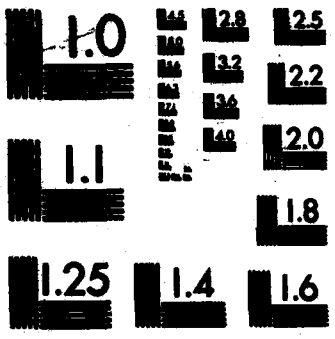
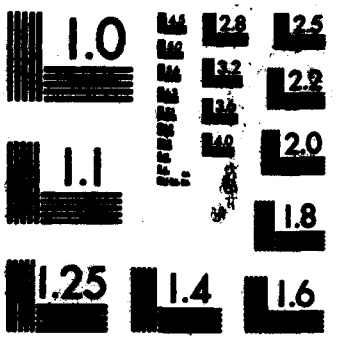
AD-A120 812 PROCEEDINGS OF THE 1982 ARMY SCIENCE CONFERENCE HELD AT 1/5
THE UNITED STATES..(U) DEPUTY CHIEF OF STAFF FOR
RESEARCH DEVELOPMENT AND ACQUISITIO.. 18 JUN 82
UNCLASSIFIED F/G 5/2 NL

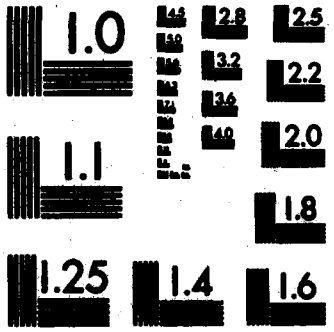




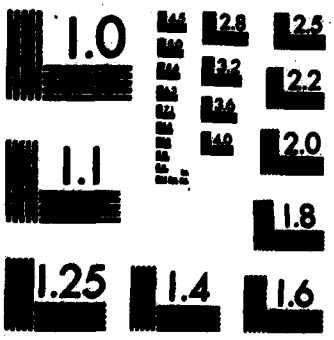
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A



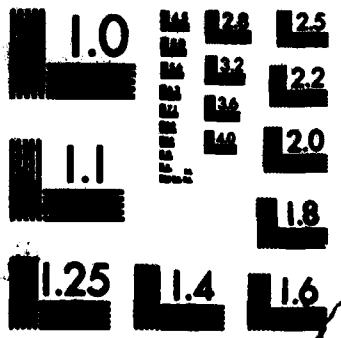
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

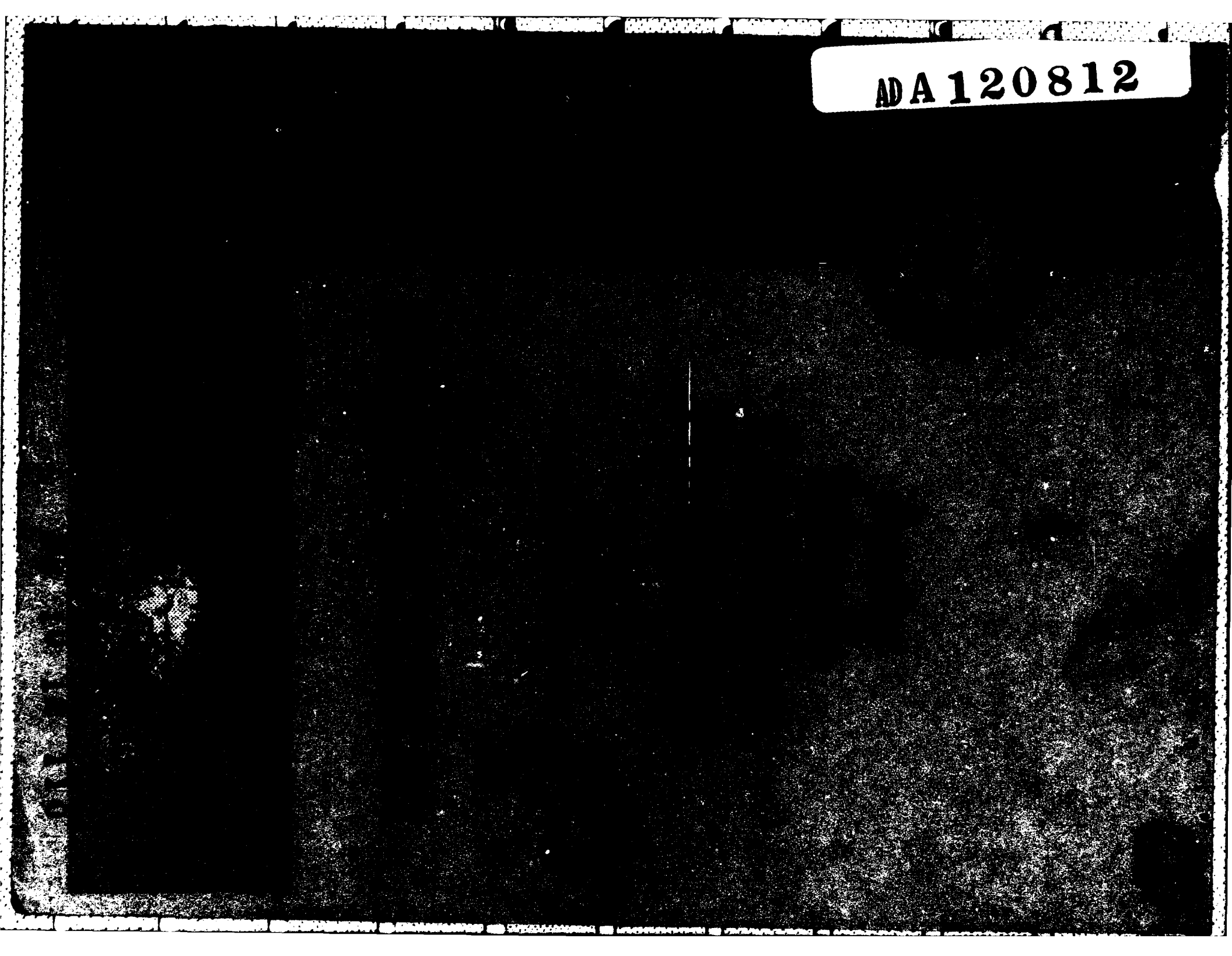


MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AD A120812



DEPARTMENT OF THE ARMY
OFFICE OF THE DEPUTY CHIEF OF STAFF
FOR RESEARCH, DEVELOPMENT, AND ACQUISITION
WASHINGTON, D.C. 20310

REPLY TO ATTENTION OF
DAMA-ARZ-D

18 June 1982

SUBJECT: Proceedings of the 1982 Army Science Conference

SEE DISTRIBUTION

1. The thirteenth in a series of Army Science Conferences was held at the United States Military Academy, 15-18 June 1982. The conference presented a cross section of the many significant scientific and engineering programs carried out by the Department of the Army and provided an opportunity for Department of the Army civilian and military scientists and engineers to present the results of their research and development efforts before a distinguished and critical audience.

2. These Proceedings of the 1982 Army Science Conference are a compilation of all papers presented at the conference and the supplemental papers that were submitted. The Proceedings consist of four volumes, with Volumes I through III unclassified, and Volume IV classified.

3. Our purpose for soliciting these papers was:

a. To stimulate the involvement of scientific and engineering talent within the Department of the Army.

b. To demonstrate Army competence in research and development.

c. To provide a forum wherein Army personnel can demonstrate the full scope and depth of their current projects.

d. To promote the interchange of ideas among members of the Army scientific and engineering community.

4. The information contained in these volumes will be of benefit to those who attended the conference and to others interested in Army research and development. It is requested that these Proceedings be placed in technical libraries where they will be available for reference.

James H. Merryman

JAMES H. MERRYMAN
Lieutenant General, GS
Deputy Chief of Staff for Research,
Development, and Acquisition

DISTRIBUTION:
Office, Under Secretary of Defense for Rsch & Engineering, Wash, DC 20310
HQDA, ODCSRDA, ATTN: DAMA-ARZ-A, Washington, DC 20310
HQDA, ODCSRDA, ATTN: DAMA-ARZ-D, Washington, DC 20310
HQDA, Deputy Chief of Staff for Logistics, Washington, DC 20310
HQDA, Deputy Chief of Staff for Operations, Washington, DC 20310
HQDA, Deputy Chief of Staff for Personnel, Washington, DC 20310
HQDA, Asst Chief of Staff for Automation and Communication, Wash, DC 20310
HQDA, Assistant Chief of Staff for Intelligence, Washington, DC 20310
HQDA, Office of the Chief of Public Affairs, Washington, DC 20310
Office of the Surgeon General, ATTN: DASG-RDZ, Washington, DC 20310
Office, Chief of Engineers, ATTN: DAEN-RDZ-A, Washington, DC 20314
Office, Chief of Engineers, ATTN: DAEN-RDZ-B, Washington, DC 20314
Office, Chief of Engineers, ATTN: DAEN-RDM, Washington, DC 20314
COMMANDERS/DIRECTORS
US Army Ballistic Missile Defense Systems Command, Huntsville, AL 35807
US Army Computer Systems Command, Ft. Belvoir, VA 22060
US Army Concepts Analysis Agency, Bethesda, MD
US Army Operational Test and Evaluation Agency, Falls Church, VA 22041
US Army Mat Dev & Readiness Cmd, ATTN: DRCDRA-ST, Alexandria, VA 22333
US Army Mat Dev & Readiness Cmd, ATTN: DRCLD, Alexandria, VA 22333
US Army Armament R&D Cmd, Dover, NJ 07801
Ballistic Research Lab, Aberdeen Proving Ground, MD 21005
Chemical Systems Lab, Aberdeen Proving Ground, MD 21010
Large Caliber Weapons Systems Lab, Dover, NJ 07801
Fire Control & Small Caliber Weapons Systems Lab, Dover, NJ 07801
US Army Aviation R&D Cmd, St. Louis, MO 63166
USARTL, Ames Research Center, Moffett Field, CA 94035
USARTL, Aeromechanics Lab, Moffett Field, CA 94035
USARTL, Applied Technology Lab, Ft. Eustis, VA 23604
USARTL, Propulsion Lab, Cleveland, OH 44135
USARTL, Structures Lab, Hampton, VA 22665
US Army Avionics R&D Activity, Ft. Monmouth, NJ 07703
US Army Aviation Engineering Flight Activity, Edwards AFB, CA 93523
US Army Communications & Electronics Cmd, Ft. Monmouth, NJ 07703
Center for Communications Systems, Ft. Monmouth, NJ 07703
Center for Tactical Computer Systems, Ft. Monmouth, NJ 07703
Center for Systems Engineering & Integration, Ft. Monmouth, NJ 07703
US Army Electronics R&D Command, Adelphi, MD 20783
Atmospheric Sciences Lab, White Sands Missile Range, NM 88002
Combat Surveillance & Target Acquisition Lab, Ft. Monmouth, NJ 07703
Electronics Technology & Devices Lab, Ft. Monmouth, NJ 07703
Electronics Warfare Lab, Ft. Monmouth, NJ 07703
Ofc of Missile Electronic Warfare, White Sands Missile Range, NM 88002
Harry Diamond Laboratories, Adelphi, MD 20783
Night Vision & Electro-Optics Lab, Ft. Belvoir, VA 22060
Signals Warfare Lab, Vint Hill Farms Station, VA 22186
US Army Missile Command, Redstone Arsenal, AL 35898
US Army Missile Lab, Redstone Arsenal, AL 35898
US Army Mobility Equipment R&D Command, Ft. Belvoir, VA 22060

US Army Natick Laboratories, Natick, MA 01760
US Army Tank-Automotive Command, Warren, MI 48090
Tank-Automotive Systems Lab, Warren, MI 48090
Tank-Automotive Concepts Lab, Warren, MI 48090
US Army Test & Evaluation Command, Aberdeen Proving Ground, MD 21005
US Army Aberdeen Proving Ground, Aberdeen Proving Gnd, MD 21005
US Army Dugway Proving Ground, Dugway, UT 84022
US Army Electronic Proving Ground, Ft. Huachuca, AZ 85613
US Army Tropic Test Center, APO Miami 34004
US Army White Sands Missile Range, White Sands Missile Range, NM 88002
US Army Yuma Proving Ground, Yuma, AZ 85364
US Army Research Office, Research Triangle Park, NC 27709
US Army Materials & Mechanics Research Center, Watertown, MA 02172
US Army Human Engineering Lab, Aberdeen Proving Ground, MD 21005
US Army Mat Systems Analysis Activity, Aberdeen Proving Ground, MD 21005
US Army Foreign Science & Tech Center, Charlottesville, VA 22901
US Army Research, Development & Standardization Group (Europe)
US Army Satellite Communications Agency, ATTN: DRCPM-SC-11, Fort Monmouth, NJ 07703
US Army Training & Doctrine Cmd, ATTN: ATCD, Ft. Monroe, VA 23651
US Army Health Services Command, Ft. Sam Houston, TX 78234
Institute of Surgical Research, Ft. Sam Houston, TX 78234
US Army Medical R&D Command, Fort Detrick, Frederick, MD 21701
Aeromedical Research Lab, Ft. Rucker, AL 36362
Institute of Dental Research, WRAMC, Washington, DC 20012
Letterman Army Inst of Research, Presidio of San Francisco, CA 94129
Medical Bioengineering R&D Lab, Frederick, MD 21701
Medical Rsch Inst of Chemical Defense, Aberdeen Proving Gnd, MD 21010
Medical Rsch Inst of Environmental Medicine, Natick, MA 01760
Medical Rsch Inst of Infectious Diseases, Frederick, MD 21701
Walter Reed Army Inst of Research, Washington, DC 20012
Walter Reed Army Medical Center, Washington, DC
Armed Forces Radiobiology Rsch Inst, Bethesda, MD 20814
US Army Cold Regions Rsch & Eng Lab, Hanover, NH
US Army Construction Eng Rsch Lab, Champaign, IL
US Army Engineer Topographic Labs, Ft. Belvoir, VA 22060
US Army Engineer Waterways Experiment Station, Vicksburg, MS 39180
US Army Rsch Inst for the Behavioral & Social Sciences, Alex, VA 22333
ARI Field Unit, Ft. Benjamin Harrison, IN 46216
ARI Field Unit, Ft. Sill, OK 73503
ARI Field Unit, Presidio of Monterey, CA 93940

COMMANDANTS:
Academy of Health Sciences, ATTN: HSHA-CDM, Ft. Sam Houston, TX 78234
US Army Air Defense School, ATTN: ATSA-CDM-A, Ft. Bliss, TX 79916
US Army Armor Center, ATTN: ATZK-CD-SD, Ft. Knox, KY 40121
US Army Chemical School, ATTN: ATZN-CM-CS, Ft. McClellan, AL 36205
US Army Engineer School, ATTN: ATZA-CD, Ft. Belvoir, VA 22060
US Army Field Artillery School, Ft. Sill, OK 73503
US Army Infantry School, Ft. Benning, GA 31905
US Army Intelligence Center and School, Ft. Huachuca, AZ 85613

US Army Logistics Center, Ft. Lee, VA 23801
US Army Ordnance Center and School, Aberdeen Proving Gnd, MD 21005
US Army Signal Center, Ft. Gordon, GA 30905
US Army Transportation School, ATTN: ATSP-CD-CS, Ft. Eustis, VA 23604

SUPERINTENDENT:

US Army Military Academy, ATTN: Technical Library, West Point, NY 10996

COPIES FURNISHED:

Defense Advanced Research Projects Agency, Arlington, VA 22209
Defense Technical Information Center, Alexandria, VA 22209
The Army Library, ATTN: ANRAL, Washington, DC
Chief, US Army Field Office, HQ AFSC/SDOA, Andrews AFB, MD 20331
Cdr, HQ Fort Huachuca, ATTN: Tech Reference Div, Ft. Huachuca, AZ 85613
Cdr, 1st COSCOM, HHC/SOTI, Ft. Bragg, NC 28303
Naval Air Systems Command (Code 310-A), Washington, DC 20361
Naval Research Library (Code 2627), Washington, DC 20361
Office of Naval Research (Code 102), Arlington, VA 22217
HQ US Marine Corps (Code RD-1), Washington, DC 20380
Air Force Systems Command, Andrews AFB, Washington, DC 20331
Lawrence Livermore Lab, Univ of California, Livermore, CA 94550
Los Alamos Scientific Lab, Los Alamos, NM 87544
Southwest Research Institute, San Antonio, TX 78228
United Nations Library, New York, NY 10017

PROCEEDINGS

OF THE

1982 ARMY SCIENCE CONFERENCE

UNITED STATES MILITARY ACADEMY

WEST POINT, NEW YORK

15-18 JUNE 1982

VOLUME II

Principal Authors H through N

Accession For
NTIS GRA&I
DTIC TAB
Unannounced
Justification
By
Distribution/
Availability Codes
Dist | Avail and/or Special
A

DTIC COPY INSPECTED 2

TABLE OF CONTENTS

PROCEEDINGS OF THE 1982 ARMY SCIENCE CONFERENCE

Author	Title	Vol	Page
Adams, Nelson L.	See Rickett, Daniel L.	3	117
Albright, Ronald	See Patton, John F.	3	1
Albritton, Gayle E.	See Cost, Van T.	1	267
Alexander, Donald	See Patton, John F.	3	1
Allan, Barry D.	A Gelled Oxidizer for Tactical Missiles	4	1
Allan, Barry D.	See Ayers, O. E.	1	17
Alster, Jack	See Gilbert, Everett E.	4	87
Althouse, Mark L.	See Cox, Christopher S.	4	61
Antle, Lloyd G.	See Goicoechea, Ambrose	1	475
Ashley, Paul R.	See Duthie, Joseph G.	1	341
Ashman, William P. Lewis, James H. Poziomek, Edward J.	A Decision Tree for Chemical Detection Application	1	1
Atchison, Valerie L.	See Voorhees, James W.	3	425
AuCoin, Thomas R.	See Malik, Roger J.	2	309
Ayers, O. E. Allan, B. D. Melvin, W. S. Murfree, J. A. Wharton, W. W. Marcucci, P. J.	Laser Induced Polymerization Reactions in Solid Propellant Binders	1	17
Baba, Anthony J. Share, Stewart	Thermal Radiation Effects on Fiber Optics and Conventional Metallic Communications Cable	4	15
Babbitt, Richard W.	See Borowick, John	1	113

Author	Title	Vol	Page
Bagby, D. Gordon	The Engineer Modeling Study	1	31
Barditch, Irving F.	See Cox, Christopher S.	4	61
Barr, Thomas A., Jr. McKnight, W. B.	A New Midinfrared Laser in Hydrogen	1	47
Barsam, Helena F.	See Simutis, Zita M.	3	273
Basso, Michael J.	See McCreery, M. J.	2	357
Bates, Calvin	See Fischer, Paul	1	399
Baussus von Luetzow, H.	Gravity Vector Determination from Inertial and Auxiliary Data and Potential Utilization of Generated Vector Component Information	1	61
Bayha, William	See Borowick, John	1	113
Beatrice, Edwin S.	See Lund, David J.	2	279
Beatrice, Edwin S.	See Randolph, David I.	3	95
Beatrice, Edwin S.	See Schuschereba, Stephen	3	173
Beatrice, Edwin S.	See Stuck, Bruce E.	3	371
Beatrice, Edwin S.	See Zwick, H.	3	449
Bedynek, Julius	See Patton, John F.	3	1
Berg, Richard	See Lunardini, Virgil J.	2	263
Berkhimer, Karl	See Sturdivan, Larry M.	4	209
Bertin, John J. Goodyear, Richard L.	Multiple Launch Rocket System (MLRS) Fuze	1	77
Bexon, Roy	See Sturdivan, Larry M.	4	209
Bhatt, Ramakrishna T.	Thermal Fatigue Behavior of FP Alumina/Magnesium Composites	1	89
Bieberich, M. J.	See Singler, R. E.	3	297
Binder, Michael	See Gilman, Sol	1	447

Author	Title	Vol	Page
Bingham, Gene J.	The Aerodynamic Influences of Rotor Blade Taper, Twist, Airfoils and Solidity on Hover and Forward Flight Performance	1	99
Binn, Leonard N.	See Lemon, Stanley M.	2	249
Black, Edward D.	See Porter, William L.	3	31
Bloom, Kenneth R.	See Zwick, H.	3	449
Borowick, John Bayha, William Stern, Richard A. Babbitt, Richard W.	Dielectric Waveguide Antenna	1	113
Bowden, Charles M.	MICOM Program in Light Control by Light	1	125
Brandt, W. E.	See Henchal, Erik A.	2	61
Brodman, Bruce W. Devine, Michael P.	Microbial Attack of Nitrocellulose	1	141
Brown, Douglas R.	An Empirical Model for Near Millimeter Wave Snow Extinction and Backscatter	1	147
Buser, R. G.	See Rohde, R. S.	4	177
Bushell, M. Manriquez, R. Merkel, G. Scharf, W. D.	Aurora Flash X-Ray Facility as a Source-Region EMP Simulator	1	159
Campana, Joseph E.	See Friedman, Joseph E.	1	415
Campbell, Donn V.	Control of Parasitic Currents on Radiating Systems	1	175
Campi, Morris	Design of Microstrip Linear Array Antennas by Computer	1	187
Carchedi, Steven	See Groff, John N.	4	97
Caslavsky, Jaroslav L. Viechnicki, Dennis J.	Resolution of Factors Responsible for Difficulty in Growing Single Crystals of YAG	1	201

Author	Title	Vol	Page
Chandra, S.	See Rohde, R. S.	4	177
Chapin, Charles C. Conley, James H. Jamison, Robert G.	Generation of a System Which Exhibits an Isogenic in the Conversion of Army Vehicles to Silicone Brake Fluid	1	209
Chu, Shih C. Steiner, James	A More Rational Approach to the Stress Analysis of Projectiles	1	225
Chubb, Douglas W.J.	Natural Language Translation in a Noisy Domain Using Occam's Razor as a Control Paradigm	4	31
Cohn, Stephen L. Pena, Ricardo	Munition Expenditure Model Verification: KWIK Phase I	1	239
Collins, John G. Link, Lewis E.	A Quantitative Evaluation of Groundwater Resources in Southwest Asia	4	47
Commings, Hubert	Hardening of Armored Vehicle Suspension Components	1	255
Conklin, James J.	See McCreary, M. J.	2	357
Conley, James H.	See Chapin, Charles C.	1	209
Cost, Van T. Albritton, Gayle E.	Response of [illegible] Shelter Models to Static and Dynamic Loading	1	267
Cox, Christopher S. Althouse, Mark L. Gerditch, Irving F.	Point and Remote Detection of Biological Aerosols Using Infra-Red	4	61
Crumley, Lloyd M. Schwein, Robert C.	Modeling Weapons Crew Performance	1	283
Davenport, Otis A.	CCM-Alternatives to Soviet Lasers Targeted Against the Helicopter Force	4	77
DeBellis, William S.	See Poston, Alan M.	3	47
DeLuca, Eugenio	See Petfat, Joseph J.	4	139
Deems, A. J.	See Singler, R. E.	3	297
DeRenzo, Frank Zegna, Angelo I.	The Development and Evaluation of Hemoglobin Solution as a Blood Substitute	1	299

Author	Title	Vol	Page
DeVenuto, Frank Zegna, Angelo I.	Field Production of Purified Sterile Water from Available Water Sources by Using a Portable Apparatus	1	315
Devine, Michael P.	See Brodman, Bruce W.	1	141
Dixon, R. E.	See Rohde, R. S.	4	177
Drolet, Anne M.	See Porter, William L.	3	31
DuBois, John	See Powanda, Michael C.	3	63
Dunn, D. A.	See Singler, R. E.	3	297
Dunn, William P.	A New Analysis of the Unwinding Ribbon as a Delayed Arming Device	1	325
Duthie, Joseph G. Ashley, Paul R. Upatnieks, Juris Liu, H. K.	Photonic Seeker Development	1	341
Elder, Alexander S. Walbert, James N. Zimmerman, Kathleen L.	Stationary and Traveling Loads in a Hollow Cylinder	1	353
Emery, Clarence E.	See Wolfe, Alan D.	3	435
Espeland, Richard H.	See Schwering, Felix K.	3	201
Essenwanger, Oskar M.	Turbulence Analysis by Use of the Fast Fourier Transform	1	367
Figucia, F. Williams, C. Kirkwood, B. Koza, W.	Mechanisms of Improved Ballistic Fabric Performance	1	383
Fischer, Paul Bates, Calvin Hartley, Joseph	A Novel Beam Bunching Concept for Millimeter Wave Tubes	1	399
Friedman, Melvin H. Campana, Joseph E. Yergey, Alfred L.	New Viewpoints in Mass Filter Design	1	415
Gall, Kenneth J.	See Rickett, Daniel L.	3	117

Author	Title	Vol	Page
Garinther, John M.	See Poston, Alan M.	3	47
Garn, Lynn E.	See Sharp, Edward J.	3	227
Garnett, Lamont W.	See Squire, Walter H.	3	341
Garrett, Paul F., Jr.	See Poston, Alan M.	3	47
Gauss, Arthur, Jr.	A New Type of EM Wave Absorbing Coating	1	431
Gentry, M. K.	See Henchal, Erik A.	2	61
Gilbert, Everett E. Sollott, Gilbert P. Alster, Jack Sandus, Oscar Slagg, Norman	Toward More Powerful Explosives--Polynitro Polyhedranes	4	87
Gilman, Sol Wade, William, Jr. Binder, Michael	High Energy Sulfuryl Chloride Batteries	1	447
Gladen, Curtis L. Link, Lewis E.	Thermal Camouflage of Fixed Installations: Project TIREX	1	459
Goicoechea, Ambrose Krouse, Michael R. Antle, Lloyd G.	An Approach to Risk and Uncertainty in Benefit-Cost Analysis of Water Resources Projects	1	475
Goicoechea, Ambrose Krouse, Michael R.	A Computer Based Interactive Model for Industrial Land Use Forecasting	1	489
Goldman, Ralph F.	Microclimate Cooling for Combat Vehicle Crewmen	1	503
Goodrick, Thomas F.	Development of Methods for Assessment of Gliding Parachute Applications	1	517
Goodyear, Richard L.	See Bertin, John J.	1	77
Gregory, Frederick H. Pearson, Richard J.	Analytical and Experimental Studies of the Response of a Cylinder to Nuclear Thermal/Blast Loads	1	531
Griffith, James R.	See Klinger, Liliana	2	223

Author	Title	Vol	Page
Groff, John N. Mirabelle, Rosemary M. Carchedi, Steven	Armored Combat Vehicle Technology Using the HITPRO/DELACC Methodology	4	97
Groves, Michael G.	See Twartz, John C.	3	411
Hagman, Joseph D.	Maintaining Motor Skill Performance	2	1
Hardy, G. David, Jr.	See Hiller, Jack H.	2	75
Harley, Samuel F.	Data Compression for Transient Measurements	2	17
Harris, Paul Presles, Henri-Noel	The Shock Front Rise Time in Water	2	33
Hartley, Joseph	See Fischer, Paul	1	399
Heavey, Karen R.	See Nietubicz, Charles J.	2	425
Heberlein, David C.	Detonation of Rapidly Dispersed Powders in Air	2	47
Heise, Carl J.	See Mando, Michael A.	4	111
Henchal, Erik A. McCown, J. M. Gentry, M. K. Brandt, W. E.	Rapid Identification of Dengue Virus Serotypes Using Monoclonal Antibodies in an Indirect Immunofluorescence Test	2	61
Herren, Kenneth A.	See Johnson, John L.	2	199
Hiller, Jack H. Hardy, G. David Meliza, Larry L.	Design of a Small Unit Drill Training System	2	75
Hoidale, Glenn B.	See Walters, Donald L.	4	239
Hollenbaugh, D. D.	Quantification of Helicopter Vibration Ride Quality Using Absorbed Power Measurements	2	91
Howe, Philip M. Kiwan, Abdul R.	A Theoretical Study of the Propagation of a Mass Detonation	2	109
Hsiesh, Jen-Shu	See McCreery, M. J.	2	357
Hubbard, Roger W. Mager, Milton Kerstein, Morris	Water as a Tactical Weapon: A Doctrine for Preventing Heat Casualties	2	125

Author	Title	Vol	Page
Hutchings, Thomas D.	Analysis of Small Caliber Manuever-able Projectile (SCMP) Concepts for Helicopter and Air Defense Applications	2	141
Huxsoll, David L.	See Twartz, John C.	3	411
Hynes, John N. Jimarez, David S.	Heuristic Information Processing as Implemented in Target Motion Resolution Analysis of Radar Data	2	157
Hynes, Thomas V.	Attenuation of High Intensity Reradiated Light by Photochromic Glass	2	171
Iafrate, Gerald J.	Utilization of Quantum Distribution Functions for Ultra-Submicron Device Transport	2	177
Jamison, Keith A.	See Thomson, George M.	3	385
Jamison, Robert G.	See Chapin, Charles C.	1	209
Jenkins, Thomas	See Lunardini, Virgil J.	2	263
Jenkinson, Howard A. Zavada, John M.	CO_2 Laser Waveguiding in GaAs MBE Layers	2	189
Jimarez, David S.	See Hynes, John N.	2	157
Johnson, John L. Herren, Kenneth A. Morgan, Robert L. Tanton, George A.	Active Imaging of Range Targets at 1.2 Millimeters	2	199
Johnson, Robert A.	See Schwering, Felix K.	3	201
Kane, P. J.	See Singler, R. E.	3	297
Kapsalis, John G.	See Porter, William L.	3	31
Kerstein, Morris	See Hubbard, Roger W.	2	125
Kirkwood, B.	See Figucia, F.	1	383
Kittleson, John K. Yu, Yung H.	Holographic Interferometry Technique for Rotary Wing Aerodynamics and Noise	2	209

Author	Title	Vol	Page
Kiwan, Abdul R.	See Howe, Philip M.	2	109
Klinger, Liliana Griffith, James R.	Fluoropolymer Barriers to Stress Corrosion in Optical Fibers	2	223
Knudson, Gregory B.	See Mikesell, Perry	2	385
Koulouris, T. N.	See Singler, R. E.	3	297
Koza, W.	See Figucia, F.	1	383
Kronenberg, Stanley	Tactical Gamma and Fast Neutron Dosimetry with Leuko Dye Optical Waveguides	2	235
Krouse, Michael R.	See Goicoechea, Ambrose	1	475
Krouse, Michael R.	See Goicoechea, Ambrose	1	489
Kunkel, Kenneth E.	See Walters, Donald L.	4	239
LeDuc, James	See Lemon, Stanley M.	2	249
Lee, H.	See Singler, R. E.	3	297
Lemon, Stanley M. LeDuc, James Binn, Leonard N.	Isolation of Hepatitis A Virus from the New World Owl Monkey: A New Animal Model for Hepatitis A Infections	2	249
Lewis, Danny H.	See Setterstrom, Jean A.	3	215
Lewis, James H.	See Ashman, William P.	1	1
Lieberman, Michael M.	See Powanda, Michael C.	3	63
Link, Lewis E.	See Collins, John G.	4	47
Link, Lewis E.	See Gladen, Curtis L.	1	459
Liu, H. K.	See Duthie, Joseph G.	1	341
Lunardini, Virgil J. Berg, Richard McGaw, Richard Jenkins, Thomas Nakano, Yoshisuke Oliphant, Joseph O'Neill, Kevin Tice, Allan	The Mobility of Water in Frozen Soils	2	263

Author	Title	Vol	Page
Lund, David J. Beatrice, Edwin S. Schuschereba, Steven	Bioeffects Data Concerning the Safe Use of GaAs Laser Training Devices	2	279
Lund, David J.	See Stuck, Bruce E.	3	371
Lund, David J.	See Zwick, H.	3	449
Machuca, Raul	Computer Detection of Low Contrast Targets	2	293
Mager, Milton	See Hubbard, Roger W.	2	125
Malik, Roger J. AuCoin, Thomas R. Ross, Raymond L. Savage, Robert O.	The Planar Doped Barrier: A New Class of Electronic Devices	2	309
Mando, Michael A. Heise, Carl J.	Development of Armature Insulation Technique for Compact, High Power Alternators	4	111
Manriquez, R.	See Bushell, M.	1	159
Marchese, Vincent P.	IR Algorithm Development for Fire and Forget Projectiles	2	325
Marchionda, Kristine M.	See Voorhees, James W.	3	425
Marcucci, P. J.	See Ayers, O. E.	1	17
Martel, C. James	Development of a New Design Procedure for Overland Flow System	2	341
McCown, J. M.	See Henchal, Erik A.	2	61
McCreery, M. J. Swenberg, Charles E. Basso, Michael J. Conklin, James J. Hsieh, Jen-Shu	Biologic Dosimetry for Nuclear Environments by Electron Paramagnetic Resonance (EPR) Methods	2	357
McDonough, John H.	The Effects of Nerve Agents on Behavioral Performance and Their Modification with Antidotes and Antidote Combinations	2	371
McGaw, Richard	See Lunardini, Virgil J.	2	263

Author	Title	Vol	Page
McKnight, W. B.	See Barr, Thomas A., Jr.	1	47
Meliza, Larry L.	See Hiller, Jack H.	2	75
Melvin, W. S.	See Ayers, O. E.	1	17
Merkel, G.	See Bushell, M.	1	159
Meyers, William E.	See Setterstrom, Jean A.	3	215
Mikesell, Perry Knudson, Gregory B.	Plasmids of Legionella Species	2	385
Miller, Miles C.	Flight Instabilities of Spinning Projectiles Having Non-Rigid Payloads	2	393
Mirabelle, Rosemary	See Groff, John N.	4	97
Morgan, Robert L.	See Johnson, John L.	2	199
Murfree, J. A.	See Ayers, O. E.	1	17
Murphy, Newell R.	Armored Combat Vehicle Technology (ACVT) Program Mobility/Agility Findings	2	409
Nakano, Yoshisuke	See Lunardini, Virgil J.	2	263
Nietubicz, Charles J. Sturek, Walter B. Heavey, Karen R.	Computations of Projectile Magnus Effect at Transonic Velocities	2	425
Nolan, Raymond V.	Explosives Detection Systems Employing Behaviorally Modified Rats as Sensory Elements	2	441
Nomiyama, N. T.	See Rohde, R. S.	4	177
Norton, M. C.	See Rohde, R. S.	4	177
Obert, Launne P.	See Ratches, James A.	4	163
O'Connell, Robert L.	Soviet Land Arms Acquisition Model	4	125
O'Neill, Kevin	See Lunardini, Virgil J.	2	263
Oliphant, Joseph	See Lunardini, Virgil J.	2	263

Author	Title	Vol	Page
Patton, John F. Vogel, James A. Bedynek, Julius Alexander, Donald Albright, Ronald	Aerobic Power and Coronary Risk Factors in 40 and Over Aged Military Personnel	3	1
Pearson, Richard J.	See Gregory, Frederick H.	1	531
Pena, Ricardo	See Cohn, Stephen L.	1	239
Phelps, Ruth H.	Expert's Use of Information: Is It Biased?	3	17
Porter, William L. Kapsalis, John G. Wetherby, Ann Marie Drolet, Anne M. Black, Edward D.	A Rationale for Evaluation and Selection of Antioxidants for Protection of Ration Items of Different Types	3	31
Poston, Alan M. Garrett, Paul F., Jr. DeBellis, William B. Reed, Harry J. Garinther, John M.	Human Engineering Laboratory Aviation Supply Class III/V Materiel (HELAVS III/V) Field Test	3	47
Powanda, Michael C. DuBois, John Villarreal, Ysidro Lieberman, Michael M. Pruitt, Basil A., Jr.	Biochemical Indicators of Infection and Inflammation in Burn Injury	3	63
Poziomek, Edward J.	See Ashman, William P.	1	1
Presles, Henri-Noel	See Harris, Paul	2	33
Prichard, Dorothy A.	See Wolfe, Alan D.	3	435
Prifti, Joseph J. DeLuca, Eugenio	Development of Ballistic Spall-Suppression Liner for M113 Armored Personnel Carrier	4	139
Pruitt, Basil A., Jr.	See Powanda, Michael C.	3	63
Ramsley, Alvin O. Yeomans, Walter G.	Psychophysics of Modern Camouflage	3	79
Randers-Pehrson, Glenn	Nonaxisymmetric Anti-Armor Warheads	4	153

Author	Title	Vol	Page
Randolph, David I. Schmeisser, Elmar T. Beatrice, Edwin S.	Laser Flash Effects: A Non-Visual Phenomenon?	3	95
Randolph, Thomas C.	See Rickett, Daniel L.	3	117
Ratches, James A. Obert, Luanne P.	FLIR/MMW Radar vs FLIR Alone	4	163
Reed, Harry J.	See Poston, Alan M.	3	47
Reed, Lockwood W.	Voice Interactive Systems Technology Avionics (VISTA) Program	3	107
Rickett, Daniel L. Adams, Nelson L. Gall, Kenneth J. Randolph, Thomas C. Rybczynski, Siegfried	Differentiation of Peripheral and Central Actions of Soman-Produced Respiratory Arrest	3	117
Rohani, Behzad	Probabilistic Solution for One-Dimensional Plane Wave Propagation in Homogeneous Bilinear Hysteretic Materials	3	131
Rohde, R. S. Buser, R. G. Norton, M. C. Dixon, R. E. Nomiyama, N. T. Chandra, S.	Laser Technology for Identification on the Modern Battlefield	4	177
Rokkos, Nikolaus	See Schwering, Felix K.	3	201
Ross, Raymond L.	See Malik, Roger J.	2	309
Roth, John A.	Measured Effects of Tactical Smoke and Dust on Performance of a High Resolution Infrared Imaging System	3	147
Rybczynski, Siegfried	See Rickett, Daniel L.	3	117
Salomon, Mark	Properties of $SOCl_2$ Electrolyte Solutions	3	163
Sandus, Oscar	See Gilbert, Everett E.	4	87
Saunders, J. Peter	See Twartz, John C.	3	411

Author	Title	Vol	Page
Savage, Robert O.	See Malik, Roger J.	2	309
Scharf, W. D.	See Bushell, M.	1	159
Schmeisser, Elmar T.	See Randolph, David I.	3	95
Schuschereba, Stephen Beatrice, Edwin S.	Autoradiography of Primate Retina After Q-Switched Ruby Laser Radiation	3	173
Schuschereba, Stephen	See Lund, David J.	2	279
Schwalm, Robert C.	See Crumley, Lloyd M.	1	283
Schwartz, Paul M.	Methods for Evaluating Gun-Pointing Angle Errors and Miss Distance Parameters for an Air Defense Gun System	3	189
Schwering, Felix K. Johnson, Robert A. Rokkos, Nikolaus Whitman, Gerald M. Violette, Edmond J. Espeland, Richard H.	Effects of Vegetation and Battlefield Obscurants on Point-to-Point Transmission in the Lower Millimeter Wave Region (30-60 GHz)	3	201
Selvaraju, G.	See Twartz, John C.	3	411
Setterstrom, Jean A. Tice, Thomas R. Lewis, Danny H. Meyers, William E.	Controlled Release of Antibiotics from Biodegradable Microcapsules for Wound Infection Control	3	215
Share, Stewart	See Baba, Anthony J.	4	15
Sharp, Edward J. Garn, Lynn E.	Electrical Properties of Heated Dielectrics	3	227
Sheldon, William J.	The Development and Production of the Tungsten Alloy M74 Grenade	4	193
Shirai, Akira	See Twartz, John C.	3	411
Shuely, Wendel J.	A New Interactive, Computer-Controlled Method for Investigating Thermal Reactions for the 'Thermodetoxification' and Resource Recovery of Surplus Chemicals	3	243

Author	Title	Vol	Page
Silverstein, Joseph D.	Near Millimeter Wave Radiation from a Gyromonotron	3	259
Simutis, Zita M. Barsam, Helena F.	Terrain Visualization by Soldiers	3	273
Sindoni, Orazio I.	Calculation on Optical Effect of Matter from First Principles Using Group Theoretical Techniques	3	283
Singler, R. E. Koulouris, T. N. Deome, A. J. Lee, H. Dunn, D. A. Kane, P. J. Bieberich, M. J.	Synthesis and Evaluation of Phosphazene Fire Resistant Fluids	3	297
Slagg, Norman	See Gilbert, Everett E.	4	87
Smith, Alvin	A Method of Polymer Design and Synthesis for Selective Infrared Energy Absorption	3	309
Sollott, Gilbert P.	See Gilbert, Everett E.	4	87
Spellicy, Robert L.	See Watkins, Wendell R.	4	249
Spoonamore, Janet H.	CAEADS--Computer Aided Engineering and Architectural Design System	3	325
Squire, Walter H. Garnett, Lamont W.	Visco-Elastic Behavior of Incendiary Compositions Under Ballistic Loading	3	341
Steiner, James	See Chu, Shih C.	1	225
Sterling, Bruce S.	The Relationship Between Company Leadership Climate and Objective Measures of Personnel Readiness	3	357
Stern, Richard A.	See Borowick, John	1	113
Stuck, Bruce E. Lund, David J. Beatrice, Edwin S.	Ocular Flash Effects of Relatively "Eye Safe" Lasers	3	371
Sturdivan, Larry M. Bexon, Roy Berkhimer, Karl	General Bullet Incapacitation and Design Model	4	209

Author	Title	Vol	Page
Sturek, Walter B.	See Nietubicz, Charles J.	2	425
Swenberg, Charles E.	See McCreery, M. J.	2	357
Tanton, George A.	See Johnson, John L.	2	199
Thomson, George M. Jamison, Keith A.	In-Bore Propellant Media Density Measurements by Characteristic X-Ray Radiography	3	385
Throop, Joseph F.	A Fracture and Ballistic Penetration Resistant Laminate	3	397
Tice, Allan	See Lunardini, Virgil J.	2	263
Tice, Thomas R.	See Setterstrom, Jean A.	3	215
Turetsky, Abraham L.	Advances in Multispectral Screening	4	225
Twartz, John C. Shirai, Akira Selvaraju, G. Saunders, J. Peter Huxsoll, David L. Groves, Michael G.	Doxycycline Prophylaxis of Scrub Typhus	3	411
Upatnieks, Juris	See Duthie, Joseph G.	1	341
Verdier, Jeff S.	See Wolfe, Alan D.	3	435
Viechnicki, Dennis J.	See Caslavsky, Jaroslav L.	1	201
Villarreal, Ysidro	See Powanda, Michael C.	3	63
Violette, Edmond J.	See Schwering, Felix K.	3	201
Vogel, James A.	See Patton, John F.	3	1
Voorhees, James W. Marchionda, Kristine Atchison, Valerie L.	Speech Command Auditory Display System (SCADS)	3	425
Wade, William L., Jr.	See Gilman, Sol	1	447
Walbert, James N.	See Elder, Alexander S.	1	353
Walters, Donald L. Kunkel, Kenneth E. Hoidale, Glenn B.	Optical Turbulence within the Convective Boundary Layer	4	239

Author	Title	Vol	Page
Watkins, Wendell R. White, Kenneth O. Spellicy, Robert L.	Simulated Plume Radiative Transfer Measurements	4	249
Wetherby, Anne Marie	See Porter, William L.	3	31
Wharton, W. W.	See Ayers, O. E.	1	17
White, Kenneth O.	See Watkins, Wendell R.	4	249
Whitman, Gerald M.	See Schwering, Felix K.	3	201
Williams, C.	See Figucia, F.	1	383
Wolfe, Alan D. Emery, Clarence E. Verdier, Jeff S. Prichard, Dorothy A.	Studies on Butyrylcholinesterase Inhibitors	3	435
Yeomans, Walter G.	See Ramsley, Alvin O.	3	79
Yergey, Alfred L.	See Friedman, Melvin H.	1	415
Yu, Yung H.	See Kittleson, John K.	2	209
Zavada, John M.	See Jenkinson, Howard A.	2	189
Zegna, Angelo I.	See DeVenuto, Frank	1	299
Zegna, Angelo I.	See DeVenuto, Frank	1	315
Zimmerman, Kathleen L.	See Elder, Alexander S.	1	353
Zwick, H. Bloom, Kenneth R. Lund, David J. Beatrice, Edwin S.	Laser Ocular Flash Effects	3	449

MAINTAINING MOTOR SKILL PERFORMANCE (U)

JOSEPH D. HAGMAN, Ph.D.
U.S. ARMY RESEARCH INSTITUTE
ALEXANDRIA, VA 22333

The Army's primary peacetime mission is to maintain combat readiness (5). To be combat ready, soldiers must first become proficient in their performance of assigned job tasks, and then retain this proficiency over what can be prolonged periods of no practice. One way to enable soldiers both to reach and maintain combat readiness is through the use of training methods that promote effective task acquisition and retention. But first, these methods must be identified and compared.

Most training methods have been investigated within the context of laboratory experiments. In these experiments, training of motor tasks has involved the execution of presentation (p) and test (t) trials. At p-trials persons perform the criterion movement to be learned, whereas during t-trials they attempt to reproduce (i.e., recall) the criterion from memory. P- and t-trials can be distinguished procedurally by the presence or absence of a movement constraint. During the training of straight arm (i.e., linear positioning) movements, for example, persons perform p-trials by moving a sliding mechanism along a horizontal track until contacting a stop positioned by the experimenter to define criterion movement distance (extent) and location (end position). At t-trials, an attempt is made to recall the criterion with the stop removed.

The number and sequential arrangement of p- and t-trials performed during training has differed across experiments. Some investigators (16) have alternated p- and t-trials during training, whereas others (2, 4) have emphasized one type of trial over another through repetition. Unfortunately, for those of us responsible for training motor tasks in the operational environment, the relative acquisition and retention benefits associated with each method have not been identified. Thus, it is difficult to select the method(s) that is most effective.

The question of which method(s) is most effective remains an empirical one because of uncertainty regarding the relative contribution of p- and t-trials to motor acquisition and retention. While most researchers and trainers would agree that p-trials contribute to both acquisition and retention, either because they provide reinforcement (2) or because they gen-

erate movement feedback information needed to form a reference for movement recall (1), the contribution of t-trials is still a matter for debate. On the one hand, t-trials may be detrimental because they are different from the criterion as a result of error in recall. Thus, they are capable of causing interference (6) and depressing performance. From this perspective, training methods that avoid t-trials and emphasize p-trials should be most effective. On the other hand, t-trials may be beneficial to movement acquisition and retention because they provide trainees with an opportunity to define their own movements during training. Researchers have shown that movements performed under a learner-defined execution mode, i.e., unconstrained by a stop, are retained better than those performed under an experimenter-defined mode, i.e., constrained by a stop (7, 8). Thus, training methods that include learner-defined t-trials may encourage trainees to rely on their retention of t- rather than p-trials, and thereby, promote movement recall. From this perspective, training methods should not only include but emphasize t-trials to be most effective.

The present paper describes three laboratory experiments. The first two compare the effects of p- and t-trials on movement distance (Experiment 1) and location (Experiment 2) within the context of three different training methods. The general approach in these experiments was to repeat or alternate p- and t-trials during training and to examine the effect of this manipulation on movement recall during acquisition as well as after short- and long-term retention intervals. The third experiment examines the validity of alternative interpretations offered to explain the results obtained in the first two experiments. The laboratory task chosen for training was linear positioning.

Experiment 1

Method

Subjects. Forty-five employees (33 men and 12 women) of the Army Research Institute (ARI) volunteered to serve as subjects in the experiment.

Apparatus. Subjects performed movements from their left to right using a metal slide that ran along a linear track consisting of two stainless steel rods 35 inches (88.9 cm) in length. Two Thompson Ball Bushings supported the slide on the rods which were mounted in parallel on a metal frame 4.25 inches (11 cm) apart and 11 inches (27.94 cm) above the base of the frame. The base rested on a standard table top 31 inches (78.74 cm) from the floor. A second slide was used by the experimenter to stop the first slide along the track during p-trials. A pointer attached to the experimenter's side of each slide ran along a meter stick to indicate respective slide position. Additional apparatus included; a chin rest to control head and body position, earphones through which subjects received tape-recorded procedural commands, and a blindfold to eliminate visual cues.

Design. The experiment contained an acquisition and a retention segment, as shown in Figure 1. The acquisition segment consisted of 18 train-

	ACQUISITION TRIALS																		RETENTION	
	CYCLE 1							CYCLE 2						CYCLE 3						
GROUPS	1	2	3	4	5	6/	/7	8	9	10	11	12/	/13	14	15	16	17	18	3 min	24 hr
STANDARD	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	T	T
TEST	P	T	T	T	T	T	P	T	T	T	T	T	P	T	T	T	T	T	T	T
PRESENTATION	P	P	P	P	P	T	P	P	P	P	P	T	P	P	P	P	P	T	T	T

Figure 1. Acquisition and retention trial sequences for each training method group.

ing trials divided into three cycles of six trials each. Each cycle contained p- and t-trials. P-trials were experimenter-defined movements constrained by the stop. The distance between movement starting position and the stop defined the criterion distance of 9.84 inches (250 mm). T-trials were learner-defined recall movements performed with the stop removed. The sequence of p- and t-trials performed within cycles differed for each of three training method groups. For the STANDARD group a cycle consisted of three p- and three t-trials administered in an alternating sequence. For the PRESENTATION group, the first five trials of each cycle were p-trials and the sixth was a t-trial. For the TEST group, the first trial was a p-trial and the next five were t-trials. The retention segment of the experiment consisted of a single t-trial performed at both 3 minutes and 24 hours after the final training trial.

Fifteen subjects were assigned randomly to each of the three training method groups with the restriction that each group contain the same proportion of men and women. Including the two retention trials, subjects performed a total of 20 movements during the experiment. Each movement began from a different starting position that varied between 20 and 30 mm, in increments of 20 mm, from the subject's left end of the apparatus. Three random sequences were developed for the 20 movement starting positions. Five subjects in each group were trained under one of the three sequences.

Procedure. At the beginning of the experiment, subjects were instructed to learn and remember movement distance and shown how distance was separated procedurally from location through changes in movement starting position. They were then shown a written copy of the entire p- and t-trial command sequence appropriate to their training group and told the meaning of each command. "Movement" was the p-trial command and "Recall Movement" was the t-trial command. Each was preceded by "Ready" and followed by "Rest." At "Ready," the experimenter grasped the subject's right hand and placed it on the handle of the slide. Five seconds later, subjects heard either "Movement" or "Recall Movement" depending on their training group membership and the specific training trial. At "Movement," they moved the slide along the track at a moderate (approximately 125 mm/second) pace until contacting the stop. At "Recall Movement," the stop was removed and subjects moved the slide along the track and stopped it when they felt they had moved it the criterion distance. Five seconds were allowed for p- and t-trial execution. During this interval, white noise was delivered through the earphones to eliminate auditory cues resulting from displacement of the slide. "Rest" marked the start of a 10-second rest interval during which subjects removed their hand from the slide and placed it in a predetermined resting position on the table. Concurrently, the experimenter recorded recall accuracy (only for t-trials) and repositioned either the slide alone or both it and the stop in preparation for the next trial. After "Rest," subjects heard "Ready" and the sequence of commands began for the next trial. During the retention segment of the experiment, intervals of 3 minutes and 24 hours were inserted between "Rest" and "Ready." Before training began, subjects were asked not to count while moving the slide. They

then donned their blindfold and earphones and received an opportunity to move the slide to get a feel for its movement characteristics.

Results and Discussion

Although both algebraic (signed) and absolute (unsigned) error were recorded for all t-trials, algebraic error revealed no significant effects of interest. Therefore, only absolute error results are reported.

