \right| = \left| \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ikt} (e^{-iht} - 1) f(t) dt \right|$$

$$\leq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |e^{-iht} - 1| |f(t)| dt \quad 1.7$$

Now the integrand is bounded by the integrable function $2|f(t)|$;
 so letting $h \rightarrow 0$ through a sequence h_n allows us to use the Lebesgue
Dominated Convergence Theorem to conclude that

$$\lim_{h \rightarrow 0} |\hat{f}(y+h) - \hat{f}(y)| \leq 0, \text{ since } \lim_{h \rightarrow 0} |e^{-iht} - 1| = \lim_{h \rightarrow 0} 2 \left| \sin\left(\frac{ht}{2}\right) \right| = 0.$$

We get uniform continuity since the last term in (1.7) is independent
 of y .

By the Rie-mann-Lebesgue lemma if $f \in L^1(-\infty, \infty)$ then

$$\lim_{|y| \rightarrow \infty} \hat{f}(y) = 0 \quad \text{we may show that}$$

$$\hat{f}(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-iyt} |f(t) - f(t + \frac{\pi}{y})| dt$$

it follows that

$$|\hat{f}(y)| \leq \frac{1}{\sqrt{2\pi}} \omega\left(\frac{\pi}{|y|}, f\right) \rightarrow 0 \quad 1.8$$

We thus see that if f is integrable, then \hat{f} is uniformly continuous and
 tends to 0 as its argument becomes arbitrarily large. However \hat{f} need
 not be integrable.

We define the convolution of f and g by

$$(f * g)(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x-t) g(t) dt$$

if f and g are both integrable, then $(f * g)^\wedge(y) = \hat{f}(y) \hat{g}(y)$

also if f and g are both integrable, then

$$\int_{-\infty}^{\infty} \hat{f}(t) g(t) dt = \int_{-\infty}^{\infty} f(t) \hat{g}(t) dt$$

each integral exists by (1.7) and (1.8) using Fubini's Theorem gives directly

$$\begin{aligned} \int_{-\infty}^{\infty} \hat{f}(t) g(t) dt &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(u) e^{-iut} g(t) du dt \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} e^{-iut} g(t) dt \right\} f(u) du \\ &= \int_{-\infty}^{\infty} \hat{g}(u) f(u) du \end{aligned}$$

CESARO SUMMABILITY AND UNIQUENESS AND INVERSION

In attempting to recapture f from \hat{f} , it would be natural to follow (1.1) and try to show that

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(y) e^{-iyx} dy \quad 1.9$$

although we do not know whether the integral converges, we do know from (1.7) that \hat{f} is uniformly continuous on the real number line. Hence for every A ,

$$\frac{1}{\sqrt{2\pi}} \int_{-A}^A \hat{f}(y) e^{-iyx} dy \quad \text{exist}$$

Guided by the Fejer-Lebesgue Theorem, we consider the mean value of (1.9) considered as a function of A . The results will be given later, first we must define (C.1) Summability in general for the continuous case. Instead of working with $\int_{-A}^A g(t) dt$, we look instead to averaging $\int_{-T}^T g(t) dt$ over the interval $0 \leq T \leq A$. The mean value, which is the limiting case of the arithmetic mean is by definition

$$\begin{aligned}
& \frac{1}{A} \int_0^A \left[\int_{-T}^T g(t) dt \right] dT = \frac{1}{A} \int_{-A}^A \int_{|t|}^A g(t) dT dt \\
& = \int_{-A}^A \left[\int_{|t|}^A \frac{1}{A} dT \right] g(t) dt \\
& = \int_{-A}^A \left(1 - \frac{|t|}{A} \right) g(t) dt
\end{aligned} \tag{2.0}$$

we say that $\int_{-\infty}^{\infty} g(t) dt$ is summable (C,1) to S if

$$\lim_{A \rightarrow \infty} \int_{-A}^A \left(1 - \frac{|t|}{A} \right) g(t) dt = S$$

thus (C,1) summability is discrete.

We now show that (C,1) summability in the continuous case just as in the discrete case is a regular method of convergence.

If $g \in L^1(-\infty, \infty)$ then $\int_{-\infty}^{\infty} g(t) dt$ is summable (C,1) to $\int_{-\infty}^{\infty} g(t) dt$ 2.1

Let $g_A(t)$ be $(1 - \frac{|t|}{A}) g(t)$ if $|t| \leq A$ and 0 otherwise.

Then by the Lebesgue dominated convergence theorem

$$\lim_{A \rightarrow \infty} \int_{-\infty}^{\infty} g_A(t) dt = \int_{-\infty}^{\infty} g(t) dt$$

which implies the results.

Consider one form of the inversion formula (1.9) which is the parallel of the Fejer-Lebesgue Theorem.

Let $f \in L^1(-\infty, \infty)$. Then for almost all values of x

$$\lim_{A \rightarrow \infty} \frac{1}{\sqrt{2\pi}} \int_{-A}^A \left(1 - \frac{|y|}{A} \right) e^{ixy} \hat{f}(y) dy = f(x) \tag{2.2}$$

Using the definition of \hat{f} and Fubini's theorem, which is justified since $f \in L^1(-\infty, \infty)$ implies that the double integral is absolutely convergent we may write the integral in (2.2) as

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} f(u) \int_{-A}^A \left(1 - \frac{|y|}{A} \right) e^{iy(x-u)} dy du$$

Writing the inner integral as $\int_{-A}^0 + \int_0^A$ and using $\exp(it) + \exp(-it) = 2 \cos t$ to calculate the inner integral, then integral in (2.2) becomes

$$\frac{1}{\pi} \int_{-\infty}^{\infty} f(u) \frac{1 - \cos A(x-u)}{A(x-u)^2} du = \frac{1}{\pi} \int_0^{\infty} [f(x+t) + f(x-t)] \frac{1 - \cos At}{At^2} dt$$

But

$$\begin{aligned} \int_0^{\infty} \frac{1 - \cos At}{At^2} dt &= \int_0^{\infty} \frac{1 - \cos At}{A^2 t^2} A dt \\ &= \int_0^{\infty} \frac{1 - \cos W}{W^2} dW = \frac{\pi}{2} \end{aligned}$$

denoting the integral in (2.2) by I , we have

$$I = f(x) = \frac{1}{\pi} \int_0^{\infty} [f(x+t) + f(x-t) - 2f(x)] \frac{1 - \cos At}{At^2} dt$$

The argument now proceeds as in the Fejer-Lebesgue Theorem and is valid for any point x in the Lebesgue set of f .

To conclude the conversion theorem

- (a) if $f \in L^1(-\infty, \infty)$ and f is continuous at x , then (2.2) is valid
- (b) if f and \hat{f} are both in $L^1(-\infty, \infty)$, then 1.9 is valid
- (c) since $\hat{f} \in L^1(-\infty, \infty)$ (1.9) converges; in fact it converges absolutely, for each x , however by (2.2), (1.5) is summable $(C,1)$ to $f(x)$.

[b] above is one form of the so-called inversion theorem which expresses a duality between f and \hat{f} , namely

$$\begin{aligned} \hat{f}(y) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-iyx} dx \\ f(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(y) e^{ixy} dy \end{aligned}$$

The L^1 theory presented so far is proper but unsymmetrical. For example, if we define T by $Tf = \hat{f}$ for $f \in L^1 \rightarrow C_0$, the latter being the space of continuous functions which tends to 0 as $|x| \rightarrow \infty$. Also, the exact image of T in C_0 is not known. Further, most of the ideas presented so far have different hypotheses for f and \hat{f} . We shall show that there is a more satisfactory symmetrical theory for L^2 .

In order to do so, the first thing we must do is to define \hat{f} for $f \in L^2(-\infty, \infty)$. Suppose we had $L^2[0, 2\pi] \subset L^1[0, 2\pi]$, but on the real line we have neither this conclusion nor the reverse one. In fact $f(x) = \frac{1}{x}$ for $|x| \geq 1$ and 0 otherwise defines a function in $L^2(-\infty, \infty)$ but not in $L^1(-\infty, \infty)$, while $g(x) = x^{-1/2} \sin x$ for $|x| \geq 1$ and 0 otherwise defines a function in $L^1(-\infty, \infty)$ but not in $L^2(-\infty, \infty)$. In order to begin our definition of \hat{f} for f in L^2 , we first suppose f is in both L^1 and L^2 .

If $f \in L^1 \cap L^2$, then $\hat{f} \in L^2$ and $\|f\|_2 = \|\hat{f}\|_2$. In this case, the definition of \hat{f} is free since $f \in L^1$. Let $g(x) = \overline{f(-x)}$ and $h = f * g$. Since f and g are in L^2 , h is continuous and since f and g are in L^1 , we also have $h \in L^1$. By (2.2) is valid for $x = 0$, which means

$$\lim_{A \rightarrow \infty} \frac{1}{\sqrt{2\pi}} \int_{-A}^A \left(1 - \frac{|y|}{A}\right) \hat{h}(y) dy = h(0) \quad 2.3$$

Now $\hat{h}(y) = \hat{f}(y) \hat{g}(y) = |\hat{f}(y)|^2 \geq 0$. Define $H_A(y) = (1 - |y|/A) \hat{h}(y)$ for $|y| \leq A$ and 0 for $|y| \geq A$, so that $H_A(y)$ increases monotonically to $\hat{h}(y)$ as $A \rightarrow \infty$.

The Lebesgue Monotone Convergence Theorem applied to (2.3) gives

$$h(0) = \lim_{A \rightarrow \infty} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} H_A(y) dy = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{h}(y) dy \quad 2.4$$

in particular, $\hat{f} \in L^2$. Also by definition

$$h(0) = (f * g)(0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) g(t) dt \quad 2.5$$

$$(2.4) \quad (2.5) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(-t) \frac{1}{f(-t)} dt = \frac{1}{\sqrt{2\pi}} \|f\|_2$$

combining 2.3 and 2.4 gives the desired results.