Acquisition. Mean absolute error scores for acquisition t-trials are shown on the left side of Figure 2. The acquisition curves for the STANDARD and PRESENTATION groups were negatively accelerated and similar in appearance when considering only those t-trials that coincided temporally for both groups, i.e., Trials 6, 12, and 18. The TEST group's curve was above those of the other two groups and had a serrated appearance. A Groups x Trials analysis of variance (ANOVA) performed on the scores for t-trials common to all three groups (i.e., Trials 6, 12, 18) revealed a significant main effect of trials, $F(2,84) = 4.08$, MSe = 529.24, and groups $F(2,42) = 4.48$, MSe = 663.45, with the rejection region for this and all other analyses being .05. Individual comparisons, using the least significant difference method (3), showed that the trials effect was caused by a decrease in error between Trials 6 and 12 for all groups and that the groups effect resulted from TEST group error scores being greater than those of the STANDARD and PRESENTATION groups, while the scores for the latter two groups did not differ significantly.

Visual inspection of all TEST group acquisition t-trial scores shown in Figure 2 revealed that a build up error within cycles was responsible for the difference found between the TEST and the other two groups. A subsequent Cycles (1-3) x Trials (1-5) ANOVA on all TEST group t-trial scores revealed a significant effect of trials, $F(4,56) = 4.24$, MSe = 292.54, with individual comparisons indicating that error was greater on the last t-trial of each cycle than on the first t-trial of each cycle. This within-cycle error increase is consistent with past motor research findings (4) and indicates that t-trial repetition during training adversely affects acquisition of movement distance information when measured within cycles.

Further inspection of TEST group scores revealed that error decreased considerably between cycles. This decrease offset the error increase that occurred within cycles and eliminated group differences found at common t-trials. A separate Groups x Trials ANOVA comparing STANDARD and PRESENTATION group scores on the last t-trial of each cycle, i.e., Trials 6, 12 and 18, with TEST group scores on the first t-trial of each cycle, i.e., Trials 2, 8 and 14, still showed a significant effect of trials, $F(2,84) = 4.78$, MSe = 593.49, but no significant effect of groups. Failure to find a groups effect indicated that TEST group performance at Trial 14, for example, after only 3 p-trials did not differ from Trial-18 performance of the STANDARD and PRESENTATION groups after 9 and 15 p-trials, respectively. Thus, t-trials appear to have potentiated or increased the effectiveness of p-trials and indirectly contributed to the acquisition of movement

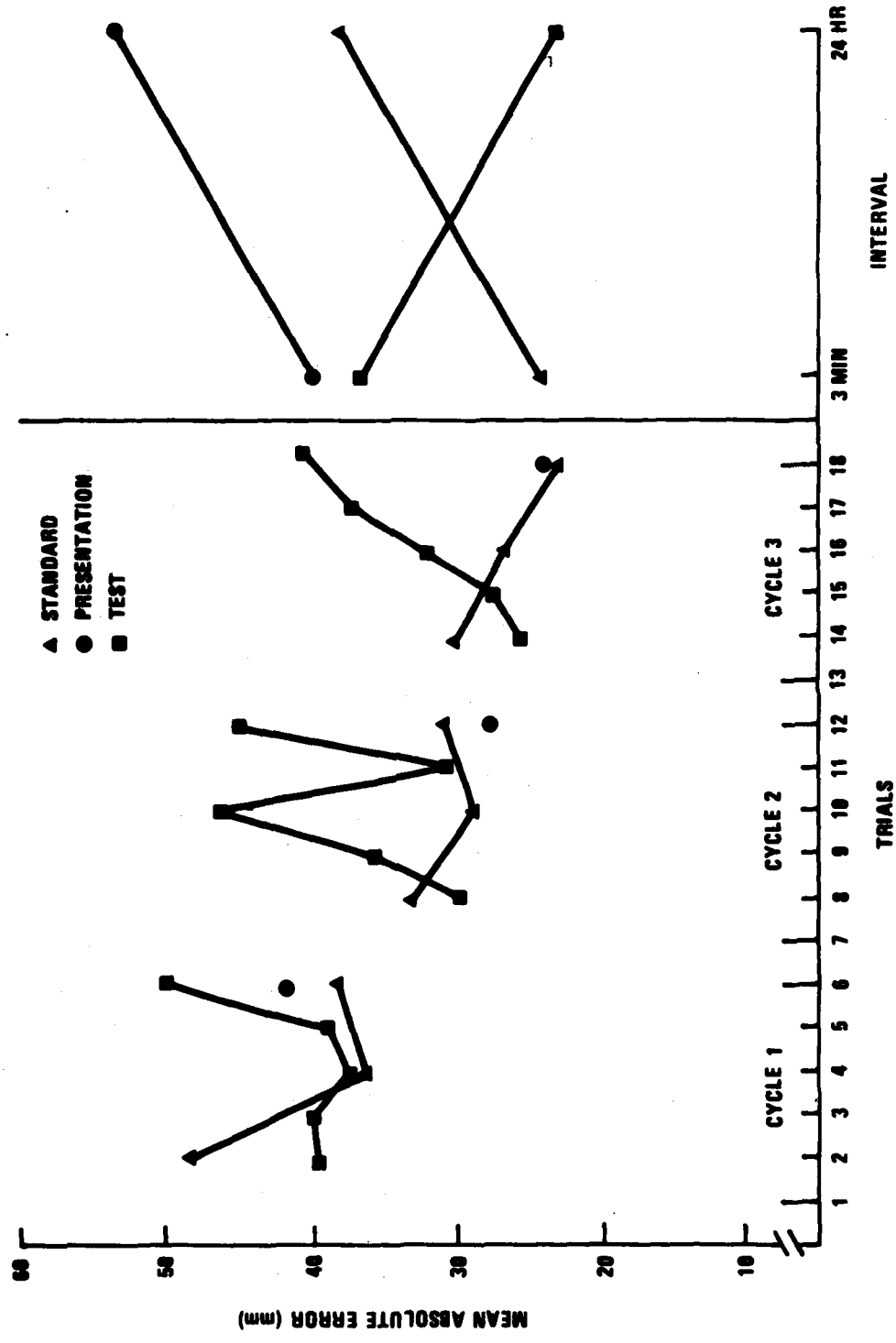


Figure 2. Mean distance absolute error at acquisition and retention t-trials for each training method group.

distance.

In summary, the acquisition data demonstrate that: (a) p-trials contribute positively to acquisition; (b) t-trials produce both positive between-cycle and negative within-cycle effects on acquisition; and that (c) p-trial repetition and p- and t-trial alternation during training produce a higher level of acquisition than t-trial repetition, but that this difference depends on the specific trials at which training methods are compared.

Retention. Mean absolute error retention scores for distance are shown on the right side of Figure 2. Scores were analyzed using a Groups (STANDARD, PRESENTATION, TEST) x Retention Interval (Immediate, 3 minutes, 24 hours) mixed factorial ANOVA with Trial-18 scores included to reflect immediate recall at the end of training. This ANOVA revealed a significant Groups x Retention Interval interaction, $F(4,84) = 4.10$, MSe = 456.68, resulting from an increase in error over time for the STANDARD and PRESENTATION groups and a decrease in error for the TEST group. Simple effect comparisons indicated that Trial-18 error for the TEST group was greater than that for both the STANDARD and the PRESENTATION groups, and that the error for these latter two groups did not differ (i.e., TEST > PRESENTATION = STANDARD). Three minutes after training, there was no difference between TEST and PRESENTATION group error scores, although their combined mean error was greater than the error for the STANDARD group (i.e., TEST = PRESENTATION > STANDARD). At 24 hours after training, TEST group error was less than STANDARD which, in turn, was less than PRESENTATION group error (i.e., TEST < STANDARD < PRESENTATION).

In summary, the retention results indicate that: (a) p-trial repetition during training results in both rapid and extensive forgetting; (b) t-trial repetition prevents forgetting and promotes increased recall proficiency over time; and (c) alternation of p- and t-trials prevents short-term forgetting and reduces long-term forgetting.

Experiment 2

The second experiment was a replication of the first except that location rather than distance was the movement cue examined. Separate investigation of location was done to test whether the results obtained for distance would generalize to another cue thought to underlie movement acquisition and retention.

Method

Subjects. An additional 45 ARI employees (33 men and 12 women) volunteered to serve as subjects in this experiment.

Apparatus. The same linear positioning and support apparatus were used as in Experiment 1.

Design. The general design was identical to that of Experiment 1. Movement starting positions varied between 20 and 380 mm, in increments of 20 mm, from the subject's left of the criterion location, set at 520 mm

from the subject's left end of the apparatus. Three random starting position sequences were generated; five subjects from each group were trained under one of the three sequences.

Procedure. The procedure was identical to that of Experiment 1 except that subjects were instructed to learn location rather than distance.

Results and Discussion.

Analysis of algebraic error revealed no significant differences of interest. Thus, only the results of absolute error analyses are reported.

Acquisition. Mean absolute error acquisition scores are shown on the left in Figure 3. The pattern of performance found for location was similar to that found earlier for distance. A Groups x Trials ANOVA performed on error scores at common t-trials revealed significant main effects of trials, $F(2,84) = 6.16$, MSe = 272.02, and groups, $F(2,42) = 8.07$, MSe = 539.72. Individual comparisons revealed that the trials effect was caused by a decrease in error between Trials 6 and 18 for all groups and that the groups effect resulted from error being greater for the TEST group than for the STANDARD and PRESENTATION groups, with no difference in error present between these latter two groups. A separate Cycle x Trials ANOVA on all TEST group t-trial scores showed that a significant increase in error occurred within cycles, $F(4,56) = 5.28$, MSe = 178.64. Individual comparisons revealed that the combined mean error for the last four t-trials of each cycle differed from that of the first t-trial of each cycle.

Location TEST group scores also showed a decrease in between-cycle error. This decrease offset the increase in error within-cycles and eliminated group differences found at common t-trials. A Groups x Trials ANOVA comparing STANDARD and PRESENTATION group scores on the last t-trial of each cycle with TEST group scores on the first t-trial of each cycle showed a significant effect of trials, $F(2,84) = 13.83$, MSe = 217.68, but no effect of groups. Thus, the p-trial potentiation effect of t-trials found for distance was also evident for location.

In summary, the acquisition data for location demonstrate that: (a) p-trials contribute positively to acquisition; (b) t-trials produce both positive between-cycle and negative within-cycle effects on acquisition; and that (c) p-trial repetition and p- and t-trial alternation during training produce a higher level of acquisition than that produced by t-trial repetition, but that this difference depends on the specific trials at which training methods are compared.

Retention. Mean absolute error retention scores for location are shown on the right side of Figure 3. A Groups x Retention Interval mixed factorial ANOVA revealed a significant main effect of retention interval, $F(2,84) = 12.53$, MSe = 312.86, and a significant Retention Interval x Group interaction, $F(4,84) = 7.03$, MSe = 312.86. Individual comparisons of simple effects confined to the interaction revealed that at Trial 18 TEST group error was greater than either STANDARD or PRESENTATION group error, and that no error difference existed between these latter two groups (i.e.,

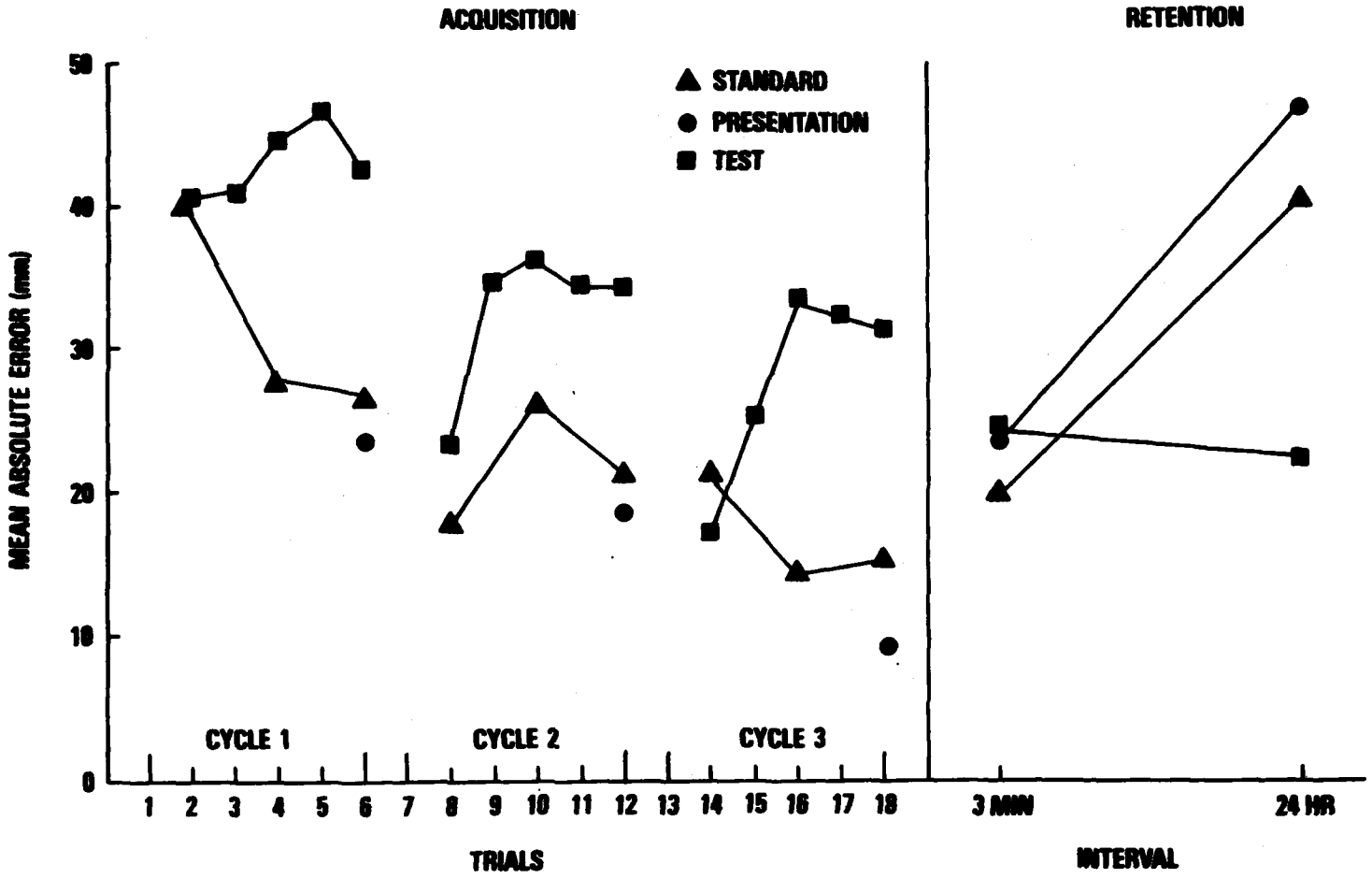


Figure 3. Mean location absolute error at acquisition and retention t-trials for each training method group.

TEST > PRESENTATION = STANDARD). Three minutes after training, no differences were present among groups (i.e., TEST = PRESENTATION = STANDARD). At 24 hours after training, TEST group error was less than STANDARD and PRESENTATION group error, while the error for these latter two groups did not differ (i.e., STANDARD = PRESENTATION > TEST).

In summary, the retention results indicate that: (a) p-trial repetition during training results in both rapid and extensive forgetting over time; (b) t-trial repetition prevents forgetting; and (c) p- and t-trial alternation produces the same effects as p-trial repetition.

In general, the results of both experiments demonstrate that: (a) both p- and t-trials affect motor task acquisition and retention, and (b) training method effectiveness depends on the type of trial repeated. P-trial repetition improves acquisition. This improvement presumably is derived from at least two sources; prevention of errors and repetition of the criterion movement. By preventing errant movement at t-trials, the possibility of interference during training is avoided. At the same time, an accurate memorial representation of the criterion is developed through repeated practice.

In contrast to the consistent beneficial effects of p-trial repetition, t-trial repetition produces mixed acquisition effects in the form of increased error within cycles and decreased error between cycles. Increases in within-cycle error result from interference generated by repetition of t-trials. These t-trials are different from the criterion because of imperfect recall by the subject, and therefore possess interference capabilities (6). In contrast, the decreases in between-cycle error are the result of p-trial potentiation following t-trial repetition. Following t-trial repetition, subjects presumably know more about their recall than after p-trial repetition. Thus, they are better able to discriminate their recalled t-trial movement from the p-trial criterion movement, and are more capable of making appropriate adjustments needed to increase post p-trial recall proficiency. Although speculative, the general notion that movement discrimination ability increases with repeated recall is consistent with past findings (10).

P- and t-trials also had different effects on retention. P-trial repetition resulted in rapid and extensive forgetting, whereas t-trial repetition prevented forgetting of both distance and location. T-trials may benefit retention for at least two reasons. First, they may be easier to remember than p-trials because of superior encoding characteristics resulting from being performed under a learner-defined movement mode, i.e., unconstrained by a stop (7,14). Learner-defined movements allow subjects to predict or anticipate movement consequences (e.g., proprioception) via corollary discharge prior to movement initiation. Prediction provides a context for receiving movement produced cues which enhances their encoding and subsequent retrieval (13). In contrast, prediction is more difficult at p-trials because movements are experimenter-defined i.e., constrained by a stop, and are not under total control of the subject. Of course, even though t-trials may be easier to remember than p-trials, subjects still

must be encouraged to rely on them for later recall. In the present experiments, this encouragement stemmed from manipulating training such that most of the trials performed during training were t-trials.

A second reason why t-trials aid retention relies on the notions of movement variability and motor schema. The motor schema is an abstraction of task and environmental characteristics that develops through repeated and varied movement during training (12), and serves as a rule or concept for movement generation. As the schema develops, the abstracted information becomes more resistant to forgetting than that resulting from specific movement instances (11). In the first two experiments, variability was generated at t-trials because subjects were variable in their recall. In contrast, no variability was generated at p-trials because every p-trial was identical. Thus, it could be argued that schema strength was greater for the TEST group than for the other two groups and this promoted superior TEST group retention.

Experiment 3

This experiment was performed to determine the relative validity of the two aforementioned reasons for enhanced retention following t-trial repetition during training. It involved adding yoked p-trial groups to the t-trial distance and location TEST groups of Experiments 1 and 2. Yoking required the use of a stop to ensure that yoked groups performed p-trials that were identical to the t-trials performed by the two TEST groups. Thus, yoking afforded the means of equating p- and t-trials for variability during training, but allowed the distinction between p- and t-trial execution modes (i.e., experimenter- versus learner-defined) to remain. If variability per se during training is the key to enhanced retention, then one would expect no difference in retention scores between the two t-trial TEST groups and their yoked p-trial counterparts. If, on the other hand, movement execution mode is the key, then the two TEST groups should display better retention than that of the two yoked groups.

Method

Subjects. Thirty additional ARI employees (22 men and 8 women) volunteered to serve as subjects in the experiment. Fifteen subjects were randomly assigned to the yoked distance group and the yoked location group with the restriction that each group contain the same proportion of men and women.

Apparatus. The apparatus were identical to those used in Experiment 1 and 2.

Design. Four training method groups were included in the design, i.e., yoked distance, yoked location, and the two TEST groups from Experiments 1 and 2. The only difference between the yoked and TEST groups was that yoked groups performed p-trials identical to the t-trials performed by the TEST groups. Data from the two yoked groups were collected in this experiment

and compared with those collected in Experiments 1 and 2 for the two TEST groups.

Procedure. Subjects were instructed to learn either movement distance or location depending on their group. Those in the yoked groups were told of the yoking procedure. The rest of the procedure was identical to that of Experiments 1 and 2.

Results and Discussion

Acquisition. Because yoked groups performed the same movements as their TEST group counterparts, no differences in acquisition scores were possible.

Retention. Separate Groups x Retention Interval mixed factorial ANOVAs were performed on absolute error retention scores for the two location and two distance groups. The means for the distance groups are shown on the left and those for the two location groups are shown on the right in Figure 4.

For distance, a significant Groups x Retention Interval interaction was found $F(1,28) = 6.85$, $MSe = 454.02$. Individual comparisons revealed that no group differences occurred at 3 minutes after training, but that the TEST group was superior to the yoked group 24 hours later. The Groups x Retention Interval interaction for location was not significant $F(1,28) = 3.11$, $MSe = 385.56$, $.05 < p < .10$. However, a priori expectations about the beneficial retention effects of t-trials based on the significant findings for distance justified further examination of simple effects. Individual comparisons indicated that group performance did not differ 3 minutes after training, but did differ 24 hours after training with the TEST group being superior to the yoked group. These data are consistent with the interpretation that differences in p- and t-trial movement execution mode, and not the increased variability associated with t-trial repetition, caused the superior long-term retention following t-trial repetition in Experiments 1 and 2. Yoking increased p-trial variability during training, yet was unable to prevent forgetting of either distance or location.

If one is to accept the execution mode interpretation of t-trial retention benefits, then it needs to be explained why the error level was so low for the distance and location TEST groups 24 hours after training. If long-term retention is based on the retention of learner-defined t-trials, then one might expect an error level equal to that of either the mean error of all training t-trials or the error displayed on the last training t-trial. Neither of which was the case. Figures 2 and 3 show that the error obtained 24 hours after training was similar to that displayed at the first t-trial of Cycle 3 (i.e., Trial 14). To account for this, it is suggested that t-trial repetition produces a form of within-cycle serial learning task where subjects attempt to recall each preceding t-trial of the series during training. Because of the strong primacy effect usually found for serial motor tasks (9), they are able to remember their recall from the first t-trial of the series better than their recall from later t-trials of

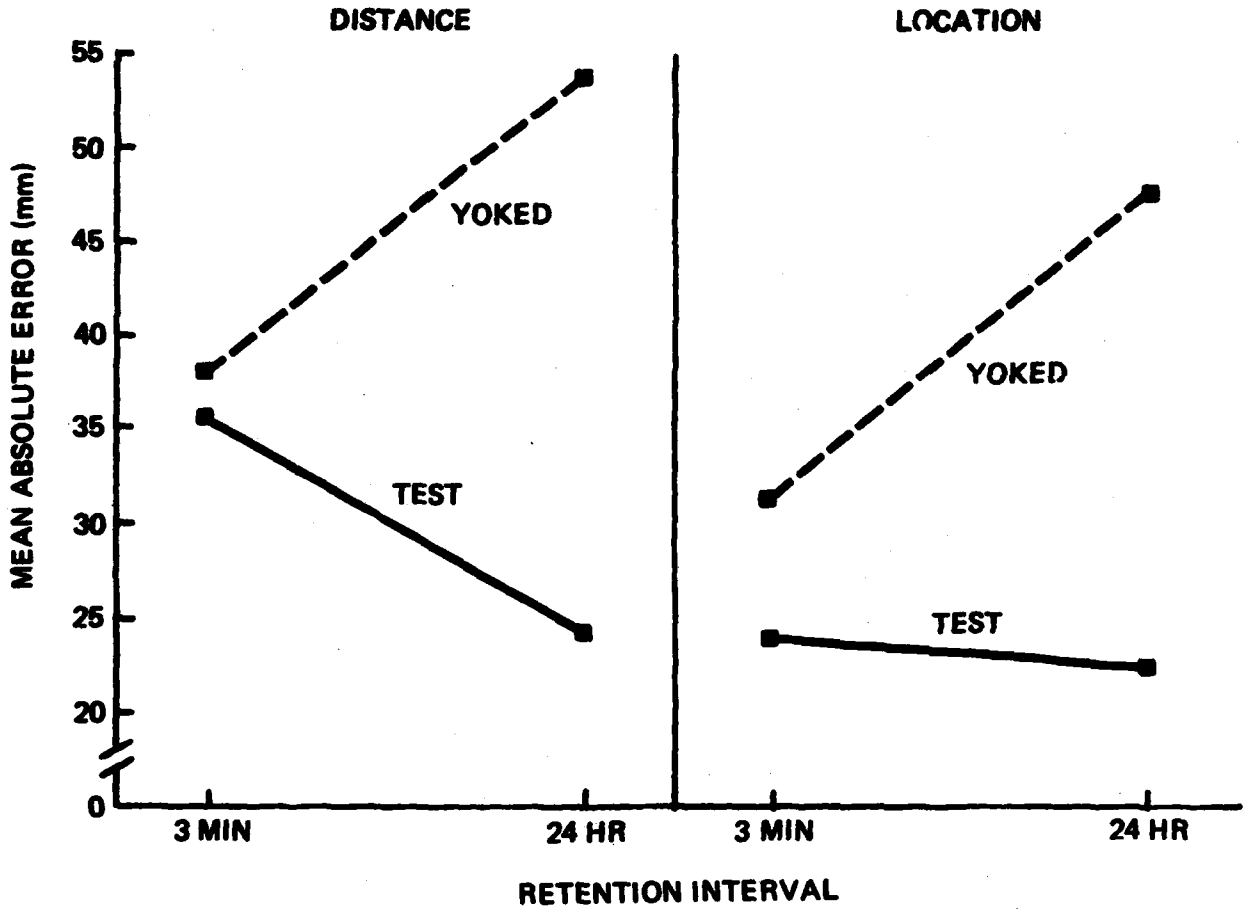


Figure 4. Mean distance and location absolute error on retention t-trials for TEST and YOKED training method groups.

the series at the time of long-term retention testing. The influence of these later t-trials, however, appears in the form of increased within-cycle error during training and retroactive interference during retention testing. If true, then the unusual decrease in error found between 3 minutes and 24 hours after training for the distance TEST group in Experiment 1 may have been caused by dissipation of retroactive interference generated by the last four training t-trials of Cycle 3 (i.e., Trials 15-18). At the same time, the rate of dissipation may not have been rapid enough to produce a decrease in error between the end of training and the 3- minute retention test. It is also possible that differences in the rate of dissipation for distance and location cues may have been responsible for the lack of a significant decrease in error over time for the location TEST group of Experiment 2. Finally, with p- and t-trial alternation, retroactive interference did not build up because t-trials were not repeated during training. Thus, both short- and long-term retention benefits were found, at least for distance. Of course the above explanation remains speculative and requires considerable additional research. However, the general notion that retroactive interference dissipates over time for motor tasks has already received support (15).

General Conclusions

The results of the present experiments clarify certain training issues regarding the relative contributions of presentation and testing to motor task performance. In doing so, they answer the Army's question of which training methods most effectively promote the highest levels of skill acquisition and retention. First, the results indicate that the goal of training should dictate the training method used. If the goal is a consistent and high level of acquisition, then training should emphasize repeated presentation or alternation of presentation with testing, with the latter method being preferred because of enhanced short- and long-term retention. If effective long-term retention is the goal, then training should emphasize repeated testing. Second, testing should not only be viewed merely as a means of evaluation, but also as a means of improving motor skill performance. And third, the enhanced retention associated with repeated testing can be achieved by merely changing the emphasis of training from presentation to testing. This could be done without the usual negative aspects of additional expenditures in training time, money and personnel.

References

1. Adams, J. A. A closed-loop theory of motor learning. Journal of Motor Behavior, 1971, 3, 111-150.
2. Adams, J. A., & Dijkstra, S. Short-term memory for motor responses. Journal of Experimental Psychology, 1966, 71, 314-318.

3. Carmer, S. G., & Swanson, M. R. An evaluation of ten pair-wise multiple comparison procedures by Monte Carlo methods, Journal of the American Statistical Association, 1973, 58, 66-74.
4. Duffy, T. M., Montague, W. E., & Laabs, G. J. The effect of overt rehearsal on motor short-term memory. Journal of Motor Behavior, 1975, 7, 59-63.
5. Guthrie, J. R. Supporting the Army in the eighties. Army Logistician, September-October 1979, 11(5), 22-25.
6. Hagman, J. D. Specific-cue effects of interpolated movements on distance and location in short-term motor memory. Memory and Cognition, 1978, 6, 432-437.
7. Kelso, J. A. S. Planning and efferent components in the coding of movement. Journal of Motor Behavior, 1977, 9, 33-47.
8. Kelso, J. A. S., & Wallace, S. A. Conscious mechanisms in movement. In G. E. Stelmach (Ed.), Information processing in motor control and learning, New York: Academic Press, 1978.
9. Magill, R. A., & Dowell, M. N. Serial-position effects in motor short-term memory. Journal of Motor Behavior, 1977, 9, 319-323.
10. Newell, K. M. Knowledge of results and motor learning. Journal of Motor Behavior, 1974, 6, 235-244.
11. Newell, K. M., & Shapiro, D. C. Variability of practice and transfer of training: Some evidence toward a schema view of motor learning. Journal of Motor Behavior, 1976, 8, 233-243.
12. Schmidt, R. A. A schema theory of discrete motor skill learning. Psychological Review, 1975, 82, 225-260.
13. Stelmach, G. E., Kelso, J. A. S., & Wallace, S. A. Preselection in short-term memory. Journal of Experimental Psychology: Human Learning and Memory, 1975, 1, 745-755.
14. Summers, J. J. Levey, A. J., & Wrigley, W. J. The role of planning and efference in the recall of location and distance cues in short-term motor memory. Journal of Motor Behavior, 1981, 13, 65-76.
15. Wrisberg, C. A. The serial-position effect in short-term motor retention. Journal of Motor Behavior, 1975, 7, 289-295.
16. Wrisberg, C. A., & Schmidt, R. A. A note on motor learning without post-response knowledge of results. Journal of Motor Behavior, 1975, 7, 221-225.

DATA COMPRESSION FOR TRANSIENT MEASUREMENTS (U)

*SAMUEL F. HARLEY, PhD
US ARMY ABERDEEN PROVING GROUND
ABERDEEN PROVING GROUND, MD. 21005

Introduction

Several years ago an effort to replace obsolete data acquistion instrumentation at Aberdeen Proving Ground's Materiel Testing Directorate was initiated. The thrust of this effort was to develop a networked, distributed acquisition and processing system to provide a higher data integrity, reduce the time frame from experiment to report, and to enhance productivity.

One type of testing that is performed within MTD involves the capture of high speed transient phenomena, such as internal weapon chamber pressure or blast overpressure.

The requirements encountered in the measurement of blast overpressure are particularly strident, and consist of up to 16 data channels, each with a recording rate of 400,000 samples per second, over a recording interval of up to 200 milli-seconds. Each repetition of a set of blast overpressure measurements generates in excess of 1.25 million words of data. This type of experiment is at times performed at approximately five minute intervals, and the amount of data involved can seriously impair the effectiveness of the data acquisition system. Other classes of testing impose data loads that can approach these figures.

To reduce the impact of these large quantities of data, a means of reducing the amount of data while maintaining the level of information has been developed and implemented in APG's Automatic Sampling Rate Digitizer.

Background

Data compression can be defined as the elimination of redundant data samples. Such techniques as the extraction of the mean and standard deviation, the root-mean-square (RMS), or other statistical measures such as histograms are certainly compressive in nature, but for the purposes of this presentation only those processes which allow the reconstruction of the time history of the original signal will be considered.

The concept of data compression is not a recent development, and in fact a family of data compression algorithms have been applied to telemetry data over the past several years. Table 1 is a listing of eight standard compression algorithms (1,4).

ORDER	CORRIDOR	RECORDED SAMPLE	DESCRIPTOR
Zero	Fixed	Nonredundant	ZFN
Zero	Variable	Preceeding	ZVP
Zero	Variable	Adjusted	ZVA
First	Fixed	Nonredundant	FFN
First	Fixed	Preceeding	FFP
First	Fixed	Adjusted	FFA
First	Variable	Preceeding	FVP
First	Variable	Adjusted	FVA

Table 1. Standard Data Compression Algorithms

The simplest of these algorithms is ZFN, demonstrated in Figure 1. In this compression scheme, the first sample is recorded and a programmed delta corridor is established with that value at the midpoint of the corridor. Each subsequent data sample that falls within the delta corridor is rejected until a sample falls outside the corridor. That sample is recorded and the corridor is moved so that it is centered on the nonredundant sample.

The maximum error produced by ZFN processing is equal to one half of the corridor value. Compression ratios (number of sampled data points divided by the number of recorded data points) in the range of 7:1 to 50:1 have been documented.

The other compression algorithms vary in complexity and effectiveness. Table 2 (1,4) indicates the peak error for the standard algorithms. Also included in this table is an indication of the range of compression ratios for some of the algorithms from an experiment using both simulated and real data.

Algorithm	Peak Error	Compression Ratio Range	
		Simulated Data	Real Data
ZFN	+-1/2 Delta	11:1	7:1-50:1
ZFP	+-1/2 Delta		
ZVA	+-1/2 Delta	19:1	10:1-65:1
FFN	+-Full Scale		
FFP	+-Delta		
FFA	+-Delta		
FVP	+-1/2 Delta	10:1	6:1-40:1
FVA	+-Delta	9:1	7:1-45:1

Table 2. Comparison of Errors and Compression Ratios

ASRD Compression Algorithm

All of the standard compression algorithms are asynchronous in nature, and as such each recorded data sample must be accompanied by the time of occurance. This effectively reduces by one-half or one-third the savings gained by data compression. In addition these schemes are primarily designed for application to stationary data.

In the ASRD a compression algorithm somewhat related to ZFN is implemented. Instead of establishing a delta corridor, however, the absolute value of the rate of change (first derivative) of the signal is calculated, and based upon this value data samples are recorded at power of two multiples of the basic sampling interval of 1.25 microseconds (2). The equation governing the recording interval is:

$$I = 1.25 \times 2^{(15 - RC)} \text{ microseconds} \qquad (1.)$$

Where

$$RC = 7 + \log_2 \left(\sum_{J=1}^{N} |\Delta_j| / M\right) - \log_2 (80/fmax) \qquad (2.)$$

In addition to the establishment of recording intervals in accordance with equation 1 the ASRD algorithm is modified by the establishment of minimum recording rates, dependent on the signal bandwidth, and by the use of recording rate hysteresis. Recording rate hysteresis implies that the recording rate is maintained at an elevated level for a programmed length of time after the required rate, as determined by equation 2, decreases. Both of these features reduce the compression ratio, but serve to enhance the fidelity of the reconstructed signal.

Table 3 is a compilation of the parameters used in the ASRD. These parameters are chosen to generate summations with values ranging from 0 to 255. The function for RC is designed to cause the recording of at least 10 samples per cycle of any full scale signal within the programmed bandwidth.

f MAX (KHz)	DIVISOR	SUMMATION INTERVAL	DECREASE DELAY (MICROSEC)	MINIMUM RECORDING RATE (SAMPLES PER SEC)
80	1	4	200	6,250
70	2	9	200	6,250
60	2	11	200	6,250
50	2	13	200	6,250
40	1	8	300	6,250
30	2	21	300	6,250
20	1	16	500	3,125
10	1	32	900	1,562.5

Table 3. ASRD Compression Parameters

Hardware Implementaion

A block diagram of the data acquisition systems of which the ASRD is a part is shown in Figure 2 (3). This system consists of assorted computer peripherals, an event interface which allows the recording of external timing events, and signal conditioning and digitizer for 16 independent data channels.

The ASRD is shown in block diagram form in Figure 3. Control information is communicated to each ASRD from the computer via the control bus, and data are transferred to the computer via the data bus. Each ASRD has in addition an internal data and control bus. Each channel has a resident microprocessor to oversee bus communications with the computer and to control the various elements within the ASRD, as shown in Table 4.

Trigger Mode	Internal/External Slope/Absolute Value
Threshold Level	+ - Full Scale
Memory Partitions	Separation of memory into pre- and post- trigger areas on any 1K boundary
Start Delay	Length of Time (0 to 4096 msec) after channel trigger before data recording begins
Stop Delay	Length of time (0 to 4096 msec) after the enable goes away before data recording ceases
Rate Calculation	Sum interval, division factor, decrease delay and "RC" from Table 2 and equation 1

Table 4. ASRD Programmable Parameters

The rate calculation circuitry, which implements the compression algorithm, is shown in block diagram form in Figures 4 and 5. Its operation proceeds in the following manner:

a. Sample to sample differences are determined.

b. The absolute values of the differences, delta, are calculated.

c. Alternately, the complement of the variable length buffer and the new delta are added to the running sum of deltas.

d. The sum is shifted right (division by a power 2).

e. This value is used as an address into a random access memory in which are stored the values calculated using equation 2.

f. If the value in the addressed location is larger than the current value, the data recording rate is immediately increased, but if it is smaller the rate is not decreased until an interval defined by the decrease delay has passed.

The implementation of the compression algorithm as a functional look-up table allows other data recording schemes to be programmed. Any functional relationship between recording interval and the absolute value of the signal's first derivative could be used. In actual practice this unit is at times programmed to record data at a constant rate.

To enable a reconstruction of the time history of the data, a four bit rate code is appended to the digitized signal amplitude. In addition, when the recording rate is increased the contents of the skipped sample counter is recorded, as is the IRIG time value at the instant the channel triggers. Each of these types of data has a unique identifier added to make up a 16 bit word.

Application

Before the construction of the first ASRD extensive computer modeling was performed (2). The results of this simulation exercise indicated that compression ratios as large as 41:1 could be attained while maintaining sufficient reconstructed signal fidelity.

After construction of the ASRD's testing was performed to verify its proper functioning (3). Shown in Figure 6 is a single cycle of a triangle wave generated by a function generator and digitized by the ASRD. Instead of connecting the data points with line segments only the points are shown. This figure shows the initial low rate increasing by a factor of 32, thus allowing the transient waveform to be accurately digitized. Notice also that the recording rate increase occurs before the onset of the transient. This figure also demonstrates the effect of recording rate hysterises, where by the recording rate is maintained at its elevated level through the decrease delay interval.

A comparison test was carried out in which data were simultaneously recorded on analog magnetic tape and digitized and recorded by an ASRD. The results of this comparison showed good agreement between the two systems, and during these actual field tests compression ratios of 20:1 were achieved.

In other tests, data were digitized using two ASRD channels: one programmed with the data compression algorithm while the other was programmed to digitize at an elevated, fixed rate. The plots are shown in Figures 7 and 8. These figures demonstrate the reconstructed signal fidelity.

Conclusion

Experimenters can be very protective of data collected during their experiments and often resist any effort to eliminate redundant data samples. The inclusion of nonintelligence bearing data, however, can impose unacceptable burdens on data acquisition and processing systems, and can thus slow the testing process and inhibit productivity. The compression algorithm described here provides a modest compression ratio while maintaining the fidelity of the reconstructed signal. The use of these techniques decreases the requirements for both dedicated off-board data memory and processor memory, for test site mass storage, and for archival storage.

REFERENCES

1. Baker, B. R., "Data Compression and the Use of Compressed Data". EMR Telemetry, Unpublished.

2. Cunningham, H. V., Harley, S. F., and Paules, P. L., "Automated Data Acquisition And Processing Technology (ADAPT)", Report No. APG-Mt- 5292, November 1979.

3. Francis, C. L., "Ballistic Test Site Terminal", Minutes of the Test and Evaluation Command Instrumentation Conference, 28-31 April 1981.

4. Simpson, R. S., Blackwell, C. A., and Frost, W. O., "Compendium of Redundancy Removal Processes", IEEE Transactions on Aerospace and Electronic Systems, Vol ASE-4, No. 4, 1966.

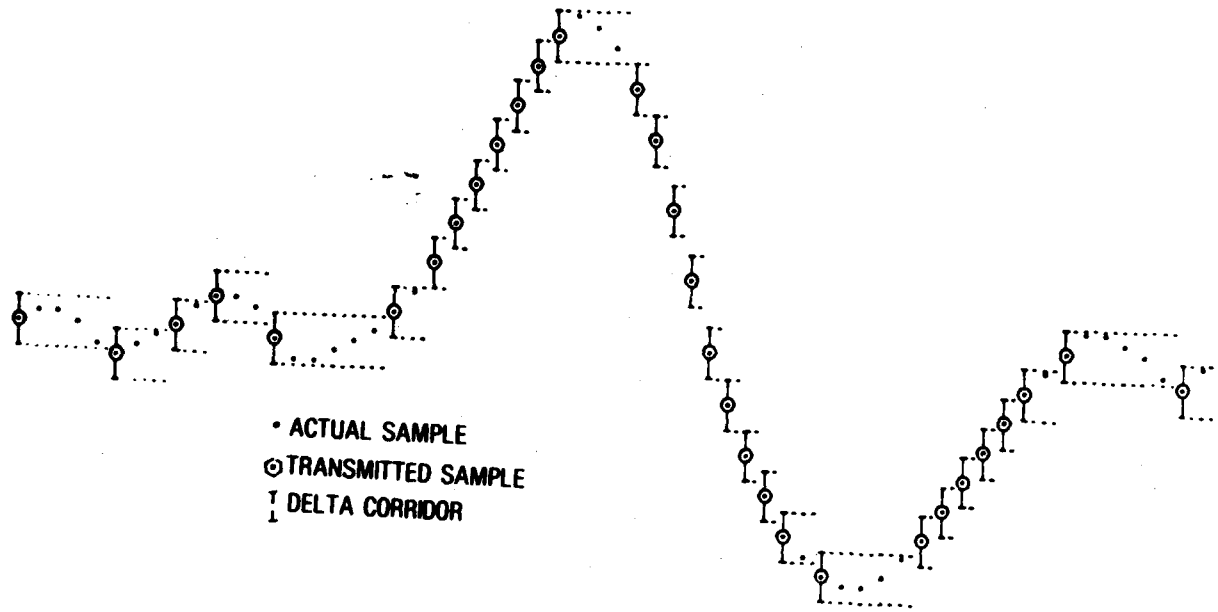


FIGURE 1 ZERO ORDER: FIXED CORRIDOR: NON—REDUNDANT SAMPLE TRANSMITTED (ZFN)

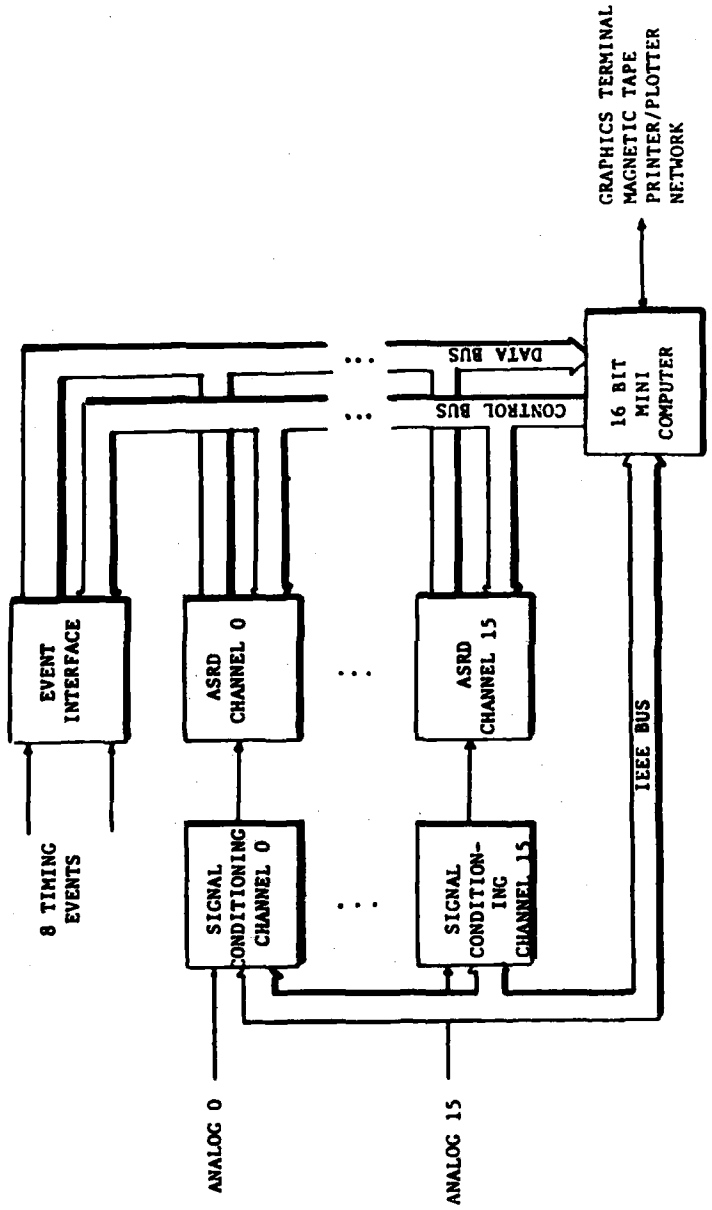


FIGURE 2 BALLISTIC TEST SITE TERMINAL BLOCK DIAGRAM

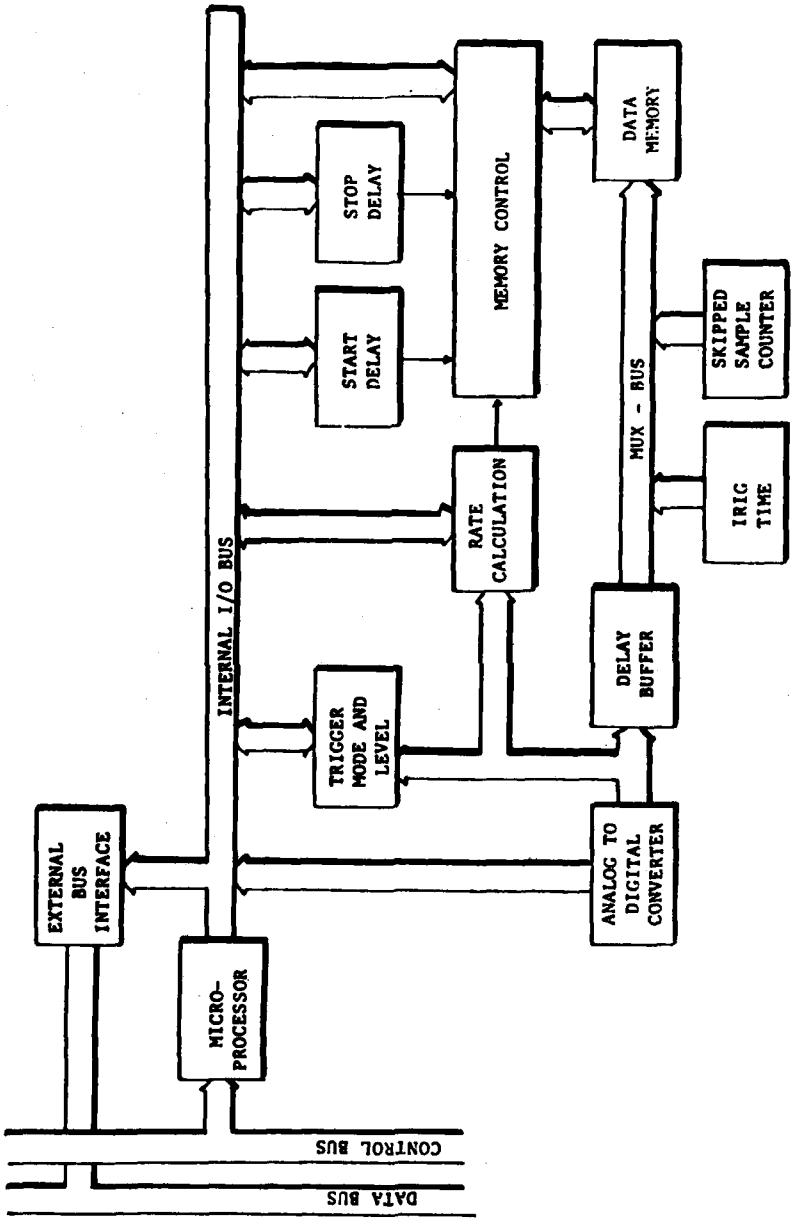
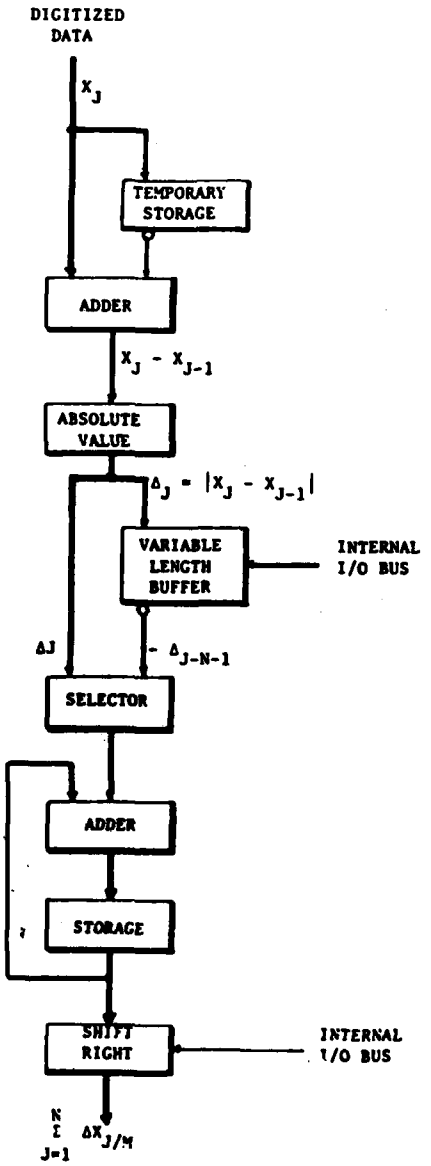


FIGURE 3 AUTOMATIC SAMPLING RATE DIGITIZER BLOCK DIAGRAM

FIGURE 4 RECORD RATE DETERMINATION — PART 1



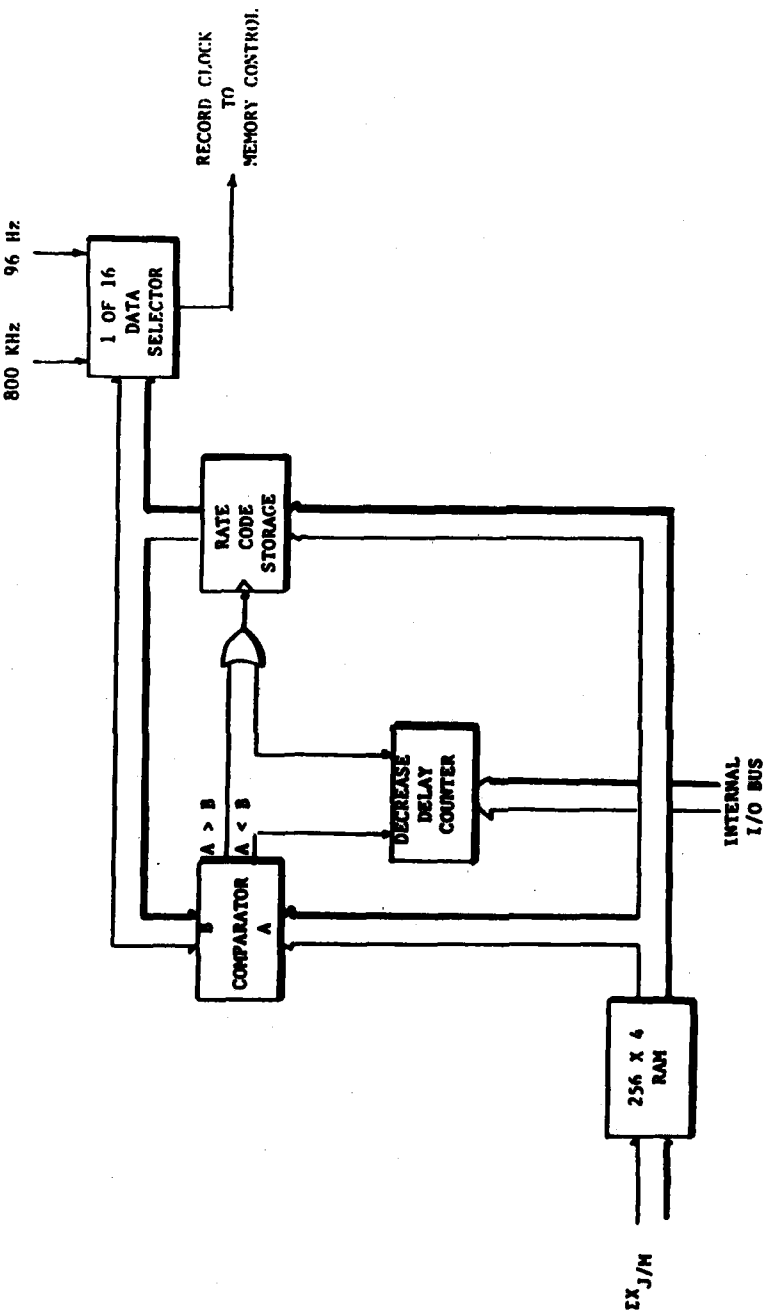


FIGURE 5 RECORD RATE DETERMINATION— PART 2

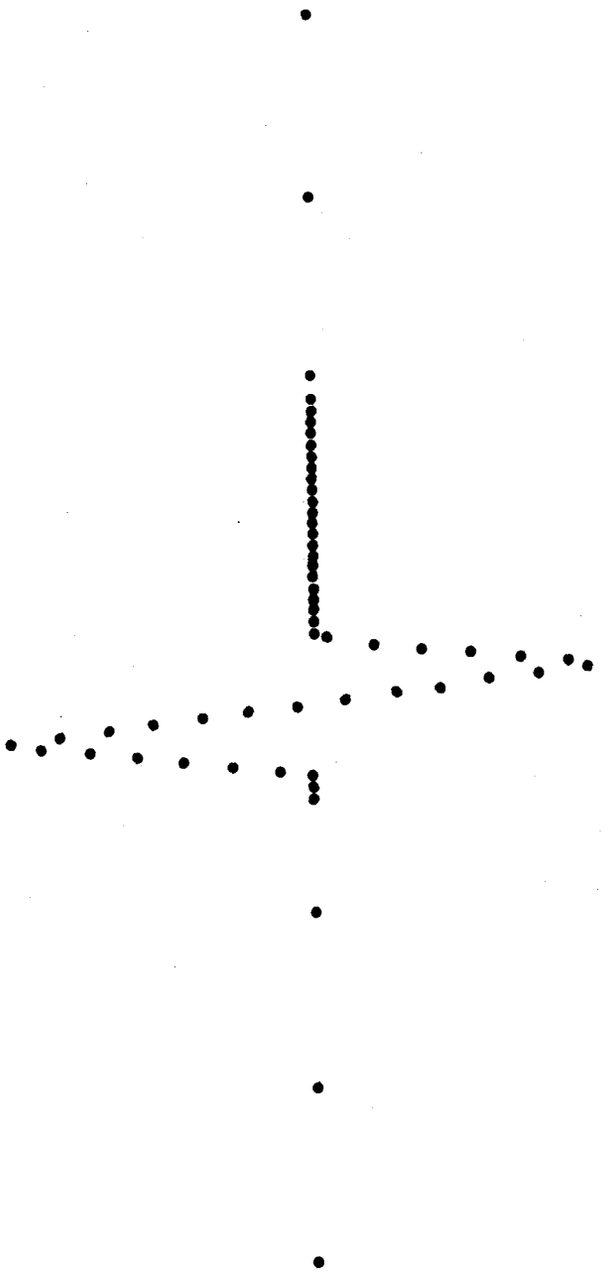


FIGURE 6. TRIANGLE WAVE DIGITIZED BY ASRID

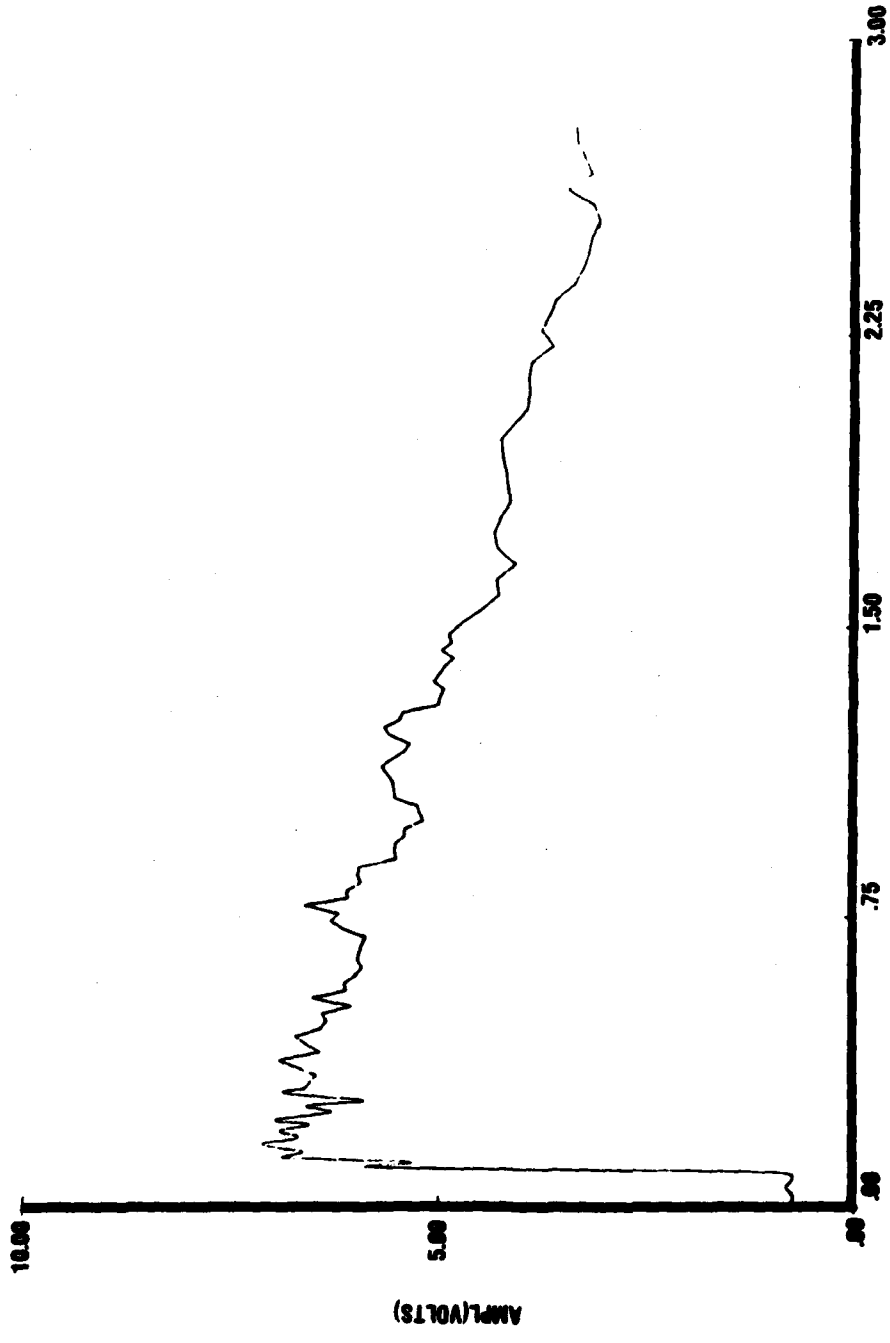


FIGURE 7. BLAST OVERPRESSURE RECORD WITH VARIABLE SAMPLING RATE

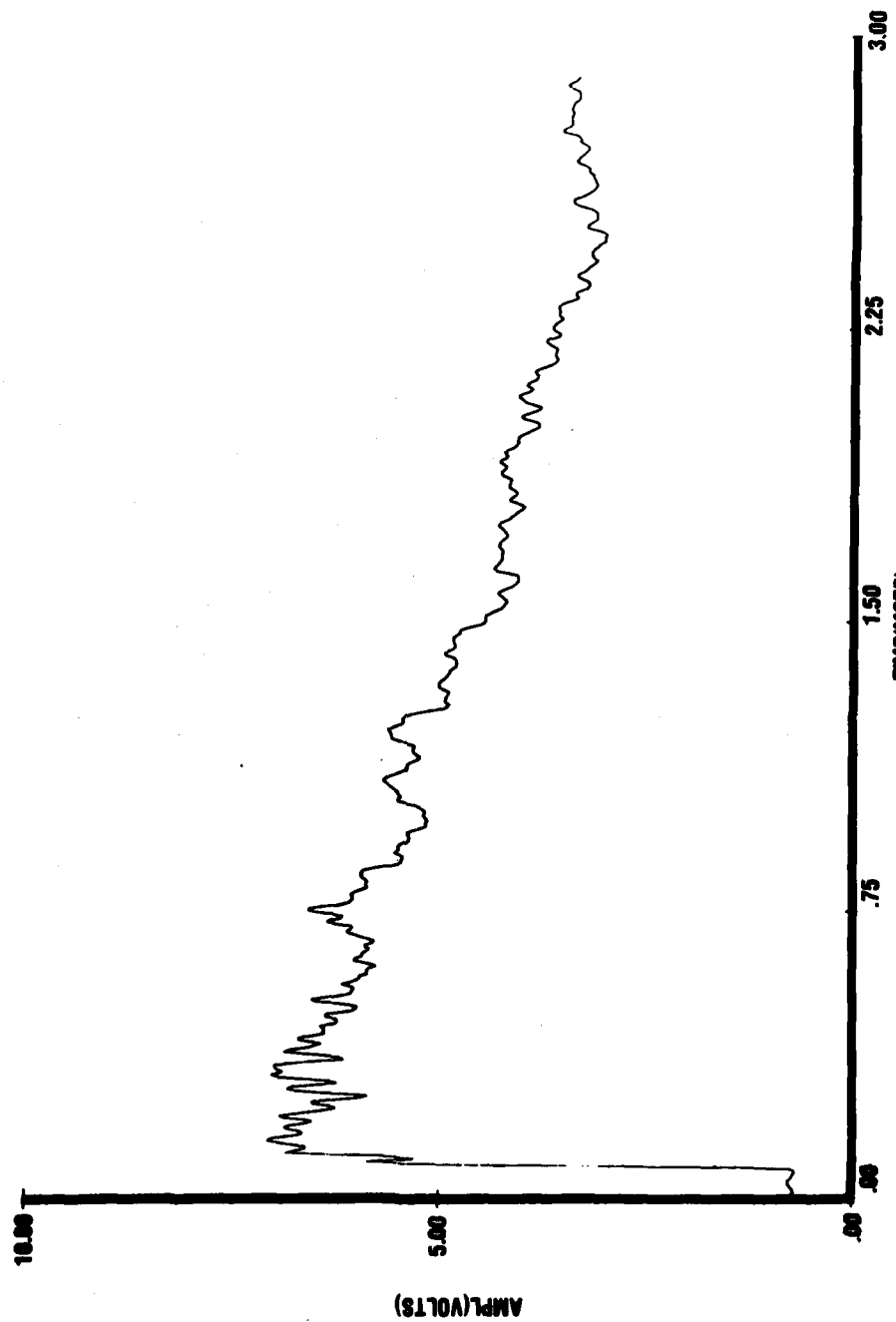


FIGURE 8 BLAST OVERPRESSURE RECORD WITH FIXED SAMPLING RATE

THE SHOCK FRONT RISE TIME IN WATER (U)

*PAUL HARRIS, DR.
ARRADCOM, DOVER, NJ 07801
HENRI-NOEL PRESLES, DR.
UNIVERSITY OF POITIERS, POITIERS, FRANCE

INTRODUCTION

The question of shock front rise time is fundamental to the physics of dynamic high pressure and high strain rate effects in condensed media. The rise time question is currently THE basic question in shock wave physics. Applications range from a better understanding of the shock wave initiation of explosives to the design of radically new sensors for impact fuzing.

In a classic paper[1] Cowan and Hornig presented a study of the optical reflectivity of a shock front in a gas. The theory presented was for the linear case of small reflectivity and constant angle of incidence (within the shock front). The question of different shock structures (i.e., non-constant density gradients) was also considered. In a subsequent paper[2] Flook and Hornig attempted to measure the reflectivity of shock fronts in a number of transparent liquids including water. For the water experiments the reflectivity was too small to be measured.

This paper reports on successful shock front optical reflectivity theory and experiment for water at 5.8 kbar (0.58 GPa), and complimentary shock induced electrical polarization experiment and theory for the pressure range 20 kbar to 100 kbar. The LCWSL - Universite de Poitiers program has experiments performed in France (funded by the Commission Nationale Recherche Scientifique), while the ILIR program within the LCWSL funds the theoretical work.

The consistent interpretation of the reflectivity and shock polarization results is (a) a shock front rise time of approximately 10^{-12} sec throughout the pressure range of 0 to 100 kbar, (b) a transition in the source of the dominant polarization voltage to the alignment of unbonded dipole moments at approximately 35 kbar, and (c) in the unbonded region the average contributing dipole is perfectly aligned in the direction of shock propagation.

SHOCK FRONT REFLECTIVITY

The shock front reflectivity theory and experiment have already been reported on elsewhere[3-5], and consequently that work will only be summarized here.

Figures 1 and 2 show both experimental and theoretical results for the reflectivity of a (laser) optical pulse from a propagating 5.8 kbar shock wave front in water. The experiments represent the first time that optically polarized beams (both parallel and perpendicular to the plane of incidence) were employed, and the theory, again for the first time, utilized an index of refraction and angle of incidence which varied within the thickness of the shock front.

The theory can be summarized via[4]

$$\text{Reflectivity} = \left| \frac{E_r}{E_o} \right|^2 \tag{1a}$$

$$\left| \frac{E_r}{E_o} \right| = \sin\left\{ \frac{1}{2n_o} \int_{\infty}^{\infty} \frac{\partial n}{\partial z} (1 \mp \tan^2\theta) \cos\left[2\pi(\gamma_A - \gamma_B)\right] dz \right\}. \tag{1b}$$

$$(\gamma_A - \gamma_B) = \frac{2}{\lambda_o} \int_{-\infty}^{z} (n^2 - n_o^2 \sin^2\theta_o)^{\frac{1}{2}} dz, \tag{1c}$$

where the integration in Eq. (1c) is to a depth z within the shock front thickness. E is the electric field intensity, n and θ are local index of refraction and angle of incidence respectively, γ_A and γ_B are fractional optical phase angles, and the subscript "o" denotes preshock values while the subscript "r" denotes a reflected parameter. The - and + in Eq. (1b) denote parallel and perpendicular to the plane of incidence respectively.

Both constant gradient[3,4] (i.e. $\frac{\partial n}{\partial z}$ = const) and hyperbolic tangent[5] models for the index of refraction were used in Eqs. (1) with essentially the same analytic result. Aside from being the first accurate observation and analysis of shock front optical reflectivity for condensed media to appear in the literature, the results shown in Figs. 1 and 2 are important in that they provide a shock front thickness upper bound of approximately 400 A at 5.8 kbar.

The 400 A upper bound is extremely important in that it pins the shock polarization results; to now we have only been successful in obtaining valid shock polarization data down to 20 kbar.

SHOCK POLARIZATION

The shock polarization results, including theoretical predictions at low pressures, are shown in Fig. 3. A modified "rigid four-point-charge" model[6] has been used in analyzing the data, and in making the low pressure predictions.

In homogeneous media computer molecular dynamics (CMD) calculations[6] each four-point-charge molecule interacts with neighboring molecules via an electrostatic potential (as well as via an oxygen-oxygen interatomic potential). In the derivation outlined below a single molecule is considered in the presence of a shock front induced torque field. The charges are used to calculate the molecular dipole moment. The model is illustrated in Fig. 4. Because only a single molecule is considered here (i.e. no neighboring molecules), inertial effects associated with the surrounding water structure is very crudely represented by placing hydrogen atom (proton) masses at those point charge positions. CMD or other detailed theoretical justification does not exist for such mass placement.

The fixed on the molecule primed coordinate system of Fig. 4 is imbedded in the laboratory (x, y, z) system where the +z direction is the direction of shock propagation (see Fig. 5). The calculation proceeds by using Euler angle transformations in order to find the torque about the line of nodes. The torque results in an equation of motion

$$\ddot{\theta} = \frac{2\ell_o f_n}{3^{\frac{1}{2}} I} \sin\theta - \frac{k}{m a_{OH}^2} \delta\theta , \qquad (2)$$

where the sin θ term is exact (if f_n is exact) and the second term is a crude restoring force with m being a proton mass, a_{OH} the oxygen-hydrogen intramolecular distance (1A), k is the hydrogen bond bending force constant, and I is the moment of inertia of the model molecule about the line of nodes with the oxygen atom as center.

The shock induced effective force (in the +z direction) on each hydrogen atom, f_n, used here is given by

$$f_n = \frac{\alpha \sigma}{n_o L_s} . \qquad (3)$$

σ is the peak stress amplitude, L_s the shock front thickness, n_o is the pre-shocked state number density of water molecules, and α is the fraction of the molecule force which acts on a hydrogen atom ($\alpha = 1/18$ for H_2O). Eqs. (2) and (3) represent a three dimensional version of a model first attempted by Horie[7].

Equation (2) has two interesting limiting solutions. One solution is where the shock induced torque is just balanced by the hydrogen bond bending restoring force so that $\ddot{\theta} = 0$. The second solution corresponds to the bond broken regime and thus has $k = 0$.

For $\delta\theta$ assumed small the $\ddot{\theta} = 0$ solution can be shown to lead to

$$\langle\cos\theta\rangle = -\frac{\pi \ell_o m a_{OH}^2 f_n}{(12)^{\frac{1}{2}} k\, I}. \tag{4}$$

Similarly the $k = 0$ solution can be put in the form

$$\langle\cos\theta\rangle = -\frac{\pi \ell_o f_n t^2}{(48)^{\frac{1}{2}} I} \tag{5}$$

In arriving at Eqs. (4) and (5) the angle average of $\cos\theta$ over a unit sphere has been used. Eqs. (4) and (5) give the average shock induced projection of the modified rigid four-point-charge along the direction of shock propagation.

Equations (4) and (5) are employed to arrive at the shock induced polarization P

$$P = n p_o \langle\cos\theta\rangle \tag{6}$$

where n is the number density of molecules which contribute to the observed shock polarization signals, and p_o is the four-point-charge molecule electric dipole moment ($p_o = 2.27 \times 10^{-18}$ esu cm from Fig. 4 with $|\vec{V}_1| = |\vec{V}_2| = 1$A and $|\vec{V}_3| = |\vec{V}_4| = 0.8$ A). Because $L_s = U_s \tau_{sh}$ (with τ_{sh} the shock front rise time), Eqs. (5) and (6) predict P proportional to τ_{sh} for $t = \tau_{sh}$.

In turn P is used to calculate the prompt pulse current denisty[7,8]

$$J = \frac{K_o}{K} \frac{P\,(U_s - u_p)}{L_o} \tag{7}$$

where K is dielectric constant, L_o is the sample thickness, and u_p the particle velocity. The form of Eq. (7) is appropriate to cgs units.

The t = 0+ voltage time derivative is given by[9]

$$\dot{V}(0+) = \frac{V(0+)}{\tau_{sh}} = \frac{P(U_s - u_p)}{\varepsilon_o} \qquad (8)$$

where ε_o is the vacuum permittivity (rationalized mks units).

POLARIZATION DATA ANALYSIS

@ Unbonded (i.e. k=0) molecules.

For $\mathcal{P}$ = 100 kbar (10^{11} cgs), $t = \tau_{sh} = 10^{-12}$ sec, $n_o = 3.47 \times 10^{22}$ cm^{-3}, $L_s = U_s \tau_{sh}$, $U_s = 5.05 \times 10^5$ cm/sec (from Ref. 10), and $I = 3.50 \times 10^{-40}$ gm cm^2

$$\langle \cos\theta \rangle = -0.82, \qquad (9)$$

which is equivalent to saying that all of the contributing shock rotated molecules are almost perfectly aligned with the direction of shock propagation at $\tau_{sh} = 10^{-12}$ sec.

For V(0+) = 0.2 Volt one has agreement between the $\tau_{sh} = 10^{-12}$ sec voltage rise time and the right hand side of Eq. (8) providing that

$$\frac{n}{n_o} = 2.7 \times 10^{-3} \qquad (10)$$

The Eq. (10) result is within a factor of 2 of the 100 kbar and corresponding Hugoniot temperature[10] ionic concentrations obtained by using the temperature dependent dissociation constants from Eigen and Maeyer[11] and the empirical conductivity versus hydrostatic pressure results of Holzappel[12].

The Eq. (10) result is also within a factor of 2 agreement with the 0.5% ion concentration at 130 kbar quoted by A.C. Mitchell el al[13] as a consequence of behind the shock front electrical conductivity measurements. It is also within a factor of 2 when compared with the ion concentrations obtained from behind the shock electrical conductivity measurements of Hamann and Linton[14] and an ion mobility of 3×10^{-3} cm^2 V^{-1} sec^{-1} (the atmospheric pressure 25°C value) from Eigen and Maeyer[11].

Applying L_o = 1mm, and $(U_s - u_p) = 3.06 \times 10^5$ cm/sec to Eq. (7) gives 0.2 Volt for the voltage predicted across a 50 Ω load by the experimental cell of 0.5 cm^2 cross section providing that $K = K_o$ and

$$\frac{n}{n_o} = 1.5 \times 10^{-3} \qquad (11)$$

are used. Although $K=K_o$ was used because these authors do not know better, it should be pointed out that in so far as it produces charges which can be separated in an electric field, a high degree of ionicity will increase the dielectric constant. An increased K would result in better agreement between Eqs. (10) and (11).

Using $\tau_{sh} = 10^{-12}$ sec and the above k=0 formulae yields the same very good agreement between theory and experiment down to 35 kbar. Near 35 kbar the slope of the polarization voltage versus pressure changes (see Fig. 3). The implication is a changing shock polarization mechanism.

(b) Hydrogen bonded molecules.

Applying the previously used numerical values at 10 kbar, and a k value[15] of 3.78×10^{-13} erg/radian^2, to Eq. (5) gives

$$\langle\cos\theta\rangle = -1.5 \times 10^{-4}. \qquad (12)$$

In turn Eq. (12) yields, via Eq. (7) and the very reasonable bonded regime assumption that on the average each shock front enveloped water molecule contributes equally to the shock polarization current density (i.e. $n = n_o$), for the predicted prompt peak polarization voltage across a 50Ω load

$$V(50\Omega) = 15 \text{ m Volts}. \qquad (13)$$

$K=K_o$ has been assumed.

For an extrapolation of the Fig. 3 experimental data to ∿ 6mV at 10 kbar, the left hand side of Eq. (8) becomes 6×10^9 Volts/sec while the right hand side is (from Eq. (12)) 8.7×10^9 Volts/sec. It is truly amazing that the very crude bending angle restoring force model utilized in Eq. (2) can give such excellent agreement between theory and experiment.

DISCUSSION

Employing the well known shock velocity values[10] and $\tau_{sh} = 10^{-12}$ sec one finds L_s (100 kbar) = 50A. That shock front thickness value is within a factor of 2 agreement with the computer molecular dynamics results[16] of Tsai and Trevino for shocks in dense monatomic liquids. 50A is consistent with the 5.8 kbar laser reflectivity value upper limit of 400A.

Because of the lack of experimental data in the low pressure region (i.e. <20 kbars), the 5.8 kbar reflectivity work becomes the glue which holds the work over the entire pressure range together. It can be shown that the 400A upper limit prevents the unbonded molecule mechanism from

being the explanation for the observed low pressure polarization signals.

USSR researchers[17-20] have claimed a shock front related viscosity, η, value in the 100 kbar region of approximately 10^4 P. By using

$$\eta \sim L_s \rho U_s \tag{14}$$

from Gilman[21], which should be valid in the 100 kbar range (ρ is mass density), one calculates $\eta \sim 0.2$ P at 100 kbar (from $L_s = 50$A).

There has been considerable discussion in the literature relating to the USSR value of 10^4P. While the reasons for the disagreement are not clear, it should be noted that Hamann and Linton[22], and Hamann[23], have consistently claimed that the USSR η values are incorrect.

The $\tau_{sh} = 10^{-12}$ sec and the corresponding claim with respect to molecular orientation within the shock front (100 kbar to 35 kbar range) represents a major advance in the interpretation of shock front physics. Second, the detail and accuracy of the experiment and theory for the shock front optical reflectivity serves as the low pressure anchor for the entire program reported on here. Further it should be noted that the 400A upper limit reflectivity result limits the many USSR claims that such reflectivity measurements can predict molecular dimension shock front thicknesses; the thickness is too small in comparison to the optical wavelengths to yield sufficient theoretical accuracy for those claims.

That the bond broken regime dominates the polarization signals at high pressures is not surprising; Graham[24] has advanced the concept that shock induced bond scission is the agent responsible for shock polarization and increased electrical conductivity in polymeric solids. At the moment, however, it does not appear that bond breaking is the polarization producing agent in the 10 kbar region.

ACKNOWLEDGEMENTS

The authors are deeply indebted to Y. Sarazin and H. Simonnet (Univ. de Poitiers) for experimental assistnace. They also gratefully acknowledge useful discussions with R.A. Graham (Sandia National Labs), S.D. Hamann (CSIRO, Australia), Arnold Karo (Lawrence Livermore National Laboratories), and F.H. Stillinger (Bell Telephone Laboratories).

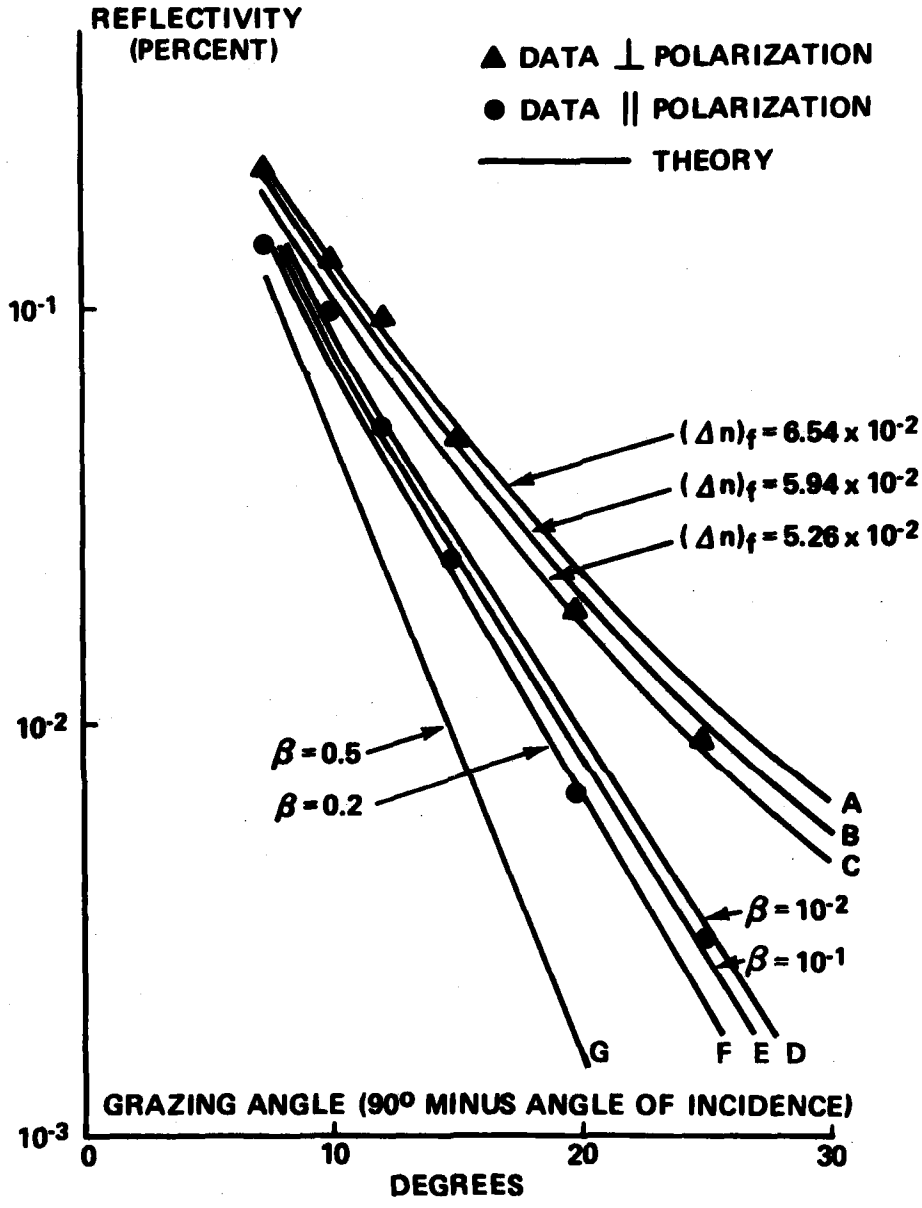


Figure 1. Theory and experimental data for shock front reflectivity at 5.8 kbar in water. Curves (A), (B), and (C) are for $\perp$ polarization and $\beta = 0.1$. Curves (D) through (G) for $\|$ polarization and $(\Delta n_f) = 5.26 \times 10^{-2}$. $\beta \equiv \frac{n_0 L_0}{\lambda_0}$, and $(\Delta n)_f$ is the total change in index of refraction across the shock front.

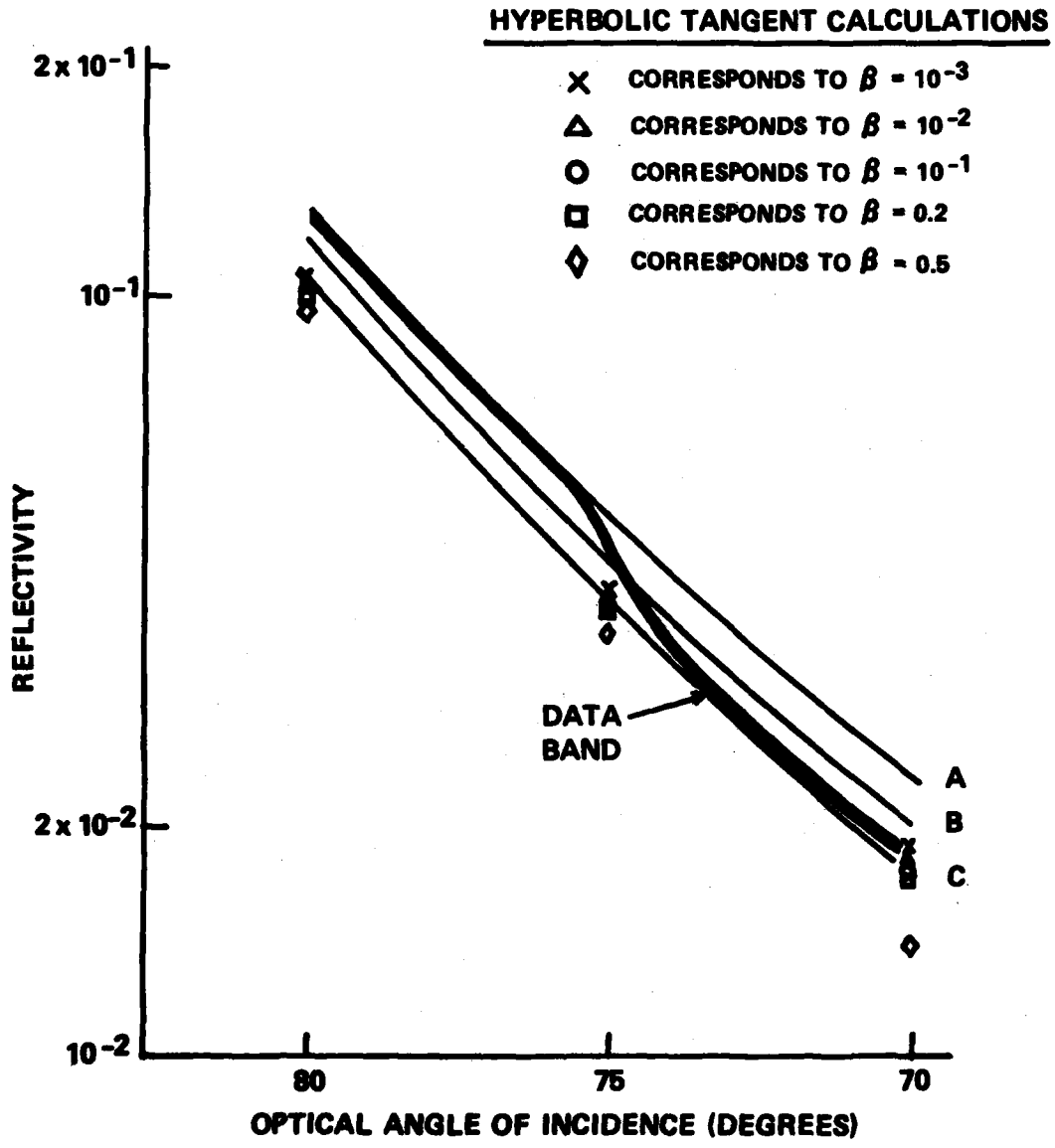


Figure 2. Hyperbolic tangent shock front calculations for perpendicular optical polarization at 5.8 kbar. Curves A, B, and C are constant gradient model results (as in Fig. 1). All hyperbolic tangent results are for $(\Delta n)_f = 5.26 \times 10^{-2}$. $\beta = Ln_o/\lambda_o$.

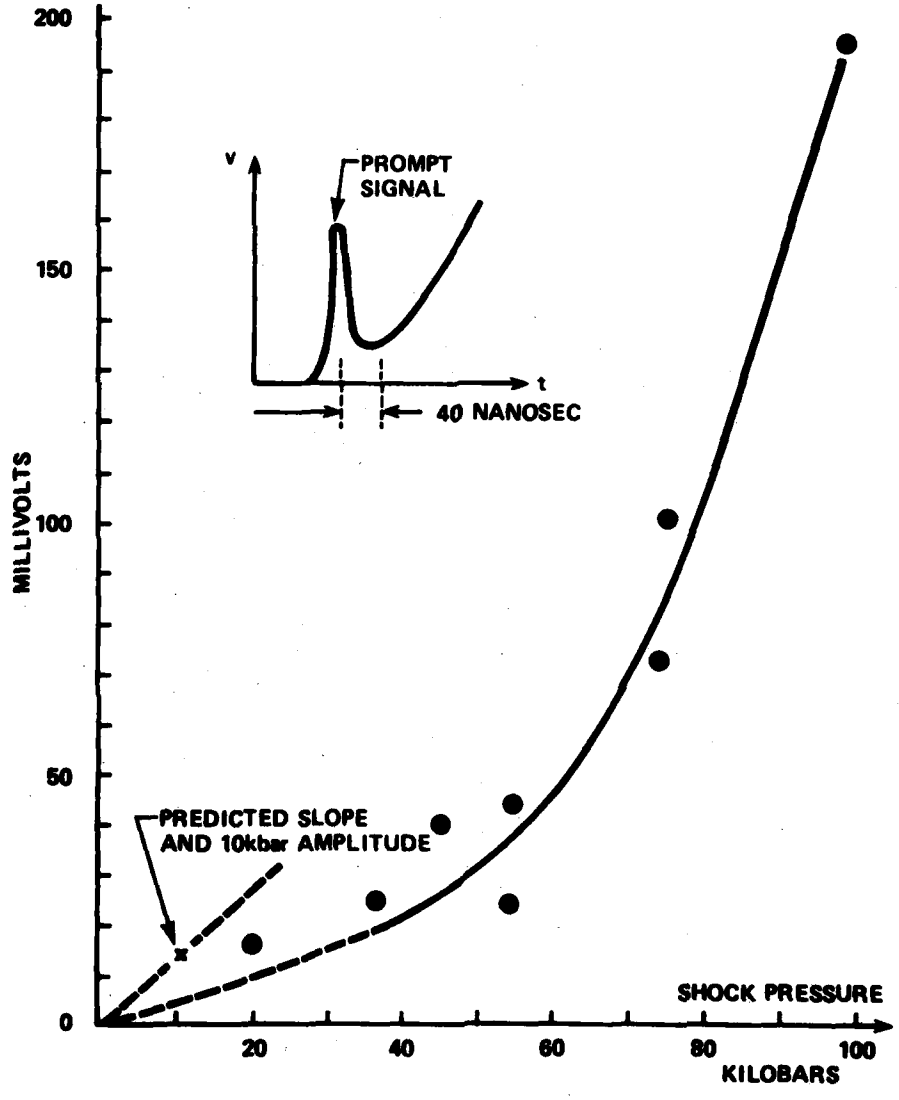


Figure 3. Averaged experimental data. The predicted slope and amplitude at 10 kbar is for net alignment of hydrogen bonded dipoles.

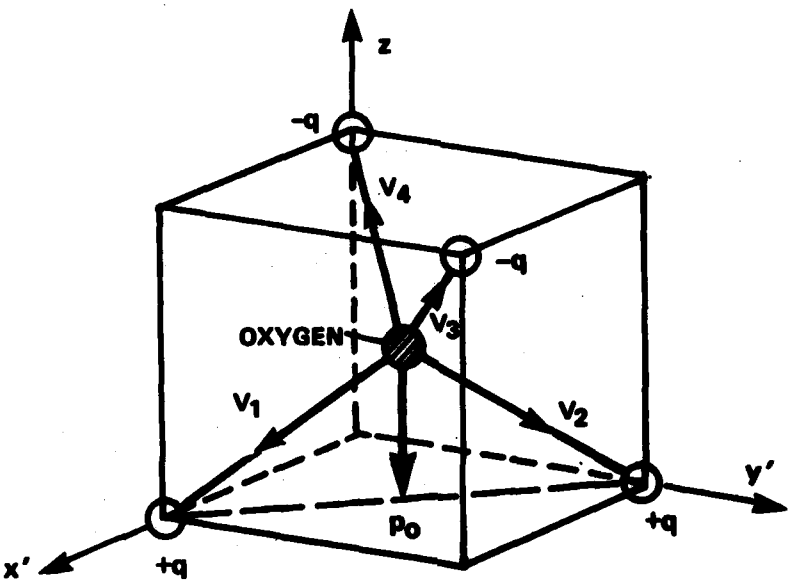


Figure 4. The tetrahedral geometry of the rigid four-point-charge model. p_o is the net molecular polarization, and the $\vec{V}_i$ are the coordinate vectors of the charges (and associated hydrogen masses). q = 0.236 electron charges.

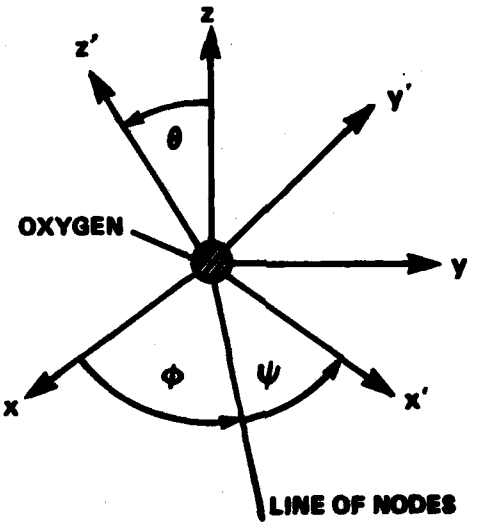


Figure 5. The Euler angle corodinates. The rigid coordinate system attached to the rigid molecule of Figure 4 is allowed to rotate within the fixed system (x, y, z). (θ, ϕ, ψ) are Euler angles, and +z is the shock propagation direction.

REFERENCES

1. G.R. Cowan and D.F. Hornig, J. Chem. Phys. 18, 1008 (1950).

2. W.M. Flook, Jr. and D.F. Hornig, J. Chem. Phys. 23, 816 (1955).

3. P. Harris and H.-N. Presles, "The Optical Reflectivity of a Shock Front", Tech. Report of the Laboratoire d'Energetique et de Detonique, Univ. de Poitiers, 86034 Poitiers, France. Jan. 1981.

4. P. Harris and H.-N. Presles, J. Chem. Phys. 74, 6864 (1981).

5. P. Harris and H.-N. Presles, "Shock Front Rise Time in Water", ARRADCOM Tech. Rept. ARLCD-TR-81032, ARRADCOM, Dover, NJ 07801. Oct. 1981.

6. F.H. Stillinger and A. Rahman, J. Chem. Phys. 60, 1545 (1974).

7. Y. Horie, Brit. J. Appl. Phys. (J. Phys. D) 1, 1183 (1968).

8. F.E. Allison, J. Appl. Phys. 36, 2111 (1965).

9. H.A. Monterrubio, thesis "Analisis de los Resultados Experimentales para la determinacion del Salto de Polarizacion ΔP Inducida por Choque". Universidad Nacional Autonoma de Mexico (1981).

10. M.H. Rice and J.M. Walsh, J. Chem. Phys. 26, 824 (1957).

11. M. Eigen and L. De Maeyer, Proc. Roy. Soc. A 247, 505 (1958).

12. W.B. Holzappel, J. Chem. Phys. 50, 4424 (1969).

13. A. C. Mitchell, M.I. Kovel, W.J. Nellis, and R.N. Keller in High Pressure Science and Technology, edited by B. Vodar and Ph. Marteau (Pergamon, New York, 1980). Vol. 2.

14. S.D. Hamann and M. Linton, Trans. Faraday Soc. 62, 2234 (1966).

15. J.A. Pople, Proc. Roy. Soc. A 205, 163 (1951). See also D. Eisenberg and W. Kauzmann, The Structure and Properties of Water (Oxford Univ. Press, Oxford, 1969). Section 5.3.

16. D.H. Tsai and S.F. Trevino, Phys. Rev. A 24, 2743 (1981).

17. O.N. Brensov, A.N. Dremin, O.K. Rosanov, V.N. Kochnev and V.V. Yakushev, Soviet Electrochemistry 5, 669 (1969).

18. O.N. Brensov, A.N. Dremin, V.N. Kochnev, O.K. Rosanov, and V.V. Yakushev, Soviet Electrochemistry 7, 395 (1971).

19. V.V. Yakushev and A.N. Dremin, Soviet Phys. Chem. 45, 50 (1971).

20. V.M. Shunin, A.N. Dremin, and V.V. Yakushev, Soviet Phys. Chem. 251, 246 (1980).

21. J.J. Gilman, J. Appl. Phys. 50, 4059 (1979).

22. S.D. Hamann and M. Linton, J. Appl. Phys. 40, 913 (1969).

23. S.D. Hamann in the 1979 Novel Berzelius Symposium on "Chemistry and Geochemistry of Solutions at High Temperature and Pressure" (Pergamon, New York, 1981). Also private communications.

24. R.A. Graham, J. Phys. Chem. 83, 3048 (1979).

DETONATION OF RAPIDLY DISPERSED POWDERS IN AIR

DAVID C. HEBERLEIN, PhD
US ARMY MOBILITY EQUIPMENT R&D COMMAND
FORT BELVOIR, VIRGINIA

INTRODUCTION

The future battlefield will have a high density of sophisticated, lethal weapons. Mines will be used on this battlefield together with nuclear, biological and chemical weapons to retard and restrict the advance of enemy forces across wide land areas. A wide variety of remotely deliverable, air scatterable mines will be available to interdict, delay and destroy enemy forces. For US forces to remain effective in this hostile environment, friendly forces must be able to maintain their tempo of operations under fire. These unique requirements dictate the development of real time countermine systems, effective against a wide variety of fuze types, that quickly clear large land areas. Explosive technology is one of a small number of technological areas capable of meeting these special mission requirements.

Traditionally, explosive weapons such as line charges have been used to produce single impulses sufficient to deflect mine pressure plates and thereby cause mine detonation. Recent advances in the development of fuel air explosive (FAE) weapons has lead to mine neutralization systems that are effective against pressure plate mines over wide land areas. Present FAE weapons explosively disseminate and detonate a liquid fuel over a wide land area. A small central explosive burster disperses the liquid fuel as small droplets suspended in a large air mass. A detonator, also dispensed into the fuel-air cloud, is detonated after an appropriate delay. While these liquid fuel systems work well for mine neutralization applications, there are shortcomings to this approach: (1) the maximum pressures obtainable for stoichiometric combustion; (2) approximately 10% of the liquid fuel rounds result in burning and not in detonation; (3) long term storage of certain fuels results in polymerization of the liquid. The use of solid powders in wide area mine neutralization weapons obviates many of the difficulties

intrinsic to liquid fuels. Solid particle sizes can be controlled during manufacture so that explosive dispersal can be accurately and reliably controlled. In addition, solids have long shelf lives reducing the problems associated with long term storage of munitions. Most importantly, however, the terminal weapon effects for liquid fuel weapons are limited by the pressures they can produce.

The vulnerability of a structure to blast overpressures is given by a critical load curve in which peak pressures are plotted versus impulse. The changing values of P, I given by each curve is shown schematically in Figure 1 together with an indication of the performance of FAE weapons against these mines. The detonation of dispersed powders does not share the limitation in maximum overpressures found for the liquid fuel-air systems. Solid powders such as explosive powders and combinations of flaked aluminum and ammonium perchlorate, carry their own oxygen at the molecular level for efficient combustion during detonation.

The pressure P, obtainable in detonation of a powder-air mixture is a function of the heat in a particular solid available to be released in detonation, Q, the density of explosive, ρ , and the detonation velocity, D, such that

$$P = P(\rho, Q, D). \qquad (1)$$

The maximum pressures attainable in powdered, dispersed solid systems that carry molecular oxygen is easily related to the density of explosive, the energy available for combustion and the detonation velocity. Therefore, dispersed powdered systems can be developed in which the pressures obtained are determined by the density of the dispersed powder-air mixture and the combustion energy intrinsic to the selected solid.

THEORETICAL BASIS FOR THE DETONATION OF POWDERS DISPERSED IN AIR

A theoretical model of the detonation of fine powders dispersed in air must accurately describe the exothermic release of chemical energy, the compressibility of the reaction products and chemical reaction rate effects. If it can be assumed that the reaction products are gaseous and rapidly mixed and that reaction rate effects are manifested in a detonation wave propagating through the unreacted mixture of a powder suspended in air, then details of the detonation process at the particle scale will not substantially affect detonation parameters. The dispersed powder-air mixture can be represented as a molecular, heterogeneous medium at an averaged thermodynamic state. The Chapman-Jouguet (CJ) detonation wave model is appropriate for this medium if the reaction

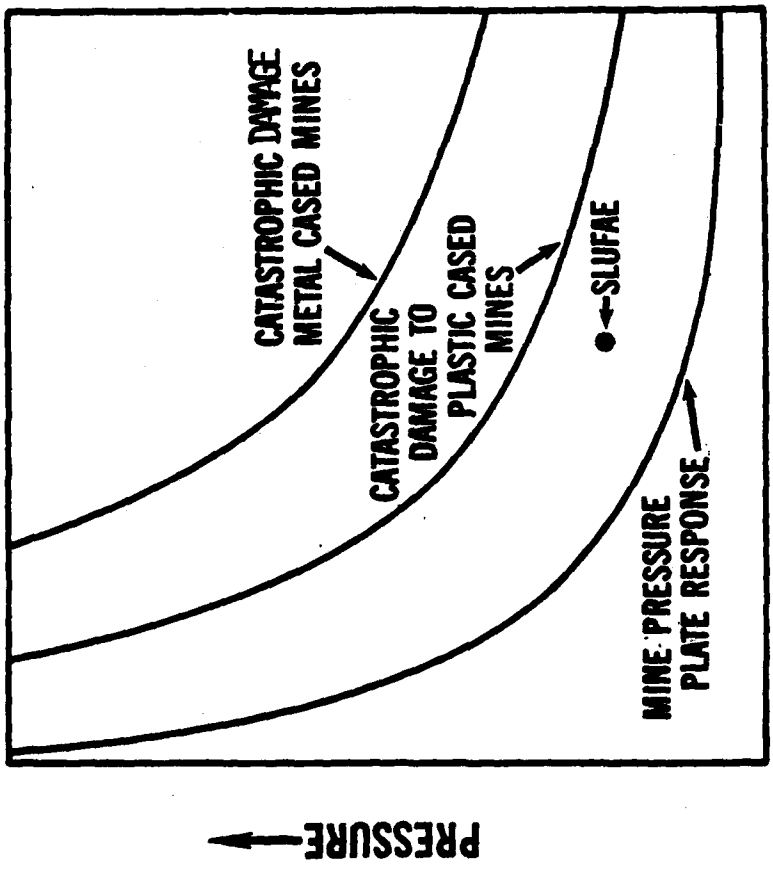


FIGURE 1. LAND MINE STRUCTURE YIELD CURVES

front is handled as a discontinuity across which an average instantaneous energy is released. Flickett and Davis[1] have shown from energy, mass and momentum relations at the CJ discontinuity that

$$P_o = 2(K-1)\rho_\infty E \,, \tag{2}$$

$$\rho_o = \frac{K+1}{K}\rho_\infty \,, \tag{3}$$

$$D = \left[\frac{(K+1)P_o}{\rho_\infty}\right]^{1/2} \,, \text{ and} \tag{4}$$

$$\rho_\infty = \rho_{air} + \left\{1 - \frac{\rho_{air}}{\rho_{fuel}}\right\} C \approx \rho_{air} + C, \tag{5}$$

where Po is the detonation pressure, K the ratio of the specific heats of the combustion products, ρ_∞ the loading density, E the exothermic reaction energy, ρ_o the density and the CJ state, D the detonation velocity and C the concentration.

The amount of energy released depends on the chemical and thermodynamic processes involved in detonation. A fundamental assumption in the Chapman-Jouguet theory of detonation is that chemical equilibrium exists among the reaction species in the reaction zone of the detonation, i.e. all of the chemical reaction energy contributes to the propagation of the detonation. It is convenient to evaluate the energy, E, for a constant volume adiabatic process so that the result can be interpreted as independent of such combustion properties as the nature and distribution of solid particles. The constant volume pressure, Pv, is given by

$$P_v = (K-1)\rho_\infty E \,. \tag{6}$$

If the values for K and E are approximately the same[2], then

$$P_o = 2P_v. \tag{7}$$

The constant volume pressure, Pv, is easily observed experimentally as the pressure observed for a detonation in a fixed volume "closed bomb".

---

[1] Flickett, W. D., "Detonation", University of California Press, Berkeley, CA, 1979.

[2] Zeldovitch, I. B., and Dompaneets, A. S., "Theory of Detonation", Academic Press, New York, 1960.

To compare this experimental value for the pressure given in equation 7 with that theoretically expected for a specific powder, the theoretical pressure is obtained for the temperature of the detonation. The temperature attained depends on the heat evolved in chemical reactions less that consumed in physical phase changes and the heat capacities of the reaction products obtained. The mathematical statement for chemical equilibrium of multi-reaction is a minimum. The Gibbs free energy is minimized using a La Grange multiplier technique developed by Naphtali[3]. A table of stoichiometric coefficients is formed in which each column represents a reaction and each row a component. The energy in row i, column j is

$$\nu_{ij} = (\partial n_i)/(\partial \xi_j), \tag{8}$$

where n_i is the number of moles of component i and ξ_j is the extent of reaction j. A change in composition is given by

$$dn_i = \sum_j \nu_{ij} d\xi_j \tag{9}$$

According to thermodynamics, $G = \sum_i \mu_i n_i$ where μ_i is the standard free energy of formation per mole. At a fixed temperature and pressure, we obtain

$$dG = \sum_i \mu_i \left\{\sum_j \nu_{ij} d\xi_j\right\} = \sum_j \left\{\sum_i \mu_i \nu_{ij}\right\} d\xi. \tag{10}$$

The free energy change for reaction j can be defined as

$$\Delta G_j = \sum_i \mu_i \nu_{ij}, \tag{11}$$

so that equation 10 becomes

---

[3]Naphtali, N. M.. "Calculate Complex Chemical Equilbira", Ind. Engr. Chem., 53, 387, 1961

$$dG = \sum_j \Delta G_j \, d\xi_j \, . \tag{12}$$

For equilibrium dG=0 requires that $\Delta G_j = 0$. The condition ΔG_j forms a set of non-linear, equlibrium equations. The difficulty in minimizing $\Delta G_j = 0$ increases as the number of equations increases.