Our next step in the definition of \hat{f} is to eliminate the requirement that $f \in L^1$.

Let $f \in L^2$ and $f_A(x) = f(x)$ for $|x| \leq A$ and 0 for $|x| > A$, then $f_A \in L^1 \cap L^2$, $\hat{f}_A \in L^2$, and \hat{f}_A converges in L^2

Since $|f_A(x)| \leq |f(x)|$ clearly $f_A \in L^2$. By Holder's inequality

$$\|f_A\|_1 = \int_{-A}^A |f(x)| dx \leq \|f\|_2 \cdot \sqrt{2A}, \text{ so } f_A \in L^1 \text{ also. By (2.3)}$$

$\hat{f}_A \in L^2$ and if $A < B$ then

$$\begin{aligned} \|\hat{f}_B - \hat{f}_A\|_2^2 &= \|f_B - f_A\|_2^2 \\ &= \frac{1}{\sqrt{2\pi}} \int_{-B}^{-A} |f(x)|^2 dx + \frac{1}{\sqrt{2\pi}} \int_A^B |f(x)|^2 dx \end{aligned} \quad 2.6$$

The last two integrals tend to 0 as A and B tend to infinity since $f \in L^2$; so L^2 being complete implies that there is a unique function $T(f) \in L^2$ such that $\|T(f) - \hat{f}_A\|_2 \rightarrow 0$ as $A \rightarrow \infty$.

This allows us to define \hat{f} given any $f \in L^2$. We simply let $\hat{f} = T(f)$ as given. In case $f \in L^1 \cap L^2$, we now have two definitions for \hat{f} , one for L^2 and previous for L^1 . Since $\|T(f) - \hat{f}_A\|_2 \rightarrow 0$, there is subsequence A_k such that

$$\begin{aligned} T(f)(y) &= \lim_{k \rightarrow \infty} \frac{1}{\sqrt{2\pi}} \int_{-A_k}^{A_k} f(x) e^{-iyx} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-iyx} dx \end{aligned} \quad 2.7$$

The last equality being true since $f \in L^1$. But the last term is just $\hat{f}(y)$ as defined previously, so the notation \hat{f} is consistent for $f \in L^1 \cap L^2$. The main results may be stated in the form

$$\lim_{A \rightarrow \infty} \int_{-\infty}^{\infty} \left| f(x) - \frac{1}{\sqrt{2\pi}} \int_{-A}^A \hat{f}(y) e^{-ixy} dy \right|^2 dx = 0 \quad 2.8$$

In Plancherel's Theorem we aim to show among other things, the inversion counterpart to (2.8), namely

$$\lim_{A \rightarrow \infty} \int_{-\infty}^{\infty} \left| f(x) - \frac{1}{\sqrt{2\pi}} \int_{-A}^A \hat{f}(y) e^{ixy} dy \right|^2 dx = 0 \quad 2.9$$

To do this, first we need the following

If $f \in L^2$, then $\|\hat{f}\|_2 = \|f\|_2$

Since $\lim f_A = f$ and f_A and f are both in L^2 , then the Lebesgue Dominated

Convergence Theorem implies that $\lim \|f_A\|_2 = \|f\|_2$. But by (2.3), if

$$\|f_A\|_2 = \|\hat{f}_A\|_2; \text{ so } \lim \|\hat{f}_A\|_2 = \|f\|_2. \text{ By definition,}$$

$$\|\hat{f}_A - \hat{f}\|_2 \rightarrow 0, \text{ so } \lim \|\hat{f}_A\|_2 = \|\hat{f}\|_2 \text{ as expected.}$$

If f and g are in L^2 , then

$$\int_{-\infty}^{\infty} \hat{f}(t) g(t) dt = \int_{-\infty}^{\infty} f(t) \hat{g}(t) dt \quad 3.0$$

Define f_A and g as in (2.6), then each function is in L^1 , so by definition

$$\int_{-\infty}^{\infty} \hat{f}_A(t) g_B(t) dt = \int_{-\infty}^{\infty} f_A(t) \hat{g}_B(t) dt \quad 3.1$$

We let $A \rightarrow 0$ in both sides of (3.1). For the left,

$$\left| \int_{-\infty}^{\infty} \hat{f}_A(t) g_B(t) dt - \int_{-\infty}^{\infty} \hat{f}(t) g_B(t) dt \right| \leq$$

$$\|\hat{f}_A - \hat{f}\|_2 \cdot \|g_B\|_2 \rightarrow 0$$

As $A \rightarrow \infty$, where we have used Holder's inequality and the definition of \hat{f} .

We act similarly to the right side using $\|f_A - f\|_2 \rightarrow 0$.

Thus after letting $A \rightarrow \infty$, 3.0 becomes

$$\int_{-\infty}^{\infty} \hat{f}(t) g_B(t) dt = \int_{-\infty}^{\infty} f(t) \hat{g}_B(t) dt \quad \text{letting } B \rightarrow \infty$$

Now produces the results in exactly the same way as above.

(Plancherel) Let $f \in L^2$. Then the fourier transforms \hat{f} exist. $\hat{f} \in L^2$ (2.8)

and (2.9) hold. Further $\|\hat{f}\|_2 = \|f\|_2$ since $\hat{f} \in L^2$, then $\bar{\hat{f}} \in L^2$ and also $(\bar{f})^\wedge \in L^2$.

$$\begin{aligned} \text{Now } \|\bar{f} - (\bar{\hat{f}})^\wedge\|_2^2 &= \int_{-\infty}^{\infty} (\bar{f} - (\bar{\hat{f}})^\wedge)(\bar{f} - (\bar{\hat{f}})^\wedge) \\ &= \int_{-\infty}^{\infty} |\bar{f}|^2 - \int_{-\infty}^{\infty} (\bar{\hat{f}})^\wedge \bar{f} - \int_{-\infty}^{\infty} \bar{\hat{f}} \bar{f} + \int_{-\infty}^{\infty} |(\bar{\hat{f}})^\wedge|^2 \end{aligned}$$

The second term on the last line by (3.1) is $\int_{-\infty}^{\infty} \bar{\hat{f}} \bar{f} = \int_{-\infty}^{\infty} |\hat{f}|^2$

So by (2.9) the first two terms cancel, the last two also.

Therefore $\bar{f} = (\bar{\hat{f}})^\wedge$. The definition of the fourier transform of $\bar{\hat{f}}$ now

$$\text{yields } \lim_{A \rightarrow \infty} \int_{-\infty}^{\infty} \left| (\bar{\hat{f}})^\wedge(u) - \frac{1}{\sqrt{2\pi}} \int_{-A}^A \bar{\hat{f}}(t) e^{-itu} dt \right|^2 du$$

So using $(\bar{\hat{f}})^\wedge = \bar{f}$ and taking conjugates gives (2.9)

Since $f \in L^2$ is the fourier transform of $\hat{f}(-x)$, we see that $Tf = \hat{f}$ defines

$T: L^2 \rightarrow L^2$ as a one-to-one onto norm-preserving transformation. T also

preserves inner products. The results may be extended to L^p , $1 < p < 2$,

by similar techniques, but to allow $p > 2$ require the theory of generalized functions.

From the definition of \hat{f}_A , we know $\|\hat{f}_A - f\|_2 \rightarrow 0$, and so some subsequence

$\hat{f}_k \rightarrow \hat{f}$. It turns out that whether or not $\hat{f}_A \rightarrow \hat{f}$ is equivalent to Lusin's

Conjecture about convergence of the fourier series of a square summable

function. Carleson's theorem thus shows that indeed $\hat{f}_A \rightarrow \hat{f}$.

Finally we relate the fourier transform to the Paley-Wiener Theorem.

If $f(z)$ is entire, then it can be expanded in a power series of z , convergent for $z \in \mathbb{C}$. If further

$$|f(z)| \leq C e^{A|z|} \tag{3.2}$$

In order that (3.2) hold it is necessary and sufficient that the fourier transform $\hat{f}(t)$ in $L^2(-\infty, \infty)$ of $f(x)$ vanish almost everywhere for $|x| > A$.

ALGORITHMS
APPLIED TO THE FOURIER TRANSFORM

Systems which facilitate in a vibrational motion have some degree of noise that affects the performance of the system. The noise can be measured through a process called Cross Spectral Analysis. The noise spectrum is computed from a time series that has been through an inversion process where discrete coefficients are formulated. These discrete coefficients are known as transform. An analysis is applied on the transform to develop the original wave form or signal through an inversion process.

The power spectrum identifies the spectrum of signals measured at different points in the system with the cross-spectrum measuring a relationship between the signals. The noise spectrum is useful in detecting the frequency and amplitude of various components of the noise signals. The power spectrum is used to give information concerning the phase of the spectrum.

When a single signal is being analyzed, we consider a complete series of functions, $x(t)$ and $y(t)$; the basis forward Fourier Transform is:

$$\begin{aligned} F[x(t)] &= S_x(j\omega) = a + jb \\ F[y(t)] &= S_y(j\omega) = c + jd \end{aligned}$$

that is examined at its real and imaginary parts.

For many application of random signals of systems, the amplitude is not meaningful as the power spectrum contained within the given frequency band. The power spectrum density or the power spectrum is the function:

$$\begin{aligned} G_{xx}(\omega) &= S_x(j\omega) \cdot S_x^*(j\omega) = a^2 + b^2 \\ G_{yy}(\omega) &= S_y(j\omega) \cdot S_y^*(j\omega) = c^2 + d^2 \end{aligned}$$

which is calculated by taking the product of a Fourier transform signal and the complex conjugate of the same Fourier transform, or by adding the

squares at the real and the imaginary parts of the Fourier transform. Where several signals are involved, the cross-spectrum of the two data signals is defined as follows:

$$\begin{aligned}
 G_{ys} &= S_y S_x^* \\
 &= (c-jd) \cdot (a-jb) \\
 &= (ab + bd) + j(ad-bc) \\
 &= c_0 + j \text{ quadrant}
 \end{aligned}$$

this is formed by taking the product of the Fourier transform of one signal and the complex conjugate of the Fourier transform of a second signal. An allocation is made to enhance the signal to noise, at the same time signals are elevated which are not mutually contained in the input.

In determining the gain and phase characteristics the transfer function is defined,

$$H(j\omega) = \frac{\text{Output}}{\text{Input}} = \text{transfer function}$$

Consequently the transfer function is the ratio of the Laplace transform of a driving function and response function. However a relation between the Fourier transform output signal to the Fourier transform of the input signal implies that:

$$\begin{aligned}
 H(j\omega) &= \frac{S_y(j\omega)}{S_x(j\omega)} \\
 &= \frac{S_y(j\omega) \cdot S_x^*(j\omega)}{S_x(j\omega) \cdot S_x^*(j\omega)} \\
 &= \frac{G_{yx}(j\omega)}{G_{xx}(j\omega)} \\
 &= \frac{\text{cross-spectrum}}{\text{input power spectrum}}
 \end{aligned}$$

Since the relationship between the transfer function is known the validity of the cross-spectrum is offset by the validity in the input power spectrum where the phases are identical and gives smooth deterministic transfer function measurements.

The Fourier transform of the original signals produce the complex quantities S_x and S_y , the forward transform operation can be followed by the inverse transformation

$$F^{-1}[S_x(j\omega)] = x(t)$$

$$F^{-1}[S_y(j\omega)] = y(t)$$

A time domain signal is successively transformed and then followed by an inverse operation, the original signal is retrieved.

Hamming window function is used to alleviate inaccuracies in retrieving the original waveform. This windowing affects gives an accumulation in preventing leakage of the original waveform or signal after it has been retrieved through the inverse process. There exist several types of windowing functions, they are: the rectangular window, hamming window, and Kaiser-Bessel window.

Hamming window function is used because of its relatively simple implementation and it can be repeatedly employed. The application of this method is done by taking the time series and multiplying by the weight function before transformations take place. The windowing affect reduces the variance of the power spectral estimations.

Since we know that power spectral analysis is used in the recapturing of the original waveform. Different types of methods are used in determining the estimations, for example auto correlation method and the direct variance method, which reduces the resolution of the wave forms and gives statistical stability for the power spectral estimations. The convolution process is relationship between the source and the impulse response of the waveforms or signals.

IV. CONCLUSION AND RECOMMENDATIONS

While the inversion and uniqueness theorems show a duality between f and \hat{f} and leads to a symmetric inversion formula between f and \hat{f} when certain conditions are satisfied, the problem that remains is the absolute recapture of f from \hat{f} .

The present literature has an excellent set of Algorithms that definitely seems adequate in the analysis of the Fourier spectrum. However, the writers recommend that a continued search for additional Algorithms be considered until an absolute recapture is assured.

BIBLIOGRAPHY

- Biot, Maurice A.; Mechanics of Incremental Deformations, (John Wiley & Sons, Inc. New York, 1965).
- Bloomfield, Peter, Fourier Analysis of Time Series: An Introduction, (John Wiley & Sons, New York, 1976).
- Chen, C.H., Digital Waveform Processing and Recognition, (CRC Press Inc., USA, 1982).
- Kaufmann, A., Theory of Fuzzy Subsets, (Academic Press, New York, 1975).
- Kreyszig, Erwin, Advanced Engineering Math, (John Wiley and Sons Inc., New York, 1962).
- Ding-Sun Fu, Kikichi Tanka, - Zadeh Lotific A., Masachichi Shimura; Fuzzy Sets and Their Applications to Cognitive and Decision Process, (Academic Press, New York, 1975).
- Kazuo Kyuma, Masahiro Nonoshita, and Shuichi Tai, "Fiber-Optic Acceleration Sensor base on Photoelectric Effect", Optical Society of America, Sept. 1982.
- Mandel, John, The Statistical Analysis of Experimental Data, (John Wiley & Sons Inc., New York, 1964).
- Pack, V.C., Potasek, M.J. and T.K. Walters, "Comparison of Wave Length For Total Dispersion"; Western Electric Engineering Brands, Princeton N.J. 1983.
- Red heffer R.M. and Sokolnikoff I.S., Mathematic of Physics and Modern Engineering, McGraw Hill Press, New York, 1958).
- Rees, Charles, S. Shaif, S.M. and Stanojevic, C.V., Theory and Application of Fourier Analysis, (Marcell-Decker, New York, 1981).
- Scanlon - Cronin Jane, Advanced Calculus, (Heath and Co., Washington, D.C. 1967).
- Smith Craig C. and Thornhill, Joe, Fourier and Spectral Analysis a Sheet Course. (Craig C. Smith and Joe Thornhill, Texas, 1980).

TEM MORPHOLOGY STUDY OF MOLECULAR COMPOSITES OF POLYMERS

by

Dr. Stephen J. Krause

ABSTRACT

Transmission electron microscopy (TEM) is used to study 30% PBT / 70% ABPBI fibers to determine if a composite has been formed at the molecular level. It is concluded from dark field images and diffraction patterns that micron-sized aggregates rich in PBT crystallites form within an ABPBI rich matrix when fiber is spun or film is cast from solution above a critical concentration. However, when fiber and film was processed from solution below a critical concentration, the PBT is dispersed as crystallites no larger than 30 \AA in size. Thus, the existence of a molecular-level composite has been confirmed by TEM.

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

TEM MORPHOLOGY STUDY OF MOLECULAR COMPOSITES OF POLYMERS

Prepared by:	Dr. Stephen J. Krause
Academic Rank:	Assistant Professor
Academic Department:	Mechanical and Aerospace Engineering
University:	Arizona State University
Research Location:	Air Force Wright Aeronautical Laboratories, Materials Laboratory, Polymer Branch
USAF Research Contact:	W. W. Adams
Date:	August 20, 1983
Contract No.:	F49620-82-C-0035

ACKNOWLEDGMENTS

The authors wish to thank the Air Force Systems Command, Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for their support of this project. The ten weeks spent in the Polymer Branch of the Materials Laboratory at Wright-Patterson Air Force Base were most productive due to the many wonderful people working there. It would be unfortunate to omit any of them, but to list all of them would appear excessive, so a collective "thanks" will have to suffice. A special thanks goes to W. W. Adams, AFWAL/MLBP research colleague, for interesting and enlightening discussions in the polymer-morphology lab.

I. INTRODUCTION

Concept of a molecular composite

A molecular composite can be defined as a material consisting of two or more components which are interspersed at the molecular level. The two components studied in this investigation are poly(p-phenylene benzobisthiazole), referred to as PBT, and poly-2,5(6)benzimidazole, referred to as ABPBI. PBT is a rigid rod, extended chain polymer of high strength, high modulus, and excellent environmental and thermal resistance. ABPBI is an aromatic heterocyclic polymer like PBT, but ABPBI has a flexible-coil structure in contrast to the rod-like PBT structure (see Figure 1). The purpose of forming a molecular composite of these two polymers is to reinforce the coil-like ABPBI with the rod-like PBT. Processing of PBT in the molecular composite form allows its orientation to be controlled so that mechanical properties can be custom-tailored. The composite has been processed in the form of vacuum cast film, spun film and spun fiber. The study examines the 30% PBT / 70% ABPBI fiber spun from two concentrations of solution, as well as the two component homopolymers as fibers.

Related studies

The major techniques which have been used to study PBT, ABPBI, and the 30 / 70 blend are mechanical testing, wide angle x-ray diffraction (WAXD), scanning electron microscopy (SEM) and, here, transmission electron microscopy (TEM). A table summarizing the types of data and information available from these techniques is given in Figure 2.

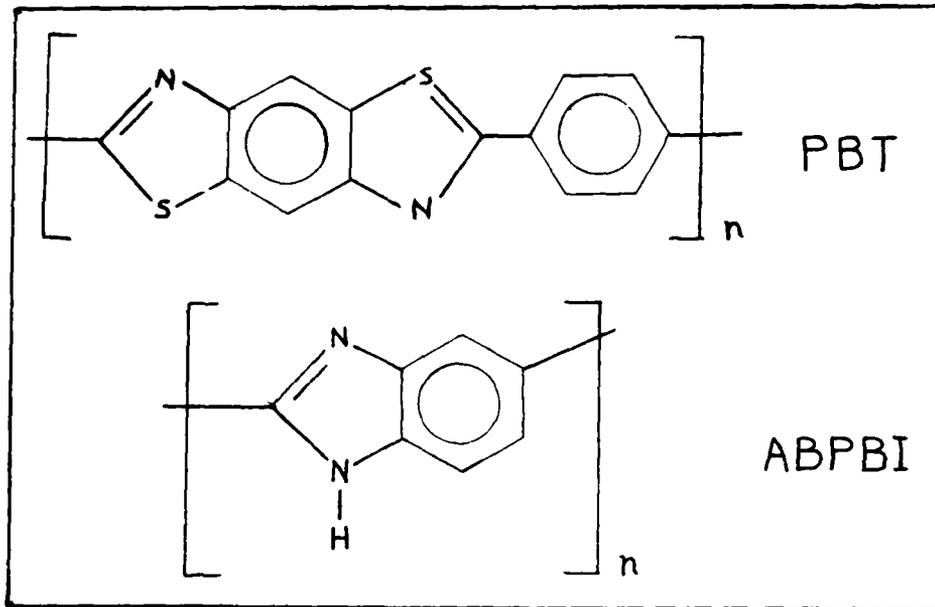


Figure 1: Schematic of PBT and ABPBI molecules

TECHNIQUE	RAW DATA	INFORMATION
Mechanical testing	stress-strain curve	modulus, tensile strength, % elongation to break
Wide angle x-ray diffraction (WAXD)	flat plate photo, equatorial line scan	d-spacing; line breadth; cryst. size & orientation
SEM liquid nitrogen fracture surface	500X - 5000X picture	fibrillarity, ductility, size & shape of phases
TEM	dark field image, selected area electron diffraction pattern	crystallite size, shape, and orientation

Figure 2: Information obtained from various techniques

The morphology of PBT has been intensively studied; major references include Thomas et al., 1980 (1), and Minter, 1982 (2). The first of these employs a variety of techniques, including TEM, to characterize the crystalline structure of PBT fibers. The latter is a dissertation which comprehensively surveys and analyzes previous PBT reports, and uses techniques such as TEM to determine structural features. It also conveniently gives an extensive set of titled references on the subject. ABPBI has not been studied as much as PBT. X-ray diffraction patterns of as-cast and stretched films of ABPBI which were used to determine d-spacings were published in 1980 by Price (3).