Any change in composition is thermodynamically possible if $dG \leq 0$. A series of these changes should lead to a minimum point, i.e. equilibrium composition. By choosing $d\xi_j = -\Delta G_j \, d\lambda$ for $d\lambda \geq 0$,

$$dG = -\left\{ \sum_j (\Delta G_j)^2 \right\} d\lambda \, . \tag{13}$$

A change in composition is then given by

$$dn_i = -\left\{ \sum_j \nu_{ij} \Delta G_j \right\} d\lambda = -\sum_k \left\{ \sum_j \nu_{ij} \nu_{kj} \right\} d\lambda . \tag{14}$$

A symmetric table of values ε_{ik} can be formed for the condition

$$\varepsilon_{ik} = \sum_j \nu_{ij} \nu_{kj} = \varepsilon_{ki} \, , \tag{15}$$

so that

$$\frac{(dn_i)}{(d\lambda)} = -\sum_k \varepsilon_{ik} \mu_k \tag{16}$$

gives the direction of decreasing free energy for any composition. Using Equation 10

$$\frac{(dG)}{(d\lambda)} = \sum_i \mu_i \left[\frac{(dn_i)}{(d\lambda)} \right] = -\sum_j (\Delta G_j)^2 , \tag{17}$$

the proximity to equilibrium can be obtained from the sum of the square for the ΔG_j's for all reactions. Near equilibirum each ΔG_j should be near zero and also the sum, $\sum_j (\Delta G_j)^2$.

For the detonation of aluminum powders in air, the potential reactants and products include the following.

TABLE I

REACTANT	PRODUCT
Al(s)	Al(s), Al(l), Al(g)
Al_2O_3(s)	Al_2O_3(s), Al_2O_3(l), Al_2O_3(g)
N_2(g)	AlO(g), AlO_2(g), Al_2O_2(g)
O_2(g)	$HAlO_2$
H_2O(g)	N_2(g), O_3(g), NO(g), O(g), N(g), H_2O(g), OH(g), H(g), NH_3(g)

Using this set of reactants and products together with thermodynamic heat capacity data for each of these species, chemical equilibrium compositions were calculated as a function of aluminum concentration in air. The result of these calculations for the Chapman-Jouguet Pressure, Equation 7, for aluminum powder detonated in air is shown as a function of aluminum concentration in Figure 2.

EXPERIMENTAL APPARATUS

The shock tube is a good diagnostic apparatus for studying one-dimensional explosive phenomena. Shock tubes can be used to investigate extrinsic properties of powder-air systems such as pressure, temperature and concentration as well as intrinsic properties such as particle size and ignition induction time. Homogeneous powder-air systems were investigated with a detonation tube with the provision for the rapid formation of powder-air mixtures. To assess the powder distribution homogeneity, the concentration of the powder-air mixture was measured through laser

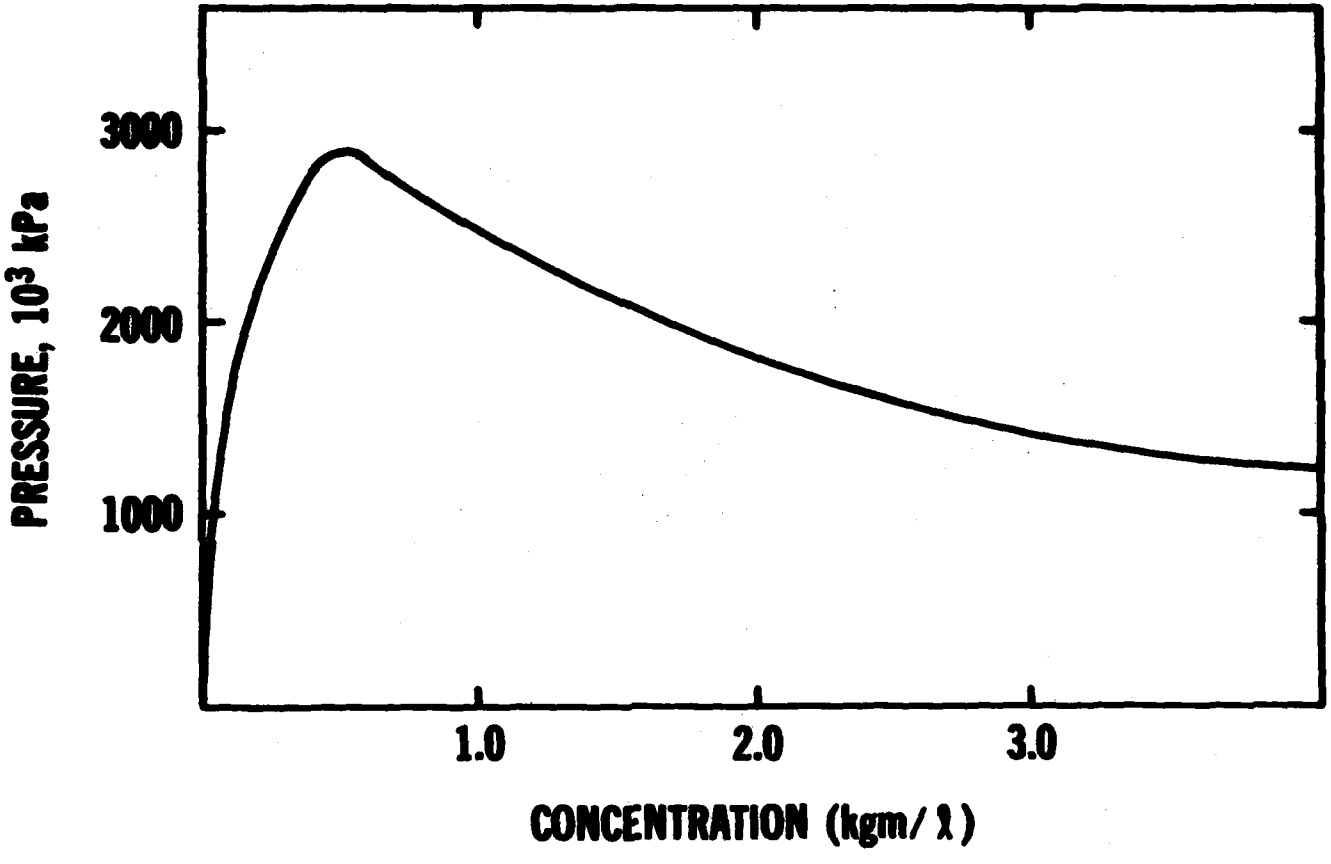


FIGURE 2. CALCULATED CHAPMAN–JOUGUET PRESSURE AS A FUNCTION OF ALUMINUM CONCENTRATION IN AIR

beam attenuation.

The detonation tube consisted of three separate sections, each 1.8m in length. The shock tube was mounted vertically for a total height of 5.4m. Each steel tube section had an internal diameter of 152mm and a 25mm wall thickness. The shock tube was designed for static pressures of 14.8GPa. Special clamps were used to hold the three sections of the tube together. Both ends of the detonation tube were open. Test powder was introduced into the bottom and vented at the top of the tube. A blower was placed at one end of the tube to pass air through the tube continuously. A manometer was used to monitor the air flow through the shock tube. The entire shock tube assembly is shown schematically in Figure 3.

Pressure measurements were obtained from piezoelectric gages mounted in the center test section. These gages were mounted in Nylon-6 to dampen acceleration effects from the shock tube. A thin coating of silicone grease containing powdered graphite was used to prevent a thermal response from the gages. All gages were initially calibrated and routinely recalibrated in situ using known shock pressures. Pressure gage responses were amplified by Tektronic AM502 amplifiers and recorded on a Bell and Howell 3700E tape recorder. The concentration of the powder was monitored by measuring the attenuation of the He-Ne laser (632mm) from a Spectra-Physics Model 132 laser. The attenuated beam was measured by a Hewlett Packard Photodiode, Model 5082-4220.

The detonation tube was calibrated using homogeneous gas phase detonations. Calibration tests used near stoichiometric ethylene-air gas phase mixtures. These tests resulted in detonation velocities of 1859 ± 15 m/sec and a detonation pressure of 298 ± 200 psi, consistent with reported Champan-Jouguet values of 1750 m/sec and 270 psi respectively.

EXPERIMENTAL RESULTS

The detonation tube was used to assess the deflagration and detonation characteristics of various fuel powders. The results of these tests are shown in Table 2.

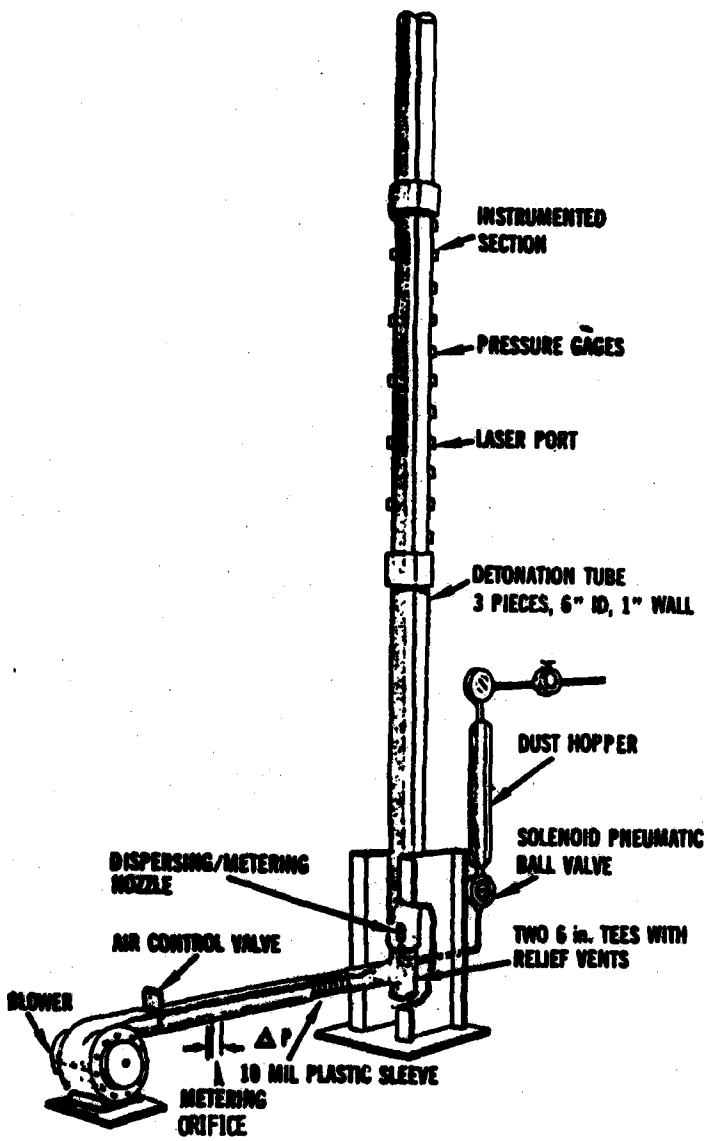


Figure 3. SCHEMATIC DRAWING OF SHOCK TUBE APPARATUS.

TABLE II

Powder	Concentration kg/m^3 Stoich	Exp	Propagation Vel (km/sec)	Overpressure (kPa)
Coal	.10	.36	.58±0.02	371
Microthene^R	.12	.19	.58±0.04	392
Starch	.12	.46	.63±0.03	427
Magnesium	.45	.38	.61±0.06	434
Aluminum	.33	.40	1.64±0.11	5097
Aluminum	.33	.60	1.62±0.12	6352
Aluminum	.33	.50	1.46±0.09	5759
Aluminum	.33	.80	1.50±0.10	5434

All of these powder dispersions propagated a shock wave disturbance based on film records and pressure measurements. However, events involving coal dust, Microthene^R, starch and magnesium were considered to be deflagrations. The flaked aluminum powder did detonate. Sedgwick et. al.[4] gives the variation of the detonation velocity with concentration of aluminum in air shown in Figure 4. It is possible that the variance of the experimental data (approximately 1.6km/sec) from the Chapman-Jouguet velocity (1.8km/sec) was due to variation from stoichiometric concentration, although wall losses due to inadequate tube size could also account for a decreased detonation velocity. The extended reaction zone found in two phase heterogeneous systems (aluminum in air) also contributes to differences in experimental data from that calculated for Chapman-Jouguet conditions. Sichel and Schaaff[5] have shown that it is

---

[4]Sedgwick, R. T., Kratz, H. R. and Baker, M., "Concepts for Improved Fuel Air Explosives", Systems, Science and Software Report SSR-R-77-3005, September 1976.

[5]Sichel, M. and Schaaff, J., "Fuel Air Explosives - Theoretical Considerations", AFATL-TR-71-171, 1, 109, December 1971.

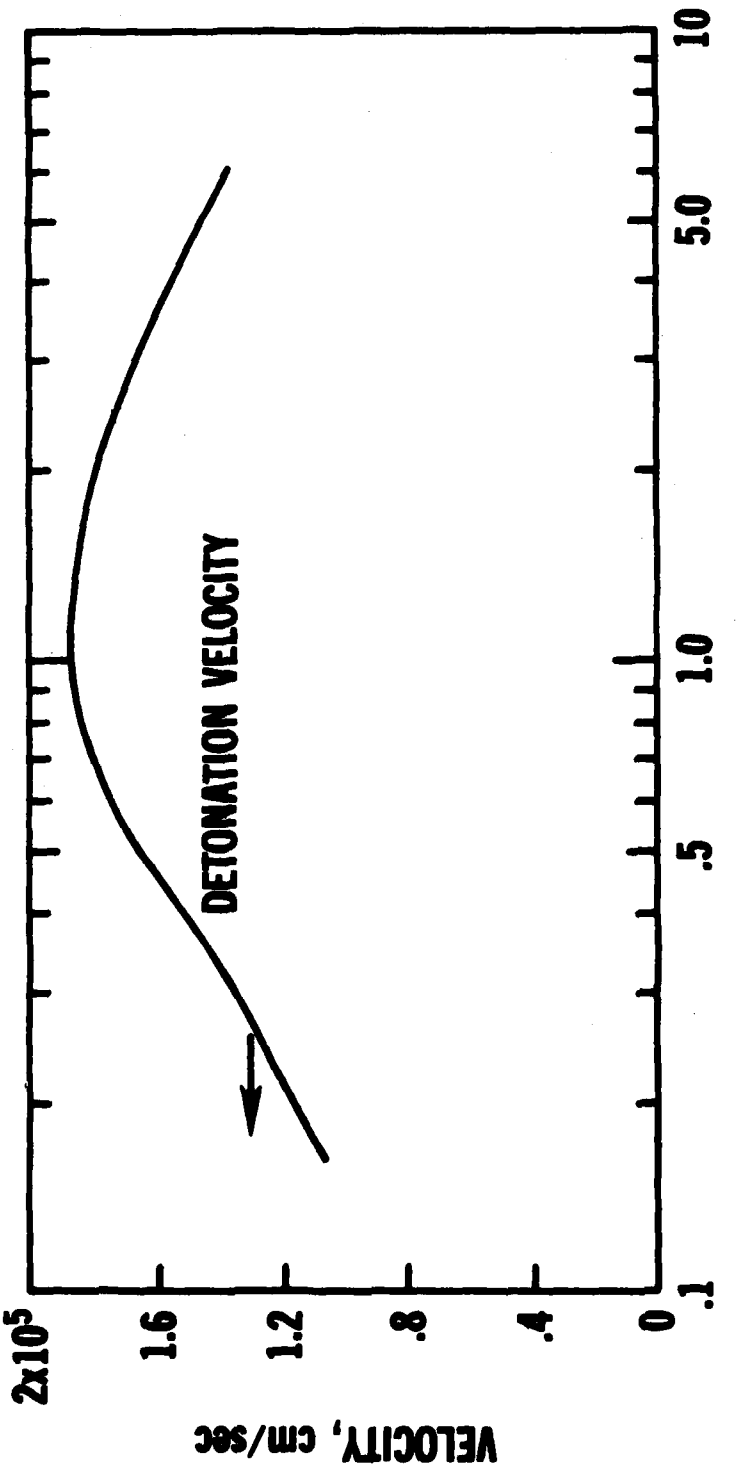


FIGURE 4. CALCULATED DETONATION VELOCITIES FOR CHAPMAN–JOUGUET DETONATION IN ALUMINUM-AIR MIXTURES AS A FUNCTION OF CONCENTRATION

necessary to consider the fact that fuel particles do not contribute to the pressure or volume occupied by the unreacted fuel (powder) and air mixture. When this aspect of two-phase detonations is considered, then it has been shown that the detonation velocity of the two-phase system is lower and the detonation pressure higher than that which could be calculated for a purely gaseous detonation.[6, 7]

Figure 2 gives the calculated Chapman-Jouguet pressures for the detonation of various aluminum concentrations in air. The calculated Chapman-Jouguet pressure nowhere exceeds 3,000 kPa, yet stable pressures were observed experimentally up to twice this value. The experimental pressure responses are believed to be accurate to the degree indicated in Table II. Gidaspow[8] has pointed out that in two-phase flow there are also complex effects that can cause instability effects that may be manifested in the pressure and other characteristics of heterogeneous detonations. It has been demonstrated that disproportionation exists in the induction zone of propagating two-phase detonations. This implies unequal phase velocities of the two phases and a more complex analysis of the instability effects. Alternatively spinning detonation of two-phase systems also increases the observed pressure while lowering the detonation tube at a lower linear velocity (and a higher rotational velocity) and creates substantially higher pressures at the walls of the detonation tube.

---

[6]Williams, F. A., "Structure of Detonations In Dilute Sprays", Phys. Fluids, 4, 1961.

[7]Nicholls, J. A., Debora, E. K. and Ragland, K. W., "A Study of Two-Phase Detonation As It Relates To Rocket Motor Combustion Instability", NASA CR 272, 1965.

[8]Gidaspow, D., "A Thermodynamic Theory of Two-Phase Flow With Unequal Phase Velocities", NATO Advanced Study Institute on Two-Phase Flows and Heat Transfer, August 1976.

CONCLUSIONS

Fine powders of flaked aluminum have been rapidly dispersed and then detonated in air. These powder-air mixtures produce substantial overpressures. Chemical combustion reactants and products have been identified for an aluminum detonation in air. A La Grange multiplier technique was used to minimize the Gibbs Free energy associated with the chemical equilibrium combustion products. Chapman-Jouguet pressures and detonation velocities were calculated for the combustion of aluminum powders in air over a large range of powder concentrations.

Experimentally, a fluidized shock tube was constructed to study the detonation of heterogeneous powder-air systems over a wide range of cloud densities. Dynamic particle concentrations were measured as a function of cloud overpressure and detonation velocity. Experimental measurements of pressures, detonation velocities and cloud densities indicated that dispersed powders of coal dust, Microthene^{R}, starch, magnesium and aluminum propagated shock wave disturbances. The events involving coal dust, Microthene^{R}, starch and magnesium were determined to be deflagrations. The flaked aluminum powder detonated over a wide range of cloud densities. The aluminum cloud pressures observed experimentally were higher than those reported for liquid fuel air explosive clouds. The pressures observed for the aluminum-air clouds were higher than those calculated by minimizing the Gibbs Free energy. Correspondingly, lower detonation velocities were observed than those obtained from theoretical calculations. These variances are attributed to the extended reaction zone found in two-phase, heterogeneous systems.

The detonation of dispersed powders has a great potential for future weapons to generate significantly higher pressures over wide land areas than that presently available through present liquid fuel air explosives. Because powders (explosives, combinations of fuel and oxidizer such as aluminum plus teflon or ammonium perchlorate, etc.) carry their own oxygen at the molecular level, the pressures generated by the detonation of these powder-air clouds will not be restricted by the amount of atmospheric oxygen available for combustion. In fact, the pressures generated by these clouds can be controlled through the density of the powder-air clouds and the chemical energy intrinsic to the solid powders used to generate each cloud. The higher pressures and impulses from these clouds will be capable of causing catastrophic damage to land mines or other structures on the battlefield. Used in a countermine role, the development of powder-air systems represents a potential for significant new capability to destroy or neutralize mines over wide land areas irrespective of mine fuzing.

RAPID IDENTIFICATION OF DENGUE VIRUS SEROTYPES USING MONOCLONAL ANTIBODIES IN AN INDIRECT IMMUNOFLUORESCENCE TEST (u)

*ERIK A. HENCHAL, CPT, MSC
J.M. McCOWN, MR.
M.K. GENTRY, MS.
W.E. BRANDT, DR.
WALTER REED ARMY INSTITUTE OF RESEARCH
WASHINGTON, D.C.

Dengue and dengue hemorrhagic fever occur in epidemic and endemic form throughout tropical areas of the world. Dengue virus types 1, 2, and 3 have been associated with epidemics of major impact in the Western Hemisphere since the 1960's (1-3) and in 1981 dengue type 4 was identified in the Carribean for the first time (4). Extensive serological crossreactions occurring among dengue virus serotypes as well as with other flaviviruses in commonly employed serological tests (i.e., hemagglutination-inhibition, immunofluorescence, and complement fixation) frequently interfere with identification of the antigenic type of dengue virus present in epidemic or endemic areas (5-8). Presently, the only certain method of identification is through the use of rigidly standardized reference antiserum in a virus plaque-reduction neutralization assay (3,9,10). Few laboratories possess sufficient resources to perform this test with the slowly replicating dengue viruses.

Monoclonal antibodies that were produced using the recently established hybridoma technology (11,12) have been used successfully to characterize viral antigens (13,14,15). The present study had as its objective the development of highly specific monoclonal antibodies suitable for rapid serotype identification of low passaged or unpassaged dengue virus from humans or insects using an indirect immunofluorescence test.

MATERIALS AND METHODS

Cell cultures and media

LLC-MK2 (monkey kidney) cells were grown in medium 199 supplemented with 20% heat-inactivated fetal bovine serum (FBS), penicillin (100 units/ml) and streptomycin (100 ug/ml). The C6/36 clone of *Aedes*

albopictus cells was grown in Earles minimum essential medium containing 10% FBS, penicillin, and streptomycin. Maintanance medium for infected cell cultures consisted of the appropriate growth medium containing 0.4% bovine plasma albumin instead of FBS.

Virus

Prototype dengue strains used in this study included DEN-1(Hawaii), DEN-2(New Guinea C), DEN-3(Philippines H87), and DEN-4(Philippines H241). Other members of the flavivirus group included Banzi(SAH336), Ilheus, Japanese encephalitis(M1/311), Kunjin(MRM-16), Langat(TP-21), Ntaya, St. Louis encephalitis(TBH-28), West Nile(E-101), Yellow fever(French neurotropic and 17D strains), and Zika. Two Sandfly Fever viruses (213452 and Candiru) were included as non-flavivirus controls. Virus inocula for cell culture studies were derived from routinely passaged virus-infected mouse brain stocks passaged 1 to 4 times through C6/36 cells.

Virus isolates were generously provided by Dr. Esmie Rose and Dr. P. Prabhaker (University of the West Indies, Kingston, Jamaica), Dr. A. Dyer (Ministry of Health, Jamaica) and Dr. J.P. Digoutte (Institute Pasteur, Dakar, Senegal). Virus isolates from the Carribean were provided as first passage isolates (Aedes pseudoscutellaris cells, AP-61) or human serum from recent dengue virus patients. African isolates were provided as acetone-fixed AP-61 cells on 12-well fluorescent antibody slides prepared at the Institute Pasteur, Dakar, and then tested by one of us (EAH) as described below during a visit there.

Virus isolates were processed by inoculating two LLC-MK2 cell cultutres and one C6/36 cell culture with 0.2 ml of each specimen. One LLC-MK2 cell culture was reserved for direct plaquing using an agar overlay; the other flasks were observed daily for cytopathic effects (CPE). After 7 days the flask reserved for direct plaquing was stained using a neutral red agar overlay and aliquots were removed from the other flasks for plaque assay of virus yields. When fifty percent of the cells showed CPE (7 to 14 days), cell culture supernatants were assayed for virus yield and then the infected cells were processed for immunofluorescence.

Hyperimmune mouse ascitic fluids (HMAF)

The preparation of HMAF containing reference antibodies directed against each of the prototype dengue viruses as well as selected arboviruses used in this study was described previously (16).

Monoclonal antibodies

Monoclonal antibodies directed against dengue virus antigenic

determinants were prepared by fusing P3x63Ag8 mouse myeloma cells with spleen cells from mice immunized with dengue virus antigens as described previously (17). Antibodies secreted by the hybridomas were detected by solid phase radioimmunoassay (RIA) (18). RIA-positive hybridomas which were reactive by the immunofluorescence assay (IFA) and sometimes by the hemagglutination-inhibition test (HAI) (5) or the plaque-reduction neutralization test (PRNT) (9,10) were selected for cloning, ascites production, and further study. Ascitic fluids were generally used in these studies because they contained on the average 100-to-1000 times more anti-dengue antibody than hybridoma tissue culture supernatants as detected by solid phase radioimmunoassay.

Indirect immunofluorescence assay (IFA)

Uninfected and infected cells displaying CPE were harvested by scraping each monolayer into 5 ml of Hank's balanced salt solution (HBSS) and the cells were collected by centrifugation at 1500 rpm for 10 minutes at 5°C. The cells were washed again with HBSS and phosphate-buffered saline, pH 7.4 (PBS), resuspended in PBS to 100,000-500,000 cells per ml, and spotted at 80 to 100 cells per 250x microscopic field on 15-well or 12-well fluorescent antibody slides. The slides were allowed to air dry, and the cells were fixed with ice cold acetone for 5 minutes. These slide preparations were frozen at -70°C for up to 6 months before use. Before staining, each slide was allowed to thaw at room temperature for 5 minutes, washed with PBS and drained. The cells were first incubated at 35°C with an appropriate dilution of hybridoma-derived ascitic fluid, DEN-2 HMAF which reacts with all four dengue serotypes, flavivirus HMAF (equal portions of DEN-2, Yellow fever, St. Louis encephalitis, Japanese encephalitis, and Banzi virus HMAF) or normal ascitic fluid. All monoclonal antibody preparations were screened first at a dilution of 1:10 with arbovirus-infected LLC-MK2 cells. Monoclonal antibodies used for serotyping were used at a dilution equal to 4 times the last dilution giving maximum fluorescence (4 fluorescence units; generally 1:20 to 1:80). Hyperimmune fluids were used routinely at a 1:40 dilution. After 30 minutes of incubation at 35°C, the slides were washed three times with PBS for 10 minutes each. The cells were then treated with fluorescein-conjugated goat anti-mouse serum for 30 minutes at 35°C, washed as above, mounted and viewed with a 25x oil immersion objective using a fluorescent microscope with epi-illumination.

Detection of DEN-2 virus replication in cell culture

LLC-MK2 or C6/36 cells grown to 90% confluency in 25 cm^2 flasks were infected with DEN-2(New Guinea C) at 1 plaque forming unit per cell. Flasks were observed daily for CPE and harvested at 0, 12, 24, 36, 48, 96, 120, and 144 hours postinfection (p.i.). Tissue culture supernatants

Table 1.
Specificity of Monoclonal Antibody
Prepared Against Dengue Virus Antigens

Immunogen Serotype	Anti-Dengue Hybridoma			
	Type Specific	Dengue Subcomplex Specific[1]	Dengue Complex Specific[2]	Flavivirus Group Reactive
Den-1	15F3	13E7	9D12	
		5C11		
Den-2	3H5			1B8
		3H1		1B10
		2H3		1C10
		1C7		4G2
		1C12		2C4
Den-3	5D4	2F3		1B6
Den-4	1H10		2H2	4B10
				1D7

[1] DEN-1/DEN-3 cross
[2] Reacts with all four serotypes

either solid or coarsely reticular similar to that described by Cardiff et al using polyclonal antibodies (19). Infected cells which reacted with type-specific antibody generally displayed perinuclear fluorescence with fine granules which permeated portions of the cytoplasm. The fluorescence pattern observed with crossreactive antibody was perinuclear but usually involved more of the cytoplasm and was less granular in appearance. The intracellular fluorescence pattern observed with C6/36 cells was difficult to discern because of the the small size of the cells.

Temporal appearance of dengue antigens as detected by monoclonal antibodies

In order to correlate fluorescence reactions with biological events during virus replication, DEN-2 infected LLC-MK2 and C6/36 cells were harvested as described and reacted with type-specific monoclonal antibodies (3H5) and flavivirus group reactive antibodies (DEN-2 HMAF) in an indirect immunofluorescence test. These results are compared in Table 2 with the appearance of CPE, infectious virus, and viral hemagglutinin (HA). The appearance of CPE was detected in both LLC-MK2 and C6/36 cells by 36 hours p.i. with increasing degeneration of the cell monolayers throughout the course of the virus growth period. In LLC-MK2 cells the CPE consisted primarily of increasing amounts of rounded, pinocytotic, floating cells. Similar CPE as well as increasing amounts of polykarocytosis which reached a maximum (complete monolayer fusion) by 72 hours p.i. was seen in C6/36 cells. Infected, fused C6/36 cell monolayers gradually degenerated until few intact cells remained after 96 hours. Virus yields in both cell lines rapidly increased after 12 hours postinfection with maximum titers occurring 48 to 96 hours postinfection in LLC-MK2 and C6/36 cells, respectively. Virus yields in C6/36 cells averaged 3.3 $\log_{10}$ greater than LLC-MK2 cells throughout the multiple cycle replication curve. Increases in HA titers generally paralleled increases in virus titer; however, HA titers in LLC-MK2 cells decreased after 96 hours, while HA titers in C6/36 cells continued to rise. The detection of viral HA preceded the appearance of CPE by 24 hours in C6/36 cells.

Fluorescence patterns obtained using type-specific and flavivirus group reactive polyclonal antibodies were similar to those described previously. In LLC-MK2 cells the detection of viral antigens by immunofluorescence generally preceded the appearance of viral HA by 12 to 24 hours (Table 2). However, in C6/36 cells the detection of immunofluorescence was nearly simultaneous with the appearance of viral HA. In both cell lines polyclonal DEN-2 HMAF detected dengue antigens by 12 hours p.i. and with generally brighter fluorescence than that obtained using monoclonal antibodies.

TABLE 2

Temporal Appearance of Dengue-2 Antigens as Detected by Monoclonal Antibodies.

Cell Line	Hours p.i	CPE[1]	Virus Yield pfu/ml[2]	Hemagglutination Titer[3]	Immunofluorescence Reactions[4] 3H5	DEN-2 HMAF
$LLC\text{-}MK_2$	1	0	2.1×10^2	0	-	-
	12	0	0	0	+/-	1+
	24	0	9.2×10^2	0	2+	4+
	36	1+	4.8×10^3	16	2+	4+
	48	1+	2.8×10^4	32	2+	4+
	72	2+	1.2×10^4	64	3+	4+
	96	3+	1.0×10^4	32	3+	4+
	120	4+	7.5×10^3	32	2+	4+
	144	4+	2.5×10^3	16	1+	4+
C6/36						
	1	0	1.2×10^4	0	-	-
	12	0	6.0×10^3	2	-	1+
	24	0	6.4×10^5	16	2+	4+
	36	1+	3.5×10^6	96	2+	4+
	48	2+	2.3×10^7	256	3+	4+
	72	3+	3.8×10^7	256	3+	4+
	96	3+	6.0×10^7	256	3+	4+
	120	4+	5.6×10^7	512	3+	4+
	144	4+	3.2×10^7	1024	3+	4+

1 CPE: 0, no cells showing CPE; 1+, 25% of the cells; 2+, 50%; 3+, 75%; 4+, 100%
2 plaque-forming units per ml
3 Mean hemagglutination titer of clarified tissue culture supernatants.
4 Immunofluorescence: 0, no cells showing fluorescence; 1+, 25% of maximum fluorescence; 2+, 50%; 3+, 75%; 4+, maximum fluorescence.

Identification of dengue virus isolates

Cultered mosquito cells have been reported to be better hosts for the isolation of arboviruses (20,21,22). In order to compare viral isolations in LLC-MK2 and C6/36 cells in our laboratory, isolates from Jamaican patients were inoculated onto both cell lines and harvested when 50% of the cells displayed CPE. Monoclonal antibodies (15F3, 3H5, 5D4, and 1H10) were used in the indirect immunofluorescence assay to serotype these isolates. As can be seen in Table 3, generally greater yields of these low passage viruses (mean difference=2.9 $\log_{10}$) could be obtained using C6/36 cells. Moreover, in those cases where acute human sera were used as sources of virus inocula (Table 3), C6/36 cells were better able to allow replication of the viral isolates. Confirmation of each serotype identity was made using the plaque-reduction neutralization test. Representative results are presented in Table 4.

Dengue isolates from African patients (Senegal) or trapped mosquitoes (Ivory Coast and Upper Volta) were serotyped using AP-61 cells and type-specific monoclonal antibodies. Eighty-four viral isolates from trapped mosquitoes were identified as DEN-2, and three isolates from dengue fever patients (one DEN-2, two DEN-1) were also identified. Representative results are presented in Table 5. Complement fixation tests performed by the Institute Pasteur, Dakar, had previously suggested that these isolates were dengue viruses (Dr. Digoutte, personnal communication). Confirmation of the identity of these isolates was done by PRNT of representative isolates as shown in Table 4.

DISCUSSION

The monoclonal antibodies characterized in this study can be divided into 4 major groups based on their reactions in the immunofluorescence assay: (1) flavivirus group reactive, (2) dengue complex specific, (3) dengue subcomplex specific, and (4) dengue serotype specific. Categories (1), (2), and (4) are compatible with subdivisions of the flavivirus group as suggested by others (5,23,24). The possibility of a serological relationship between DEN-1 and DEN-3 has been reported previously using complement fixation and neutralization assays (25,26); however, this is the first definitive evidence of a subcomplex existing among the dengue viruses. Previously, Trent (27,28) suggested that the major envelope glycoprotein (V3) of a flavivirus contains at least three determinants: flavivirus group reactive, complex specific, and serotype specific. He proposed that an invariant segment exists in the primary sequence of the envelope glycoprotein V3 and is the crossreactive determinant shared by all flaviviruses. The varible portion of the glycoprotein, which is the major portion of the molecule, contains epitopes which determine type specificity and complex specificity (27,28). The monoclonal antibodies in this study were prepared using methods which favor the development of monoclonal antibodies directed against structural proteins (17). These

TABLE 3

Serotype Identification by Indirect Immunofluorescence and First Passage Yields of Dengue Fever Patients From Jamaica, 1980-1981

Patient #	Inoculum	First Passage Virus Yields (pfu/ml) LLC-MK2	C6/36	Serotype Identification Indirect Immunofluorescence
AP-61				
247/80	3.0×10^2	8.0×10^2	6.4×10^6	DEN-1
1229/80	$<5.0 \times 10^1$	2.0×10^2	3.0×10^7	DEN-1
1462/80	2.0×10^2	6.0×10^1	1.0×10^7	DEN-1
1296/80	4.0×10^2	4.3×10^2	1.2×10^7	DEN-1
1376/80	<5	1.9×10^4	5.0×10^3	DEN-1
1525/80	2.7×10^2	3.5×10^3	4.3×10^6	DEN-1
1532/80	<5	$3.5 \times 10_1$	3.5×10^4	DEN-1
1533/80	<5	3.5×10^1	1.0×10^6	DEN-1
58/81	1.5×10^1	1.2×10^4	9.8×10^6	DEN-1
Human Serum		2.1×10^4		
1229s/80	$<5.0 \times 10^1$	<5	6.6×10^6	DEN-1
1462s/80	$<5.0 \times 10^1$	<5	3.4×10^4	DEN-1
1296s/80	$<5.0 \times 10^1$	1.2×10^3	8.2×10^6	DEN-1
1661s/81	<5	7.5×10^4	1.6×10^5	DEN-2
1726s/81	3.0×10^1	1.4×10^4	8.7×10^5	DEN-4
1776s/81	3.3×10^2	7.0×10^4	1.5×10^6	DEN-4
1870s/81	3.4×10^2	8.1×10^4	1.3×10^6	DEN-4
1943s/81	6.0×10^1	1.8×10^4	6.4×10^5	DEN-4
1903s/81	4.0×10^1	1.3×10^3	6.4×10^5	DEN-2
1902s/81	<5	<5	2.2×10^5	DEN-4

TABLE 4

Serotype Identification of Dengue Virus Isolates by Neutralization.

Isolate	Origin	Plaque Reduction Neutralization Titer			
		DEN-1	DEN-2	DEN-3	DEN-4
247/80	Jamaica	320	<10	12	<10
1462/80	Jamaica	>320	<10	22	<10
1666s/81	Jamaica	12	>320	12	20
1903s/81	Jamaica	<10	>320	<10	<20
1776s/81	Jamaica	<10	<10	<10	19
1870s/81	Jamaica	<10	<10	<10	18
A510	Africa	32	>320	22	10
891	Africa	25	>320	24	10
20761	Africa	24	>320	22	12

TABLE 5

Serotype Identification of African Dengue Isolates by Immunofluorescence Using Monoclonal Antibodies

Isolate #	Origin *	Serotype Identification
A510	A. taylori	DEN-2
2022	A. africanus	DEN-2
2010	A. luteocephalus	DEN-2
1035	A. opok	DEN-2
891	A. opok	DEN-2
20761	A. luteocephalus	DEN-2
10674	DF patient	DEN-2
29177	DF patient	DEN-1
28790	DF patient	DEN-1

A. = Aedes
DF = Dengue Fever.
* Information provided by Dr. J.P. Digoutte, Institut Pasteur, Dakar, Senegal

antibodies also were reactive in serological tests which are based on the presence of virion antigens (hemagglutination-inhibition and neutralization tests). We propose that a fourth epitope contained within the varible region of V3 must also exist which is complex specific and common to DEN-1 and DEN-3. This would be the first definitive description of a subcomplex category among the dengue viruses. Moreover, monoclonal antibodies 9D12 and 2H2 provide serological confirmation of the dengue complex, because they were totally unreactive with the other flaviviruses which we examined.

This study demonstrated that serotype specific monoclonal antibodies prepared against high passage laboratory strains also can be used successfully to serotype low passage isolates from clinical or field environments. The ability of these unique antibodies to serotype virus strains with widely different passage histories (some separated by almost 25 years) is evidence of the immutability of the type specific determinant with which these antibodies react.

In order to determine the best host cell for demonstrating fluorescence with the dengue viruses, a comparison was made in this study between LLC-MK (monkey kidney) and C6/36 cells. Both cell lines proved suitable for these fluorescence studies using highly passaged DEN-2 (new Guinea C) virus at a multiplicity of infection equal to 1 plque forming unit per cell. Both had detectable dengue antigen (using monoclonal antibodies) within 24 hours and developed maximum fluorescence within 48 to 72 hours postinfection. However, the C6/36 cell, in agreement with reports by others (20,21,22) was a superior host for dengue viruses, yielding an average of about 3 $\log_{10}$ more infectious virus. Its superiority as a medium for dengue virus isolation was further substantiated by the significant amplification of low-titered virus inocula to easily detectable levels (Table 3).

For this study, infected cells were not harvested until 50% of the cells displayed CPE to insure a sufficient level of the single determinant against which these monoclonal antibodies were directed. This level of CPE often occurred 9 to 14 days p.i. or was difficult to detect in C6/36 cells in the absence of polykaryocyte formation. Even so, serotype identification using monoclonal antibodies often preceded PRNT identification by 14 to 21 days. Evidence obtained using DEN-2 (New Guinea C) virus suggests that type-specific monoclonal antibodies could be used to serotype isolates even in the absence of CPE or viral HA (Table 1). It is proposed that significant time and resources could be saved if type-specific monoclonal antibodies are used in conjunction with C6/36 cells. The use of monoclonal antibodies in the manner described in this report by field laboratories should make dengue virus isolation and identification a rapid and routine procedure.

REFERENCES

1. Ehrenkranz, N.J., Ventura, A.K., Cuadrado, R.R., Pond, W.L., and Porter, J.E., 1971. Pandemic dengue virus in Caribbean countries and the southern United States-past, present and potential problems. New England Journal of Medicine, 285: 1460-1469

2. Wisseman, C.L. and Sweet, B.H.,1961. The ecology of dengue, p15-46. In J.M. May (ed), Second Series of studies in medical geography. Hafner Publishing Co., N.Y.

3. McCown, J.M. and Bancroft, W.H., 1977. Identification of dengue viruses from the Carribean and the Bahamas. W.H.O. Pan Amer. Hlth Org. Newsletter, 6:11-12.

4. Dengue 4 infections in U.S. travelers to the Carribean. Morbidity and Mortality Weekly Report, (1981) 30: 249-250.

5. Casals, J., 1957. Viruses: The versitile parasites. I. The arthropod-borne group of animal viruses. Trans. N.Y. Acad. Sci., ser.II, 19: 219-235.

6. Sabin, A.J. and Young, I., 1949. A complement fixation test for dengue. Proc. Exp. Biol., 69: 478-480.

7. Russell, P.K., Chiewsilp, D., and Brandt, W.E., 1970. Immunoprecipitation analysis of soluble complement fixing antigens of dengue viruses. J. Immunol., 105:: 838-845.

8. Kuberski, T.T., and Rosen, L., 1977. A simple technique for the detection of dengue antigen in mosquitoes by immunofluorescence. Am. J. Trop. Med. Hyg., 26: 533-537.

9. Russell, P.K., Nisalak, A., Sukhavachana, P., and Vivona, S., 1967. A plaque reduction test for dengue virus neutralizing antibodies. J. Immnol., 99: 285-290.

10. Russell, P.K. and Nisalak, A.,1967. Dengue virus identification by the plaque reduction neutralization test. J. Immunol., 99: 291-296.

11. Kohler, G. and Milstein, C., 1976. Derivation of specific antibody-producing tissue culture and tumor lines by cell fusion. European J. Immunol. 6: 511-519.

12. Kennett, R.H., Denis, K.A., Tung, A.S., and Klinman, N.R., 1978. Hybrid production: Fusions with adult spleen cells, monoclonal spleen cells, monoclonal spleen fragments, neonatal spleen cells and human spleen cells. Curr. Top. Microbiol. Immunol., 81: 77-91.

13. McFarlin, D.E., Bellini, W.J., Mingioli, E.S., Behar, T.N., and Trudgett, A., 1980. Monospecific antibody to the hemagglutinin of measles virus. J. gen. Virol., 48: 425-429.

14. Flamand, A., Wiktor, T.J., and Koprowski, H., 1980. Use of hybridoma monoclonal antibodies in the detection of antigenic differences between rabies and rabies-related virus proteins. I. The nucleocapsid protein. J. gen. Virol., 48: 97-104.

15. Lostrom, M.E., Stone, M.R., Tam, M., Burnette, N.W., Pinter, A., and Nowinski, R.C., 1979. Monoclonal antibodies against murine leukemia

viruses: Identification of six antigenic determinants on the p15 (E) and gp envelope proteins. Virology, 98: 336-350.
16. Brandt, W.E., Buescher, E.L., and Hetrick, F.M., 1967. Production and characterization of arbovirus antibody in mouse ascitic fluid. Am. J. Trop. Med. Hyg., 16: 339-347.
17. Gentry, M.K., Henchal, E.A., McCown, J.M., Brandt, W.E., and Dalrymple, J.M.,1982. Identification of distinct determinants on dengue-2 virus using monoclonal antibodies. Amer. J. Trop. Med. Hyg., In press.
18. Zollinger, W.D., Dalrymple, J.M., and Artenstein, M.S., 1976. Analysis of parameters affecting the solid phase radioimmunoassay quantitation of antibody to meningococcal antigens. J.immunol., 117: 1788-1798.
19. Cardiff, R.D., Russ, S.B., Brandt, W.E., and Russell, P.K., 1973. Cytological localization of dengue-2 antigens: an immunological study with ultrastructural correlation. Infect. Immun., 7: 809-816.
20. Tesh, R.B., 1979. A method for the isolation and identification of dengue viruses using mosquito cell cultures. Am. J. Trop. Med. Hyg., 28:1053-1059.
21. Race, M.W., Fortune, R.A.J., Agostini, C., and Varma, M.G.R.,1978. Isolation of dengue viruses in mosquito cell culture under field conditions. Lancet, 1:48-49.
22. Igarashi, A., 1978. Isolation of a Singh's Aedes albopictus cell clone sensitive to dengue and chikungunya viruses. J. Gen. Virol., 40: 531-544.
23. Madrid, A.T. de, and Porterfield, J.S., 1974. The flaviviruses (group B arboviruses): a cross-neutralization study. J. gen. Virol. 23: 91-96.
24. Berge, T.O.,ed., 1975. "International Catalogue of Arboviruses", 2nd edition. DHEW Publ. No. (CDC) 75-8301. U.S. Dep. Health, Educ., Welfare, Public Health Serv., Washington, D.C.
25. Russell, P.K. and Nisalak, A.,1967. Dengue virus identification by the plaque reduction neutralization test. J. Immunol. 99: 291-296.
26 Hammon, W. McD., Rudnick, A., Sather, G.E., 1960. Viruses associated with hemorrhagic fevers of the Philippines and Thailand. Science 131: 1102-1103.
27. Trent, D.W., Harvey, C.L., Qureshi, A., and Lestourgeon, D., 1976. Solid-phase radioimmunoassay for antibodies to flavivirus structural and nonstructural proteins. Infect. Immun., 13: 1325-1333.
28. Trent, D.W., 1977. Antigenic characterization of flavivirus structural proteins separated by isoelectric focusing. J. Virol., 22: 608-618.

DESIGN OF A SMALL UNIT DRILL TRAINING SYSTEM

*JACK H. HILLER, Ph.D., G. DAVID HARDY, JR.
LARRY L. MELIZA, Ph.D.
ARMY RESEARCH INSTITUTE FIELD UNIT
PRESIDIO OF MONTEREY, CALIFORNIA 93940

Combat readiness is critically dependent on unit training programs which units have difficulty in constructing and executing in a straightforward, efficient manner. Cutbacks in TRADOC resources for resident institutional training have placed a heavy individual-soldier-training load on units, whereas units traditionally have concentrated on collective training which builds on those individual skills imparted at the schools and training centers. Furthermore, the collective-training load in units has itself effectively grown through adoption of the relatively unstructured Army Training and Evaluation Program (ARTEP), with its many mission- and task-training requirements, in place of the highly structured step-by-step Army training Programs. And all the while, the time available to prepare for and conduct training in units has effectively shrunk for a variety of reasons. This paper first depicts the major problems confronting combat unit training, and then describes a significant unit training innovation recently designed by ARI.

Background. A brief description of an earlier project (1), which aimed to develop a unit training system for individual soldier training, will provide a useful perspective. In that project, performance-oriented lesson plans were prepared for each task that each soldier, in a unit such as a rifle squad, is responsible for performing according to official Soldier's Manuals. In addition, training management guides for each set of Soldier's Manual tasks comprising a duty position, such as rifleman or radiotelephone operator, were also carefully prepared. A feature of the training management guide was explicit information and terminology which integrated this individual training subsystem into the Battalion Training Management System (BTMS), which is the official unit training system. As part of the overall research effort, workshops for both the BTMS and the individual training subsystem were developed by the Army Training Board (proponent of the BTMS and research sponsor) and ARI.

Beginning with these workshops, held for the entire chain of command in two battalions-drawn from two brigades of the host division, the BTMS along with the individual training subsystem were subjected to field testing over a three-month period. Results of this study regarding the individual skill training subsystem and materials were clear-cut. Although the workshops were highly rated by soldiers and researchers alike, the system itself was not employed by any battalion, company, platoon, or squad (2).

The major reasons for the subsystem's failure were judged by the research team to derive from unit training priorities and capabilities. First, individual skill training traditionally has had a low priority in units; unit commanders expect to receive qualified school-trained personnel, even though TRADOC schools typically provide only partial training of enlisted soldiers. The BTMS sought to remedy this problem by emphasizing the contribution of timely individual skill training to efficient collective training, and by providing charts which identify the specific Soldier's Manual tasks required for smooth performance of unit missions (the charts are descriptively named the ARTEP Soldier's Manual Interface or Crosswalk). However, the Interface was not effectively used during the field test.

The second reported problem largely concerns the training environment in units, and thus is much more difficult to handle than leader perception of the value of individual training. The experimental training subsystem held junior leaders (specifically squad leaders) responsible for diagnosing the specific individual-training needs of their soldiers, and then conducting the required performance-based training and testing. In addition, the results of testing were to be recorded in job books for each soldier to support efficient preparation of training schedules. However, the designated trainers often lacked the qualifications to serve as trainers, and were not themselves systematically prepared to conduct specific training sessions. Additionally, personnel turnover and internal turbulence made it difficult for them to diagnose training needs and then keep track of which soldiers needed what individual training. Finally, work details of various sorts frequently disrupted training plans.

The major system design issues revealed by this research effort are shown in Table 1. In sum, existing unit training programs may be characterized as attempting to satisfy two uncoordinated requirements: (a) high priority collective training and (b) low priority individual soldier training. Furthermore, the unit training environment may be characterized as difficult for any training (3).

Current Project

Despite this system design failure, ARI was requested by the infantry division hosting the field test to design a new integrated training system

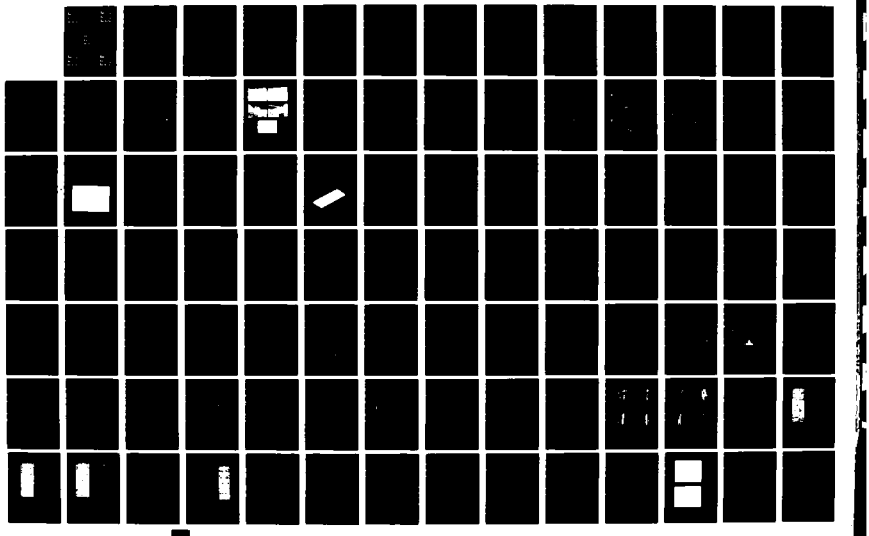
in which individual training would support collective training. In addition, the Army Training Board requested ARI to design a prototype integrated training system and corresponding system development guideline to be used by the TRADOC schools to fulfill the Chief of Staff's directive to standardize training in units whenever practical.