The fabrication of molecular composites of polymers first received attention in 1978 by Helminiak (4). Preliminary work on molecular composites was reported by Husman et al. in 1980 (5). The study emphasized various PDIAB*/ABPBI composites with discussions of how morphology (observed with SEM) and mechanical properties were related to processing. The same relationships with respect to PBT/ABPBI composites were discussed in 1983 by Hwang et al. (6).

In the study of Hwang et al., different weight compositions of PBT/ABPBI were mixed in various concentrations in a solvent. It was found that above a critical concentration (C_{cr}) of PBT/ABPBI in solvent, the solution stopped being optically isotropic and became biphasic--that is, liquid crystalline domains appeared within an isotropic phase for $C > C_{cr}$. Scanning electron microscopy photos showed that films prepared from $C > C_{cr}$

* PDIAB is poly(p-phenylenebenzodiazole)

solutions retain the phase separation. SEM backscattering indicated that the domains are composed chiefly of PBT.

This phase separation has important consequences in terms of a parameter known as the "aspect ratio." The aspect ratio is defined as $2L/D$, where L is the length and D is the diameter of the reinforcing phase within a matrix phase. For efficient reinforcement, mechanical theory states that the aspect ratio should be about 100 or more.

Phase separation as observed in $C > C_{cr}$ fibers and films results in a low aspect ratio (see Figure 3a). When fibers and films were processed from $C < C_{cr}$ solutions, no aggregates were observed with SEM to its resolution limit of 500 Å (6,7). By studying the equatorial reflections in WAXD patterns of $C < C_{cr}$ fibers, it was concluded that the PBT molecules were dispersed at the molecular level among the ABPBI molecules (6). If the PBT molecules are indeed dispersed singularly or in small bundles in an ABPBI matrix, then a high aspect ratio is achieved (see Figure 3b). The mechanical testing results show that the $C < C_{cr}$ fiber has significantly higher modulus and tensile strength than the $C > C_{cr}$ fiber (see Figure 4).

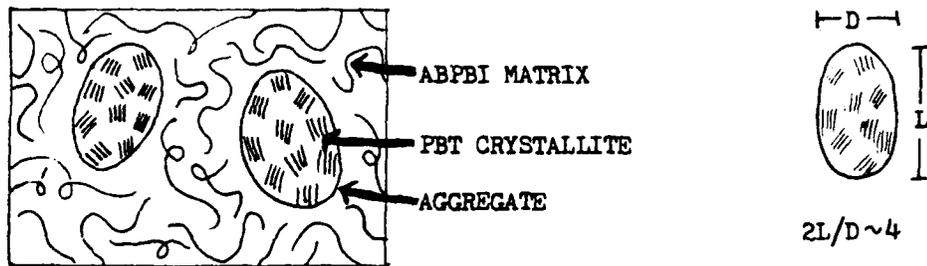


Figure 3a: Phase separated aggregates yield a low aspect ratio

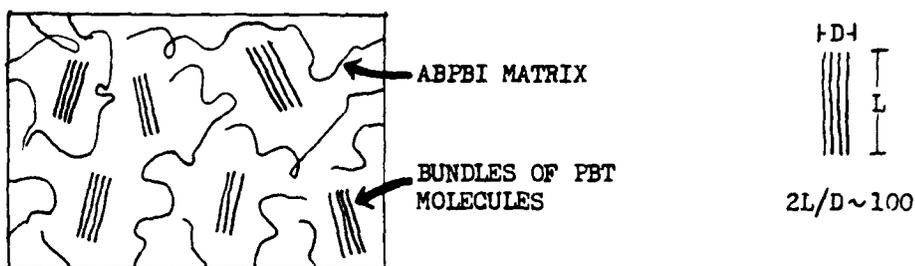


Figure 3b: Well-dispersed PBT molecules yield a higher aspect ratio

	MODULUS	TENSILE STRENGTH	% ELONGATION TO BREAK
PBT fiber	265 GPa 38 Mpsi	2800 MPa 400 ksi	1.1 %
ABPBI fiber	36 GPa 5.3 Mpsi	1090 MPa 158 ksi	5.2 %
30 PBT / 70 ABPBI C > C _{cr} fiber	11 GPa 1.6 Mpsi	310 MPa 45 ksi	13 %
30 PBT / 70 ABPBI C < C _{cr} fiber	117 GPa 17 Mpsi	1270 MPa 184 ksi	1.4 %

Figure 4: Mechanical testing results for PBT, ABPBI, and the blends above and below the critical concentration

II. OBJECTIVES

The goal of this study is to directly examine the morphologies of 30% PBT / 70% ABPBI blends, and to verify that molecular level composites have been achieved. This is done by using TEM (in dark field and selected area diffraction modes) to examine the $C > C_{cr}$ and $C < C_{cr}$ fiber forms of 30% PBT / 70% ABPBI along with PBT and ABPBI homopolymer fibers. The morphologies of these 4 fibers as determined by TEM techniques are compared in this paper. After a review of experimental procedures, the PBT and ABPBI results are given; then results and comparisons of the blends are given. Finally, conclusions and recommendations for future work are presented.

III. EXPERIMENTAL PROCEDURES

Materials processing

The details of processing are given in Hwang et al. (6). A partial summary is presented here.

The molecular weights of PBT and ABPBI were estimated at 41000 and 100000 g/mol respectively. A mixture of 97.5 vol% methanesulfonic acid (MSA) and 2.5 vol% chlorosulfonic acid (CSA) was used as the solvent. Fibers were produced by extruding solution through a spinnerette into water, then drawing between rollers and winding onto a water-immersed roll. The fibers were neutralized in NH_4OH overnight, rinsed, and heat treated in air with an additional drawing.

Films were vacuum cast as described in Husman et al. (5). Solvent from $\text{C} < \text{C}_{\text{cr}}$ solution was removed in a sublimator, resulting in $\text{C} > \text{C}_{\text{cr}}$ films; they were then dried in a vacuum oven.

TEM sample preparation

In order to use TEM, the sample thickness must be less than 1000 Å, with thinner samples giving better results. The main technique used to achieve this was detachment replication. Ion thinning (of cast films) and ultramicrotomy were used as auxiliary techniques; since these are standard machine processes, they will not be described here.

For detachment replication, a solution of about 20% collodion* in ethyl acetate (a "syrupy" consistency) is used. A layer of this solution

*obtained as Parlodion from Ted Pella, Inc.

of about 20 x 10 x 2 mm is placed by pipette onto a clean glass slide and allowed to dry for 20 to 40 minutes until it is tacky. Then pieces of the sample about 1 cm long are laid on the collodion solution with a few mm overhang (to facilitate later removal); the sample should displace a little collodion without being completely embedded (see Figure 5). After drying overnight, the sample is carefully peeled from the collodion with the hope that some microscopic fibrils remain stuck in the grooves. This can be observed under an optical microscope at a magnification of about 500 (see Figure 6). Grid-size (3mm) squares containing these fibril-filled grooves are cut and lifted off the slide. They are placed onto holey-carbon grids on a mesh screen inside a covered petri dish which is filled with ethyl acetate to the level of the screen (see Figure 7). The ethyl acetate dissolves the collodion overnight, allowing the sample to settle onto the grid. After a brief drying on filter paper, the grid is optically inspected to make sure some sample is in place and is ready for TEM study.

It is necessary for a holey-carbon film to have been formed on a copper mesh grid (see Figure 8). The carbon film provides extra support and prevents charging of the sample by the electron beam. The holey-carbon grids used in this study were prepared by Eric Matheson of Arizona State University. Typically, the film had 0.5 to 5 micron holes on a 100 to 200 mesh (per inch) copper grid.

Detachment replication works well on samples with a fibrillar nature. For less fibrillar samples such as C_{60} fiber, ordinary detachment replication gave little success. Roughing up the surface of the

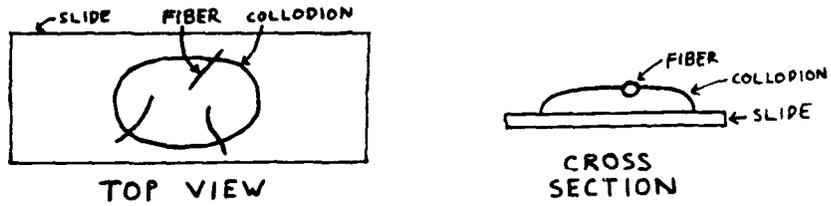


Figure 5: The placement of samples on collodion

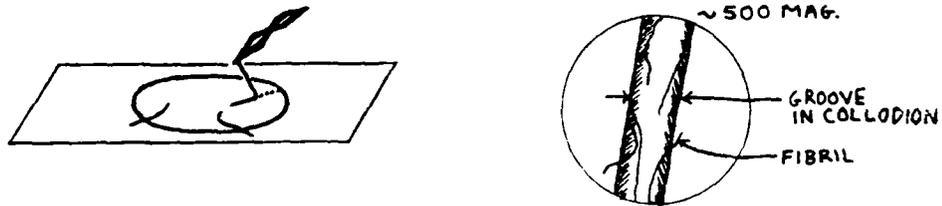


Figure 6: Peeling samples off the collodion



Figure 7: Cutting and placement on grid of collodion squares

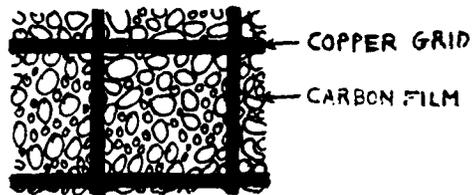


Figure 8: Magnification of typical holey-carbon grid

FIGURES 5-8: THE DETACHMENT REPLICATION PROCESS

sample with a razor blade improved the chances of detaching a thin section.

Equipment and operation

TEM of the samples was carried out on a JEOL 100CX using an accelerating voltage of 100KV or 120KV. Micrographs were taken on Kodak SO183 film.

Polymers such as PBT and ABPBI are susceptible to beam damage by the transfer of energy from the electrons as they pass through the sample. The most pronounced effect of beam damage is a loss of crystallinity. Special TEM observation techniques must be employed so that results are obtained from minimally damaged samples. The two main strategies are to keep both the beam intensity and the time the sample is exposed to the beam as low as possible.

The grids were scanned in bright field with the beam spread out at a low magnification, typically 2600X. When a promising section was located, it was examined at a higher magnification with the beam slightly converged. The selected area electron diffraction (SAED) mode was then engaged using an SAED aperture, and the sample was searched for a clear diffraction pattern. A prominent reflection of the SAED pattern was placed so that the rest of the pattern was masked by the objective aperture, and the dark field image was observed with the SAED aperture removed. In most cases the sample was translated to obtain a fresh area for dark field imaging.

IV. RESULTS AND DISCUSSION

PBT

Figure 9 is a representation of the SAED pattern of PBT heat treated spun fiber (AFTECH II). The pattern shows sharp and well-arc'd equatorial ($hk0$) spots, and high orders of meridional ($00l$) reflections. The sharpness indicates large crystallites, and the arcing and high orders indicate a high degree of orientation. PBT dark field images show crystallites typically 200 \AA by 100 \AA aligned along the fiber axis.

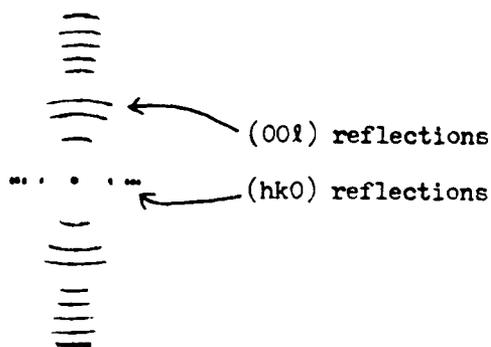


Figure 9: SAED pattern of PBT fiber

ABPBI

Figure 10 represents the SAED pattern of ABPBI heat treated spun fiber. The equatorial reflections are broader and the arcing is spread more than in PBT; there are also less meridional orders. These facts indicate that there are smaller crystallites and they are less oriented than in PBT. Dark field images confirm that ABPBI crystallites are less than 40 Å.



Figure 10: SAED pattern of ABPBI fiber

30% PBT / 70% ABPBI: C>C_{cr}

In 30 PBT / 70 ABPBI heat treated fiber which has been spun from C>C_{cr} solution, the SAED patterns vary from those with diffuse rings characteristic of unoriented polymer (Figure 11a), to those with many close meridional reflections which appear to be overlapping PBT and ABPBI patterns (Figure 11b). Similar patterns are obtained from the vacuum cast film form. Dark field images of both fiber and film show aggregates containing crystallites within another phase which may contain small (~25 Å) crystallites. This inhomogeneity causes the variation in diffraction patterns. These results are consistent with previous SEM results which found 20000 Å aggregates rich in PBT within an ABPBI-rich phase.

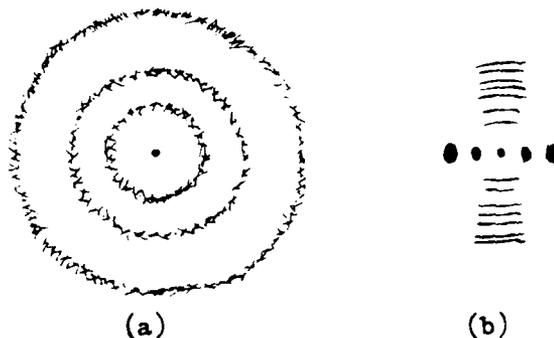


Figure 11: SAED patterns of 30 PBT / 70 ABPBI C>C_{cr} fiber

30% PBT / 70% ABPBI: C<C_{cr}

The SAED patterns observed for the C<C_{cr} heat treated fiber blend vary from distinct diffuse rings (Figure 12a) to patterns with (001) reflections that indicate both PBT and ABPBI (Figure 12b). The diffuseness of these patterns imply there are small crystallites. Dark field images shows features less than 30 Å in size which are probably PBT crystallites; however, no larger aggregates are observed. The dispersion of PBT throughout the ABPBI is much more complete in the C<C_{cr} fiber than the C>C_{cr} fiber.

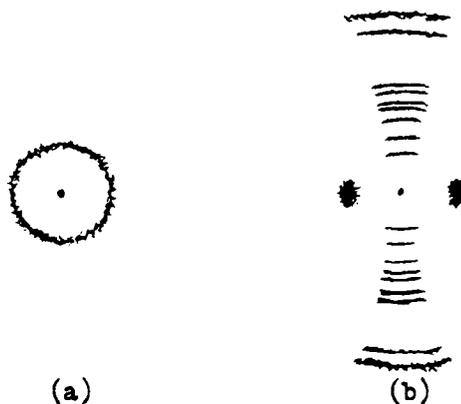


Figure 12: SAED patterns for 30 PBT / 70 ABPBI C<C_{cr} fiber

V. CONCLUSIONS

TEM dark field images and SAED patterns show that there are 20000 Å aggregates containing PBT crystallites in 30 PBT / 70 ABPBI fiber spun from C>C_{cr} solution. When fibers are spun from C<C_{cr} solution, the PBT is dispersed as crystallites less than 30 Å in size. Thus, a molecular level composite has been achieved with this fiber, which should result in higher strength and modulus than the C>C_{cr} fiber. These TEM results are in agreement with previous SEM, WAXD and mechanical testing results.

VI. RECOMMENDATIONS

A major technique which has not been used to study these blends is scanning transmission electron microscopy (STEM). STEM allows for precise control of the area of sample irradiated by the electron beam, so beam damage to the sample can be controlled. In combination with energy dispersive x-ray analysis and electron energy loss analysis, STEM could be used to analyze the composition of phase separated regions. The orientation and texture of crystallites could be obtained by using microdiffraction in STEM.

REFERENCES

1. Thomas, E.L., R.J. Farris, S.L. Hsu, S. Allen, A. Filippov, J. Minter, E. Roche, K. Shimamura, T. Takahashi and G. Venkatesh, Mechanical Properties vs. Morphology of Ordered Polymers, AFWAL-TR-80-4045, Vol. II, 1980.
2. Minter, J.R., Structural Investigations of Fibers and Films of Poly(p-phenylenebenzobisthiazole), AFWAL-TR-82-4097, 1982.
3. Price, G., X-ray Diffraction Measurements on Molecular Composites, Part I, Univ. of Dayton Research Inst.-TR-80-41, 1980.
4. Helminiak, T.E., Aromatic Heterocyclic Polymer Alloys and Products Produced Therefrom, patent application 902 525, 1978.
5. Husman, G., T. Helminiak, W. Adams, D. Wiff and C. Benner, "Molecular composites: Rodlike Polymer Reinforcing an Amorphous Polymer Matrix," ACS Symposium Series, #132, p. 203, 1980.
6. Hwang, W.F., D.R. Wiff, C.L. Benner and T.E. Helminiak, "Composites on a Molecular Level: Phase Relationships, Processing, and Properties," J. Macromol. Sci.-Phys., B22(2), p. 231, 1983.
7. Kulshreshtha, A.K., and G.E. Price, SEM Studies on ABPBI/PBT Cast Films, AFWAL-TR-82-4047, 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

ANALYTICAL REPRESENTATION OF AFTERBODY SURFACE
OF X-24C-10D REENTRY VEHICLE

Prepared by:	Dr. Madakasira V. Krishna
Academic Rank:	Associate Professor
Department and University:	Department of Mathematics South Carolina State College
Research Location:	Air Force Wright Aeronautical Lab., Computational Division
USAF Research:	Dr. W.L.Hankey
Date:	August 10, 1983
Contract No:	F49620-82-0035

ANALYTICAL REPRESENTATION OF AFTERBODY SURFACE
OF X-24C-10D REENTRY VEHICLE

ABSTRACT

As a first step in the generation of three dimensional grid for the X-24C-10D reentry vehicle the afterbody surface is represented analytically. The coordinate system is constructed by a series of coaxial cross-sections. A computer code is developed, which can be easily adapted for use with new generation of computers.

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 Flight Dynamics Laboratory, Wright Patterson AFB, Ohio. He would like to acknowledge the Laboratory, in particular the Computational Division, for its hospitality and excellent working conditions.

Finally, he would like to thank Dr. W.L. Hankey for suggesting this area of research and for his collaboration and guidance, and he would like to acknowledge many helpful discussions with Dr. J.S. Shang.

I. INTRODUCTION:

The understanding and analysis of three-dimensional fluid behavior around reentry vehicle like X-24C-10D is of considerable importance for flight applications. A vehicle like this can produce very complex flow fields. The flow is inviscidly dominated at small angles of attack. The flow field changes radically with lee side of the vehicle becoming dominated by large regions of crossflow separation, when the angle of attack is increased.

Vehicle designers need pressure, heat transfer and shear distributions to support design activities. Traditionally these have been obtained by experimental data. Significant reduction in design costs can be obtained if the calculated results based on finite difference procedures are used to verify and supplement the experimental results. Experimental results are available for X-24C and X-24C-10D models^{1,2,7}.