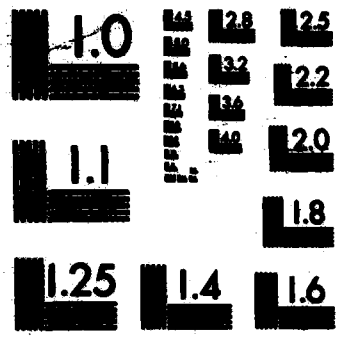
A conventional research strategy was adopted that included issue identification by problem review, problem solving by system design, and system refinement through testing. The final step was to design a general guideline to assist TRADOC training system developers. Thus, the team began by reviewing the reasons for the failure of the previous individual training system (as described above), examined the BTMS concept intended to drive integration of individual and collective training (the ARTEP Soldier's Manual Interface), and then reviewed the key unit training documents--the Soldier's Manual and the Army Training and Evaluation Program (ARTEP) for light infantry (the combat arm selected for study by the research sponsor). Regarding the ARTEP Soldier's Manual Interface, we concluded that it was not used because the number of Soldier's Manual tasks identified for each unit mission was too high to be of practical use. The Soldier's Manual was considered, at the time of review, to be an inadequate training support for inexperienced trainers. The ARTEP was similarly considered to lack the specificity required to support inexperienced trainers. Numerous training supports for individual and collective training do exist, but trainers generally lack the time to pull these diverse materials together and to construct lesson plans from them. Furthermore, individually produced lesson plans would tend away from standardization, rather than towards it. Thus, a fairly clear picture of training system needs and problems emerged from this review.

A New Drill Training System

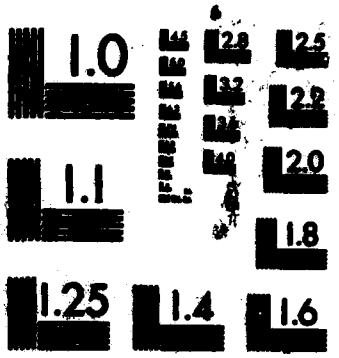
The new training system is keyed to the traditional unit training priority--collective training, as represented by ARTEP mission tasks. The system effectively defines a new form of unit task (based on an ancient training concept) which bridges the gap between Soldier's Manuals and ARTEPs--namely, the Drill. To facilitate attention to detail by trainers and evaluators, Drills were defined in concept to deal with relatively brief unit activities, and further, to apply only to relatively small units (e.g., Mech Platoon, Armor Platoon, rifle squad, fireteam, etc.). Table 2 presents the defining characteristics of Drills. Considerable effort was spent by the research team analyzing the ARTEP to identify candidate Drills based on these characteristics. After review by subject matter experts from the Army Training Board and the Infantry School, 16 Drills were selected (see Table 3) from a list of candidates. Next, the research team designed a new format for defining collective tasks as illustrated by the sample Drill training objective shown in Figure 1.

AD-A120 812 PROCEEDINGS OF THE 1982 ARMY SCIENCE CONFERENCE HELD AT 2/5
THE UNITED STATES.. (U) DEPUTY CHIEF OF STAFF FOR
RESEARCH DEVELOPMENT AND ACQUISITIO.. 18 JUN 82
UNCLASSIFIED F/G 5/2 NL

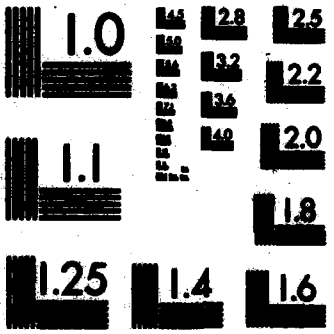




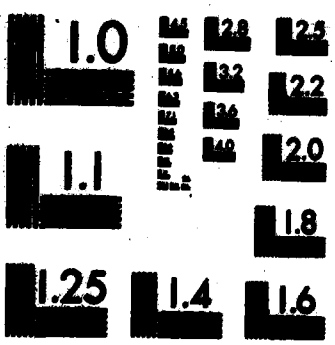
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A



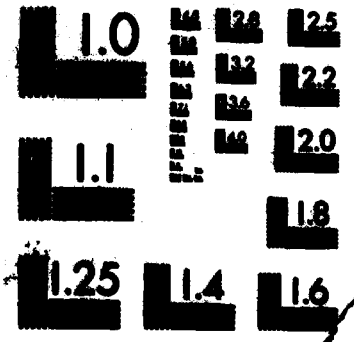
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

Individual training was systematically integrated with collective, Drill training in one of three ways. As the first preference, individual performance requirements were explicitly included in a Drill's performance requirements or standards. Individual skills which incur safety hazards or take a long time to train were identified as Drill training prerequisites; since each Drill encompasses only a few actions, the list of prerequisites was always small. Finally, certain Soldier's Manual tasks which are easy to train, but only partially covered by a Drill's performance standards, were identified for optional training in the field.

To insure that commanders understand how Drill training efficiently supports ARTEP mission training, a Drill training management guide was designed which shows the specific ARTEP mission tasks that each Drill prepares their units to perform. The Drills in the prototype Drill Training Package for Light Infantry Squads prepared by the research team were in fact defined so that most Drills served as building blocks for several ARTEP missions and mission tasks. This fact is illustrated by a portion of the training management guide shown in Table 4. The management guide also contains information required to schedule the Drills in an effective training sequence, along with resource requirements.

To assist trainer preparation, comprehensive, but easy to read lesson plans for each Drill were designed. To assist the trainer when actually conducting training in the field, the lesson plans were summarized and printed in a booklet that fits into the field jacket's pockets. A portion of one of the lesson plans is shown in Figure 2. Note that the formal training standards are well cued, and then followed by coaching tips.

Finally, to assist the evaluation of unit Drill performance by training supervisors, a booklet containing evaluation checklists for the Drills was designed. A sample checklist is shown in Figure 3.

Field Testing

The complete prototype Drill Training Package for light infantry squads was prepared by ARI with the support of the Army Training Board, the Infantry School, and a Brigade Commander of the 4th Infantry Division (COL Burba, who was independently working on a Battle Drill training system for Mechanized Infantry Squads). This training package was tried out by the rifle squads in two infantry companies during two weeks of field training in Nov 81. The field tryout identified the need for only minor changes, such as the addition of one diagram and, in a few instances, improved wording for clarity. It was clear from chain of command approval, and observed use by platoon and squad sergeants, that the prototype Drill Training Package met unit needs. Shortly after returning from the field, both the participating test battalion, and its parent brigade requested additional copies. Within

a few days, the colocated infantry brigade requested multiple copies for use. Within a few weeks, units from Forts Bragg, Carson, Campbell, and Lewis also requested a total of approximately 600 copies for use. After these requests were satisfied, TRADOC HQ decided to print and distribute another 600 copies to the MACOMs.

The research plan had been to prepare a Drill System Developer's Guideline for TRADOC based on practical experience gained from the research efforts to design only sample portions of a Drill Training Package. The initial project plan was altered as the result of three categories of events:

a. the system design effort proved to be so complex that design experience based on the research required to construct an entire training package appeared to be necessary for writing a valid Drill Developer's Guideline.

b. resource limitations at the Infantry School precluded School completion of the prototype package in time to meet TRADOC needs.

c. the enthusiastic support of the FORSCOM CDR, the Army Training Board's President, COL Hagan, and DCSOPS-Training staff, as work progressed, conveyed the feeling that a training system of great potential was taking shape which deserved to be fully developed and tested.

At the time this report was prepared, the Guideline for Designing Drill Training Systems was undergoing favorable review by TRADOC for publication in a Regulation.

References:

1. Bialek, H. M., Taylor, J. E., Melching, W. H., Hiller, J. H., and Bloom, R. D. "Continuation of Development of an Individual Extension Training System for Managing and Conducting Training in the Army Unit," ARI Technical Report TR-78-B1, 1978.

2. Bialek, H. M., Brennan, M., and Hiller, J. H. "Development and Implementation of a Performance Based Training and Evaluation System for the Combat Arms," ARI Technical Report TR-79-B4, 1979.

3. Funk, S. L., Johnson, C. A., Batzer, E., Gambell, T., Vandecaveye, G., and Hiller, J. H. "Training Detractors in FORSCOM Divisions and How They are Handled," ARI Research Report 1278, 1980.

TABLE 1

SYSTEM DESIGN ISSUES FOR INDIVIDUAL TRAINING IN UNITS

SYSTEM FUNCTION	ISSUES
MOTIVATION OF LEADERS	• EXCLUDING WEAPONS AND PHYSICAL TRAINING, INDIVIDUAL TRAINING IS PERCEIVED AS LOW PRIORITY • SKILL QUALIFICATION TESTING (SQT) COVERS TOO FEW TASKS BY PERFORMANCE TESTING TOO INFREQUENTLY TO PROVIDE ADEQUATE INCENTIVE • SQT ALSO IS FREQUENTLY REPORTED AS A DISTRACTOR TO UNIT TRAINING, SINCE IT IS EXTERNALLY IMPOSED • EXCLUDING SQT, WEAPONS AND PHYSICAL TRAINING, THERE IS NO EFFECTIVE ACCOUNTABILITY
MOTIVATION OF TRAINERS/SUPERVISORS	• PERCEIVED LACK OF COMMAND EMPHASIS CONVEYS LITTLE INCENTIVE • TIME MUST BE SPENT DURING OFF-DUTY HOURS TO PREPARE • FREQUENT DISRUPTIONS OF SCHEDULED TRAINING MAKE TIMELY PREPARATION DIFFICULT. • LACK OF PERFORMANCE TESTIN[illegible] TRAINERS PRECLUDES REWARD FOR SUCCE[illegible]

TABLE 1 (CONT.)

IDENTIFICATION OF TRAINING NEEDS	• LARGE NUMBER OF DUTY POSITIONS, INDIVIDUAL SOLDIERS AND SOLDIERS MANUAL (SM) TASKS WOULD OVERWHELM CENTRALIZED DIAGNOSTIC TESTING • THE NUMBER OF SM TASKS LISTED FOR MISSIONS IN THE ARTEP SM INTERFACE IS TOO LARGE FOR PRACTICAL USE BY DECENTRALIZED MANAGEMENT • DIAGNOSTIC TESTING IS VERY EXPENSIVE, ESPECIALLY FOR TIME • RAPID TURNOVER/TURBULENCE OF SOLDIERS DEFEATS DIAGNOSTIC TESTING & LOCAL RECORD KEEPING • SKILL/KNOWLEDGE TIME DECAY COMPLICATE RECORD KEEPING
PLANNING AND SCHEDULING OF TRAINING	• SHORTAGES OF RESOURCES MAKE SCHEDULING DIFFICULT • LOW PRIORITY INDIVIDUAL TRAINING COMPETES POORLY AGAINST OTHER REQUIREMENTS • TURNOVER/TURBULENCE OF TRAINEES, TRAINERS, AND TRAINING MANAGERS (LEADERS & SUPERVISORS) FRUSTRATES CAREFUL SCHEDULING • UNPLANNED REQUIREMENTS/DETAILS WRECK SCHEDULES

TABLE 1 (CONT.)

TRAINER PREPARATION AND CONDUCT OF TRAINING	• SHORTAGES OF QUALIFIED TRAINERS (E.G., SQUAD LEADERS) ARE COMMON • GIVEN LOW PRIORITY OF INDIVIDUAL TRAINING, SMALL EFFORT IS MADE TO PREPARE/ REHEARSE TRAINERS • EXISTING TRAINING SUPPORT DOCUMENTS ARE NOT CONVENIENTLY AVAILABLE TO JUNIOR TRAINERS, AND OFTEN HARD TO READ • STANDARDIZED LESSON PLANS ARE NOT PREPARED FOR TRAINERS • JUNIOR TRAINERS TYPICALLY LACK THE KNOWLEDGE, SKILL AND TIME TO PREPARE LESSON PLANS • TRAINING TYPICALLY IS CONDUCTED AS LARGE GROUP LECTURE
TRAINEE TESTING	• GIVEN LECTURE TRAINING MEDIUM, TRAINEES ARE NOT PERFORMANCE TESTED: INSTEAD ATTENDANCE IS RECORDED • STANDARDIZED PERFORMANCE TESTS DO NOT EXIST (EXCEPT FOR A FEW FROM THE SQT)
TRAINING EVALUATION AND QUALITY CONTROL	• GIVEN LACK OF PERFORMANCE TESTING, EVALUATION IS TYPICALLY PROCESS, INSTEAD OF PRODUCT, BASED • GIVEN LACK OF PRODUCT MEASUREMENT, ACCOUNTABILITY IS LOW, AND QUALITY CONTROL IS NON-EXISTENT OR IRRELEVANT

TABLE 2

BATTLE DRILL DEFINITION

BATTLE DRILLS ARE COLLECTIVE TASKS, DEFINED THRU THE CONVENTIONAL FORMAT:

TASK, CONDITIONS, STANDARDS

EACH DRILL HAS THE FOLLOWING CHARACTERISTICS:

ARTEP RELATIONSHIP	• IS KEYED TO ONE OR MORE ARTEP MISSION TASKS • IS DEFINED TO MAXIMIZE APPLICATION ACROSS ARTEP MISSIONS
UNIT ACTIVITIES	• REQUIRES PERFORMANCE BY MOST OR ALL UNIT MEMBERS • REQUIRES RAPID UNIT REACTIONS TO ENEMY THREAT OR LEADER ORDER • MINIMIZES NEED FOR LEADER TACTICAL DECISIONS AND COORDINATION WITH OTHER UNITS--SELF-CONTAINED
DRILL FEATURES	• REQUIRES A RELATIVELY STANDARD SET OF ACTIONS IN A VARIETY OF SITUATIONS--DRILL-LIKE ACTIVITY • HAS NATURAL STARTING AND STOPPING POINTS (I.E., IS A MEANINGFUL CHUNK OF BATTLE) • CALLS OUT OPFOR DISPOSITION, AND TERRAIN

TABLE 3

DRILLS FOR LIGHT INFANTRY SQUADS

1ST TRAINING SET	1.	FIRETEAM MOVES IN WEDGE FORMATION
	2.	SQUAD MOVES IN TRAVELING OVERWATCH
	3.	SQUAD CROSSES A DANGER AREA
	4.	SQUAD TAKES ACTION ON CONTACT (TRAVELING OVERWATCH)
	5.	SQUAD MOVES IN BOUNDING OVERWATCH
	6.	SQUAD TAKES ACTION ON CONTACT (BOUNDING OVERWATCH)
2ND TRAINING SET	7.	SQUAD EMPLOYS DIRECT FIRE
	8.	SQUAD PREPARES TO PROVIDE COVERING FIRE
	9.	SQUAD PROVIDES COVERING FIRE
	10.	SQUAD OCCUPIES POINT AMBUSH POSITION AND EXECUTES AMBUSH AND SEARCH
3RD TRAINING SET	11.	FIRETEAM MANEUVERS USING HIGH/LOW CRAWL
	12.	FIRETEAM ADVANCES MOVING IN SHORT RUSHES
	13.	SQUAD EXECUTES FIRE AND MANEUVER
	14.	SQUAD CLEARS A ROOM/BUILDING
	15.	SQUAD DISENGAGES
	16.	SQUAD EXECUTES DEFENSE/WITHDRAWAL TO SUPPLEMENTARY POSITIONS (URBAN AREA)

TABLE 4

SAMPLE SQUAD ARTEP MISSIONS & TASKS SUPPORTED BY EACH DRILL

		DRILL NUMBER														
MISSIONS	TASKS	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.
RAID	PREPARE TO CONDUCT RAID																
	MOVE TO OBJECTIVE AREA	X	X			X											
	EXECUTE RAID								X								
	CONDUCT WITHDRAWAL	X														X	
AMBUSH PATROL	PREPARE FOR AMBUSH PATROL																
	MOVE TO AMBUSH SITE	X	X			X											
	EXECUTE AMBUSH							X			X						
	SEARCH OBJ AREA & RETURN										X						
ANTI-ARMOR AMBUSH	PREPARE ARMOR AMBUSH																
	SELECT AMBUSH																
	OCCUPY AMBUSH SITE		X			X											
	EXECUTE AMBUSH																
	CONDUCT WITHDRAWAL & REORG	X						X								X	

DRILL TITLE: SQUAD PROVIDES COVERING FIRE

TASK: SQUAD PROVIDES COVERING FIRE FOR AN ASSAULTING SQUAD.

CONDITIONS:

FRIENDLY: PLATOON IS CONDUCTING A HASTY ATTACK.

ENEMY: ENEMY IS LOCATED ON THE PLATOON OBJECTIVE ROUGHLY 200 METERS FROM THE OVERWATCHING SQUAD.

SET-UP DIRECTIONS:

TRAINING SITE: TERRAIN PROVIDES FOR:

- PLATOON OBJECTIVE TO BE OCCUPIED BY SQUAD-SIZED OPFOR ELEMENT.
- PROTECTED ROUTE OF APPROACH TO OBJECTIVE FOR ASSAULTING SQUAD.
- COVERED AND CONCEALED POSITIONS FOR OVERWATCHING SQUAD TO COVER THE OBJECTIVE BY OBSERVATION AND FIRE.

OPFOR: EMPLACE TWO OR THREE OPFOR RIFLEMEN IN COVERED/CONCEALED POSITIONS ON OBJECTIVE. INSTRUCT OPFOR TO REMAIN IN POSITION AND OPEN FIRE ON SIGNAL FROM TRAINER. FIRING RATES ARE AS FOLLOWS:

- INITIALLY DELIVER HIGH RATE OF FIRE (EACH MAN FIRES ONE ROUND PER SECOND) FOR ABOUT 10 SECONDS.
- THEN FIRE AT REDUCED RATE IN RESPONSE TO OVERWATCHING SQUAD'S FIRE UNTIL ASSAULT BEGIN.

FIGURE 1. SAMPLE DRILL TRAINING OBJECTIVE

- INCREASE RATE OF FIRE IN RESPONSE TO ASSAULT ELEMENT'S FIRE, OR ON SIGNAL FROM TRAINER.

UNIT: SQUAD IS IN PLACE AFTER HAVING SUCCESSFULLY EXECUTED THE DRILL TASK "SQUAD RESPONSE TO PROVIDE COVERING FIRE."

START/FINISH DIRECTIONS:

- START DRILL WHEN SQUAD IS IN POSITION BY SIGNALING OPFOR TO COMMENCE FIRING.
- END DRILL WHEN SQUAD HAS SHIFTED FIRE IN RESPONSE TO SIGNAL THAT THE ASSAULT HAS BEGUN. (WHEN THIS DRILL IS RUN CONCURRENTLY WITH "ASSAULT" DRILL, END DRILL AFTER ASSAULTING ELEMENT HAS SWEPT THE OBJECTIVE.)

PERFORMANCE STANDARDS:

1. SQUAD RETURNS HEAVY VOLUME OF FIRE FROM COVERED AND CONCEALED POSITIONS WHEN THEY BEGIN OPFOR FIRE.
2. SQUAD REDUCES FIRE WITHIN 10 SECONDS AFTER ENEMY REDUCES THEIR RATE OF FIRE.
3. SQUAD MAINTAINS APPROPRIATE RATE OF FIRE, AND AVOIDS LULLS CAUSED BY RELOADING, MALFUNCTIONS, ETC.
4. SQUAD INCREASES VOLUME OF FIRE WHEN ITS LEADER OR THE ASSAULTING SQUAD SIGNALS.
5. SQUAD SHIFTS ITS FIRE ON SIGNAL FROM ITS LEADER BEFORE THE ASSAULT ELEMENT REACHES ITS OBJECTIVE.

FIGURE 1. (CONTINUED) SAMPLE DRILL TRAINING OBJECTIVE

(2) During the walk-thru:

(a) Signal OPFOR to commence firing. When OPFOR begins firing, order squad to open fire, and apply the Standard:

- "SQUAD RETURNS HEAVY VOLUME OF FIRE FROM COVERED AND CONCEALED POSITIONS WHEN THE ENEMY OPENS FIRE."

Coach soldiers to

a Remain in position
b Cover entire sector of fire
c Use high individual rate of fire
d Attend to Fireteam Leaders for directions

(b) When OPFOR reduces rate of fire, order squad to reduce its rate of fire and apply the Standard:

- "SQUAD REDUCES FIRE WITHIN 10 SECONDS AFTER ENEMY REDUCES THEIR RATE OF FIRE."

Coach soldiers to

a Pay attention to Fireteam Leaders
b Relay Fireteam Leader orders
c Reduce individual rate of fire (conserve ammunition)

(c) During period of reduced rate of fire, apply the Standard:

- "SQUAD MAINTAINS APPROPRIATE RATE OF FIRE AND AVOIDS LULL CAUSED BY RELOADING, MALFUNCTIONS, ETC."

1 Explain to soldiers that

a They must maintain fire superiority (keep enemy suppressed).
b Maneuvering element is advancing to assault position.

2 Coach soldiers to

a Vary individual rate of fire
b Correct malfunctions immediately
c Pick up fire for teammate as required

FIGURE 2. SAMPLE PORTION OF A DRILL LESSON PLAN

SQUAD PROVIDES COVERING FIRE

TASK: SQUAD PROVIDES COVERING FIRE FOR AN ASSAULTING SQUAD.

CONDITIONS:

FRIENDLY: PLATOON IS CONDUCTING A HASTY ATTACK.

ENEMY: ENEMY IS LOCATED ON THE PLATOON OBJECTIVE 200 METERS FROM THE OVERWATCHING SQUAD.

PERFORMANCE STANDARDS:

		GO	NO GO
(a)	• SQUAD IMMEDIATELY RETURNS HEAVY VOLUME OF FIRE WHEN HIT.		
	• SQUAD MAINTAINS COVERED/CONCEALED POSITIONS.		
(b)	• SQUAD REDUCES FIRE WITHIN 10 SECONDS AFTER ENEMY REDUCES THEIR RATE OF FIRE.		
(c)	• SQUAD MAINTAINS APPROPRIATE RATE OF FIRE.		
	• SQUAD MEMBERS PROMPTLY RELOAD AND CLEAR MALFUNCTIONS.		
(d)	• SL/TL SIGNALS TO INCREASE VOLUME OF FIRE PRIOR TO ASSAULT.		
	• SQUAD INCREASES VOLUME OF FIRE.		
(e)	• SL/TL SIGNALS TO SHIFT FIRE FOR ASSAULT.		
	• SQUAD SHIFTS FIRE.		

FIGURE 3. SAMPLE DRILL EVALUATOR'S CHECKLIST

QUANTIFICATION OF HELICOPTER VIBRATION RIDE QUALITY USING ABSORBED POWER MEASUREMENTS

D. D. HOLLENBAUGH
APPLIED TECHNOLOGY LABORATORY
USARTL (AVRADCOM)
FORT EUSTIS, VIRGINIA 23604

INTRODUCTION

One of the concerns in helicopter design is the specification of acceptable vibration levels, mainly for the purpose of providing a reasonable work environment for the crew. In the multifrequency, multi-axis vibration environment of the helicopter, the determination of what is "acceptable" can be difficult. The accepted practice for many years has been the measurement and specification of vibration levels in terms of acceleration. This method is thoroughly documented in ISO 2631, "Guide for the Evaluation of Human Exposure to Whole-Body Vibration," (1) established by the International Organization for Standardization. This has come to be the standard reference for evaluation of vibrations. Even this thorough treatment has its limitations when applied to helicopters, as expressed by Kidd in a recent paper surveying the methods available for assessment of helicopter vibration (2).

In an effort to provide a more comprehensive method of evaluating helicopter ride quality, a preliminary investigation was conducted into the use for helicopter ride quality assessment of a ride quality assessment method developed for ground vehicles by the US Army Tank Automotive Command (TACOM) (3). This concept uses the amount of energy absorbed by the human body when subjected to vehicle vibrations as a measure of ride quality. Vibration measurements were taken on five US Army helicopters, and these vibration data were converted to absorbed power. This paper presents initial results of the investigation underway and describes plans for future research.

ABSORBED POWER

TACOM researchers, working from extensive ride simulator tests, first postulated that the perceived severity of a vibration is proportional to the rate at which the human body is absorbing energy. This energy absorption, referred to as absorbed power, is mathematically defined as:

$$P_{avg} = \lim_{T \to \infty} \frac{1}{T} \int_0^T F(t)\, V(t)\, dt$$

where P_{avg} = average power absorbed by the subject

$F(t)$ = input force on the subject

$V(t)$ = velocity of the subject

T = averaging time interval

If the input force is written as

$$F(t) = \sum_{i=0}^{n} F_i \sin (w_i t + \phi_i)$$

and the velocity as

$$V(t) = \sum_{i=0}^{n} V_i \sin w_i t$$

then the average power absdorbed becomes

$$P_{avg} \sum_{i=0}^{n} \frac{G(jw_i) \sin \phi_i}{w_i} A_i^2 = \sum_{i=0}^{n} K_i A_i^2$$

where A_i = rms acceleration of the subject at frequency w_i

w_i = frequency of vibration

ϕ_i = phase angle between force and velocity

$G(jw_i)$ = Transfer function that relates force to acceleration
$G(jw) = F(jw)/A(jw)$

j = imaginary number, $j = \sqrt{-1}$

Note that the units of absorbed power are watts.

The transfer functions G(jw) were developed for the three axes of whole-body vibration in extensive ride simulator tests and may be defined as body equivalent mass in each direction, for an average subject (male, 28 years, 150 pounds). Detailed information on these tests and the

resulting transfer functions may be found in (4). The product of this extensive research is a model of vibration severity that correlates well with subjective response (5).

Two points should be noted about absorbed power. First, it is a measurable scalar quantity, and absorbed power in different axes may be added to produce total absorbed power from multidegree of freedom vibrations. Secondly, it can be measured or calculated for periodic or random vibrations. An example of these two points will be shown later.

Another interesting facet of absorbed power may be illustrated by conversion of the ISO data (1) into absorbed power, as shown in Figure 1. This is the ISO 8-hour fatigue/decreased proficiency boundary for vibration in the vertical direction. A large peak of absorbed power appears at 4-5 Hz. This large rise in absorbed power at an essentially constant acceleration is generally attributed to a resonating condition of the organs in the upper abdomen, with the diaphragm acting as a spring (6). A second, smaller peak appears around 12 Hz. This is believed to be caused by a resonating condition in the spinal column (6). These absorbed power peaks are due to the influence of the vertical body transfer function.

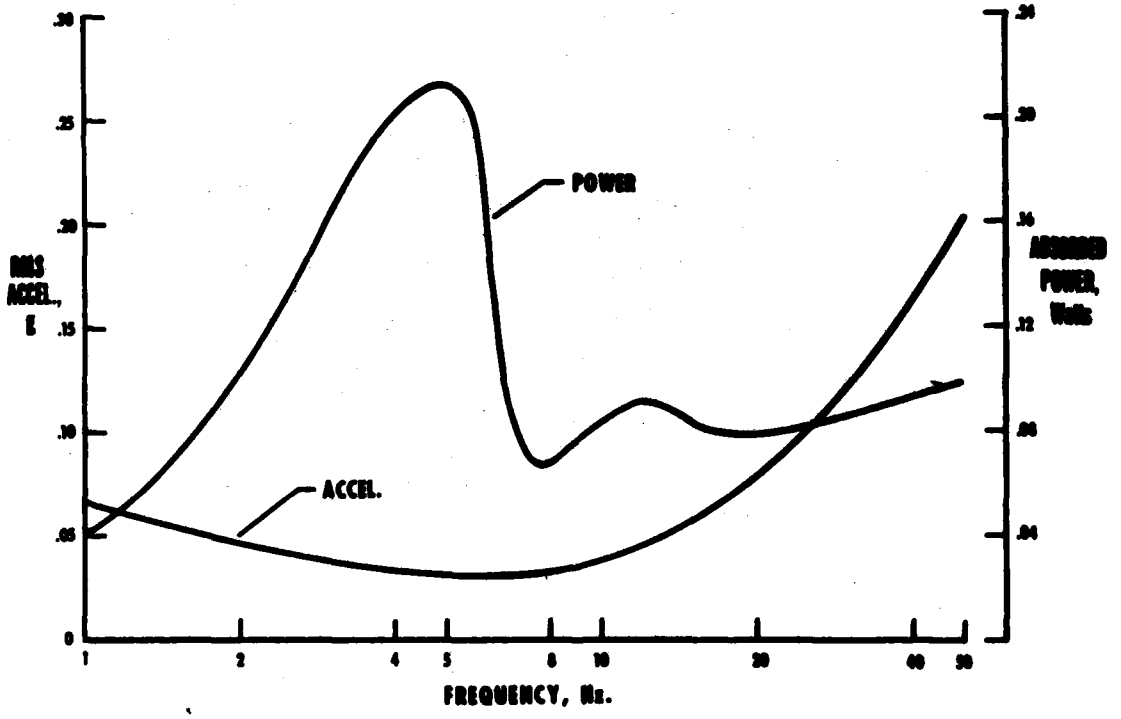


Fig. 1 - ISO 2631 8-hour fatigue/decreased proficiency boundary converted to absorbed power

More to the point, it should be noted that these two frequency bands are primary regions of rotor-induced vibration for many helicopter designs, as can be seen in Table 1. Rotor-induced vibrations occur at the rotational speed of the main rotor (1 per revolution, or 1P), at the blade passage frequency (2P for a two-bladed rotor, 3P for a three-bladed rotor, etc.) and at the harmonics of the blade passage frequency.

Table 1. Significant frequencies for vibration considerations on helicopters tested

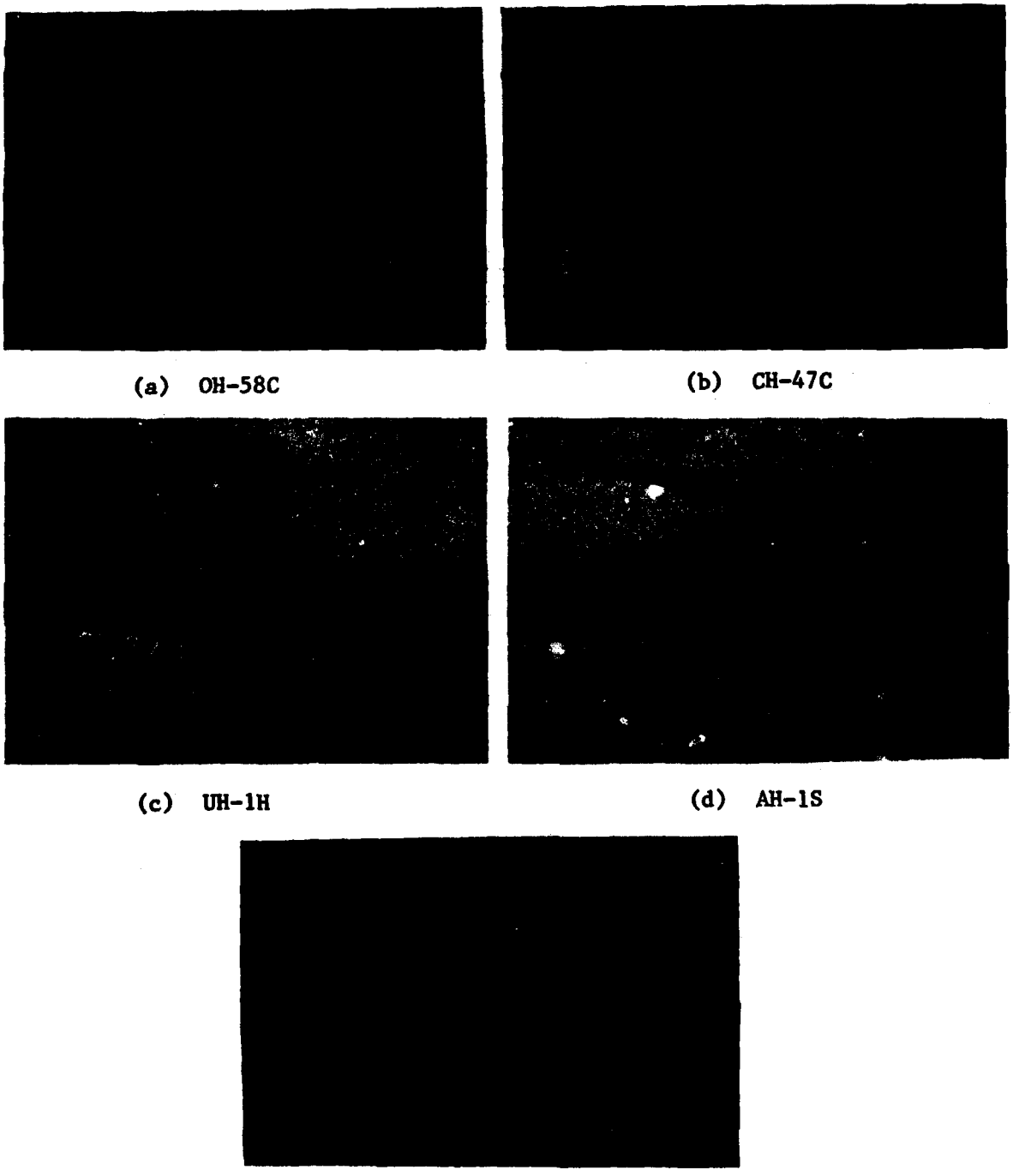
	FREQUENCY, Hz					
AIRCRAFT	1P	2P	3P	4P	6P	8P
CH-47C	3.9	-	11.8*	-	23.5	-
UH-60A	4.4	-	-	17.5*	-	35.1
UH-1H	5.4	10.8*	-	21.6	-	43.2
AH-1S	5.4	10.8*	-	21.6	-	43.2
OH-58C	5.9	11.8*	-	23.6	-	47.2

*Blade passage frequency

MEASUREMENTS

The test aircraft chosen for this investigation were representative of each helicopter type presently in use by the active US Army. Included were a Bell UH-1H Iroquois, a Bell OH-58C Kiowa, a Bell AH-1S modernized TOW Cobra, a Boeing Vertol CH-47C Chinook, and a Sikorsky UH-60A Blackhawk. Aircraft and flight crews were provided by the Aviation Maintenance Management Training Division of the US Army Transportation School, Fort Eustis, Virginia. These aircraft are shown in Figure 2. It should be pointed out that these were operational fleet aircraft, and only one example of each type, in one configuration, was flown for measurements. Therefore, the data shown should not be taken as a thorough vibration survey of any of the aircraft types.

Two sets of instrumentation were carried on board each test flight. The first consisted of a triaxial accelerometer package linked to a seven channel analog tape recorder for recording vibration levels. The second set consisted of an acoustical tape recorder and two microphones for recording aircraft internal noise. All instrumentation was provided by NASA-Langley Research Center.



(a) OH-58C

(b) CH-47C

(c) UH-1H

(d) AH-1S

(e) UH-60A

Fig 2. - Test aircraft for absorbed power measurements

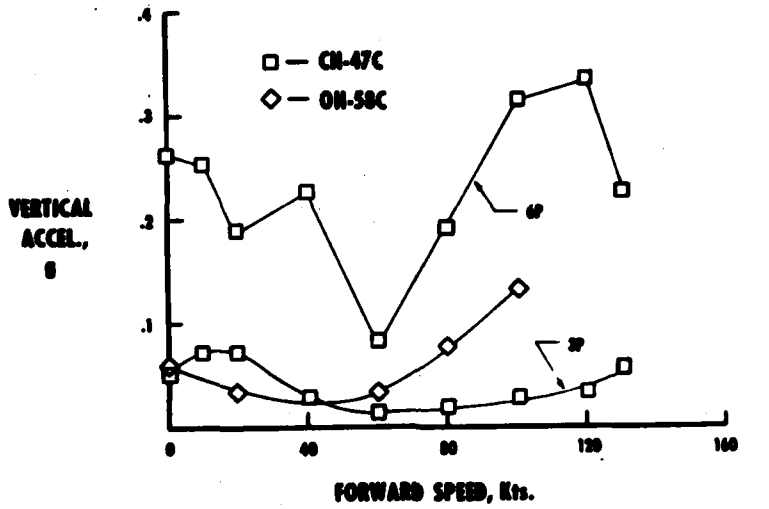
The accelerometer package was placed as close as possible to the base of the aircraft pilot's seat. The acoustical microphones were placed near the pilot's and copilot's heads. Vibration and sound recordings were then taken for a period of roughly 30 seconds at each of the following conditions: hover in ground effect (IGE), hover out of ground effect (OGE), rearward flight, left and right sideward flight, standard climb at cruise velocity, and forward flight at speeds from 10 knots to maximum level flight speed.

It was originally intended that absorbed power measurements would be made directly from the vibration recordings, using an instrument developed by TACOM for that purpose (6). However, during the data reduction process it was discovered that the instrument electronics were optimized for the vibration levels and frequencies characteristic of ground vehicles, and, as a result, the instrument proved inaccurate for absorbed power measurements on helicopter vibration. Therefore, the absorbed power measurements were obtained from the recorded vibrations using a computer implementation of the absorbed power equations.

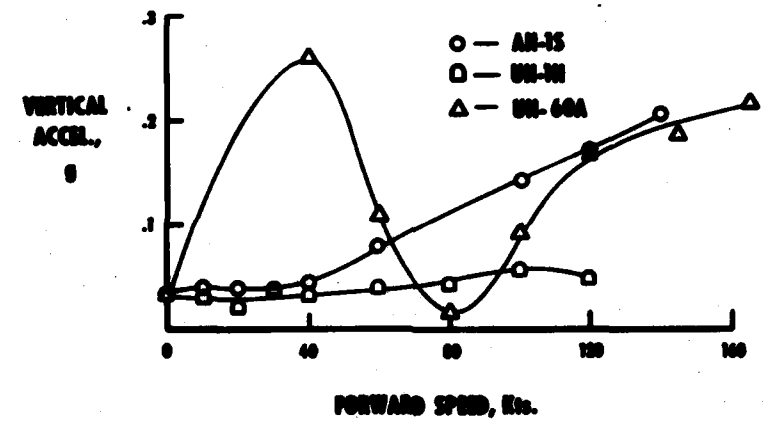
RESULTS AND DISCUSSION

The vertical acceleration levels measured on test aircraft are shown as a function of forward flight speed in Figure 3a and 3b. The data shown are peak amplitude at blade passage frequency, which is shown in Table 1. This is normally the highest vibration level in the helicopter spectrum. In the case of the CH-47C, the blade passage frequency (3P) and its second harmonic (6P) are shown to display the dominance of the second harmonic in this particular type.

These data are shown to indicate the general trend of vibration with airspeed, which came out as expected. It also shows the wide variation of vibration level between aircraft types. Once again, it should be remembered that these data are for a particular aircraft on a particular day, to determine general trends, and data on a particular aircraft should not necessarily be taken as representative in a statistical sense of the type.



(a)



(b)

Fig. 3(a) & (b) - Variation of vertical vibration with airspeed

The evolution of a particular absorbed power data point is shown in Figures 4-6. It begins with a Fourier transform of the raw acceleration data, a portion of which is shown in Figure 4. This is the vertical acceleration data recorded on an AH-1S in forward flight at 135 knots. Only the portion from 0-40 Hz is shown but, as will be shown, this is the most important area. Figure 5 is a vibration spectral density plot, which pinpoints the major vibratory frequencies and vibration levels.

The clearly dominant vibration, at 10.5 Hz, is the main rotor blade passage frequency (2P). It is interesting to compare Figure 5 with the absorbed power density graph of Figure 6. Figure 6 is created by multiplying Figure 5 by the vertical transfer function. Here the effect of the body transfer function can be fully appreciated. The low vibration level at 5.4 Hz (1P) becomes a major producer of absorbed power, while the higher vibration levels at 1 Hz and at 20 Hz become relatively less important. Also, any vibration above 30 Hz becomes almost negligible when expressed as absorbed power. This graph is then integrated to produce an absorbed power number, which in this case is .624. By this method both periodic and random vibration at all frequencies are taken into account, properly weighted and quantified by a single ride quality (absorbed power) number.

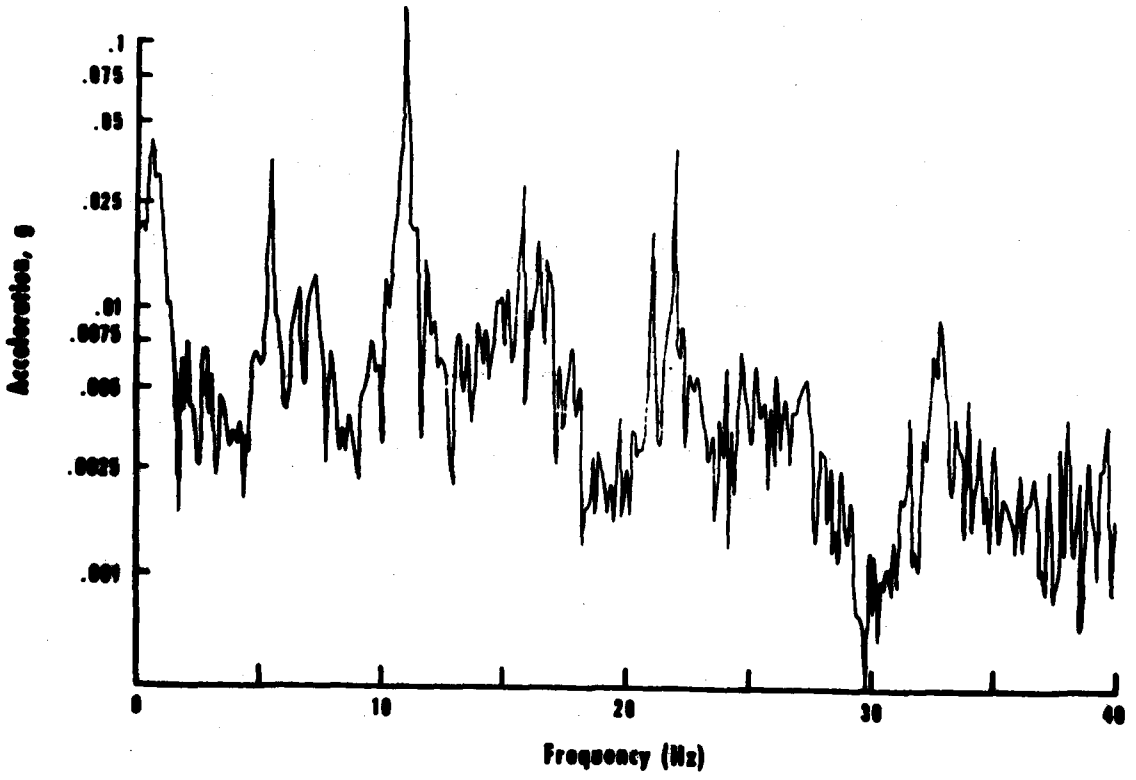


Fig. 4 - Vertical acceleration levels vs frequency, AH-1S, 135 knots

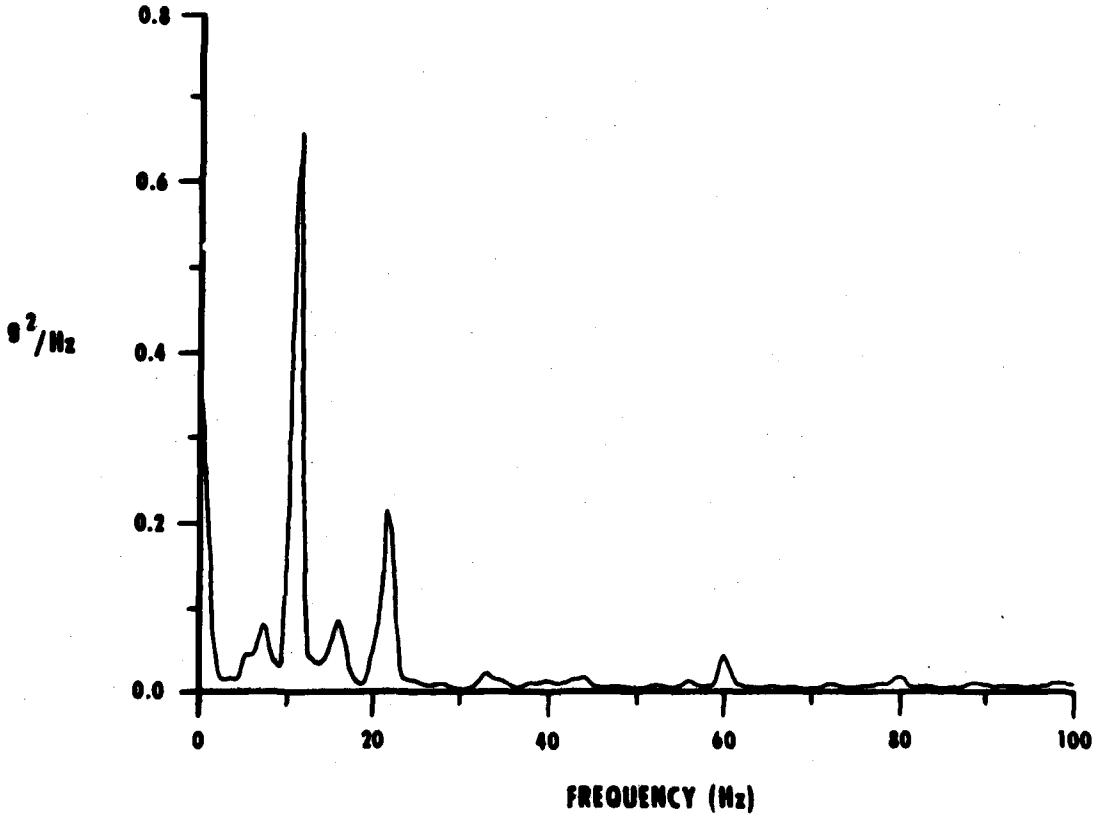


Fig. 5 - Power spectral density, vertical axis, AH-1S, 135 knots

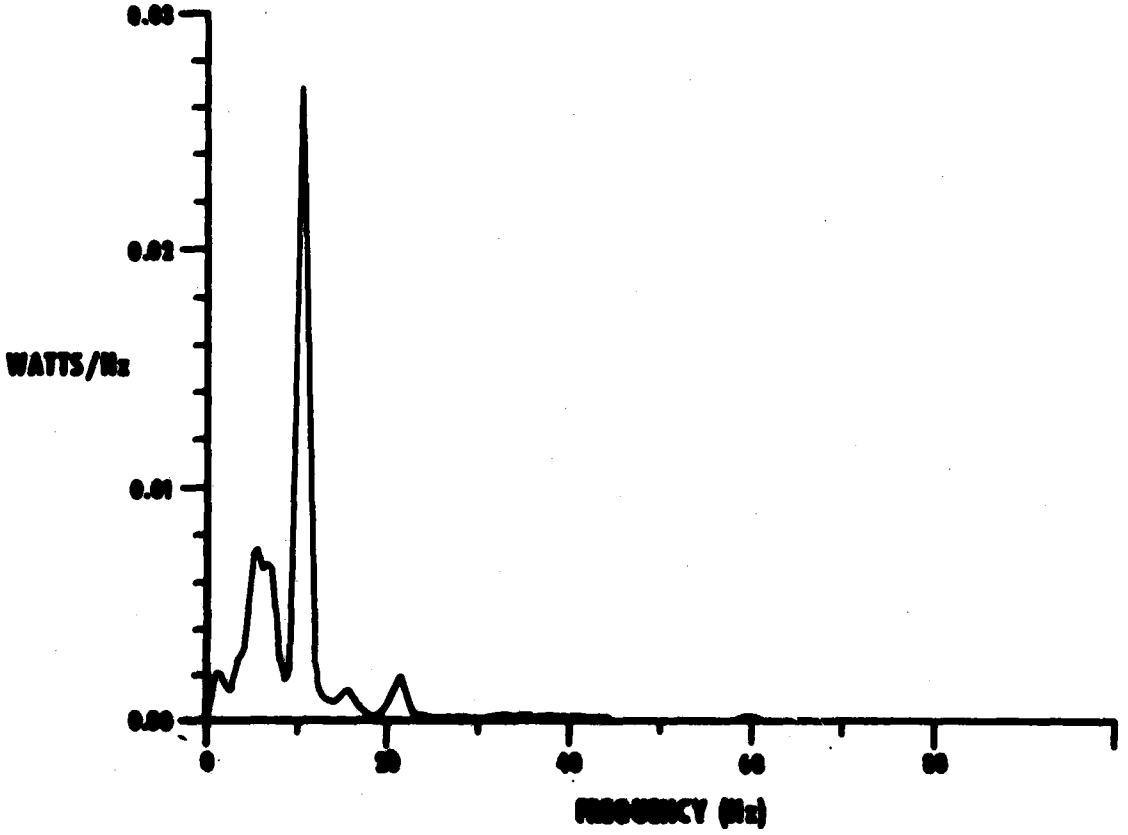


Fig. 6 - Absorbed power density, vertical axis, AH-1S, 135 knots

The collected absorbed power data is shown versus airspeed in Figures 7-9 for vertical, fore-aft and side-to-side axes, and in Figure 10 the three axes are summed to produce total absorbed power versus airspeed. The points shown for 0 airspeed are hover IGE. It can be seen that the side-to-side and fore-aft absorbed power levels are generally lower than the absorbed power levels from vertical vibrations. This can be attributed to the fact that measured vibration levels in these directions are generally lower, and to the transfer functions, which show less body vibration response in the fore-aft and side-to-side