The forward section of X-24C-10D is same as that of X-24C. The flow field grid for the forward section of X-24C has been developed³ for use with parabolized Navier-Stokes code that already exists. The object of this project is to represent the aft section of X-24C-10D analytically as a first step in the development of three-dimensional grid for the vehicle. The available grid for the forward section of X-24C can be combined with grid for the aft section of X-24C-10D for use with Navier-Stokes code using MacCormack's explicit-implicit method⁴.

The grid generation process must achieve the following:

1. Develop accurate surface definition
2. Distribute grid points on the body surface
3. Generate a clustered smoothly varying interior mesh.

In this project the surface of the aft section of the vehicle is represented analytically using a blueprint supplied by Martin Marietta Corp., Denver, Colorado. The coordinate system is constructed by a series of

coaxial cross sections. In each cross section plane a two-dimensional grid system can be established between two control surfaces. The inner surface is to depict the body contour and the outer surface should depict the enveloping shock wave. The field points may be generated by an interpolation function. The computer code for the body surface is written such that it can be easily adopted for use with the new generation of computers such as Cyber205, Cray XMP, Cray 2.

II. OBJECTIVES

The main objective of this project was to define the afterbody portion of the surface of X-24C-10D model analytically and to write a computer code. We did not attempt to generate the complete three-dimensional grid. This is left for later investigation. In order to define the body accurately, it is split into seven sections:

1. Top wing
2. Flat top side
3. Round top of the crown
4. Middle side wing
5. Flat side
6. Side Strake
7. Bottom surface

Equations that define the surface explicitly in the blueprint have been used. Portions of the surface not defined are fitted with cubic or quadratic splines wherever necessary. Transitions from one section to the next are made smooth using filler circles to depict the surface accurately. A computer code that defines any section normal to the body axis is developed.

III. ANALYTICAL REPRESENTATION OF THE BODY

Analytical representation of the afterbody surface of the vehicle is started by choosing X axis along the body axis. The $X = \text{Constant}$ plane in which Y and Z vary is chosen as a plane normal to the body axis. The afterbody of the vehicle

is described from X = 432 inches to 641 inches.

1. Top Wing

The center vertical plane of the top wing is described by Y = 4. An ellipse is fitted to match the curve in the round part of the top from X = 502 to X = 582.88.

Slope of top = 0.6750742 (Figure 1)

$$\begin{aligned} \theta &= 34.022271 \\ x_2 &= 4.8261375 \\ x_1 &= 1.50998721 \end{aligned}$$

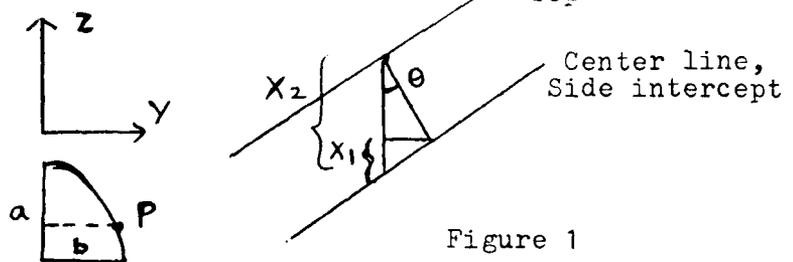


Figure 1

Intercept of 4 inch circle with side occurs at (Y,Z)=(4,4)
This corresponds to P = (4,1.5099872) on the ellipse:

$$\frac{z^2}{a^2} + \frac{y^2}{b^2} = 1 \quad \dots (1)$$

with a = 4.8261375. Coordinates of point P are used in Eq (1) to determine b. Equation of the ellipse is found to be

$$\frac{z^2}{23.291603} + \frac{y^2}{17.73623538} = 1 \quad \dots (2)$$

Equations of the top and the center of the ellipse are found using points (502,72.4) and (582.88,127), and the point slope form of the equation of a line. The equations are:

$$Z_{\text{Top}} = 0.6750742 X - 266.48724 \quad \dots (3)$$

$$Z_{\text{Center}} = 0.6750742 X - 271.31338 \quad \dots (4)$$

Using equations (2), (3) and (4), Z as a function of Y for the top of the wing from X = 502 to 582.88 is found to be:

$$Z = \sqrt{23.291603 \left\{ 1 - \frac{y^2}{17.73623538} \right\}} - 4.8261375 + Z_{\text{Top}} \quad \dots (5)$$

The top wing starts to round off at X = 579.5 and has flat section at the top. The ellipse is decreased into a filler circle of radius 1.5 inches at X = 582.88 and also its center is moved out simultaneously. A subroutine (QUADS) is used to

find the coefficients of the quadratic equation which represents the top of the wing. QUADS subroutine uses two points and the slope at one point to find coefficients for quadratic equations. Equation of top and center of ellipse are:

$$Z_T = -0.044309851 X^2 + 51.69794 X - 14954.15 \quad \dots (6)$$

$$Y = 3.7702448 - (-0.6716701 X + 393.00304) \quad \dots (7)$$

$$Z_C = -0.3240281 X^2 + 377.8226 X - 110013.3 \quad \dots (8)$$

Equation of the ellipse is:

$$\frac{Z^2}{A^2} + \frac{Y^2}{B^2} = 1 \quad \dots (9)$$

with $A = Z_T - Z_C$

and $B = -0.6716701 X + 393.00304$

$$Z = \sqrt{\left(1 - \frac{Y^2}{B^2}\right) A^2} + Z_C \quad \dots (10)$$

From $X = 582.88$ to 620.975 , $Z = 127.0$

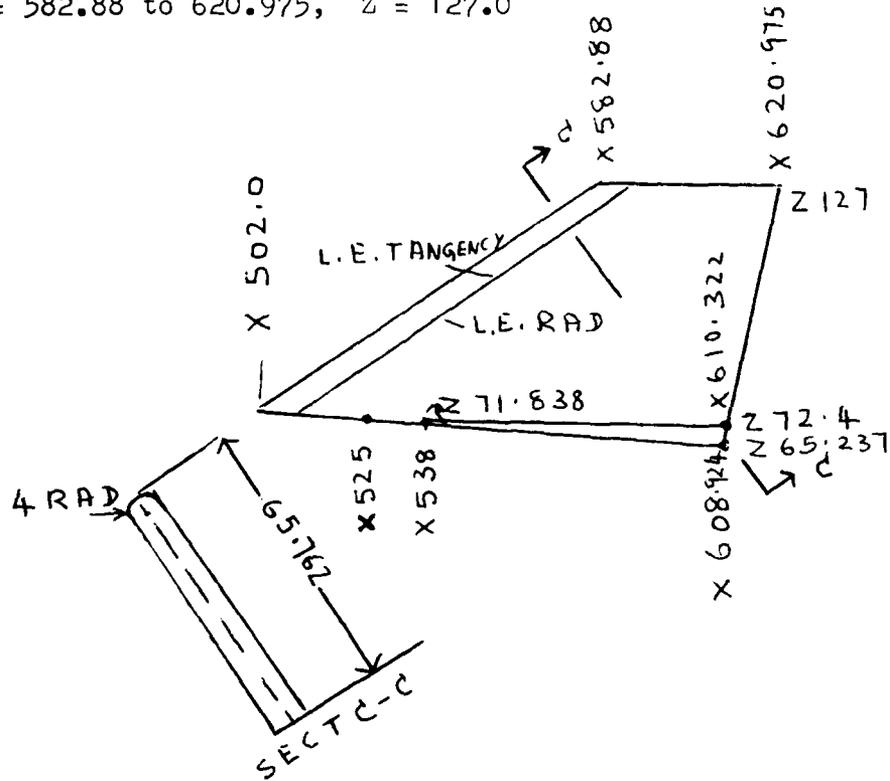


Figure 2 - The Top Wing

2. Flat Top Side

The height of the top is given by:

$$Z = 72.4 \quad \text{from } X = 432 \text{ to } 525 \quad \dots (11)$$

$$Z = -0.0033274 X^2 - 3.4937205 X - 844.70162 \quad \dots (12)$$

from $X = 525$ to 538.105

$$Z = -0.09308 X + 121.91537 \quad \text{from } X = 538.105 \text{ to } 582.88 \quad \dots (13)$$

Equations (11) and (13) are from the blueprint, coefficients of equation (12) are calculated.

Intersection of top with crown from $X = 432$ to 538.105 was found with the help of subroutine(ICSICU), that computes an interpolatory approximation to a given set of points by cubic splines with arbitrary second derivative end conditions.⁵ Points used were defined at X values: $432, 455, 480.4, 502, 525$ and 538.105 . Intersection of top with crown from $X = 578.105$ to 582.88 is a line given by:

$$Y = -0.0326626 X + 40.733813 \quad \dots (14)$$

Inner side intersection started either at $Y = 0$ or at $Y = 4$ the top wing - flat top intersection.

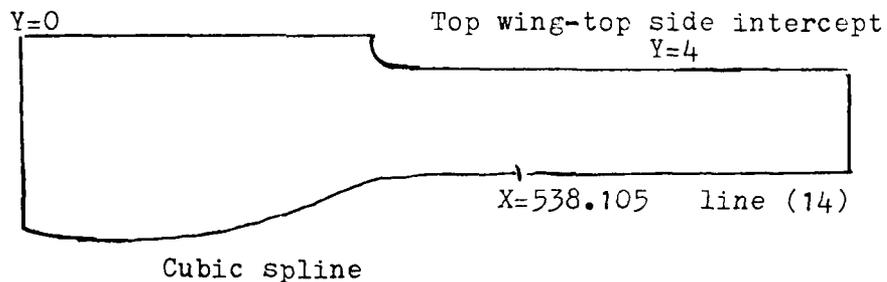


Figure 3 - Top view of Flat Top Side

3. Round Crown

From $X = 432$ to 525 the curved crown was defined using cubics generated by a subroutine(IQHSCU). This subroutine computes the coefficients for a set of cubic polynomials to be used to interpolate a set of points from a single valued function. The spline produced is continuous and

has continuous first derivative, and does not have a continuous second derivative⁶.

Four constant Y curves are first defined by IQHSCU routine with values at X = 432,450,468,485,503 and 525. Then the same routine is used again at the X location using six points, the four constant Y curves(Y =40,50,60 and 70) and Z at 49.271 and 72.4. Thus a cubic in the Y-Z plane is obtained. Points used to get the curvature are listed below.

X	Y	Z
432	70	116.3158
	60	60.6316
	50	67.7895
	40	71.3684
450	70	52.2105
	60	63.1579
	50	69.0526
	40	71.7895
468	70	54.9474
	60	64.2105
	50	69.4737
	40	71.7895
485	70	56.6316
	60	64.2105
	50	68.8421
	40	71.3684
503	70	58.8947
	60	64.00
	50	67.7895
	40	70.7365
525	70	58.7368
	60	62.9474
	50	66.1053
	40	68.6316

From X = 525 to 602 the crown is defined differently. The flap and flat region of the crown was represented in the blueprint by the equation:

$$Z = -0.096364 X - 0.26795 Y + 130.11 \quad \dots (15)$$

The region where crown was still curved was not defined in the blueprint. Subroutine IQHSCU was used again to fit the area that went flat. This area was split into two regions denoted by 'Top' and 'Bottom' in figure-4. Points used in the IQHSCU routine to fit the TOP and BOTTOM curve are listed below.

Top:	X	Y
	523.78947	40
	524.2105	36.2105
	528.42105	30.0
	557.87474	26.526316
	556.5	25.473684
Bottom:	X	Y
	523.78947	40.0
	524.21053	44.631519
	525.2	48.631579
	534.0	58.526316
	545.47368	66.947368
	549.05263	69.684211
	550.275	72.031579

Coefficients of cubic representing Y as a function of X are:

Top flap Curve-

X	Coefficients		
< 524.2105	-12.76363	2.2911569	15.77820
[524.2105, 528.42105)	-2.441058	0.2060667	5.5511636E-03
[528.42105, 537.89474)	-0.4105266	-1.6316578E-2	2.210993E-3
[537.89474, 556.5)	-0.1243674	-1.0461481E-3	2.5206586E-4

Bottom flap curve-

X	Coefficients		
[524.2105, 525.2)	6.098326	-1.363094	-0.7221882
[525.2, 534.0)	1.279444	8.6709168E-3	-2.9874141E-3
[534.0, 545.47368)	0.7398423	-1.224276E-3	7.5754062E-5
[545.47368, 549.05263)	0.7398423	5.4589706E-3	4.1642934E-4
[549.47368, 550.275)	0.7949184	1.970840	-0.5306804

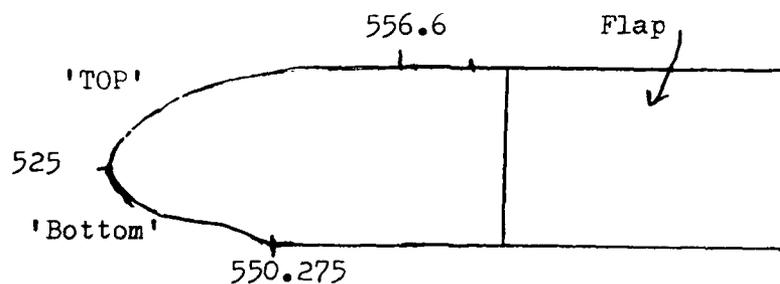


Figure 4 - Top view of Round Crown

Subroutine QUADS is used to fit quadratic curve to smooth together the flat top-crown intercept to the area where the crown goes flap. The same Subroutine is used to fit a quadratic curve to the area where flap curves down to the side or the middle wing, shown in figure 5 by arrows.

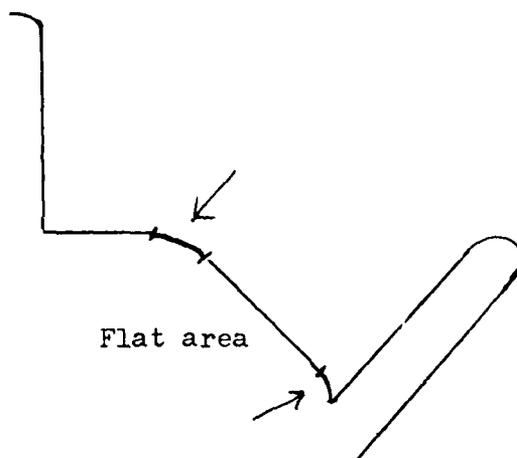


Figure 5 - Cross section showing the flat area of crown

4. Middle Wing

Inner portion of middle wing was defined using five lines marked A,B,C,D and leading edge in figure 6. For a given X, the Y and Z values were input into the subroutine IQHSCU, so that a cubic spline could fit the shape of the wing.

The points used to define the five lines are measured

off the blueprint. The equations of the five lines are:

<u>LINE</u>	<u>EQUATION</u>
LE	Y = .2472702 X - 48.68334 Z = .6750823 X - 284.30955
A	Y = .2472702 X - 53.041057 Z = .6750823 X - 287.02178
B	Y = .34 X - 108.10526 Z = .8346429 X - 381.71429
C	Y = .5143541 X - 210.02188 Z = 1.3977273 X - 710.33885
D	Y = 1.4473684 X - 753.7867 Z = 5.125 X - 2882.5447 ... (15)

Inner surface of the middle wing from line D to the trailing edge is calculated using points on the trailing edge whose equations are:

$$Z = 5.125 X - 3055.5$$

$$Y = 1.4473684 X - 794.52632 \quad \dots (16)$$

Equation of this portion ($X > 581.91$) of the inner surface was found to be:

$$Y = .2401745 X + .2355515 Z - 74.80315$$

$$Z = -1.0196263 X + 4.2453561 Y + 317.96601 \quad \dots (17)$$

To find where the mid-wing intercepts the body, Quadratic subroutine was used. Equations of segments of the intercept denoted by Q1, Q2, Q3, Q4, Q5 and Q6 in figure 5 are given below.

$$Q1 : Y = .0026577541 X^2 - 2.80692 X + 812.4464$$

$$Q2 : Y = .001240315 X^2 - 1.267643 X + 394.9114$$

$$Q3 : Z = -0.012609578 X^2 + 13.20499 X - 3398.32$$

$$Q4 : Z = -0.0045087412 X^2 + 4.848958 X - 1244.25$$

$$Q5 : Y = 0.0013911215 X^2 - 1.320870 X + 388.7149$$

$$Q6 : Z = 2.426139289392 X^2 - 2417.252317199 X + 602189.155098$$

... (18)

The mid-wing outer surface and side surface intercept is found by solving their equations.

$$\text{Outer Surface of wing: } Y = -0.06993 X + 0.26795 Z + 29.0$$

... (19)

Rear side plane: $Y = -0.36397 Z + 94.3$... (20)

Solving equations (19) and (20) equation of side-mid wing intercept was found to be,

$Z = -0.1106627 X + 103.33587$... (21)

For the outer top curved part of the wing an ellipse that was rotated 20.11681 degrees and offset by 0.6 inches was used. The equation of ellipse was found to be,

$$\frac{z^2}{a^2} + \frac{y^2}{b^2} = 1$$

... (22)

with $a = 4.1448197$ and $b = 3.1279155$

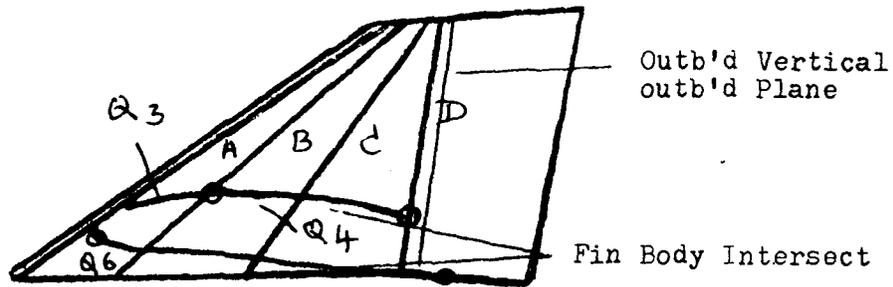
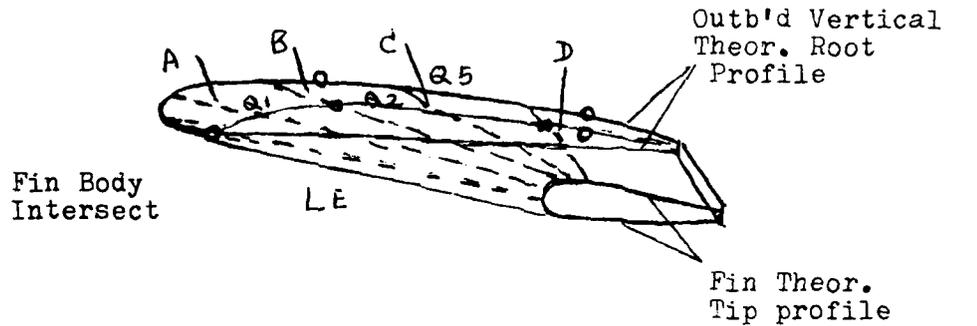


Figure 6 - The Middle Wing

The mid wing rounded off like the top wing and a routine similar to the one used for the top wing was used. The round off region is from X = 565 to 571.5. For the inner side the leading edge equation changed to a quadratic and IQHSCU was used to connect all points as before. The outer top ellipse turned into a 1.5 inch filler circle linearly and top began to flatten off. From 571.5 to 616 the top was flat with two 1.5 inch filler circles.

5. Side of Body

From X = 432 to 525 the side is assumed to be straight and is determined by using two point intercept form of equation for a line at each X cross section. Points at Z = 49.271 line and the side strake intercept are used in writing the equation of lines. From X = 525 to 580.5 an equation is given in the blueprint for the side:

$$Y = - 0.36397 Z + 94.3 \quad \dots (23)$$

A three inch filler circle is used to connect the side and the strake. A subroutine D2CR (which finds the intersection of two lines using Cramer's rule) is used to get the side-filler circle intercept, top of the strake-filler circle intercept and the center of the filler circle. When X is less than 440.99833 a cubic is used to fit the side and the strake together.