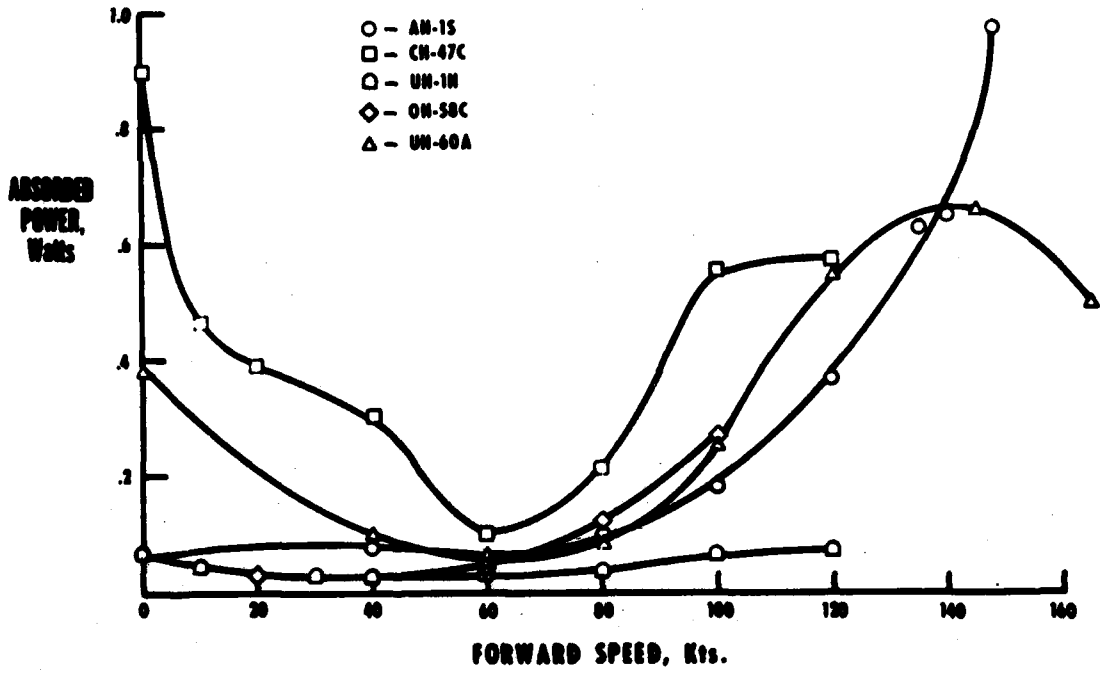


Fig. 7 - Absorbed power from vertical acceleration vs airspeed

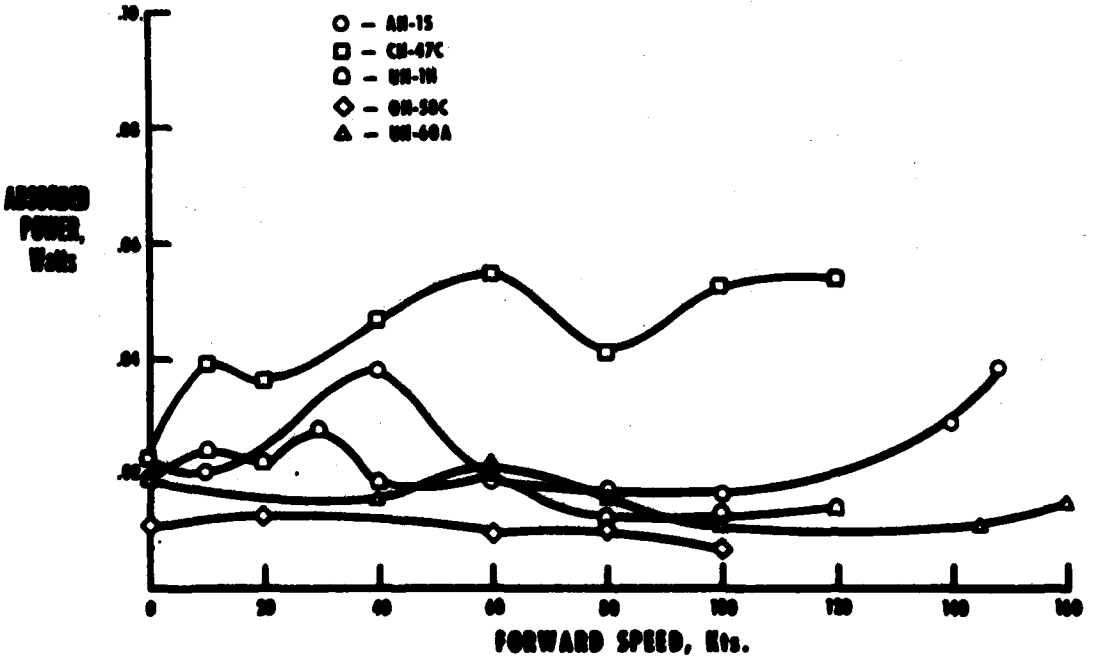


Fig. 8 - Absorbed power from fore-aft acceleration vs airspeed

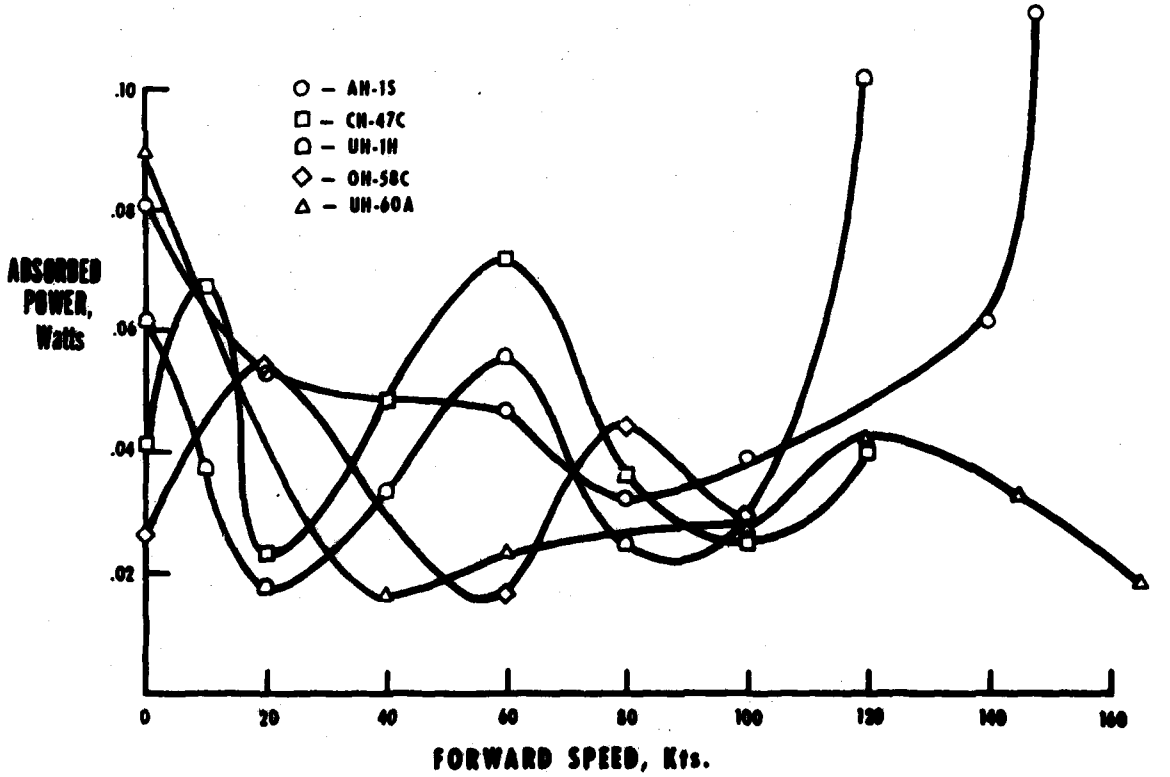


Fig. 9 - Absorbed power from side-to-side acceleration vs airspeed

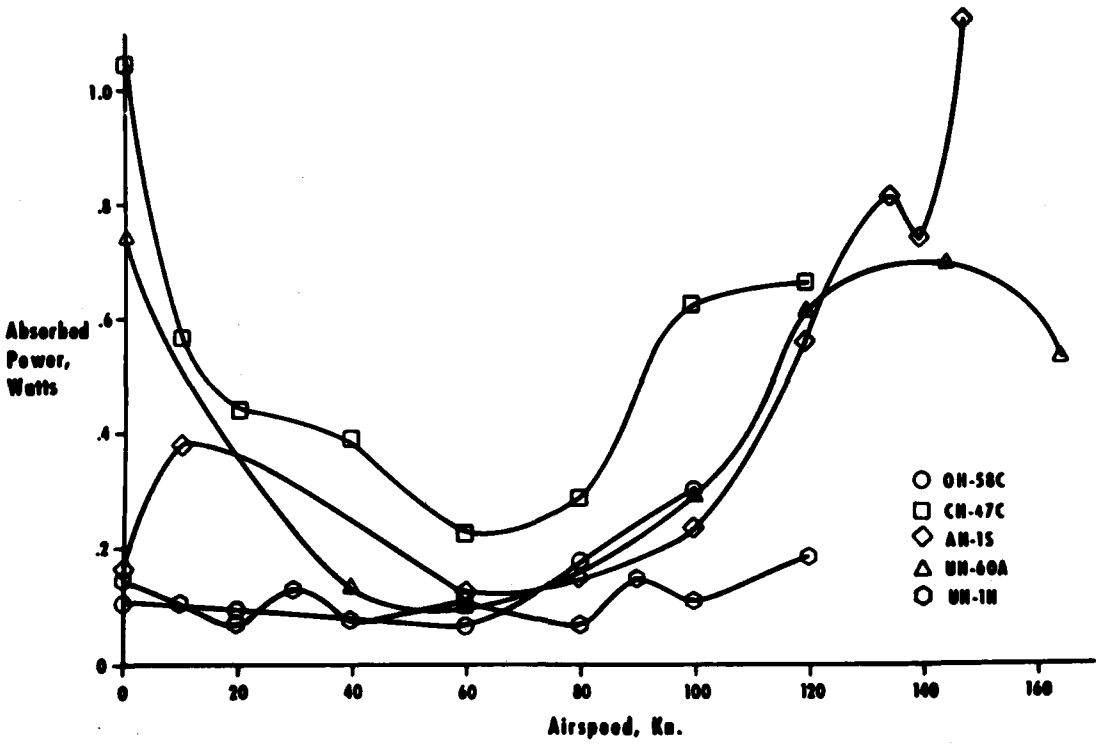


Fig. 10 - Total absorbed power vs airspeed

axes. There is an exception to the dominance of the vertical, however, and it illustrates a further strength of the absorbed power concept. Figures 11 and 12 show the relative contributions to absorbed power of vibration in the three axes for two of the aircraft tested. In the case of the UH-60A the fore-aft and side-to-side are relatively minor, and the fore-aft is essentially constant with airspeed. For the AH-1S the fore-aft contribution is larger than for the UH-60A and fluctuates with airspeed, and the side-to-side component contributes a much larger percentage of total absorbed power than in the case of the UH-60A. This can be explained by an examination of the designs of the two rotor systems. The UH-60A uses a fully articulated rotor, with lead-lag and flapping hinges and a pitch bearing on each blade, while the AH-1S has only a pitch bearing. Without lead-lag and flapping hinges to alleviate in-plane rotor induced vibrations, this vibration is transferred to the aircraft. Thorough study of all the absorbed power data showed this difference in relative contributions to absorbed power of the three axes due to rotor system design to be the trend. This ability of the absorbed power concept to properly weigh and integrate the various axes of vibration in this manner is viewed as another strength of the concept.

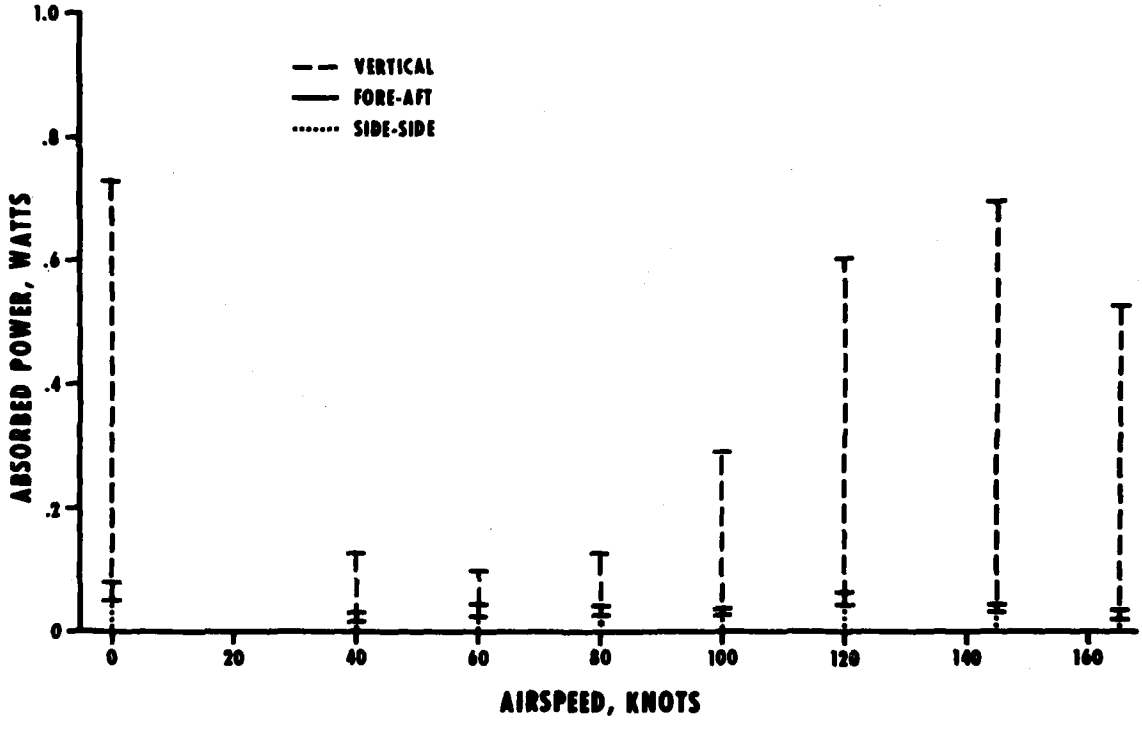


Fig. 11 - Relative contributions of absorbed power in three axes vs airspeed - UH-60A

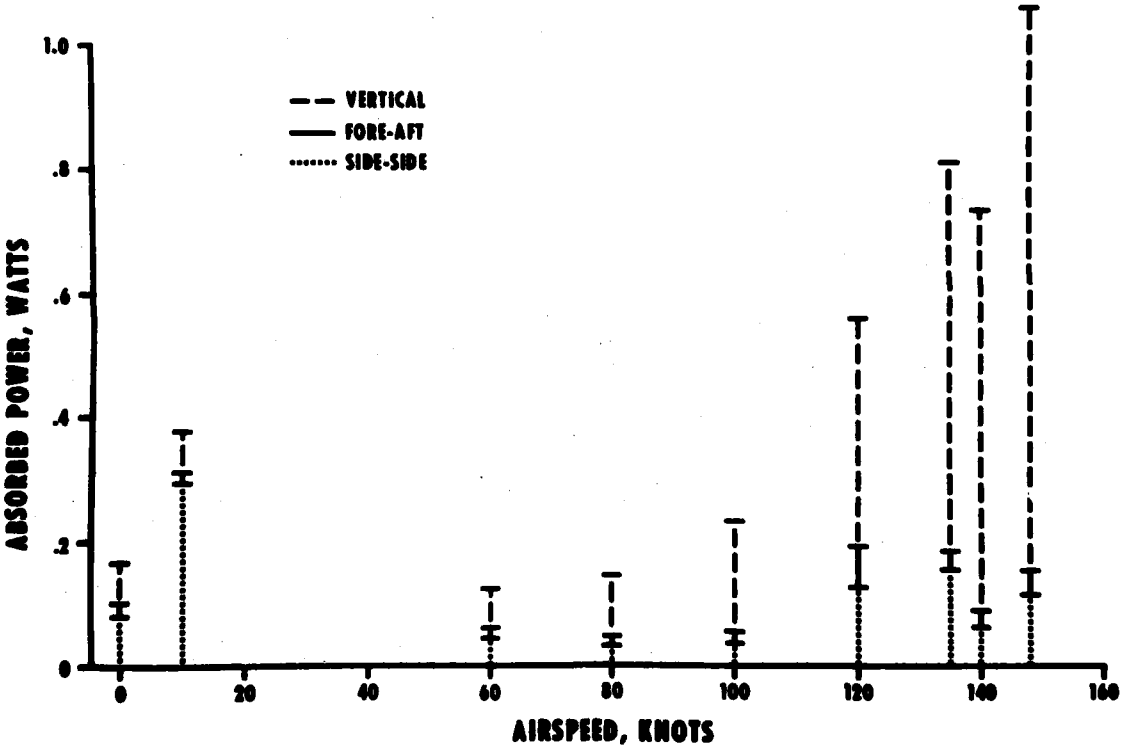


Fig. 12 - Relative contributions of absorbed power in three axes vs airspeed - AH-1S

SUBJECTIVE RESPONSE EXPERIMENTS

Currently underway is a program to correlate the collected absorbed power data with subjective response data, to aid in determining the validity of the absorbed power model for rotorcraft. The apparatus used to conduct these experiments is the Passenger Ride Quality Apparatus (PRQA) at NASA-Langley Research Center (Figure 13). The PRQA is a three-axis hydraulically driven motion simulator configured as a portion of an airliner cabin. A thorough description of the PRQA is given in (7). The PRQA is equipped with four airline-type seats and speakers for repeating recorded aircraft sounds. For the purpose of this experiment, the recorded noise and vibration data for hover IGE and forward flight at normal cruise velocity in each of the five Army aircraft are used. Noise levels are lowered using attenuation levels for the standard SPH-4 Army flight helmet (8). To produce various combinations of noise and vibration levels, noise levels are played with SPH-4 helmet attenuation, reduced by 7 dB, reduced by 14 dB, and eliminated altogether. Vibration levels are played as recorded, reduced by 3 dB, and by 9 dB. All different combinations of noise and vibration levels are run for each of the ten ride segments, producing 120 data points.

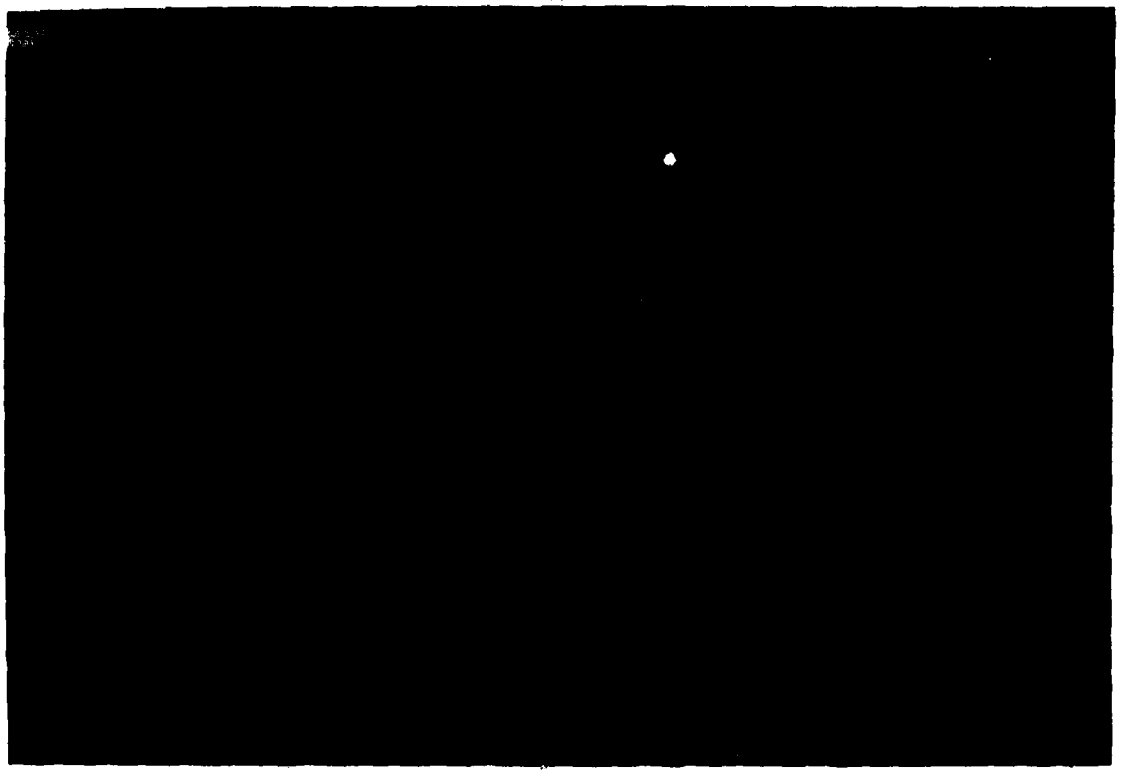


Fig. 13 - NASA Passenger Ride Quality Apparatus (PRQA)

The purpose of varying noise levels is to provide subjective response data for comparison with the NASA-Langley Ride Comfort Model. This is an empirical model which determines ride quality as a function of vibration and noise levels. In concert with the evaluation of absorbed power described in this paper, NASA-Langley researchers are studying the applicability of the NASA Ride Comfort Model to rotary wing aircraft. The results of this effort up to this time are covered in (9). For each of the 120 data points, the test subjects, Army and Navy helicopter pilots, rate each ride segment on a comfort scale. As of this writing, too little data has been produced to draw any conclusions.

CONCLUDING REMARKS

The absorbed power concept offers certain advantages over pure acceleration for helicopter ride quality evaluation. First, it takes into account multi-frequency, multi-axial vibration across a broad frequency range. Second, it provides proper weighting functions for all frequencies and axes according to body response. Third, it is applicable to random as well as periodic accelerations.

A larger data base and further experimentation is required for full validation of absorbed power as a means for quanitfying helicopter vibration ride quality. Further in-flight data will be collected to expand the vibration/noise environment data base, repeating measurements on the aircraft mentioned here, and add further aircraft. Already underway is a program to take measurements on US Navy aircraft types. Data have been recorded on a Sikorsky RH-53D and a Kaman SH-2F. The subjective response experiments will be continued for correlation and validation of the absorbed power and NASA Ride Comfort Model.

REFERENCES

1. Anon., "Guide for the Evaluation of Human Exposure to Whole-Body Vibration," International Standard ISO 2631, 1978.

2. Kidd, D. L., "Assessment of Helicopter Vibration Criteria," American Helicopter Society Preprint 81-22, 37th Annual Forum of the American Helicopter Society, New Orleans, Louisiana, May 1981.

3. Lee, R. A. and Pradko, F., "Analytical Analysis of Human Vibration," Society of Automotive Engineering Congress, Detroit, Michigan, January 1968.

4. Lins, W. F., "Human Vibration Response Measurement," US Army Tank-Automotive Research and Development Command Technical Report 11551, June 1972.

5. Pradko, F., Lee, R. A., Greene, J. D., "Human Vibration Response Theory," American Society of Mechanical Engineers, Paper 75-WA/HUF-19, Winter Annual Meeting, Chicago, Illinois, November 1965.

6. Lee R. A., Cumbow, J., "Absorbed Power Ride Measurement Instrument," US Army Tank-Automotive Research and Development Command Technical Report 12415, December 1978.

7. Stephens, D. G., Clevenson, S. A., "The Measurement and Simulation of Vibration for Passenger Ride Quality Studies," Proceedings of the Technical Program, National Noise and Vibration Control Conference, 1974.

8. Megehee, J. H., "Service Tests of SPH-4 Protective Helmet," Army Aviation Test Board, Fort Rucker, Alabama, September 1969.

9. Hammond, C. E., Hollenbaugh, D. D., Clevenson, S. A., Leatherwood, J. D., "An Evaluation of Helicopter Noise and Vibration Ride Qualities Criteria," presented at American Helicopter Society Northeast Region Specialist's Meeting on Helicopter Vibration, Hartford, Connecticut, November 1981.

A THEORETICAL STUDY OF THE PROPAGATION OF A MASS DETONATION (U)

*PHILIP M. HOWE, DR.
ABDUL R. KIWAN, DR.
BALLISTIC RESEARCH LABORATORY
USAARRADCOM
ABERDEEN PROVING GROUND, MARYLAND 21005

I. INTRODUCTION

A. Mass Detonability.

Ammunition items are assigned to various hazard classes, based on the level of risk considered acceptable for stipulated exposures. The maximum amount of explosives permitted at any location is determined by the prevailing distance from that location to other explosives. United Nations Organization (UNO) CIass 1, Division 1 is composed of "mass detonating" ammunition and explosives. A "mass detonation" is defined as the "virtually instantaneous explosion of a mass of explosives when only a small portion is subjected to fire, severe concussion or impact, the impulse of an initiating agent, or to the effect of a considerable discharge of energy from without" (1). The majority of large caliber ammunition, e.g., 155 mm, 175 mm, 8" separate loading projectiles and general purpose bombs, are classified as mass detonating, and the constraints of mode of storage and transportation imposed to provide adequate safety create a significant economic and operational burden (2). By use of appropriate packaging or shielding (3) or by use of different storage configurations (4), the round-to-round propagation tendency can be reduced significantly with concomitant reduction in the tendency for mass detonation. The purpose of this effort was to determine, as a function of the munition array, how much the tendency for round-to-round propagation need be reduced to control explosion size and prevent mass detonation.

B. Round to Round Propagation.

Numerous experiments have been performed with various types of ammunition to ascertain the nature of round to round propagation (5-8). Of special interest to this effort, it was found that a straightforward criterion for round to round propagation could be developed; if a munition, in a regular array, subjected to the blast/fragment field of a detonating

neighbor munition itself "detonated," then the blast/fragment field it generated would cause the next munition in the array to detonate, also, and propagation could continue within the array (9). If, however, the munition subjected to the donor blast/fragment field reacted with subdetonation violence, the process would extinguish. No dependence upon the number of nearest neighbor munitions within the array was found; testing could be performed with a linear array (indeed, an array with one donor and one acceptor) and the results could be applied to two dimensional quadratic or hexagonal arrays. Apparently, the confinement provided by multiple nearest (second nearest, etc.) neighbors does not appreciably affect the ability of one munition to cause another munition to detonate. In the development of a model of propagation of detonation between munitions, one can thus apply a quantal response criterion, and treat the interaction probabilities between munition pairs as independent (10).

C. Model Development.

Consider a large, two dimensional array of munitions. (See Figure 1, showing a storage array for 155 mm separate loading projectiles. Here, the array

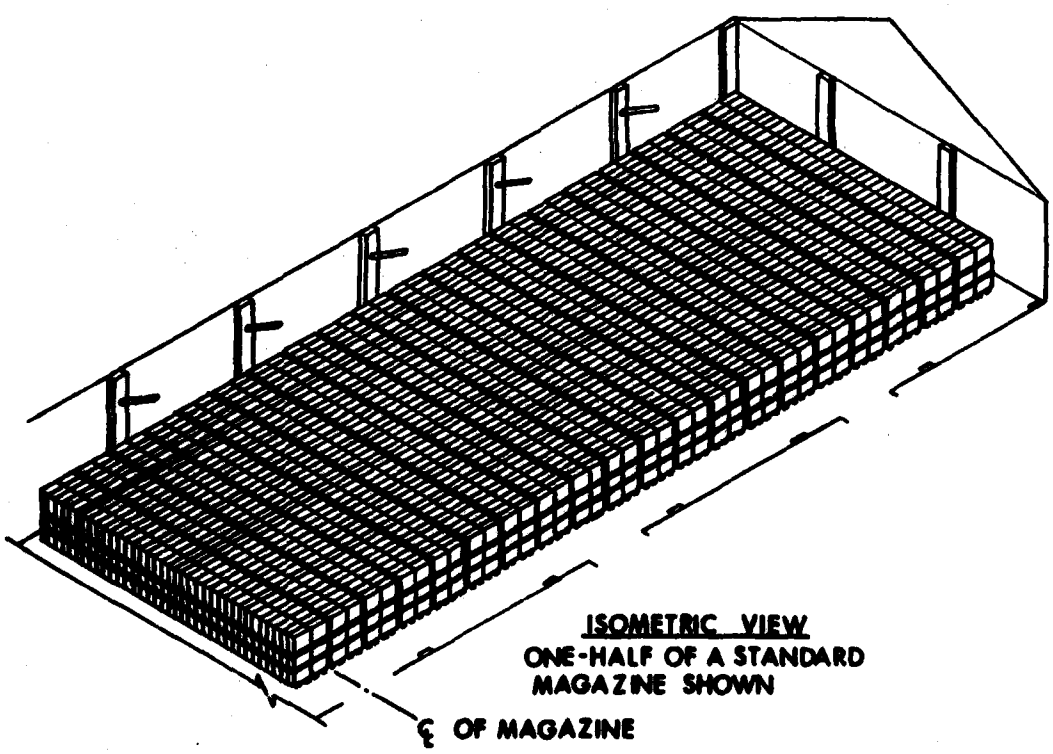


Figure 1. Storage array for 155 mm separate loading projectiles. Each box represents a pallet of 8 rounds.

is three dimensional, but little loss in generality occurs as a result of considering the two dimensional case.] Of interest is the size of the explosion (i.e., the number of participating munitions] resulting from the detonation of a single munition. Since the munitions are nearly equidistant, the array can be represented by an infinite regular lattice, the vertices of which represent individual munitions sites. Because of translational symmetry, a chosen site is typical of the rest, and the origin can be chosen arbitrarily. Interest then centers upon the cluster of sites, containing the origin, and representing the number of munitions which participate in the explosion. Three assumptions are made:

1. Propagation of detonation occurs only through nearest neighbor interactions. Experimental evidence has been obtained in support of this. The nearest neighbor munitions effectively shield next nearest neighbors from direct fragment attack.

2. The interaction probabilities (i.e., the probability that one round will detonate another) are independent.

3. The process of propagation of detonation is Markovian. Only the last state of the process (whether or not one set of rounds under consideration detonated) is relevant in determining whether or not the next set of nearest neighbors will detonate. Experimental results generally support this assumption. However, in the limit of high packing densities, large munitions, thin munition walls, and deformation sensitive explosives, it is expected that this assumption would break down.

Let p be the interaction probability, i.e., the probability that detonation of one round will cause detonation of its nearest neighbor. Experimentally, p can be measured by observing results of a large number of repetitions of an experiment involving a donor and an acceptor round, separated by a spacing identical to that in the array of interest, and noting the fraction of acceptor rounds which "detonate," according to the criterion discussed in the introduction. Clearly, $q = 1 - p$ is the probability that the interaction is too weak to cause a round to detonate, given the detonation of the donor. In a quadratic lattice (for example), the donor or source has four nearest neighbors, which comprise members of the first generation. (By definition, the source will be considered the zeroth generation.) The nearest neighbors of the munitions which detonated in the first generation comprise potential members of the second generation. The possible configurations for the zeroth through the first generation are shown in Figure 2. Note that the number of configurations for a given number of the first generation detonations is represented by the coefficients of the terms of the expansion:

$$(p + q)^4 = p^4 + 4p^3q + 6p^2q^2 + 4pq^3 + q^4$$

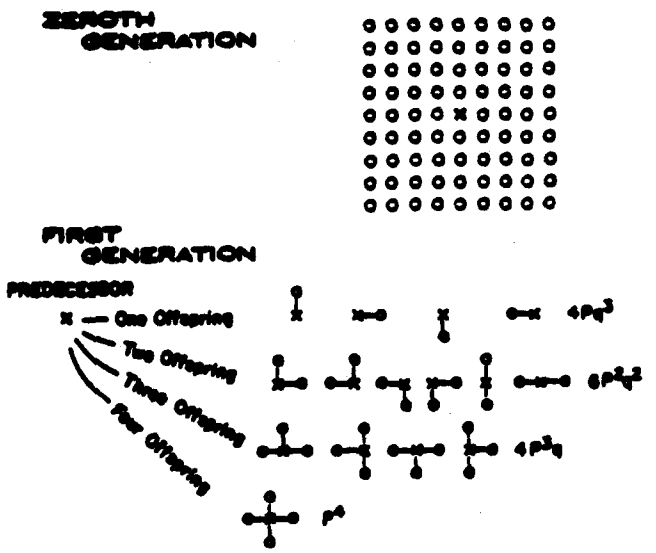


Figure 2. Simple quadriatic lattice showing all possible configurations for clusters containing the source munition (x) and possible members of the first generation. The bonds indicate an interaction has occurred. Undetonated rounds are suppressed in 2b.

The individual terms on the right hand side of this expression represent the probabilities of a given number of these neighbors being detonated. Denoting the expected or mean number of neighbors detonated by E(S) one has

$$E(S) = 4.p^4 + 3.(4p^3q) + 2.(6p^2q^2) + 1.(4pq^3) + 0.q^4$$

$$E(S) = 4p.$$

and $S(p) = 1 + E(s) = 1 + 4p.$

This is to be expected because of assumption 2.

In Figures 2 and 3, the bonds indicate that a munition has detonated. In principle, this procedure of direct enumeration can be continued through r generations, where r is arbitrarily large. In practice, direct enumeration is difficult because of the extremely rapid growth in the total number of clusters. An additional complication arises in the situation of interest here, in that there is the physical constraint that we not detonate the same round twice. In enumeration beyond the first generation, one must exclude forbidden configurations (See Figure 3). The mean explosion size S(p) is then equivalent to the mean number of bonds associated with clusters containing the source munition. Thus, in general, we can write

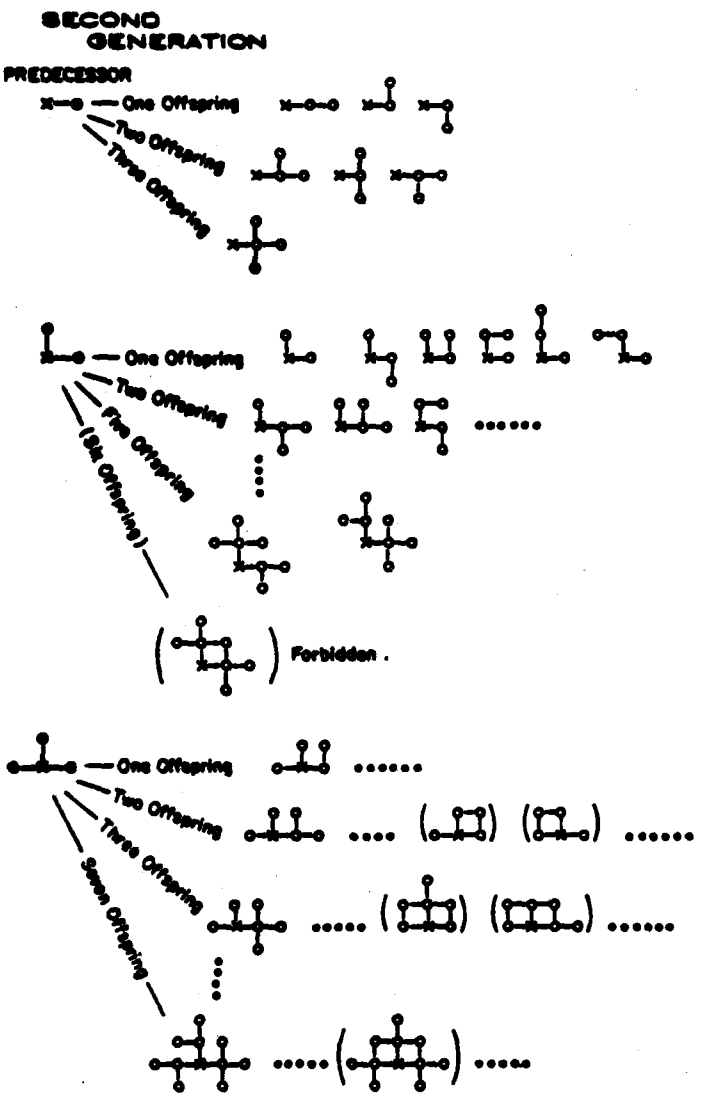


Figure 3. Some configurational members of the second generation. The configurations shown in parentheses are physically unrealizable, as they correspond to situations where the same round is caused to detonate twice.

$$S(p) = \sum_{n=0}^{\infty} a_n p^n$$

where the infinite cluster is excluded.

The description provided above of an explosion in a munitions array is a special case of a "bond" percolation problem (11, 12). It differs from the general case in that closed loops, as shown in Figure 3, are prohibited. Traditionally, the subject of percolation theory has been divided into two types of problems, the "bond" problem and the "site" problem (13). In the bond problem, each pair of neighboring lattice sites has probability p of being connected, independently of all other such pairs. In the site problem, each site has a probability p of being in state A and a probability $q = 1 - p$ of being in state B. A site is contained within a multi-site cluster if there is at least one nearest neighbor in the same state. The site problem arises, for example, in models of binary alloys (14), dilute ferromagnetic crystals (15), and thermal conductivity of disordered two-phase materials (16). The bond problem arises naturally in models of single phase dispersive flow of a liquid through a porous medium (17), the propagation of a blight through an orchard (18), or gelation of polymers (19). The site problem is not a natural choice for modeling an explosion in stacked munitions, as the site probabilities are not easily measured experimentally while interaction probabilities ($\equiv$ bond probabilities) are, at least in principle, directly measurable. However, it can be shown (20) that

$$P^{(s)}(n/p) \leq P^{(b)}(n/p) \ ,$$

where $P^{(s)}(n/p)$ refers to the probability of obtaining a cluster of size n, given an interaction probability, p, for the site problem, and $P^{(b)}(n/p)$ is the probability of getting a cluster of size n for the bond problem. Since

$$S(p) = \sum_{n} n\, P(n/p),$$

$S^{(s)}(p) \leq S^{(b)}(p)$ and we can use the mean cluster size, for the general site and bond problems, as lower and upper bounds, respectively, for the specialized bond problem of interest here. Vyssotsky, et al, has reported Monte Carlo estimates for the general bond problem in two and three dimensions for several lattices (21). Frisch, et al, have reported similar estimates for the general site problem (22). Plots of interaction probability versus cluster size for their site and bond results are shown in Figure 4, for the simple cubic lattic.

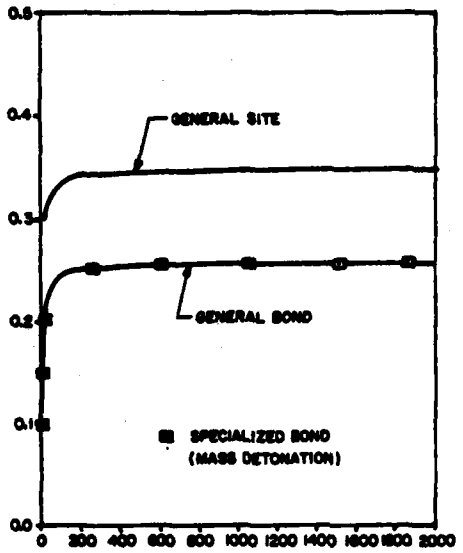


Figure 4. Mean cluster size versus interaction probability for general site and bond problems (21, 22) and specialized bond problem. Note that proscription of closed loops does not significantly change bond calculation results.

For infinite lattices, a percolation probability, P(p), can be defined as the probability that an infinite number of sites will belong to the cluster containing the source. Thus,

$$P(p) = \lim_{n \to \infty} P_n(p),$$

where $P_n(p)$ is the probability of obtaining clusters at least of size n. A critical probability, p_c, is defined as

$$P_c = \text{Supremum} \quad p/P(p) = 0 \ .$$

For $p>p_c$, there exists a nonzero probability that there will be an infinite cluster, i.e., that the detonation will propagate to an infinitely large extent. For $p<p_c$, the mean explosion size grows exponentially as $p \to p_c$ and diverges at p_c. Critical probabilities have been estimated for common lattices for site and bond problems by series expansion techniques and by Monte Carlo methods (15). Some calculated values are shown in Table I.

TABLE I

CRITICAL PROBABILITIES FOR COMMON LATTICES

Lattice	Monte Carlo				Series Method	
	$p_c^{(B)}$		$p_c^{(S)}$		$p_c^{(B)}$	$p_c^{(S)}$
Honeycomb	0.640	-	0.688	0.679	0.6527 (exact)	0.700
Kagome	-	0.435	-	0.655	-	0.6527 (exact)
Square	0.493	0.498	0.581	0.569	0.5000 (exact)	0.590
Triangular	0.341	0.349	0.493	0.486	0.3473 (exact)	0.5000 (exact)
Diamond	0.390		0.436		0.388	0.425
Simple Cubic	0.254		0.325		0.247	0.307
Body Centered Cubic	-		-		0.178	0.243
Face Centered Cubic	0.125		0.199		0.119	0.195
Hexagonal Closest Packed	0.124		0.204		-	-
	Ref. (15)					

D. Monte Carlo Estimates.

The series expansion description described above provides useful information regarding mass detonation phenomena, but it does not have the flexibility required, to address readily, certain additional issues. For example, munitions rarely have isotropic interaction probabilities: design features are usually such that nose-nose or base-base interactions are enhanced or depressed vis a vis side-side interactions. Furthermore, experiments have shown that simultaneous or near simultaneous detonation of collocated munitions can generate an extremely lethal collimated blast/fragment field with high probability of detonation of munitions within its path. Thus, if a round causes two nearest neighbors to detonate simultaneously, the probability of detonation of the next nearest neighbor in common with these two munitions is essentially unity. To address these problems and others, a Monte Carlo model was developed (23). This model is capable of handling both site and bond problems in one, two, and three dimensions. The computation is started by setting up the computational lattice as specified by the input. A site of the lattice, representing a munition round, is selected at random. The selected round is considered to be detonated. If the input specified that more than one round is initially detonated than a program subroutine is called to select the remaining rounds of the initial reaction set (ISET) from the nearest neighbors of the randomly selected site. An array (IND) in this computational model keeps record of the status of each round in the lattice. Thus, in a two dimensional bond problem, IND(i, j, 1) = 1, if the reaction propagated to the round at (i, j, 1), IND(i, j, 1) = 0, otherwise. At the beginning of a typical cycle of calculations, the bonds

emanating from all sites at the reaction front are examined to see which bonds block the reaction. This determination is achieved by using a random number generator to generate a continuous random number, r, such that $0 > r > 1$, and r has a uniform probability density distribution $f_r(r_o)$. The sample space for r is partitioned into two events; (i) the event E_1, $(r < p)$, that the bond is unblocked and propagates the reaction to a neighboring round, and (ii) the event, E_2, $(r > p)$, that the bond is blocked and does not propagate the reaction to a neighbor. Because of the assumption that a round can only be initiated by an immediate neighbor, the search process is limited to the first generation neighbors of the reaction front. The newly detonated rounds form the reaction front for the next cycle of calculations. The location of the new reaction front at the end of each cycle is saved in coordinate arrays. The calculation cycles are terminated when no new rounds are detonated. This will complete a trial and a new trial is initiated up to an input specified number of trails NTRIAL. At the end of each trial, the total number of reacted rounds in the reaction cluster for the trial is saved in an array, ND(j). At the end of the run, the mean reaction cluster size, and its standard deviation are computed and printed.

Several values of the interaction probability can be computed in a single run. The code has a number of options that can be either selected on input or achieved with a change of a few cards. The code will print out the hierarchy of the reaction branching process through the ammunition lattice if input specified. It is also possible to treat the nonisotropic case of unequal interaction probabilities p_x, p_y, and p_z. Another option treats the synergistic case of collimated blast fragments, by making the interaction probability $p = 1$, when two neighboring rounds detonate simultaneously.

The mean explosion size, for a simple cubic lattice, as determined by our Monte Carlo calculations, is juxtaposed with the results of Vyssotsky, et al, and Frisch, et al, in Figure 4. Our results are essentially identical to the results of Vyssotsky, et al, for the general bond problem. Evidently, restricting cluster configurations only to those which contain no closed loops has little effect upon calculated mean cluster size, or estimates of critical probabilities. Of special interest is the fact that the mean explosion size remains very small for $p < p_c$, and it is reasonable to take p_c as an upper bound of an acceptable interaction probability, with prevention of mass detonation the objective. As $p \to p_c$, mean explosion size grows very rapidly, approaching infinity at the critical point.

Shown in Figure 5 is the mean cluster size with and without the synergistic effect included, for the simple cubic lattice. Note that the synergistic

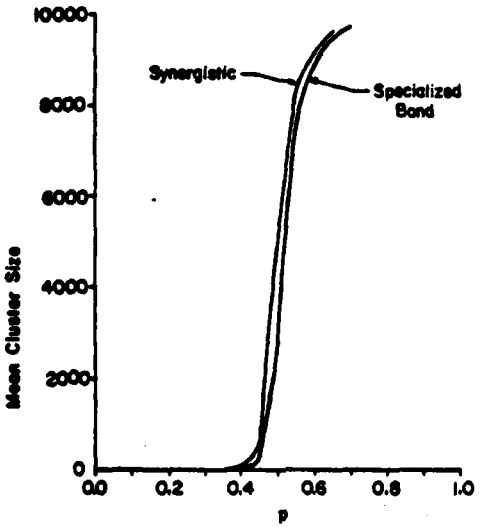


Figure 5. Mean explosion size for two dimensional square lattice, with and without synergistic effects.

effect lowers somewhat the probability required to get an explosion of any given size, but does not radically change the results.

In Figure 6, mean explosion size is plotted versus the interaction probability in the x and y directions, with P_z fixed at various values. Note that small values of P_z lead to greatly reduced explosion sizes. The roll-over at the top of each curve is due to edge effects. Not shown in Figure 6 is the curve for P_z fixed at unity. It would be to the left of the curve for $P_z = P_x = P_y$. Figure 7 shows calculations for the probability of getting an explosion of at least n rounds, as a function of P, for P_z constant and for P_z equal to a fixed fraction of p. As expected, the results for P_z equal to a constant lie to the left of the results for P_z equal to a fraction of p. Holding P_z constant simulates fixing the munition design and spacing between rounds in the z direc ion. Letting P_z vary with P_x allows one to account for variation in explosive sensitivity, as well.

Very high values of P_z are representative of shaped charge warheads, where the jet formed, when one round detonates, represents a very severe threat to the opposite round in the next layer. Very low values of P_z are representative of artillery munitions, such as 155 mm and 8" shell, where

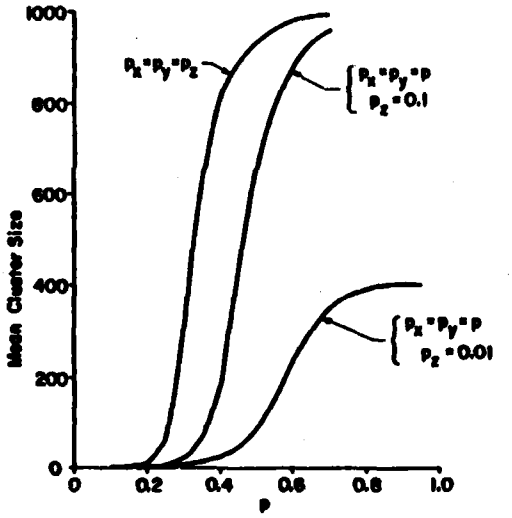


Figure 6. Mean explosion size versus interaction probability for simple cubic lattice: effects of anisotrophy.

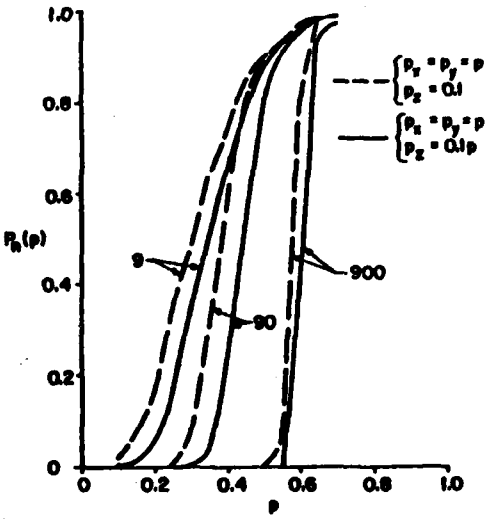


Figure 7. Probability of getting an explosion of at least n munitions, as a function of the interaction probability, for 3D cubic lattice: effects of anisotrophy.

the interaction probabilities between noses and bases are expected to be far weaker than the side-side interactions. The calculations show that this anisotropy can greatly reduce explosion size, for large three dimensional arrays. These calculations were used to design tests in which it was shown that explosion size could indeed be controlled by exploiting orientational effects. Thus, it was shown experimentally that 155 mm M107 shell (filled with TNT or composition-B) will not propagate in base-base orientation when separated by as little as 25 cm, for pallet sized units. As unit size was increased above the standard 8 round pallet, larger spacings were required, but it was shown that explosion did not propagate between units as large as 8 pallets (64 rounds, with approximately 15 pounds explosive per round) oriented base to base, and nose to nose and separated by less than 60 cm (2 feet). It follows from these results that it is advantageous to store munitions in arrays such that the z-axis, with low interaction probabilities, is the long axis of the array. For transportation on rail, for example, artillery ammunition should be oriented nose-nose and base-base, with the munition axes parallel to the train axis, in order to minimize explosion size.

It might be expected that restricting the interaction probability to low values in one direction essentially reduces the three dimensional problem to the appropriate two dimensional problem. Thus, setting $P_z = 0.01$, for example, for arrays with simple cubic symmetry would produce results nearly equivalent

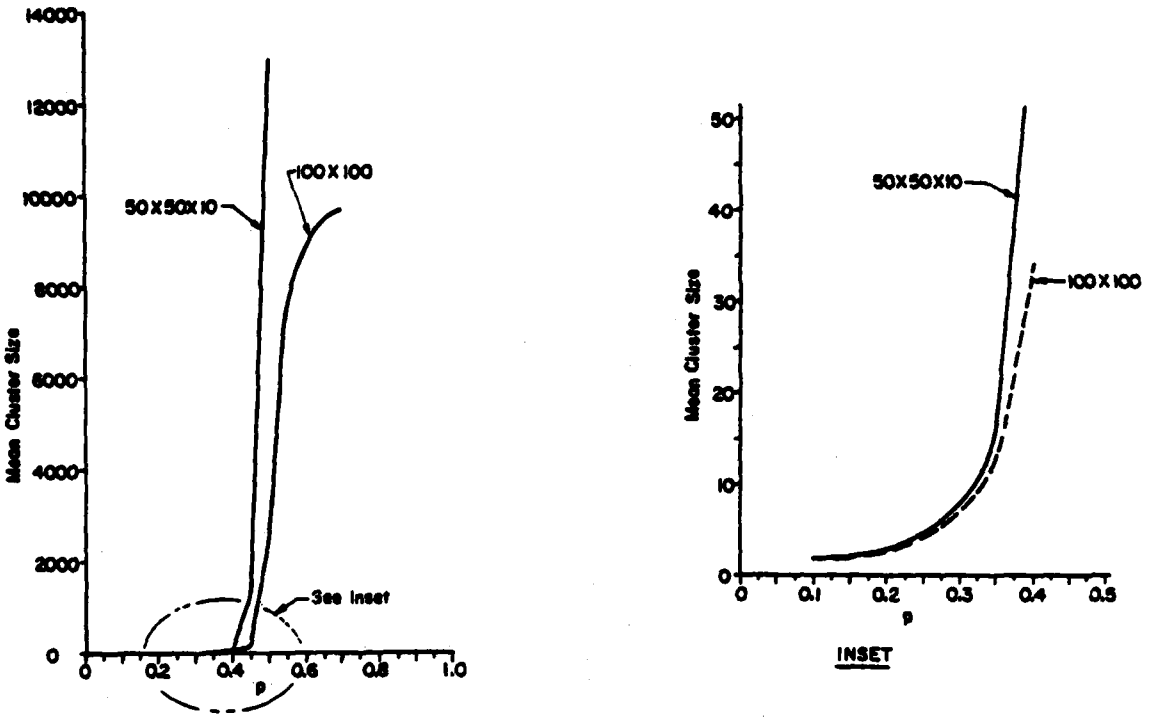


Figure 8. Comparison of results for square lattice and the simple cubic lattice, with $P_z = 0.01$. Note divergence of results at large cluster sizes.

to those for square arrays. In Figure 8, we show mean cluster size results for the simple cubic lattice, with $P_z = 0.01$, and results for the two dimensional square lattice. For small cluster sizes, the two problems are nearly equivalent. However, as the critical point is approached, the mean cluster size increases more rapidly for the three dimensional problem than for the two dimensional case. This is because propagation in the z direction depends not only on P_z, but on the number of sources, which depends on the size of clusters in the two dimensional arrays. Of considerable practical importance, it is noted that, as long as P_z is small, the same critical point criterion can be used for both two and three dimensional arrays.

II. SUMMARY AND CONCLUSIONS

The mass detonation problem has been formulated as a dynamic probablistic process, equivalent to a specialized bond propagation problem in percolation theory. A Monte Carlo model was constructed, with the flexibility of treating both bond and site percolation problems, but subject to the constraint that

no munition be allowed to detonate more than once. This constraint is equivalent to forbidding existence of closed loops in the cluster configurations, i.e., in graph theoretic terminology, trees are the only permissible configurations. Calculations were made for two and three dimensional arrays. Results of three dimensional calculations were compared with Monte Carlo calculations for the general site and bond problems as reported in the literature. The results of our specialized bond problem calculations are essentially indistinguishable from those for the general bond problem, indicating that the restriction of permissible configurations to trees has little influence on the results. Of special importance, it was found from plots of mean explosion size versus interaction probability that, as long as the immediate neighborhood of the critical region is avoided, the probability of achieving a mass detonation remains small. Thus, the critical interaction probability can be used to make estimates of the required munitions sensitivity to prevent mass detonation. The synergistic effect associated with simultaneous detonation of two rounds, causing near-unity probability of detonation of the next nearest neighbor, was treated and found to have a noticeable, but not strong, effect on the mean explosion size and critical probability.