6. Strake

Equations that define the strake are provided in the blueprint.

TOP:

$$440.99833 < X \leq 550.275 , Z=0.15831 X+0.36397Y-104.28213 \quad \dots (24)$$

$$550.275 < X \leq 641 , Z=0.06993X+0.36397Y-55.64883 \quad \dots (25)$$

Bottom:

$$Z = 0.06993 X + 0.57735 Y -94.60861 \quad \dots (26)$$

Bottom surface-strake intercept:

$$432 \leq X < 485 \quad Y = -0.2213388 X + 188.69938 \quad \dots (27)$$

$$485 \leq X < 580 \quad Y = -0.1211223 X + 140.0944 \quad \dots (28)$$

$$580 \leq X < 641 \quad Y = 69.84348 \quad \dots (29)$$

A cubic spline was used to connect the side to top of the strake using ICSICU routine, when X is not greater than 440.99833. A quadratic curve was used to connect side to strake using QUADS routine when X is inbetween 440.99833 and 447. A three inch filler circle was used when X is greater than or equal to 447.

7. Bottom

Bottom surface is flat and varies in the X direction only. Its equation is supplied in the blueprint:

$$X < 485 \quad Z = -0.05786 X + 14.337 \quad \dots (30)$$

$$485 \leq X \leq 593 \quad Z = -13.72514 \quad \dots (31)$$

IV CONCLUDING REMARKS

As a first step in the grid generation process the surface of the body is represented analytically. A computer code has been developed. The program defines the Aft section of the body from 432 to 641 along the X axis. For a given X location the program defines the profile of the body in the Y-Z plane. A few sections of the body defined by the program are presented in figure 7.

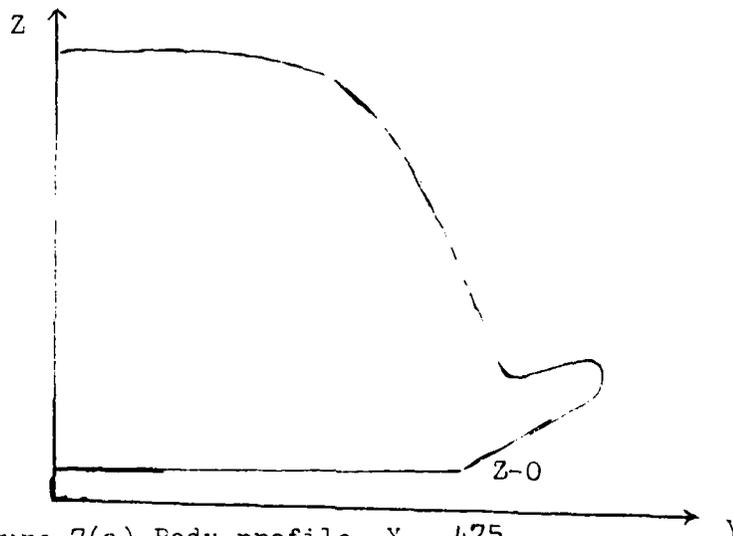


Figure 7(a) Body profile, X = 475.

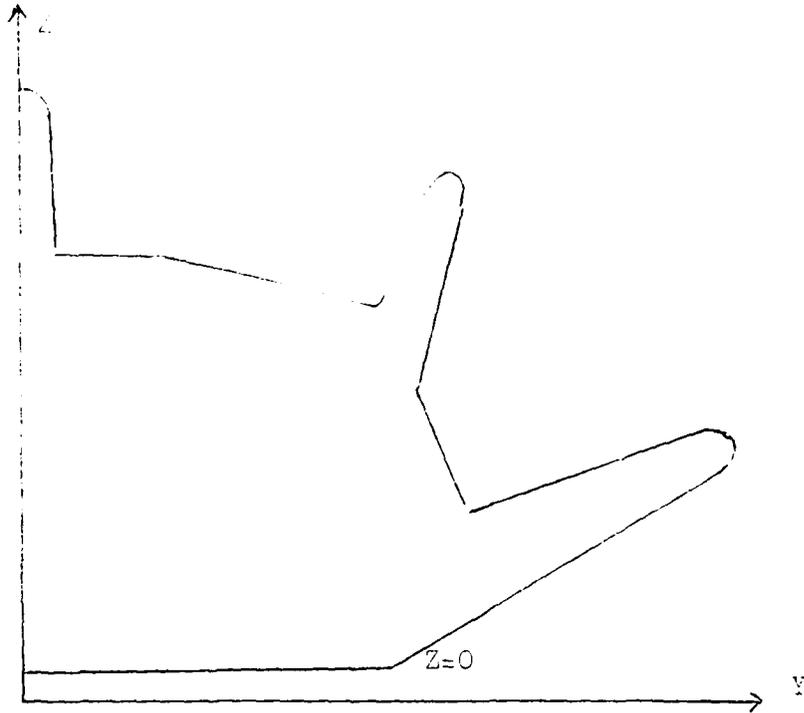


Figure 7(b), X=550

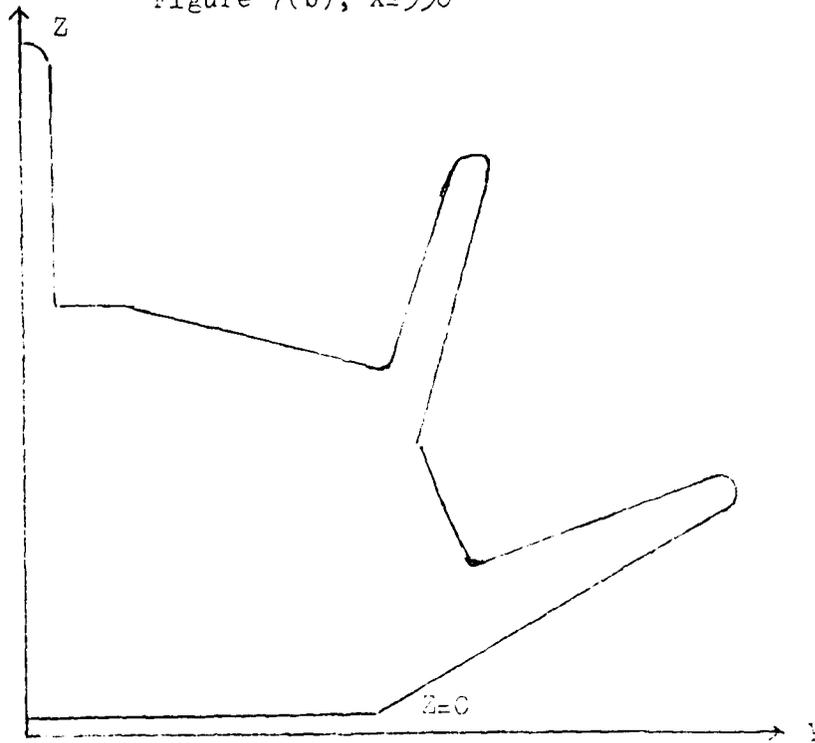


Figure 7(c) Body Profile, X=571

V. RECOMMENDATIONS

Design costs of a complex body like X-24C-10D can be significantly reduced if the available experimental results are verified and supplemented by finite difference procedures. In this project we have description of the body surface for the aft section. This can be combined with forebody description which is available. A three-dimensional grid can be generated with the bow shock as the outer surface and body surface as the inner surface, for the entire body using (1) Algebraic Method or (2) Method of Conformal grid generation. The grid generated by one of these methods can be used with Navier-Stokes code using MacCormack's explicit-implicit method. The code generated should be such that it can be used with the new generation of computers.

It is advantageous to divide the configuration into a number of circumferential segments and then grid each segment to satisfy its own particular requirements. I propose to consider one circumferential segment and generate a two-dimensional grid for that segment to demonstrate how an algebraic grid can be generated.

REFERENCES

1. D.B.Carver, " Mach Number/Flow Angle Surveys Within Bow Shock Of The X-24c Model at Mach Number 6", AEDC-TR-79-V47
2. G.D.Wannenwetsch,"Pressure Tests Of The AFFDL X-24C-10D Model at Mach Number of 1.5,3.0,5.0 and 6.0",AEDC-TR-76-92
3. S.P.Shanks,G.R.Srinivasan,W.E.Nicolet,"AFWAL Parabolised Navier-Stokes Code: Formulation and users Manual" AFWAL-TR-1982-3034
4. R.W.MacCormack," Numerical Methods for Solving the Equations of Compressible Viscous Flow," AIAA J.Vol.20 ,No.9,Sept 1980.pp.1275-1281.
5. J.Ahlberg,F.Nilson,and J.Walsh,The Theory of Splines and Their Applications,Academic Press,New York,1967.
6. H.Akima," A new method of interpolation and smooth curve fitting based on local procedures",JSCM, 17(4)1970 ,pp 589-602.
7. F.R.Gord,A.J.Brigalli,"X-24C-10D Force and Moment Test Results at Mach Numbers from 0.4 to 3.0",AFFDL-TR-78-3-FXG,Jan.1978.

MECHANICAL CHARACTERIZATION OF CARBON-CARBON
COMPOSITES FOR TURBINE ENGINE DESIGN:
A STATE-OF-THE-ART REVIEW

by

William Kyros

ABSTRACT

A review of the state-of-the-art of the mechanical characterization of carbon-carbon composites for advanced gas turbine engines was conducted. The purpose of the review was to establish a basis for defining areas of potential research in the mechanical characterization of carbon-carbon materials. The design methodology for laminated composites is delineated and clarified; the turbine engine application is presented as the context for identifying relevant materials for consideration, defining required mechanical properties and test methods as well as for analytical methods of property prediction. The general conclusion drawn from this review is that there are insufficient material properties and an inadequate data base for conducting reliable stress analyses of gas turbine engines. Further, it is observed that the computer codes currently in use to generate preliminary properties require experimental validation. Finally, areas of research are identified.