Anisotropic interaction probabilities can exert a very strong influence upon mean explosion size and probability of mass detonation. Thus, it was found that setting the interaction probability in the z direction to a high value - e.g., 0.8 -, as would be observed experimentally for stacked shaped charge warheads, led to very large explosion sizes, even for relatively low values of $p_x = p_y$. Alternatively, it was found that low values of p_z were very effective in limiting explosion size. This was verified experimentally using 155 mm projectiles, and it was found that there are significant reductions in mass detonability obtained by oriented artillery shell in nose-nose and base-base configurations.

REFERENCES

1. DARCOM Regulation 385-100, p 2-7 (17 Aug 81).

2. "Safe Transport of Munitions," MTMC Report MTT81-1 (Jun 81).

3. For example, M1 105 mm HOWITZER Ammunition is not mass detonating when packaged in its standard fashion, two to a box.

4. Howe, P., "STROM Task 10 Report," in press.

5. Howe, P., "The Response of Munitions to Impact," ARBRL-TR-02169, Ballistic Research Laboratory, Aberdeen Proving Ground, MD (1979).

6. Gibbons, G., "Multiple Round Fragmentation Hazards and Shielding," ARBRL TR-02329, Ballistic Research Laboratory, Aberdeen Proving Ground, MD (1981).

7. Thomas, J. and Howe, P., "Effectiveness Testing for Antipropagation Shields Developed for M456 HEAT Tank Ammunition," ARBRL-TR-02370, Ballistic Research Laboratory, Aberdeen Proving Ground, MD (1981).

8. Porzel, J., et al, "Naval Explosives Safety Improvement Program (NESIP): Summary and Status," NSWC TR-81-27, Naval Surface Weapons Center, Dahlgren, VA (1976).

9. "Detonation" - the quotation marks are provided, because it is not absolutely clear that the target rounds need detonate according to a rigorous definition of a detonation. For test purposes, a "detonation" was considered to occur if the target munition reacted with design mode violence, as indicated by production of numerous high velocity, small fragments, consumption of all the explosive, and perforation of a 2.5 cm thick mild steel witness plate. We currently believe that this level of violence can be obtained from a constant volume explosion, followed by very rapid case rupture. The point of importance is that an unambiguous quantal response criterion for propagation can be established.

10. In order to treat a particular situation, the "independent" condition will be relaxed later in this paper.

11. Sykes, M. and Glen, M, "Percolation Processes in Two Dimensions I: Low Density Series Expansions," J. Phys. A: Math. Gen. 9, 87, (1976).

12. Dunn, A, Essam, J, and Ritchie, D, "Series Expansion Study of the Pair Connectedness in Bond Percolation Models," J. Phys. C: Solid State Phys 8, 4219 (1975).

13. For an excellent review of percolation theory, see Essam, J., "Percolation and Cluster Size" in Domb, C. and Green, M. Phase Transactions and Critical Phenomena, 2, Academic Press, NY (1972).

14. Broadbent, S. and Hammersley, J., "Percolation Processes, I Crystals and Mazes," Proc. Cambridge Philosophical Soc. 53, 629 (1957).

15. Shante, V. and Kirkpatrick, S., "An Introduction to Percolation Theory," Ado. in Phys., 20, 325 (1971).

16. Yuge, Y., "Three Dimensional Site Percolation Problem and Effective Medium Theory: A Computer Study," J. Stat. Phys., 16, 339 (1977).

17. Fisher, M. E., "Critical Probabilities for Cluster Size and Percolation Problems," J. Math Phys., 2, 620 (1961).

18. Todd, H, "A Note on Random Associations in a Square Point Lattice,"Roy. Stat. Cos Supplement 7 79, (1940).

19. Flory, P., Principles of Polymer Chemistry, Cornell UP, Ithaca, NY (1953).

20. Hammersley, J., J. Math Phys., 2, 728 (1961).

21. Vyssotsky, V., et al, "Critical Percolation Probabilities (Bond Problem)," Phys. Review 123, 1566, (1961).

22. Frisch, H., et al, "Critical Percolation Probabilities (Site Problem)," Phys Review 124, 1021, (1961).

23. Kiwan, A., "A Monte Carlo Solution to the Problem of Survivability of Munitions Stores," ARBRL-TR-02163, Ballistic Research Laboratory, Aberdeen Proving Ground, MD (1979).

WATER AS A TACTICAL WEAPON: A DOCTRINE FOR PREVENTING HEAT CASUALTIES (U)

*ROGER W. HUBBARD, Ph.D., MILTON MAGER, Ph.D.
AND MORRIS KERSTEIN, M.D., CPT, USN-R
US ARMY RESEARCH INSTITUTE OF ENVIRONMENTAL MEDICINE
NATICK, MA 01760, AND
4TH MARINE DIVISION, NEW ORLEANS, LA

INTRODUCTION

Success in battle requires maximal operational capability, flexibility, and effectiveness under all environmental conditions. The battlefields of the future will require continuous and independent operations around both the compass and the clock with greater emphasis on maneuverability and firepower. The survivability of the modern force will depend on the total and reliable performance of both man and machine; therefore, during hot weather operations, combat crews must operate at optimal efficiency yet remain free of dehydration and heat illness (1).

A number of authors have commented on the impact of heat illness on military campaigns, (2,3) yet, in no prior conflict (WWII to Vietnam) has the true impact of environmental heat on operational effectiveness been adequately measured or documented (4,5,6). According to COL T. F. Whayne (4), Chief Preventive Medicine Division, OSG, 1951, the data on heat injury from of WWII was defective because: 1) We have data only on cases severe enough to be admitted to a medical installation, and milder heat disorders are not recorded. 2) Criteria for diagnosing heat injury were not generally well understood. 3) Heat casualty rates have a seasonal incidence which are biased by the calculation of annual rates. 4)Study of heat injury on a theater rather than unit basis gives a false picture of the potential military problem. Navy monthly morbidity reports appear to confirm that for each case of heat illness admitted to the sick list, there are more than ten unreported cases treated as outpatients (5). In the History of U.S. Marine Corps Operations in World War II (6) is found: "Although several combat commanders believed that they lost as many men to the enervating heat as to enemy fire, no definite count of such casualties existed. Most heat prostration cases were usually treated at the medical aid station close to the front, where no records were kept."

In July 1943, the heat casualty rate in the desert area of the Persian Gulf Command in the Middle Eastern Theater reached 57 per thousand per year (4). This figure suggests a mean weekly incidence rate of 11 per 10,000. According to Terrill (7), "the incidence of heat injury in Vietnam varied from a low of 0.7 per 1000 per

day in January to a high of 5.4 per 1000 per day in May. These statistics include only those who required treatment by a medical officer for heat injury and include all logistical and support troops, most of whom are never exposed to significant degrees of heat and humidity under stress. More heat casualties occur in combat units and are never documented since they are treated by company aidmen or their buddies." The May rate calculates to a weekly casualty rate of 378 per 10,000 which is nearly 35 times higher than that reported for the Persian Gulf in July 1943. If this rate underestimates the true incidence of heat illness by a factor of 10, we have a significant military heat illness problem. Furthermore, these statistics do not address performance decrements although there is evidence that combat effectiveness would be severely impaired (5).

In 1954, the Navy Bureau of Medicine evaluated a new index of heat stress, the wet-bulb-globe temperature (WBGT). This index integrates the four major climatic variables: temperature, humidity, radiant heat exchange, and air movement into a single number equivalent to the environmental stress. Thermal stress is defined as environmental heat load acting on the body. The impact of the environmental heat load on man's thermal balance is complex because metabolic heat production is another major variable. By estimating dangerous levels of environmental heat stress (WBGT) and then compensating for this heat load by reducing metabolic heat production (work rate), it was thought that the rate of heat illness could be reduced. In 1956, the WBGT index was adopted by the Training Command at Paris Island and subsequently the weekly heat casualties dropped from 39.5 per 10,000 (1952-53) to 12.5 in 1955 and 4.7 in 1956 (6). These data, therefore, indicate that a heat casualty prevention program geared to the WBGT index can be very successful for recruits training at fixed installations.

Unfortunately, heat casualty rates among units participating in combat exercises in tropical and subtropical areas (5) and in the Mojave desert (37 per 1000 per day in 1962) (6) remained high. Many factors contributed to the risk of heat injury (9) including the inability to control the level of physical exertion and the logistical problems resulting in water shortage (5). Following our participation in Brave Shield XVI in 1977, we attempted to identify factors contributing to the high heat casualty rates. An analysis of these factors suggested problems in three general areas: 1) Attitude and doctrine. 2) Water and food logistics and consumption. 3) Work-rest cycles and weather information. Clues to prevailing attitudes and problems were frequently expressed during interviews in the field such as: "You can't stop the war" (work/rest); "Dehydration toughens up the troops" (doctrine); "I need very little water" (voluntary dehydration); "Troops don't need 3 C-rations per day" (salt); "High heat is routine here" (weather) and "Heat casualties are inevitable" (attitude). Further analysis suggested that U.S. Forces could operate effectively in hot weather if Commanders had the information, training and equipment required to adjust work-rest cycles and water intake to the prevailing environmental conditions. In short, we needed simple, rational guidelines relating fluid intake and work-rest cycles to local environmental conditions during combat operations. This presupposed three essentials: 1) a suitable index (WBGT) of environmental stress, 2) a small, light, rugged and inexpensive device for measuring the index (Botsball), and 3) guidelines for relating fluid intake and work-rest cycles

to the index. Two Marine Corps training exercises, CAX 8-80 and Lancer Eagle, provided the necessary opportunity to test our doctrine. Essentially, we planned to document the number of heat casualties and determine whether the heat casualty rate of selected units could be reduced by a short briefing, use of the Botsball to evaluate the environmental conditions, and modification of work levels and water consumption.

METHODS

Casualty Definition:

A heat casualty is defined as an individual with heat symptoms who fails to function for at least an hour. This represents an individual who reports to a medical facility such as a field hospital or battalion aid station (BAS), for evaluation, treatment and may be returned to duty. In most cases, this took the better part of a day and represented a significant amount of time lost from duty (NonVertical Marine).

Casualty Reporting:

CAX-8-80. At the completion of the exercise, medical records (unit LOGS) were reproduced from the following units: Field hospital located at the Brigade Service Support Group (BSSG); Battalion Aid Stations (BAS) located at the Expeditionary Air Field (Air Combat Element, ACE); BAS located at the BSSG; BAS of the 1/24; BAS of the 3/24.

RESPHIBLEX 1-81. Each corpsman from control and test companies were given copies of a medical questionnaire card for collecting heat casualty data, which we collected daily. We recorded heat casualty data admissions to the supporting medical facilities; Boone Clinic, Little Creek, VA; Naval Regional Hospital, Camp Lejeune, NC and the field hospital.

Heat Stress Index:

Official WBGT data for the training periods was obtained from Expeditionary Air Field (EAF), 29 Palms, CA; the central weather station, Little Creek, VA; five stations at Camp Lejeune. Botsball readings were taken by the Heat Research Team (HRT) and two test companies (at Little Creek, VA) at hourly intervals, from 0800 to 1600 hours each training day.

Training Exercise Description:

The Second Marine Amphibious Brigade (2nd MAB) Combined Arms Exercise 8-80 (CAX 8-80) was carried out entirely on the open desert (Mojave) at 29 Palms, CA, from 2-16 August, 1980. There were 636 officers and 5,374 enlisted reservists (6,010 total) from 110 units in 27 states. The weather was generally hot/dry (120-130°F db) with desert floor temperatures reaching 150 to 180 F. Men were domiciled in tents with food provided in field mess kitchens or as "C" Rations. Water was supplied from stationary, prefixed watering points via tanker trucks and water buffalo.

RESPHIBLEX 1-81, Operation "Lancer Eagle", was conducted by the 43D Marine Amphibious Unit which consisted of various units with a strength of 1150 men. Approximately 700 men of the 3rd Battalion, 5th Regiment, 4th Marine Division trained at Little Creek, VA from 10-18 July 1981 under hot, humid

conditions. The training varied from lectures in air conditioned rooms to a squad assault on a fortified emplacement during the heat of the day, followed by a night amphibious raid. In other words, the intensity of training varied; periods of rest were bracketed with periods of maximum exertion. The men were domiciled in air-conditioned barracks, meals were provided at either the galley or two field mess sites. Water buffaloes were provided throughout the training area. These units were transported aboard ship (18-20 July). Remaining units of the 43D MAU trained the entire period at Camp Lejeune, NC. From the beach assualt until the conclusion of the exercise (21-23 July) men were under field conditions.

Experimental Design:

In both exercises, the experimental design was kept simple and consisted of selecting two "high risk" or lead rifle companies as the test group. The two companies selected (CAX 8-80) were Alpha Co., 1st Battalion, 24th Regiment and Kilo Co., 3rd Battalion, 24th Regiment. The remainder of the 2nd MAB served as the control group. During Operation "Lancer Eagle", Lima and Kilo Companies served as the test group. Because of the nature of the training and schedule, two additional companies (Mike and India) served as a control group. The Test Companies received a short 15 minute briefing on the prevention of heat casualties emphasizing the importance of prehydration, drinking by the clock in the absence of thirst, and eating three meals per day. A corpsman was selected to receive additional instructions on the use of the Botsball, all corpsmen were given the Medical Questionnaire cards during Lancer Eagle but not during CAX 8-80. The Commanders of the Test Companies were given provisional heat doctrine cards containing the following simple guidelines:

WATER INTAKE, WORK/REST CYCLES DURING FIELD OPERATIONS FOR HEAT ACCLIMATED UNITS

HEAT CONDITION	BOTSBALL WGT°(F)*	WATER INTAKE (QT/HR)	WORK/REST CYCLES (MIN)
1. Green	80° - 83°	0.5 to 1.0	50/10
2. Yellow	83° - 86°	1.0 to 1.5	45/15
3. Red	86° - 88°	1.5 to 2.0	30/30
4. Black	88° & Above	2.0	20/40**

*To convert WGT to WBGT add 2°F. Below 80°drink up to 0.5 qt/hr, 50/10 work rest cycles.

**Depending on condition of the troops.

To Maintain Physical Performance:

1. Drink 1 qt of water in the AM, at each meal, and before any hard work.
2. Frequent drinks are more effective than all at once. Larger men need more water.
3. Replace salt loss by eating 3 rations per day.

4. As the WGT increases, rest periods must be more frequent, work rate lowered, and loads reduced.
5. Use water as a tactical weapon, and maintain top efficiency by drinking each hour.
   a) Remember these conditions apply to where the troops are located. For example, the outside temperature is often heat condition 1, but heat condition 4 in an amtrack. Thus, water consumption must increase.

   b) Meals: The troops must eat the equivalent of 2-3 C-rats per day, as this is how they derive their salt intake. We do not advocate salt supplements unless troops have water but no food.

   c) Rest: The troops must get 6-7 hours of rest per day (24 hr). This could be obtained by 2-3 hr rest periods if the situation allowed it. For example, the more successful CO's attempted to rest the majority of their troops (70% rest-30% on watch) whenever possible. They rotated troops out of hot vehicles on a more frequent basis.

   d) The incidence of eye irritations and nose bleeds is probably related to both the drying effects of the air and the prevalence of dust. We eliminated these problems in ourselves by 1) wearing goggles and kerchiefs, 2) washing the eyes and nostrils with water cupped in the palm of one hand as needed.

The Control Companies received the normal preactive duty training which heavily-stressed an awareness of the benefits of overdrinking. In addition, "M" and "I" company corpsmen were given the Medical Questionnaire cards. Their training was guided by the information contained in TB MED 507/NAV MED P-5052-5, Prevention, Treatment and Control of Heat Injury:

Use of the WBGT Index in Control of Physical Activity

The proponents of the WBGT index have proposed the following as a standard for application of the index. IT SHOULD BE EMPHASIZED that the measurements must be taken in a location which is the same as, or closely approximates, the environment to which personnel are exposed.

a. When the WBGT index reaches 78°F (26°C), intense physical exertion may precipitate heat injury, caution should be taken.

b. When the WBGT index reaches 82°F (28°C) discretion should be used in planning heavy exercise for unseasoned personnel.

c. When the WBGT index reaches 83°F (29°C), strenuous exercise such as marching at standard cadence should be suspended in unseasoned personnel during their first three weeks of training. At this temperature, training activities may be continued on a reduced scale after the second week of training.

d. Classes in the sun should be avoided when the WBGT exceeds 85°F (29°C).

e. When the WBGT reaches 88°F (31°C), strenuous exercise should be curtailed for trainees with less than 12 weeks training in hot weather. Acclimatized personnel can carry on limited activity at WBGT of 88°F to 90°F (31°C - 32°C) for six hours a day.

f. When the WBGT index is 90°F (32°C) and above, physical training and strenuous exercise should be suspended for all personnel excluding essential commitments.

RESULTS

During the 15 days (2-16 August, 1980) of the exercise (CAX 8-80) in the Mojave Desert, 1,387 marines reported for sick call or treatment on at least one occasion (1,387 per 6,010 or 23%). In Table 1, we have classified those complaints related to environmental factors

TABLE 1

Number of Heat-Related Cases Recorded by Medical Facility

Facility	Heat Exh.	Head ache	Cramps-Nausea	Nose-Bleed	Eyes	
Field Hospital	40	6	13	14	21	
BAS-1/24	34	18	6	8	5	
BAS-3/24	2	0	0	1	1	
BAS-EAF	22	16	2	15	7	
BAS-BSSG	12	13	10	8	12	
TOTALS	110	53	31	46	46	All-286
% Recorded at BAS'S	64%	89%	58%	70%	54%	

There were 75 cases of heat illness (Table 1) diagnosed as heat exhaustion or heat reaction, and 35 additional cases of possible heat exhaustion, based on symptoms. Thus, there were 110 cases of treated heat exhaustion lost to the commander for the better part of a day (Non-Vertical Marines). Additionally, there were 176 cases of heat related disorders (Table 1) such as headache, cramps, nausea, nosebleed and eye irritations. These cases represent a large group of marines and should not be overlooked when assessing the impact of the environment on the total force. A greater percentage of casualties (approximately 67%) were treated at battalion aid stations. The total number of heat-related cases was 286 (1 man in 5 reporting to a medical facility had a heat-related disorder). By conservative estimate (4 hours per man), this represents 1,144 man hours lost to duty. The majority are considered preventable illnesses.

Figure 1 shows that the incidence of headache, cramps, and nausea peaked at day 4 (75% of the cases by day 8), followed by the peak incidence of nosebleed and eye problems at day 5 (73% of the cases by day 8) and heat exhaustion at day 6 (85% by day 8).

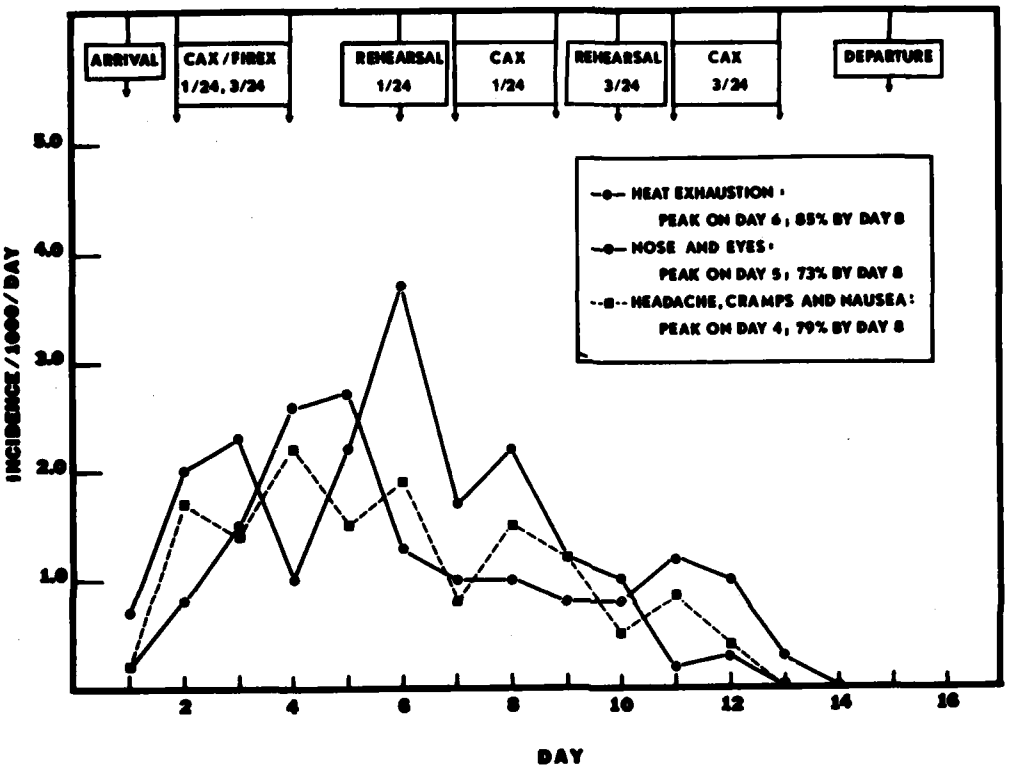


Figure 1. The incidence heat injury during CAX 8-80.

It should be noted that certain features of this exercise were repeated on separate occasions to evaluate their military proficiency. For example, the large scale military maneuver (1st Battalion, 24th Regiment, days 5-9) was repeated on the subsequent four days (3rd Battalion, 24th Regiment).

These data were further summarized as all heat related illnesses (Figure 2), including the effects of heat, low humidity and dust on the incidence of nose and eye problems. The summary of these data suggests a bell-shaped distribution with peak incidence on the 6th day. The only discontinuity in the data occurred on the 7th day of the exercise when a significant proportion of the total force was on or anticipating leave. Two features should be noted: a) the rapid rise in casualty rates through day six; b) the casualty rates for similar activities early on and later in the exercise. For example, note the 6.9 casualty rate for the rehearsal (1/24) on

day 6 versus the 2.3 casualty rate for the rehearsal (3/24) on day 10. By the same token, note the 4.7 casualty rate for the combined arms exercise (1/24) on day 8 versus the 1.7 casualty rate for the similar exercise (3/24) on day 12.

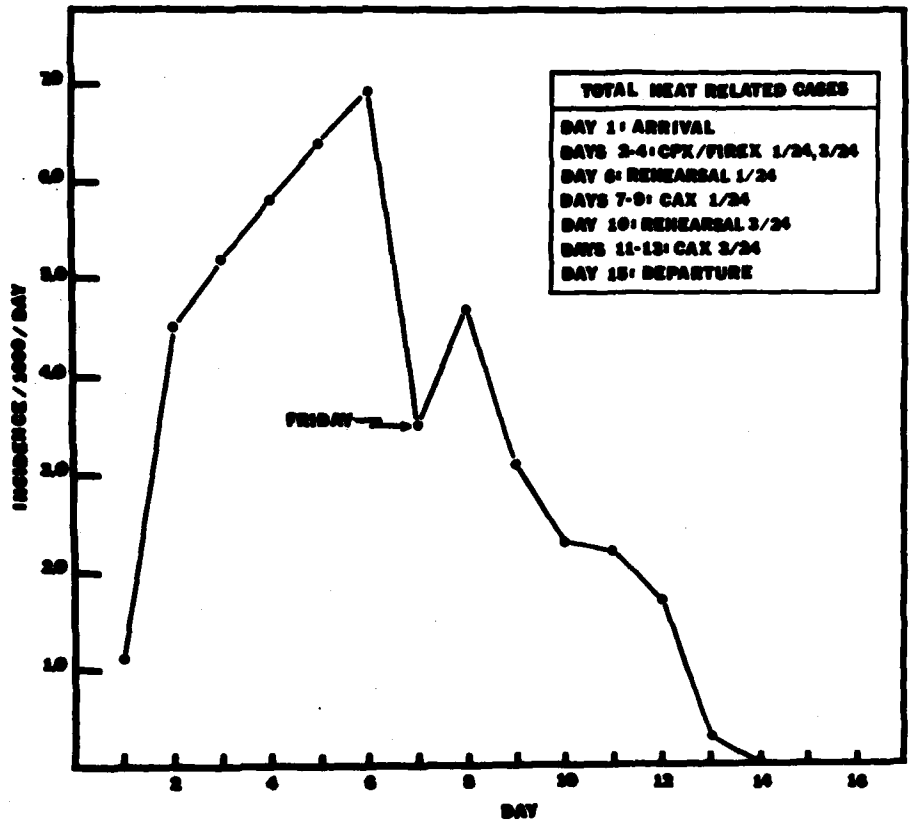


Figure 2. The combined incidence of all heat related cases during CAX 8-80.

The EAF weather station was located between the airstrip and a major road at MCB, 29 Palms while the field hospital located at the BSSG was approximately 4-5 miles down range. The WBGT readings were recorded at the field hospital by members of the HRT. Figure 3 depicts the highest daily WBGT readings taken at the same time of day at the two different locations. Note that the weather station did not begin recording WBGT readings until day 6 of the exercise and the marked variations in WBGT readings between the two sites. The average difference (mean $\pm$ S.D.) shown in this example was equivalent to 7.2 ± 2.3 degrees.

Kilo Company, 3rd Battalion; 24th Regiment was briefed on 4 August and Alpha Company, 1st Battalion, 24th Regiment was briefed on 5 August. Both companies completed their training without a single heat related casualty including nosebleeds and eye irritation. The CO of Alpha Co recounted his experience with

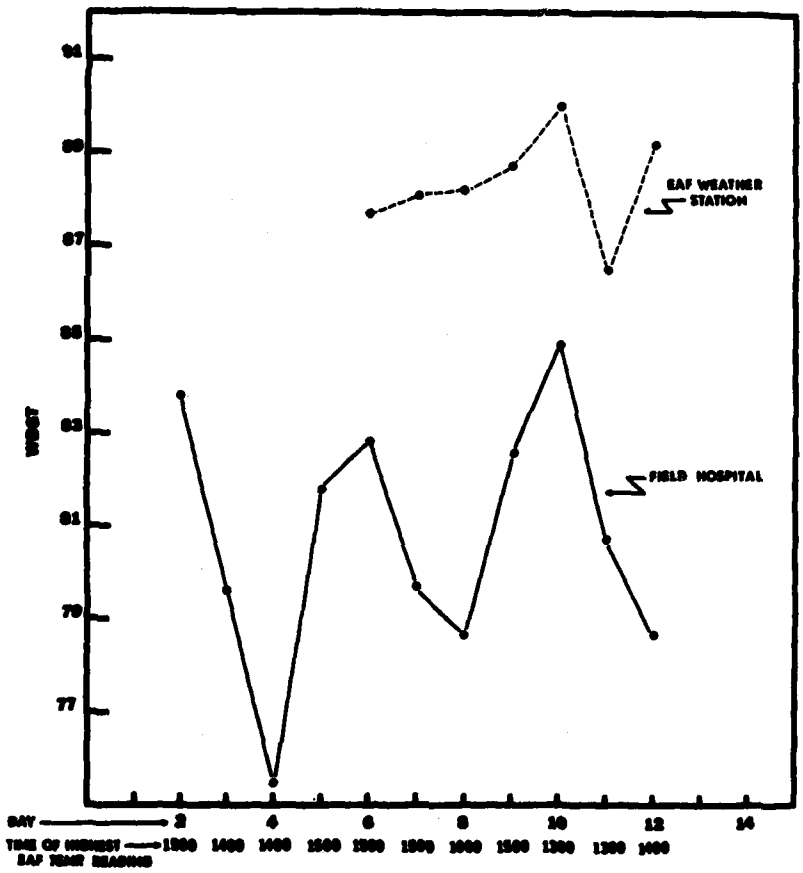


Figure 3. The highest daily WBGT readings recorded at the same time of day from two different sites.

the Botsball and said the instrument was an invaluable asset for adjusting work-rest cycles and fluid intake and had a profound psychological impact on the company. He further explained that the presence and use of the Botsball was a tremendous incentive to the men to continue drinking. It thus became an important "behavior modification" device which made command drinking an easier task to accomplish.

Operation "Lancer Eagle" - From 11-24 July, we collected heat casualty cards on 71 heat cases. There were 20 heat casualties from the "control" companies (I & M) which represents 91 cases per 1000 men (20 per 220 total). Since all casualties for the entire exercise were grouped within six training days (12-15, 21, 22 July), this represents a rate of 15/1000/day. In contrast, there were 13 heat casualties from the "test" companies, 42/1000 men (13 per 306 total). This represents a rate of 7/1000/day.

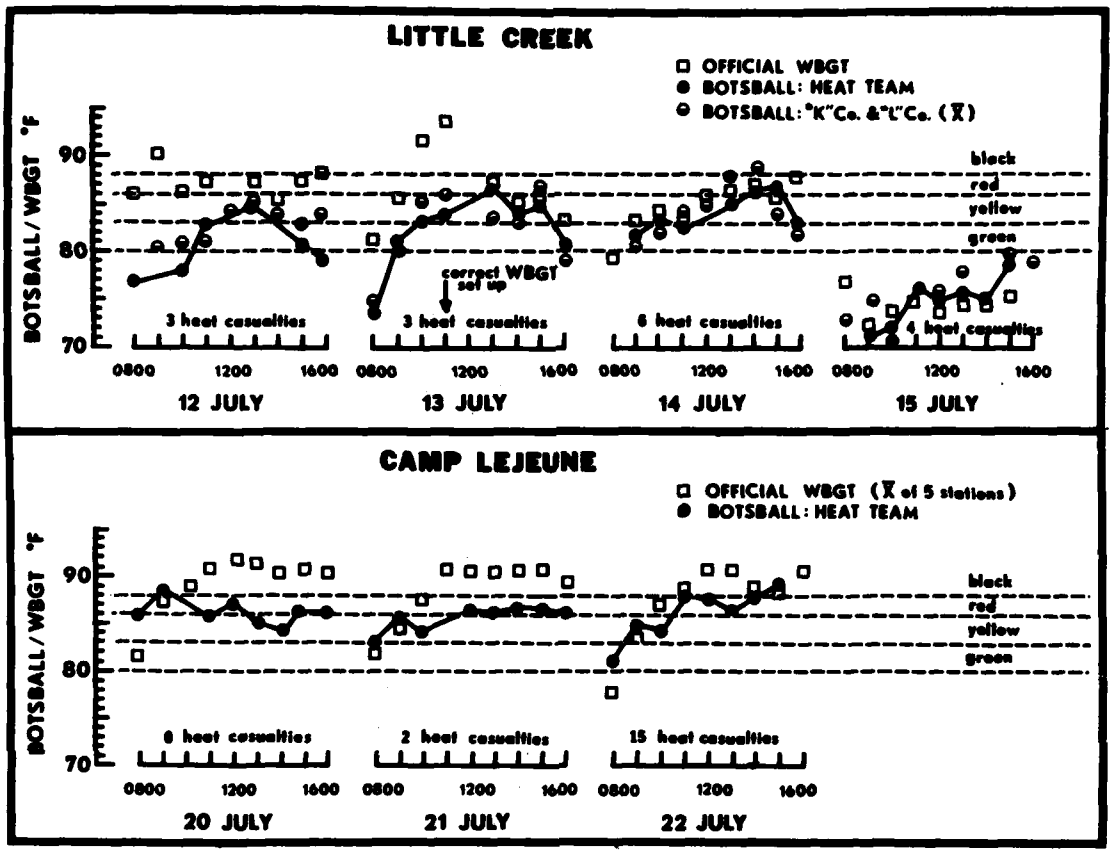


Figure 4. The daily heat casualties and temperatures during Lancer Eagle.

In addition, there were 38 heat cases among units not under our direct observation which were not included in the final statistical analysis. However, an estimate (71/1150/6 days) indicates a rate of 10.3/1000/day which is quite high compared to the desert environment (Figure 2). Approximately 50% of the heat cases within both control and test groups occurred during the training at Little Creek (Figure 4). These cases occurred during short periods of intense activity (1-2h) dictated by the training schedule. Training activities at both Little Creek and Camp Lejeune were subject to modification depending upon the WBGT index. Although the necessary operational procedures and technical facilities to collect and disseminate WBGT data were essentially in place, we discovered factors which, critical! impair the effectiveness of the system.

Th first involves consisting defining codes for categories of heat stress based on the WBGT index: Condition I, II, or III. At Little Creek, condition I was the most severe (WBGT, 88°F). However, in the operation plan for this exercise, there were four categories based on the WBGT: Condition I, II, III and IV-condition I

WBGT = 80°F. The operation plan lists the currently accepted condition codes. Nevertheless, a statement received in the field that "...we are in heat condition I" could legitimately be interpreted to mean either "we are in the least or the most dangerous heat condition."

The second involves the quality of the WBGT measurement itself. It soon became apparent to us that the official WBGT readings were significantly higher than our Botsball readings. The magnitude of this difference could not be explained by either the known 2°F difference between Botsball and WBGT readings or any reasonable site variations. A total of six WBGT stations were inspected by us: one at Little Creek and five at Camp Lejeune where we found technically incorrect set-ups of varying magnitude.

The WBGT station at Little Creek was corrected on 13 July. As shown in Figure 4, subsequent data from that station showed good correlation with the Botsball readings. Figure 4 also indicates the number of heat casualties by training day. The 15th of July at Little Creek was the coolest of the training days-- neither the official WBGT nor any of the Botsball readings exceeded 80°F. Nevertheless, four heat casualties were sustained, three of which came from "K" company which had carried out its night raid exercise on the previous night. Also, there may have been additional pressure to complete the training as quickly as possible. The results suggest acute exertion, dehydration and fatigue components in these casualties.

There were no casualties on 20 July with control and test companies still aboard ship. The amphibious landing at Camp Lejeune took place on 21 July. with "Red" Botsball conditions (hot/wet) persisting most of the day. Yet, only two heat casualties were reported during the landing. The level of physical exertion did not approach that required on the "Bridging Beach Obstacles" course at Little Creek.

The day following the landing (22 July) was the hottest day of the exercise and readings had moved into the black range by 1400 hours. Fifteen heat casualties occurred as the troops were pushing inland from the beach against the defending forces. During the assault phase at Camp Lejeune, no Botsball readings were recorded by the test companies' corpsmen. It seems likely that the immediacy of their other medical responsibilities simply took precedence over Botsball data collection.

DISCUSSION

A consistent feature of this and prior studies (5), was the pride, the motivation and the performance of the individual Marine. It is also a "given" that the combat training was realistic, that the Mojave was hot and dry in August and that the coastal regions of Virginia and North Carolina were hot and humid in July. Furthermore, the total sample size (7,150 individuals at risk) and the number of training days involved (21) appear adequate. Thus, our initial design criteria were met. However, there were differences between this and prior studies which should be borne in mind: 1) we were collecting casualty data from a lower level in the medical chain (the BAS and Corpsman) 2) there was an increased emphasis on water consumption. These two factors would increase and decrease apparent casualty rates, respectively.

In spite of statistical difficulties, a comparision of casualty rates from the desert area of the Persian Gulf and Vietnam,suggests the likelihood of more heat casualties under hot/wet conditions. There could be a number of reasons for this: 1) The WBGT index might not adequately differentiate between physiological responses to work in humid as compared to dry environments (10),2) heat dissipation would be limited, especially in the tropics, by the absorptive capacity of the atmosphere or its "cooling power" (11). In other words, sweat evaporation is more difficult under humid, "jungle" conditions and therefore, the heat dissipation rate is only rarely determined by the maximal rate of sweat production (12). It was expected that the peak casualty rate under hot/wet conditions (15 per 1000 per day) exceeded that under hot/dry (6.9 per 1000 per day). Thus, there appears to be a considerable impact of high humidity on the casualty rate. Also, 50% of the heat cases during Lancer Eagle occurred during short periods of intense activity clearly demonstrating the danger of overloading the heat dissipating mechanism under humid conditions.

There is no question that the WBGT index proposed by Yaglou and Minard (13) is effective in reducing heat casualties. Thus, we were concerned not with its theoretical soundness but its application under combat conditions. A number of problems were readily apparent. WBGT readings must be taken in close proximity to the personnel exposed due to variations in temperatures between two locations in the same area (Figure 3). However, the apparatus is cumbersome, fragile and unusable in many situations such as on the march or in tactical vehicles. As a result, the central weather station has been used. This can provide adequate information to a well defined training area, but is unsuitable for most combat scenarios. For example, the weather station did not begin recording WBGT readings until day 6 of the exercise (Figure 3).Moreover if a company commander wanted environmental temperature information he would have to radio his request through headquarters, to RLT HQ to the EAF weather station. The Company CO would receive temperature data for a different location at some prior time. Additionally, its usefulness can be compromised by confusion in the use of codes (see Results) and any errors in estimating the true heat stress index.

Another problem is suggested by examining the current guidelines for the use of the WBGT index in control of physical activity (see Methods). A rapid survey of these guidelines indicates: a) all of the examples relate to training, and presuppose: b) a training plan or prior experience with the effects of the training under cool or moderate conditions, and, therefore, c) command experience and the command control necessary to either reduce, curtail or suspend activity. When applied to combat scenarios, the following situational characteristics are not met: a) local WBGT readings at the site of the activity, b) the ability to suspend activity (the mission usually takes precedence) or c) the ability to predict activity levels for a six hour period (guideline, e).

The potential for having accurate, reliable heat stress information available to the Commander constantly made hour by hour adjusments in fluid intake and work/rest cycles truly feasible. The proposed guidelines for water intake and work/rest cycles represented estimated requirements (1) modified by the known beneficial effects of overhydration (14) and fluid ingestion during prolonged running

(15). The appropriateness of the work/rest cycles was checked by computer modelling of the predicted rectal temperatures (16).

The emphasis on "Water as a Tactical Weapon" and "Command Drinking" guided by the clock and not by thirst was required for two reasons. The first was to counter the misconception that dehydration toughened up the troops. Various forms of this misguided concept are prevalent at every level of Command. It is not known to what degree a similar Soviet doctrine impacted on the Egyptian Army via their Soviet advisors but intelligence reports indicate that during the 1967 Six-Day War with Israel, the Egyptians suffered 20,000 deaths with no visible wounds (apparently dehydration/heatstroke). During this period, Israeli heat casualties were minimal.

The second reason is based on a wealth of scientific information on the dangers of dehydration which limits sweating, adversely affects cardiovascular and thermoregulatory functions and predisposes to heat illness. The various signs and symptoms of dehydration are well-known and have been adequately described in Adolph's classic text (1). The physiological range of dehydration most important to the military commander is quite narrow, from 2-6% loss of body weight as body water. Thirst usually occurs at a 2% deficit. This is equivalent to approximately 1.5 qts of sweat in the average 75 kg soldier. Battlefield deficits of 2-3% loss of body weight as body water are very common. 5-6% losses represent an upper range, unless water is unavailable since thirst is intense (1) However, body water deficits representing 6-10% of an athlete's body weight can occur during prolonged strenuous sporting events (17). Between 3-5% occur a series of symptoms including vague discomfort, lassitude, weariness, sleepiness and apathy, all of which could clearly impact on the success of a military campaign.

As shown in the following quote by Minard in 1967, there are those within the medical community who feel that heat casualties during hot weather operations are almost inevitable. "Procedures successful in preventing undue heat stress in recruit training, chiefly by reducing metabolic heat load to offset a rise in environmental heat, cannot be applied to combat, since the level of physical exertion is dictated by the tactical situation. Also logistical problems in desert climates may lead to shortage of water and hence water rationing". Note two factors appear critical, the level of physical exertion and the supply of water. A more important factor, voluntary dehydration, was overlooked.

Dehydration, as reported by the Israelis, is one of the major operational causes of heatstroke (18). There are three principal causes of dehydration: 1) Inadequate supplies of water; 2) Depletion of the body's salt content (19). 3) The third cause of dehydration is is termed voluntary dehydration. It is due primarily to the inadequacy of the thirst mechanism to monitor or assess the severity of dehydration and to stimulate the amount of drinking required for rehydration. Adolph's research showed that voluntary dehydration increases with sweat rate and, thus, increases with higher ambient temperatures or work rates. It also increases with the temperature of the water and the distance a soldier has to walk to fill his canteen. He is less likely to walk 300 yds for a drink than he is to drink from a water bag hanging close by. Finally, a high percentage of the total deficit occurs during the interval between meals as approximately 44% of the cumulative deficit is replaced during meals (Adolph, 1947).

The cooling of battlefield water supplies to prevent voluntary dehydration is not a military concept that has been uniformly accepted or practiced. The British, in particular, have taken a stand against it (20). A careful consideration of this report suggests the author's conclusion is not justified and, in fact, provides data to support the concept (20). Cooling and flavoring of water supplies to improve palatability have been demonstrated as an effective substitute for an inadequate thirst mechanism (21). However, during CAX 8-80 and Operation Lancer Eagle, "drinking by command" was the only solution to this problem.

It has been amply demonstrated that acclimatization is extremely important in preserving the ability of troops to function in hot environments (5). The results of many exercises showed high rates of heat casualties occur in nonacclimatized infantry units operating in hot climates during the first week of the operation. The 1967 Six-Day War is a grim example of the potential for disaster. Since the members of the 2nd MAB were drawn from 110 units from 27 states, there was a lack of uniformity in the degree of individual acclimatization. This would necessarily predict, even with an "over drink" policy and with some preexisting acclimatization, a higher heat casualty rate during the first week of the exercise than in the second. As is indicated in Figures 1 and 2, there was a rapid increase in the incidence of heat casualties during the first critical days. It could be argued that the fall in the casualty rate after day six coincided with the availability of accurate heat stress information from the EAF weather station. This seems unlikely, however, since the incidence of headache, cramps and nausea had peaked on day 4 followed by nose and eye problems on day 5.

The test results from CAX 8-80 (zero heat casualties in two "high risk", highly motivated companies) establish the value of the new doctrine in desert combat scenarios. This doctrine will help to prevent heat illness during the acclimatization period of desert operations. It also provides Commanders increased flexibility to deal with environmental conditions during tactical situations. Moreover, their awaremess of and emphasis upon their own performance and casualty rates under known heat stress conditions will provide the very best training experience.

Although the test conditions during Operation Lancer Eagle were less than ideal, the heat casualty rates of test companies were reduced 50 percent. Further experience with these new guidelines under hot-wet condtiions could produce even better results.

The authors gratefully acknowledge the cooperation and enthusiatic support of the participating military personnel and their units.

The views, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official department of the Army position, policy, or decision, unless so designated by other official documentation.

REFERENCES

1. Adolph, E.F. Physiology of Man in the Desert. London, Interscience Publishers, Ltd. 1947.
2. Shibolet, S., M.C. Lancaster and Y. Danon. Heat stroke: A review. Aviat. Space Environ. Med. 47:280-301, 1976.
3. Knochel, J.P. Environmental heat illness: An eclectic review. Arch. Intern. Med. 133:841-864, 1974.
4. Whayne, T.F. History of Heat Trauma as a War Experience. Medical Service Officer Basic Course, Washington, DC, 9 June 1951, pp 1-38.
5. Minard, D. Studies and recent advances in military problems of heat acclimatization. Mil. Med. 132:306-315, 1967.
6. Garand, G.W. and T.R. Strobridge. History of U. S. Marine Corps Operations in World War II, Western Pacific Operations, vol. 4, p. 150.
7. Terrill, A.A. Troop Effectiveness and Acclimatization in the Middle East and Sub-Sahara Africa (MEAFSA). Office of the Command Surgeon, HQ USREDCOM, MacDill AFB, FL, 7 December 1979, pp 1-5.
8. Clarke, B.G. Prevention and treatment of heat casualties in the Mojave Desert. U. S. Navy Medical News Letter 45:4-5, 1965.
9. Bartlay, J. P. Heat stroke: Is total prevention possible? Mil. Med. 142:528-535, 1977.
10. Ramanathan, N.L. and H.S. Belding. Physiological evaluation of the WBGT index for occupational heat stress. Amer. Indust. Hyg. Assoc. J. 34:375-383, 1973.
11. Schickele, E. Environment and fatal heat stroke. Mil. Surg. 100:235-256, 1947.
12. Gilat, T., S. Shibolet and E. Sohar. The mechanism of heatstroke. J. Trop. Med. Hyg. 66:204-212, 1963.
13. Yaglou, C.P. and D. Minard. Control of heat casualties at military training centers. Arch. Indust. Hlth. 16:302-310, 1957.
14. Moroff, S. V. and D. E. Bass. Effects of overhydration on man's physiological responses to work in the heat. J. Appl. Physiol. 20:267-270, 1965.
15. Costill, D. L., W. F. Kammer, and A. Fisher. Fluid ingestion during distance running. Arch. Environ. Hlth. 21:520-525, 1970.
16. Givoni, B. and R.F. Goldman. Predicting rectal temperature response to work, environment, and clothing. J. Appl. Physiol. 32:812-822, 1972.
17. Buskirk, E.R. and W.C. Grasley. Heat Injury and Conduct of Athletics. In W. R. Johnson and E. R. Buskirk (Eds), Science and Medicine in Exercise and Sport (2nd ed.). New York: Harper and Row, 1974.
18. Hubbard, R.W. An Analysis of Current Doctrine in Use (USA vs IDF) for the Prevention and Treatment of Heat Casualties Resulting from Operations in the Heat, pp 1-40. Presented at the Commanding Officer's Conference, 4th MAW/MARTC, New Orleans, LA, 2-4 Oct. 1978.
19. Marriott, H.L. Water and Salt Depletion. Charles C. Thomas, Springfield, IL, pp.1-80, 1950.
20. Crowdy, I.P. Drinking Water Requirements in Hot Climates. Army Personnel Research Establishment Report No. 1/68.
21. Sohar, E., J. Kaley, and R. Adar. The Prevention of Voluntary Dehydration. Lucknow Symposium, Paper No. 5. UNESCO/NS/A2/680, Paris 27, July 1962.

ANALYSIS OF SMALL CALIBER MANEUVERABLE PROJECTILE (SCMP) COMCEPTS FOR HELICOPTER AND AIR DEFENSE APPLICATIONS

THOMAS HUTCHINGS
Systems Division, FC&SCWSL, ARRADCOM
DOVER, NJ 07801

I. INTRODUCTION

In air to air engagements with mobile threat helicopters, the possession of a round of ammunition which could depart from a normal flight trajectory to engage a maneuvering enemy target would be of great value. Maneuvering projectiles launched from an automatic cannon would have the added benefit of putting a number of rounds on target in a short time. The FC&SCWSL has initiated the development of such a projectile. In the sections that follow results of analytical investigation of the explosive thruster concept, which is only one of several maneuvering techniques, are described.

II. SCMP CONCEPTS

Recent technology developments in private industry suggest the possibility of developing a small caliber maneuvering projectile weapon system. Specific technological advances have been made in the areas of guidance system design, sensor technology, projectile maneuver control, projectile configuration design, packaging of microcomponents and g-hardening of projectile components. Several concept designs of small caliber maneuverable projectile systems, in calibers varying upwards from 30mm, have been proposed by several private industries. Significant technological advances have been made in the development of critical subsystem components. Private industry has been particularly active in the design of projectile control mechanisms and guidance techniques. Among the various proposed projectile concepts now being considered are spin-stabilized projectiles employing either lateral explosive thrusters or a tiltable nose as the maneuver mechanism and fin-stabilized projectiles as controlled by fins, spoilers, ram air jet devices, lateral explosive thrusters or lateral gas jets.

III. AERODYNAMIC ANALYSES OF EXPLOSIVE THRUSTER CONCEPTS

The analyses of the explosive thruster concept were based on a 30mm (and in some cases 40mm) projectile configuration similar in design to a conventional 30mm projectile.

The analyses consist of: 1) a preliminary study of down range deflection resulting from a lateral impulse, 2) a six degree of freedom analysis of the rigid body motions of the projectile subsequent to the application of a lateral impulse, and 3) system performance simulations of SCMP terminal miss distances.

In the preliminary impulse deflection study, a concept configuration for the band of explosive charges was developed, which consists of nine charges approximately 31mm long equally spaced around the perimeter. Each explosive charge imparts an impulse of 1.89 N.s causing a sudden change in lateral velocity of 5.83 m/s. Lateral drift of the projectile is shown to vary linearly with down range distance travelled after the impulse. In effect, the lateral impulse causes the projectile's velocity vector to rotate by a small amount. When the impulse is applied at a later time along the path the amount of velocity vector rotation is increased; therefore, the total lateral displacement depends on the time of impulse application and on the down range distance travelled after the impulse.

In the six degree of freedom rigid body dynamic simulations, the time variations in the angular orientation of the projectile's longitudinal axis and the angle of attack were determined before and after the impulse. Analytical results for the 30mm and 40mm spin-stabilized concept projectiles show that the projectile remains stable after pulse application, even when the pulse is applied after the velocity has slowed considerably. Results for both linear and helical grove configurations are presented in Figures 1 and 2. The figures contain graphs of the angle of attack and the azimuth orientation angle Θ for impulses applied 0.5 sec. after the start of the trajectory. After the impulse, the angle of attack decays, usually in about one second, to a value near the pre-impulse value; however, the helical explosive thruster arrangement results in a much faster decay time of about one-half second. The helical design produces an impulsive moment as well as an impulsive force so that the projectile's angular momentum vector gets instantaneously rotated by the same amount as the linear momentum vector. For the case of a linear explosive element, the angular momentum vector is also realligned, not instantaneously, but rather due to the action of aerodynamic moments affected by the sudden increase in the angle of attack. The precessional motion of the projectile caused by the sudden increase in the angle of attack lasts for over one second.

Off-centered impulsive loads cause the projectile to oscillate with greater amplitude and, in some cases, increase the settling time. In general, the amplitude of oscillations increases with the pulse application time.

IV. SYSTEM SIMULATIONS OF EXPLOSIVE THRUSTER CONCEPTS

System simulations of the explosive thruster SCMP concept were conducted to determine the overall accuracy of the system for different guidance schemes and to assess the effects of system errors on performance. Three types of guidance system were considered in these investigations. The first scheme determines the required amount of course correction based on repetitive trajectory calculations. Although this technique is not practical for hardware implementation, due to the large number of computations required to generate a solution, it does generate precise guidance commands. Therefore, the trajectory shooting algorithm provides a bound on the performance capabilities of the system. Another type of command guidance scheme considered is based on the AH-1S fire control computer equations. These equations determine how much rotation in the projectile's velocity vector is needed at any time to get back on course. The third guidance scheme analyzed is a target rate sensitive scheme that activates an impulsive correction when the lateral target rate exceeds some threshold. This guidance scheme could be implemented in a homing projectile that contains a sensor to measure the relative target distance and lateral rate.

Major findings of the system performance simulations of command guidance concepts are listed below:

1) One or two impulses are usually sufficient to correct for initial aim error of 5 milliradians in each axis, provided the guidance scheme is accurate and the relative distance to the target is not too small. Nine explosive line elements are sufficient to compensate for an accelerating helicopter target at ranges between 1500 and 4000 meters.

2) Initial aim errors are corrected most efficiently (fewest explosive elements fired) when impulsive firings take place at the earliest possible moment.

3) Errors caused by undetected target acceleration in the down range direction are small unless the target is initially very far down range (greater than 3500 meters). Miss distances caused by undetected target acceleration in the cross range direction are significant if the target range is beyond 1500 meters initially. In either case, explosive thrusters fired at half-second time intervals accurately compensate for the acceleration induced errors. Figure 3 demonstrates this point. In the first case, impulsive corrections are applied at only one time (0.5 seconds). In the second case, impulsive corrections are determined every half-second along the trajectory. Also presented is a simulation in which there are no impulsive corrections. In addition to the -5 milliradian initial aim error, an error occurs due to undetected target acceleration components of 2 m/s in both the down range and cross range directions. At 0.5 seconds from launch two pulses are fired regardless of the target range. Corrections made at this time effectively remove

the initial gun-pointing error contribution. Afterwards, the miss distance error grows from the effect of target acceleration.

Final miss distances on the order of one meter were obtained in the simulation of the perfect guidance scheme. As is shown in Figure 4 the inexact guidance scheme also resulted in small miss distances for targets at ranges below 4000 meters. Beyond 4000 meters the inexact guidance scheme required more than the nine available explosive charges for full correction.

4) The effects of 20 degrees error in the projectile's roll orientation and a 0.1 to 0.2 second delay in initializing the guidance command were investigated. These errors were shown to cause only a minor increase in miss distance. Figure 5 shows the effect of 20 degrees roll reference error on miss distance. In situations where corrections take place at multiple times along the trajectory the net effect of roll reference error is not large, since subsequent impulsive firings will compensate, in part, for errors induced by previous impulsive firings. The effect of roll reference error for the accelerating target simulation (described previously) in which impulsive corrections are determined every half-second is minimal. The final miss distance errors are less than two meters from the target center in all cases shown and usually are less than one-half meter.

5) Tracking errors (measured position and velocity errors of the projectile and target) significantly affect the miss distance of the SCMP. Miss distances often were increased by two to six meters when representative tracking errors were included in the simulations.

V. PERFORMANCE ANALYSIS OF A SCMP SYSTEM WITH TARGET RATE SENSITIVE GUIDANCE

For the active guidance mode, guidance computations are made primarily by a computer system located on or near the launcher vehicle. For semi-active and homing guidance, guidance computations are made onboard the projectile (in a microprocessor chip). The microprocessor determines course corrections on the basis of target state information provided by an onboard sensor.

In Figure 6 simulation results are presented for target rate homing guidance. The number of charges fired at each pulse time is allowed to vary, depending on the magnitude of the relative target cross velocity. Also, the rate homing scheme is activated only when the range between projectile and target is below some prescribed value. In this way, the simulation can account for the inability of many onboard sensor devices to detect the target beyond a given range. For the simulation shown the rate homing scheme was activated at a range of 500 meters. The system activated at 500 meters fires anywhere from three to all nine explosive thrusters initially. In this case, excellent results were obtained for targets at initial ranges below 3500 meters. The degredation in performance at longer ranges is again due to the fact that more

than nine explosive charges are needed to achieve full correction. Although the rate homing guidance works well, particularly at shorter target ranges, it is not as efficient as the active scheme in conserving explosive charges.

VI. RAM AIR JET CONCEPT DEFINITION AND ANALYSIS

Currently, a maneuverable projectile concept is being developed by Ford Aerospace & Communications Corporational for eventual use on Army helicopters. The baseline concept, shown in Figure 7, is a 40mm fin stabilized projectile that actively maneuvers by means of a ram aim mechanism. About 40 ms after being launched, the projectile will be given a unique address by a milimeter command guidance system onboard the launcher vehicle. A shared aperture guidance system allows the helicopter to track both the target and multiple rounds during the engagement.

Maneuver capabilities of the projectile will depend on a number of factors including the diameter and exit angle of the ram air jet mechanism; furthermore, the moment arm of the control force with respect to the projectile's center of mass is an important design parameter. In Figure 8, the theoretical lateral maneuver accelerations are shown for two design configurations as a function of Mach number.

System simulation of the FACC concept projectile demonstrate that this concept will perform well against enemy helicopters and surface targets at extended ranges. Figure 9 contains results of simulations for three targets -- and for two values of the ram air jet exit angle. Optimal performance is achieved for a 30 degree exit angle. Near zero miss results for targets at ranges up to 3.0 km.

VII. AIR DEFENSE CONCEPTS

A low drag projectile concept containing a hot gas generator mechanism is being considered for use in air defense applications. In Figure 10 the effects of 5.0 mrad aim error is illustrated for an evasive helicopter target. Both a ballistic trajectory and the maneuverable projectile trajectories are shown. More evasive targets were considered including jinking, fixed wing aircraft. The gas jet controlled projectile has no difficulty in intercepting the jinking target, even at ranges beyond 4500 meters.

In conclusion, analytical studies have demonstrated the high performance potential of small caliber maneuverable projectile concepts. These investigations reveal that the SCMP concepts have the performance capability to destroy airborne and surface targets regardless of any in-flight evasive tactics. In the near future, we plan to demonstrate the feasibility of small caliber maneuverable projectile concepts first by designing and testing hardware components, and eventually, by developing and testing a flyable breadboard.

FIGURE 1: ANGLE OF ATTACK VS TIME
SIX DOF TRAJECTORY SIMULATION

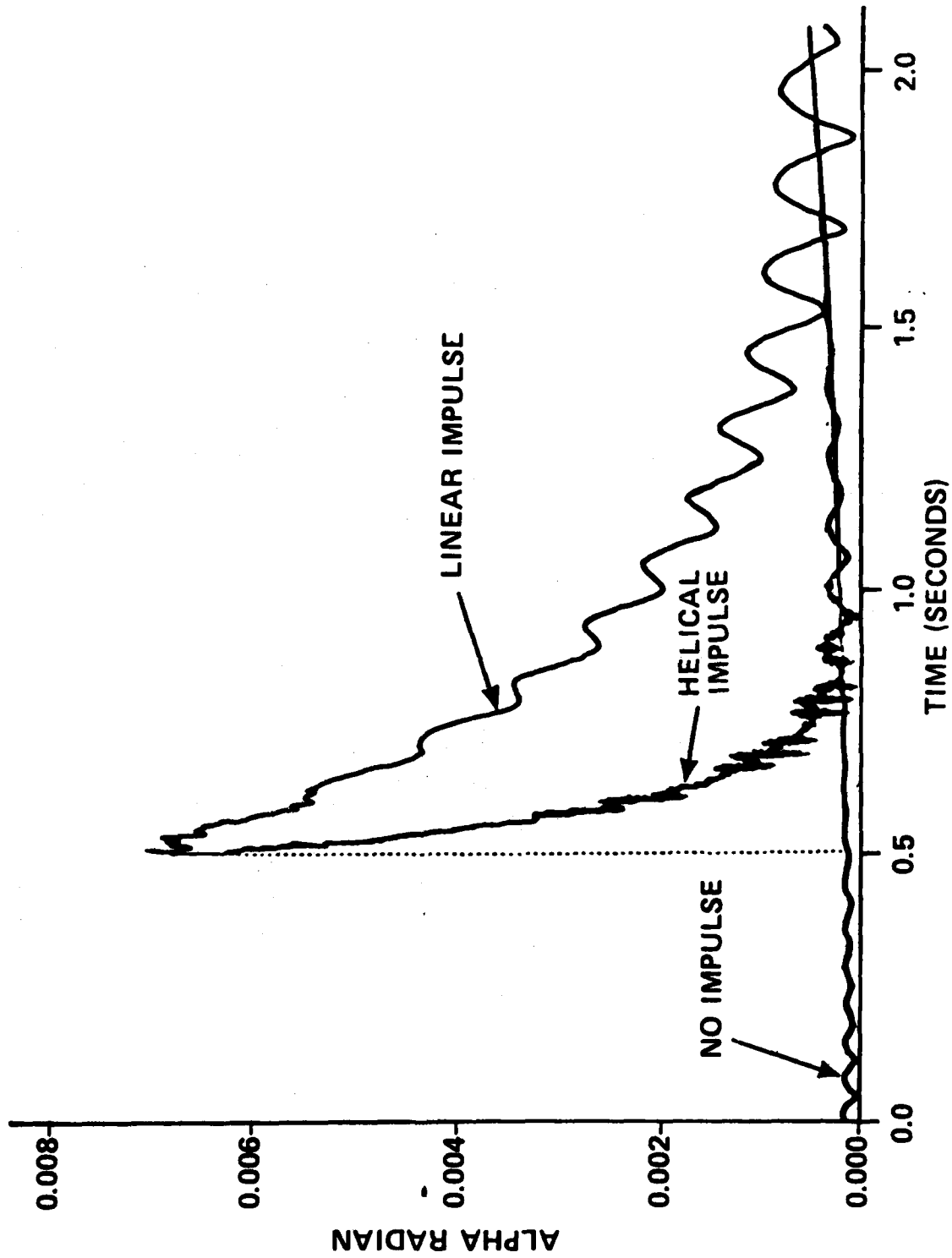
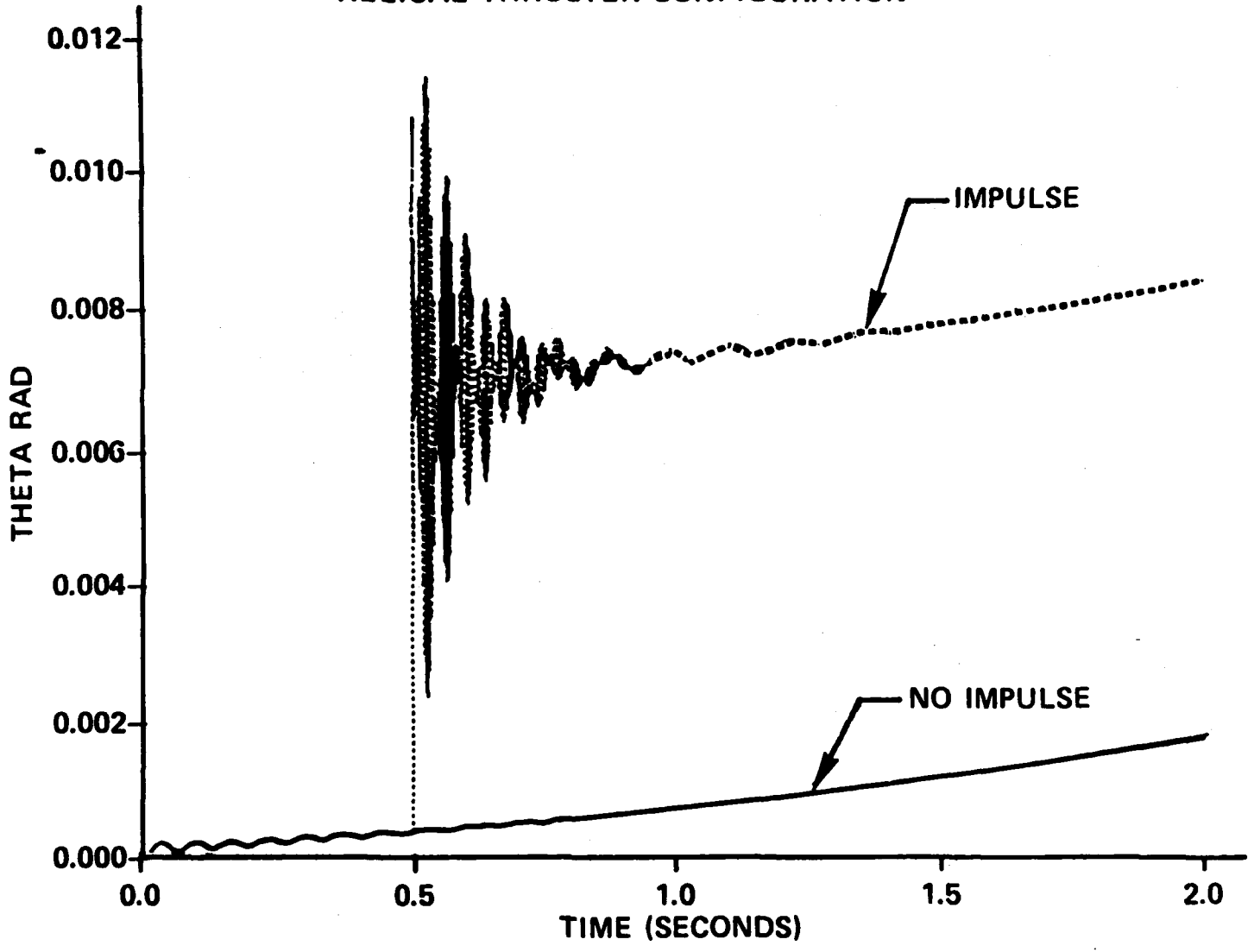


FIGURE 2:

HEADING ANGLE VS TIME
SIX DOF TRAJECTORY SIMULATION
HELICAL THRUSTER CONFIGURATION



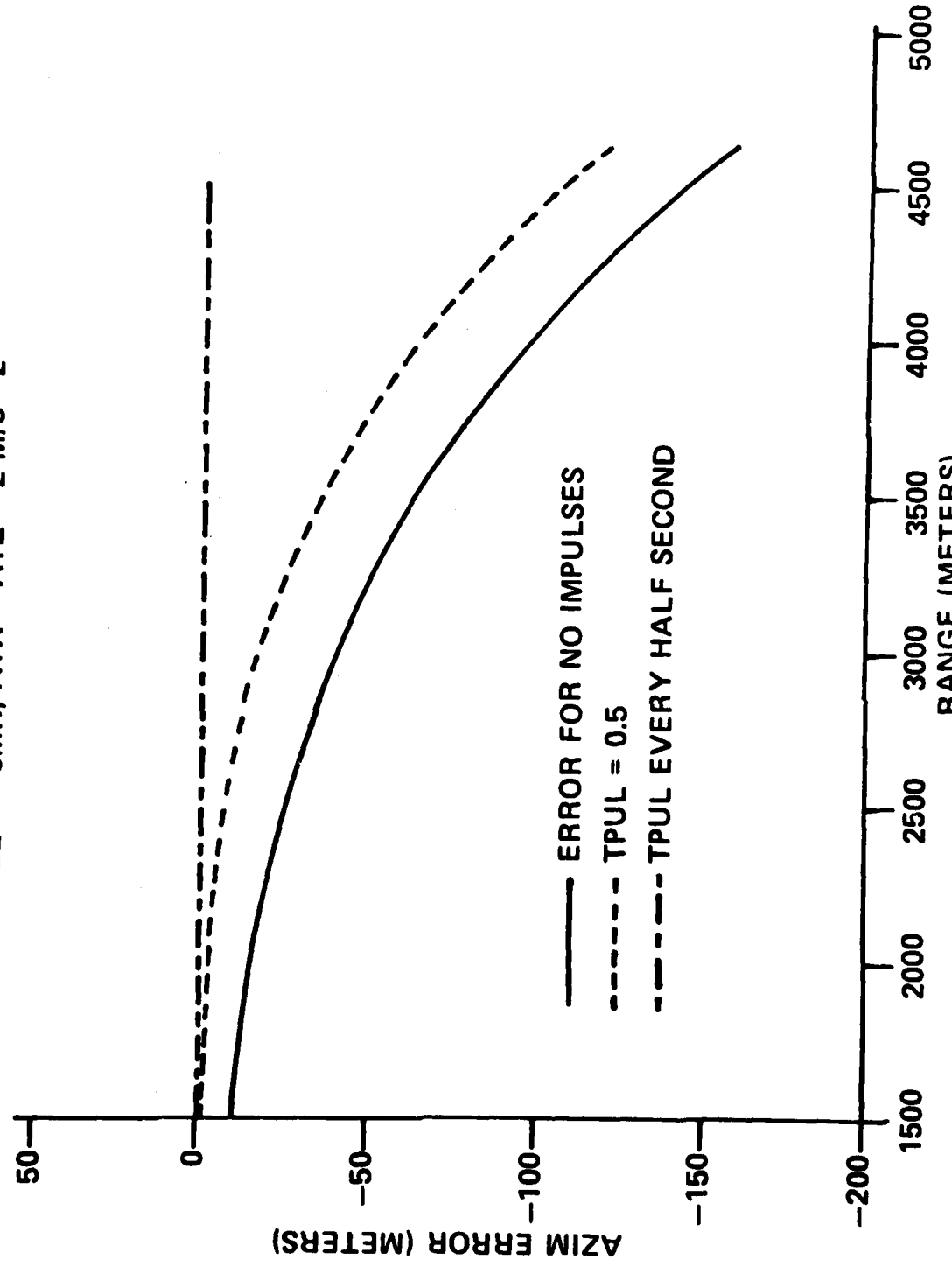

PRECISE COMMAND GUIDANCE SIMULATIONS
SPIN STAB EXPLOSIVE THRUSTER PROJECTILE
AZ = EL = −5MR; ATX = ATZ = 2 M/S**2
AZIM ERROR (METERS)
50
0
−50
−100
−150
−200
1500
2000
2500
3000
3500
4000
4500
5000
RANGE (METERS)
ERROR FOR NO IMPULSES
TPUL = 0.5
TPUL EVERY HALF SECOND

FIGURE 3:

FIGURE 4:

COMMAND GUIDANCE SIMULATIONS
SPIN STAB EXPLOSIVE THRUSTER PROJECTILE

AZ = EL = −5MR; ATX = ATZ = 2 M/S/S
TPUL = EVERY HALF SECOND

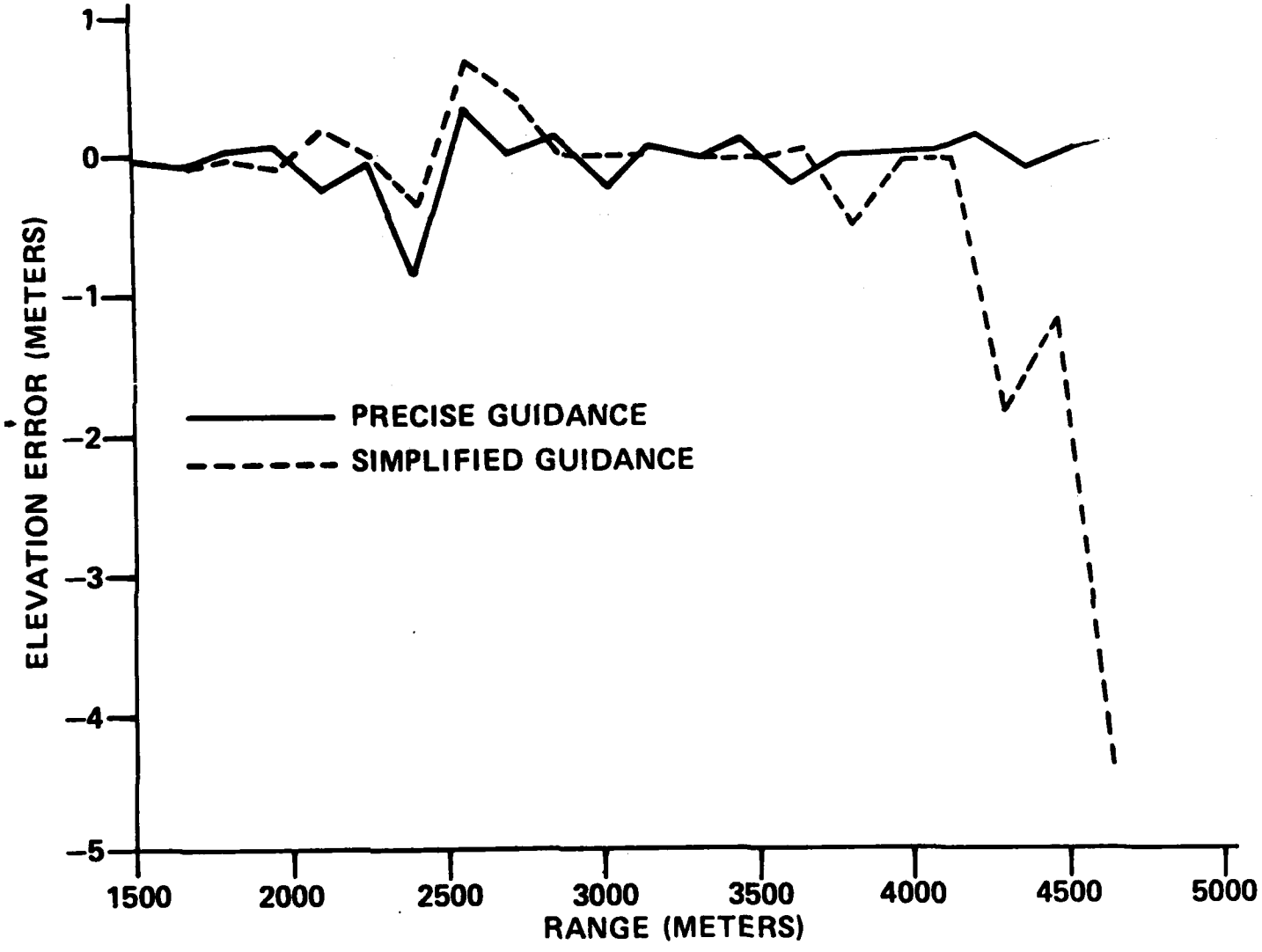


FIGURE 5:

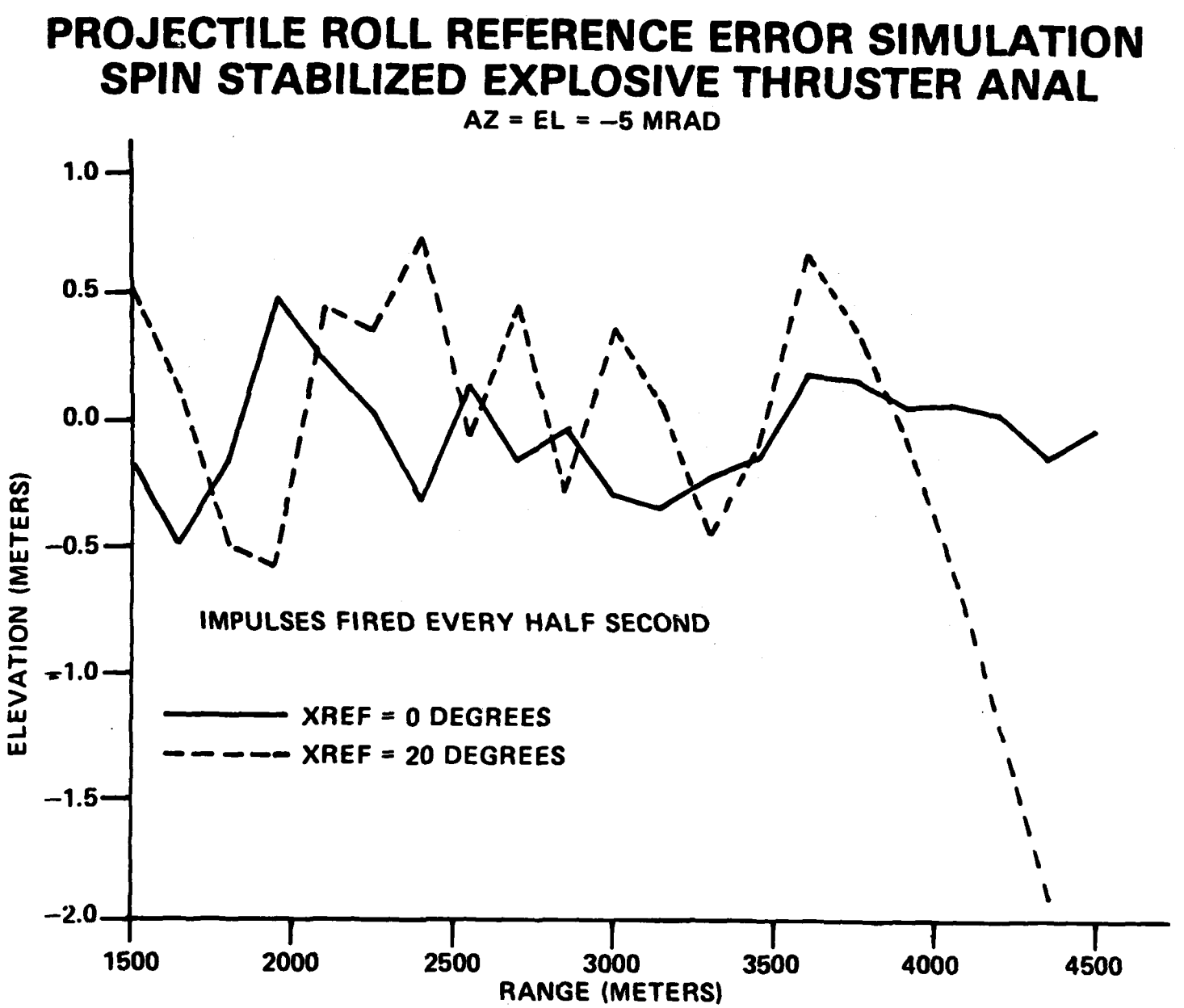

PROJECTILE ROLL REFERENCE ERROR SIMULATION
SPIN STABILIZED EXPLOSIVE THRUSTER ANAL
AZ = EL = –5 MRAD
1.0
0.5
0.0
–0.5
–1.0
–1.5
–2.0
ELEVATION (METERS)
IMPULSES FIRED EVERY HALF SECOND
XREF = 0 DEGREES
XREF = 20 DEGREES
1500
2000
2500
3000
3500
4000
4500
RANGE (METERS)

FIGURE 6:

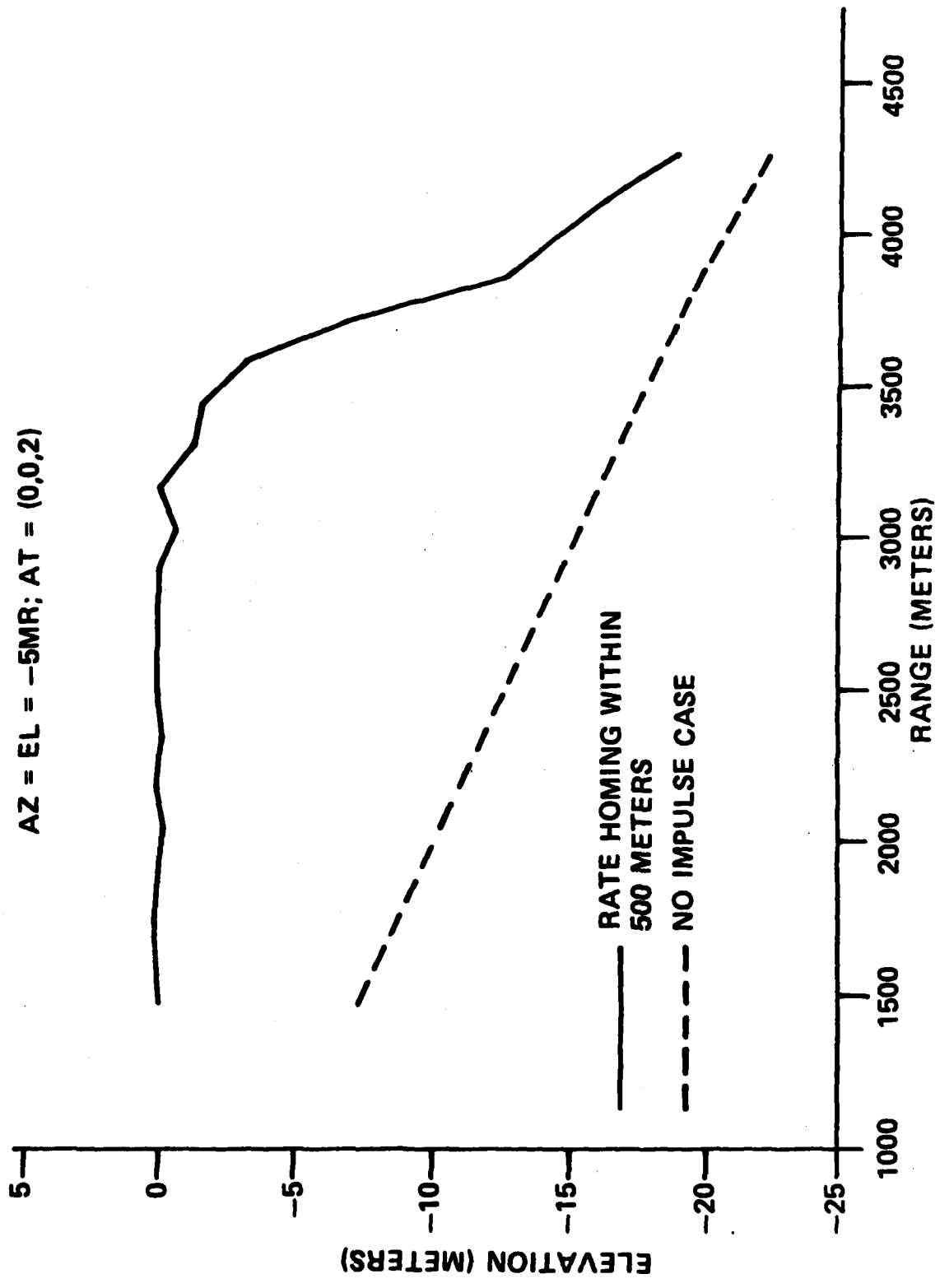

RATE HOMING GUIDANCE SIMULATION
DTP = 0.25
AZ = EL = −5MR; AT = (0,0,2)
RATE HOMING WITHIN 500 METERS
NO IMPULSE CASE
ELEVATION (METERS)
5
0
−5
−10
−15
−20
−25
1000
1500
2000
2500
3000
3500
4000
4500
RANGE (METERS)

FIGURE 7: *THE BASELINE 40mm PROJECTILE USES RAM AIR CONTROL AND FLEX FINS*

HUTCHINGS

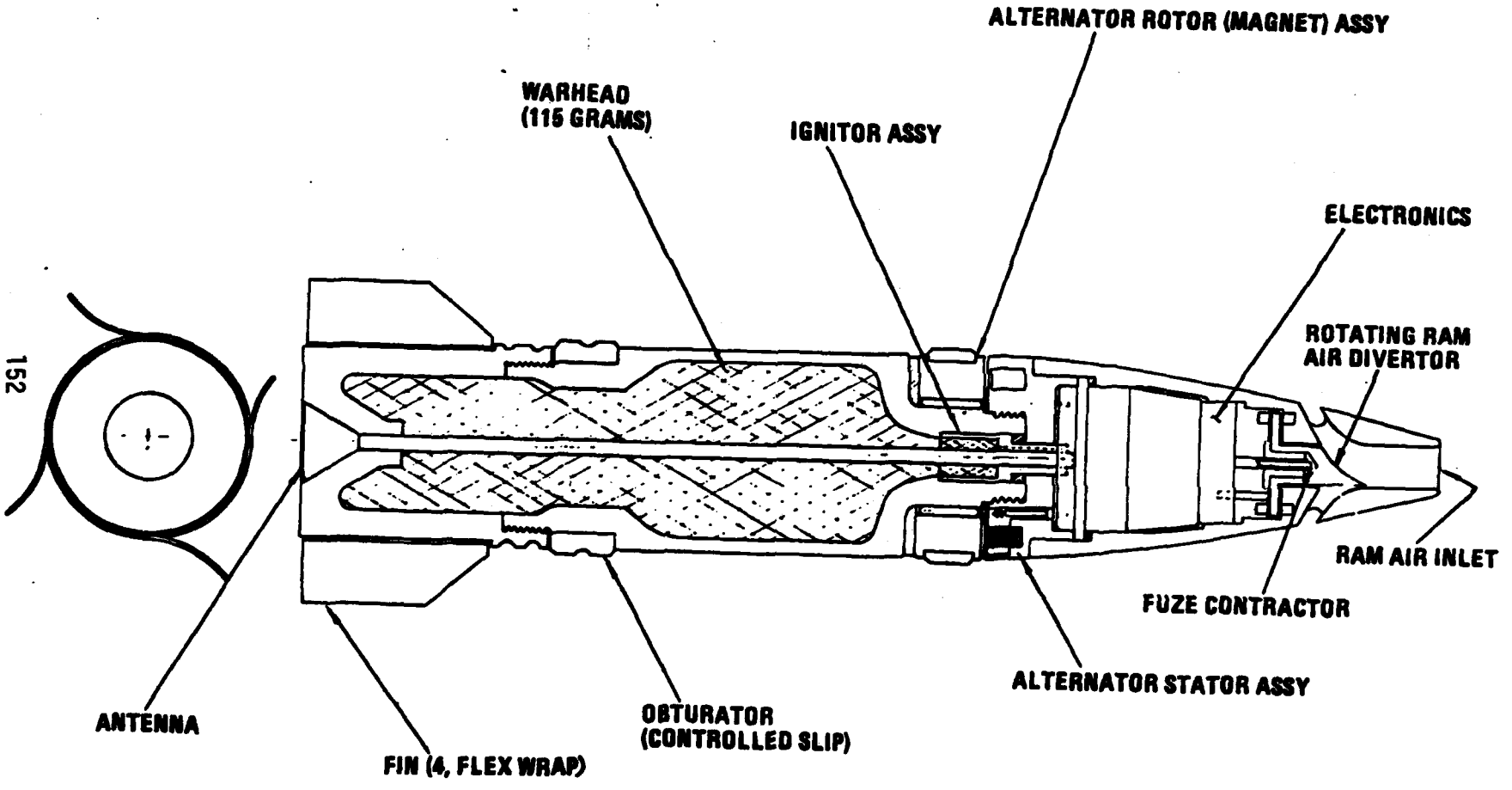
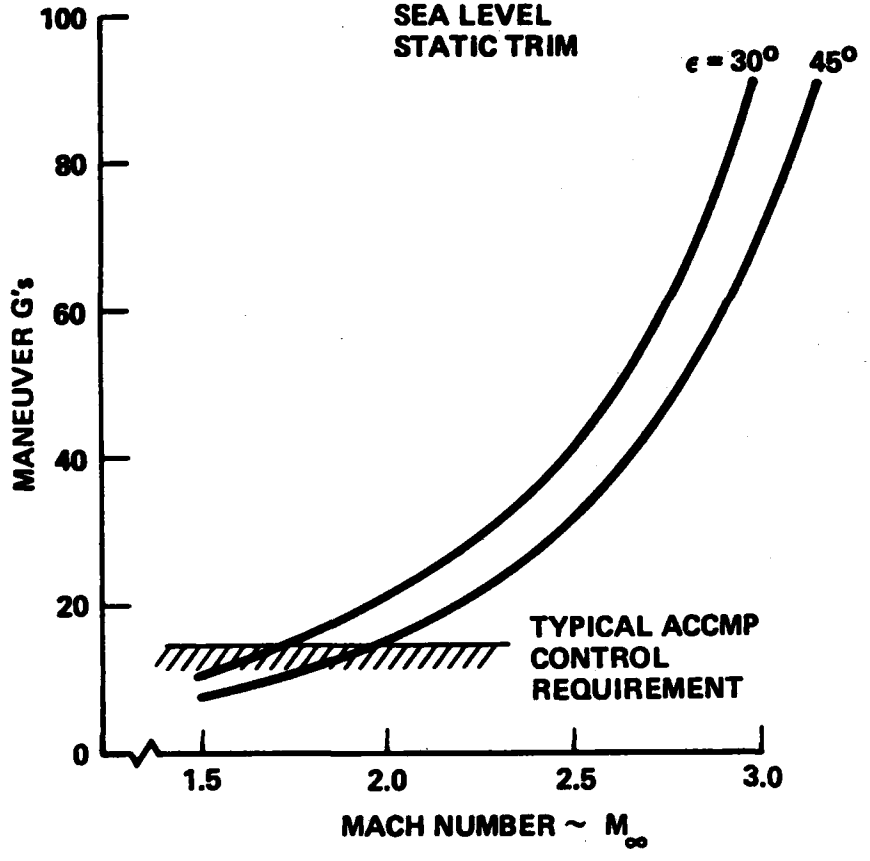


FIGURE 8: *MANEUVER CAPABILITIES*



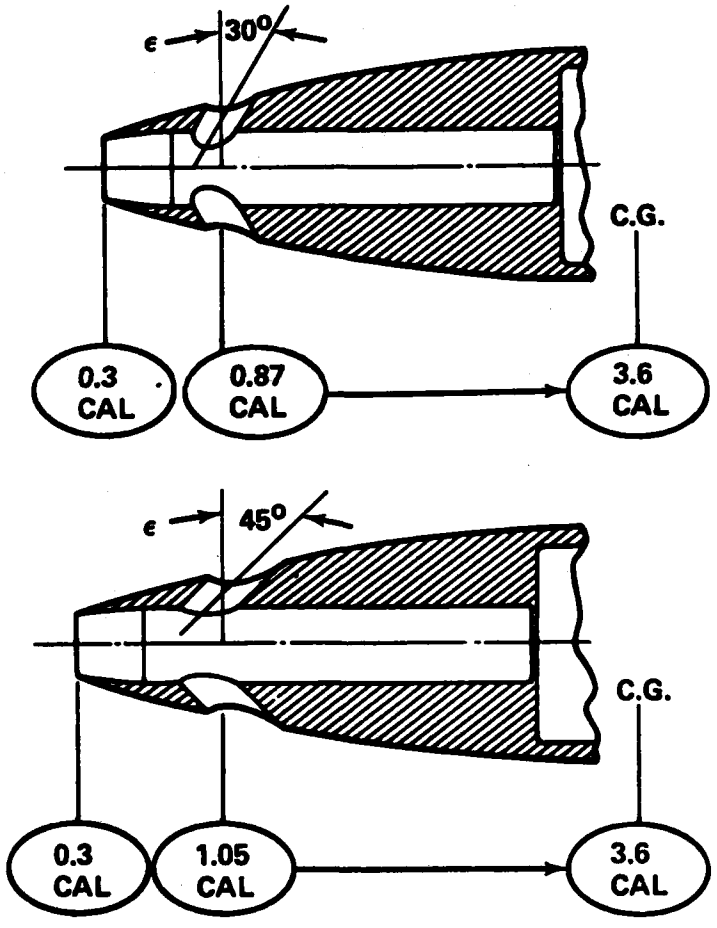
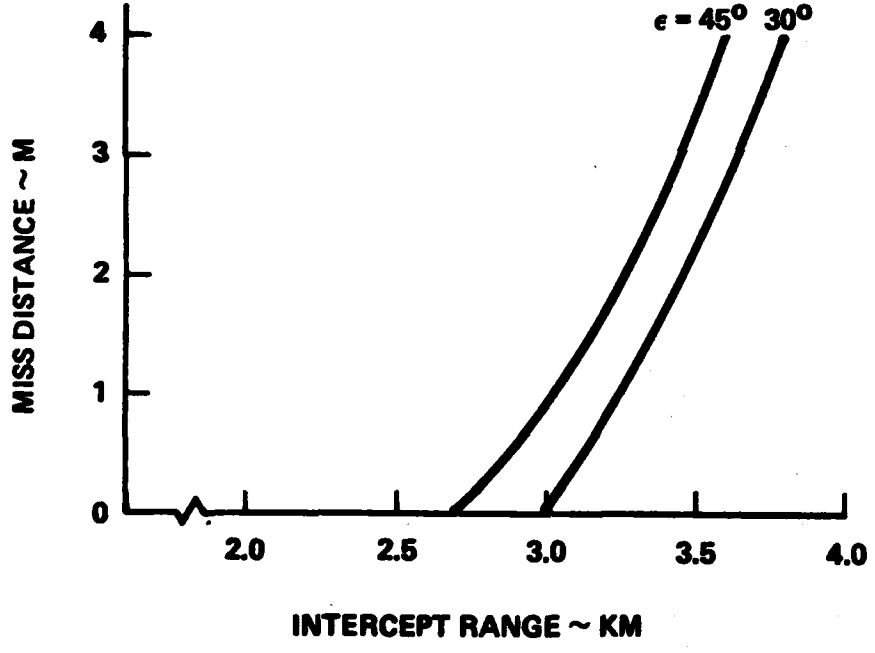


FIGURE 9: *SYSTEM ACCURACY*

HUTCHINGS

CLOSING TARGET

ALTITUDE – 30M
MUZZLE VEL – 1200 M/SEC
TARGET VEL – 200 KT
TARGET MANEUVER – 3.0 g's
AT 1.0 SEC TO INTERCEPT



CROSSING AND TAIL CHASE TARGETS

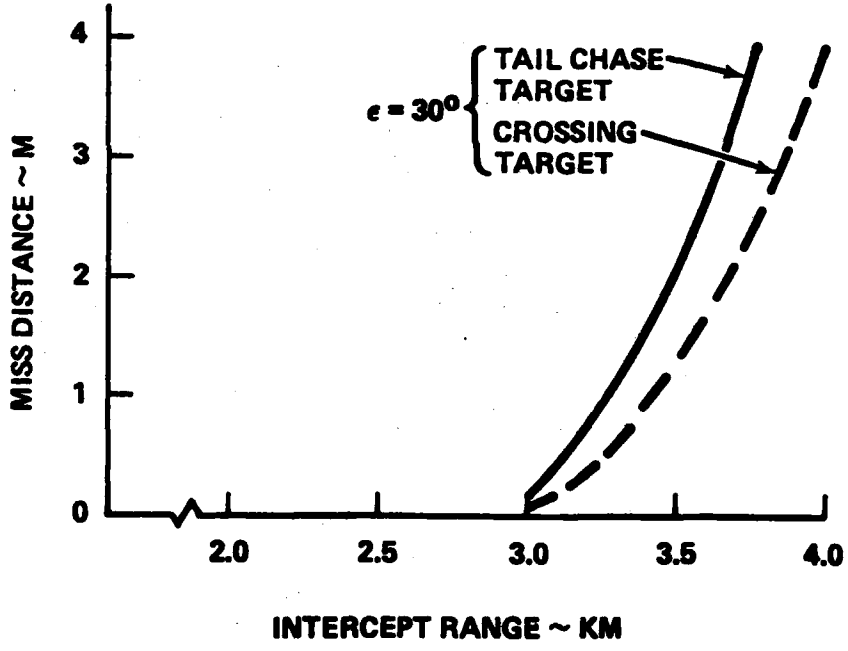
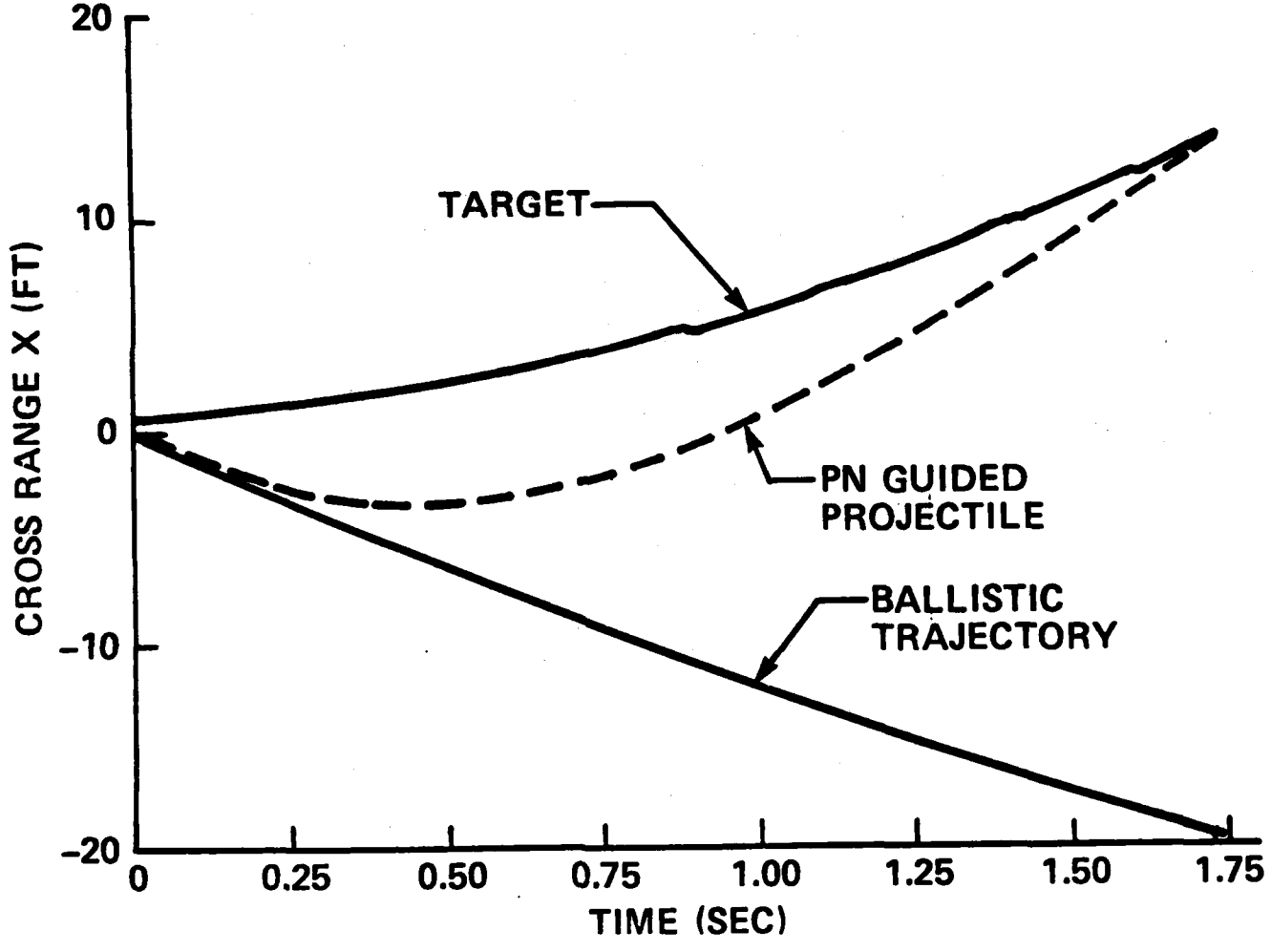


FIGURE 10: **FIN STABILIZED SCMP SIMULATION GAS JET REACTION CONTROL MECHANISM**

HELICOPTER TARGET; 5.0 MRAD AIM ERRORS



HEURISTIC INFORMATION PROCESSING AS IMPLEMENTED IN TARGET MOTION RESOLUTION ANALYSIS OF RADAR DATA

*JOHN N. HYNES, MR.
DAVID S. JIMAREZ, MR.
US ARMY WHITE SANDS MISSILE RANGE
WHITE SANDS MISSILE RANGE, NEW MEXICO 88002

INTRODUCTION

The US Army White Sands Missile Range (WSMR) has been conducting a research and development program for the production of precise trajectory parameters, center of gravity motions, and target events in support of a number of range customers. Some of the more important of these customers have included the Multiple Launch Rocket System (MLRS), Patriot, DIVADS, and Copperhead.

The process is called Target Motion Resolution (TMR) and is based on post-flight processing of the doppler-shifted returns from coherent C-band instrumentation radar systems. Since 1974, Mark Resources Incorporated, Santa Monica, California, the contractor for the development, has been extending the basic theory and producing procedures and software. They have also been implementing the developments on an interactive graphics minicomputer system operated by government data technicians at WSMR. Of all WSMR personnel associated with the development and implementation of TMR, data technicians have probably felt the greatest impact. Being a highly specialized technology, TMR requires both a large amount of processing and a fairly high level of knowledge in areas such as radar theory, digital signal processing, FM theory, spline functions, computer science, and interactive graphics minicomputer operation. Data technicians, by contrast, generally have a high school level education and are accustomed to submitting a prefab deck of cards, along with a radar tape, and awaiting results. Naturally, they have resisted the multistep, interactive graphics TMR reduction requiring highly technical interaction between steps.

In 1979, WSMR began a new effort aimed at improving the complex interface between the technician and the TMR reduction process. Studies performed included investigation of heuristic algorithms and knowledge

based system (KBS) technology by the Integrated Sciences Corporation (ISC), and human factors by the Army Research Institute (ARI), Human Factors Area. Results revealed that the system contained an inadequate level of "function" development for efficient use by data technicians. System response needed improvement, processing needed simplification, function keys were needed to facilitate interaction, and cumbersome manual procedures needed automation. In the months that followed, hardware was implemented and processing procedures simplified to meet the needs identified. Two particularly cumbersome manual processes were also significantly improved through an agreed upon heuristic approach. The methodology utilized involved developing a working model of the expert analyst's approach by observing him solve the particular TMR processing problems, generating the associated protocols used, and then placing this intelligence on the machine. The two new processes involved the development of specialized spectral trackers primarily for use in extracting spin frequency information from the MLRS. Basic measurement principles and the general TMR reduction process will be described before presentation of these trackers. Conclusions about their effectiveness compared to procedures used earlier will then be made.

MEASUREMENT PRINCIPLES

AN/FPS-16 and AN/MPS-36 C-band instrumentation radars, traditionally, make radar range and angle measurements by transmitting a pulse of energy to the target and measuring time and direction of arrival of the reradiated pulse. The coherent radars can also make average range rate measurements by tracking the relative doppler shift of the reradiated pulse with respect to the carrier signal reference (1). Unfortunately, these measurements produce only a rough target location and tell nothing about center of gravity motions. Further, the range rate trackers often prove unreliable due to problems such as glint, scintillation, and clutter. The problem of crude target location is inherent to the gross measuring stick used to make the measurement (i.e., a 1/4 microsecond pulse covers a range of about 75 meters), and precision has often been demonstrated to be only 5 to 10 meters. TMR processing, by contrast, utilizes the radar wavelength as its basic measuring stick, about 5 centimeters for a carrier frequency of about 5.8 GHz. With this type of precision, it is possible to measure the movement of one part of a target with respect to another to obtain center of gravity motions (2). In trade for the finer precision, however, the processing becomes complicated having to deal with modular phase measurements and the fact that each return is a combination of the returns from all scattering centers. The modular phase measurement causes a range ambiguity every 1/2 wavelength or about 2.5 centimeters, considering the two way path a pulse must travel. The range rate ambiguity, roughly 8 meters/second at 320 pulses per second, is obtained by dividing the range ambiguity by the time between pulses. Similarly computed, the acceleration ambiguity is about

2560 meters/second^2, which is generally large enough to cause little problem. Range and range rate ambiguities, however, require joint processing with the usual crude range data for their resolution. Fourier analysis is then used to map the average, combined returns from all scattering centers into their average, individual returns over a period of time. Fourier analysis of successive time segments would then produce a doppler history, or a history of the individual scattering center velocities, relative to the radar.

Trajectory compensation begins by spline approximating the usual unambiguous range data, and subtracting this information from the phase data. This removes the gross translational motion from the phase data, leaving only residual translational wander and center of gravity motions. Next, spline approximation and removal of the residual translational wander leaves only center of gravity motion. Precise approximation and removal of the center of gravity motion then gives the desired alignment to a particular scattering center, allowing for relative spin frequency measurement. Figures 1 through 4 show representative doppler history plots of raw, range compensated, residual translational wander compensated, and precisely aligned phase data. In interpreting each plot, a peak's position in a spectrum indicates the radial velocity of the scattering center it represents. The portion of the plot left of center represents increasing velocity of a scattering center moving away from the radar, and right of center represents increasing velocity of a scattering center moving toward the radar. Should a velocity ambiguity occur, a peak would alias, or run off the end of a line, and appear wrapped around on the other end.

BASE SPECTRAL PEAK TRACKING

Precise alignment of data, as seen in figure 4, requires alignment of figure 3 data to within ±20 Hz of zero doppler. Before the alignment process, the data would be low pass filtered in this interval to eliminate possible interference from spin frequency returns. Unfortunately, a ±20 Hz alignment is not always possible by a vernier spline adjustment of range compensated data, and requires an analyst to manually intervene. WSMR, ISC, and ARI collectively agreed that work under the heuristic approach mentioned should be initiated to automate this process (3). The problems that had to be overcome, as seen in figure 5 to 8, fall into four main classes: 1) Clutter, 2) desired return amplitude fading, 3) large spin frequency amplitudes, and 4) velocity ambiguities, seen as aliasing or wrap around in the doppler history plots. Figure 5 shows an example of clutter, considered here to be a broad class of undesired returns, appearing at any frequency, singularly or clustered, with amplitudes often larger than the desired return. Figure 6 shows how target orientation to the radar and destructive interference from other returns can cause relative fading or cancellation of the desired return.

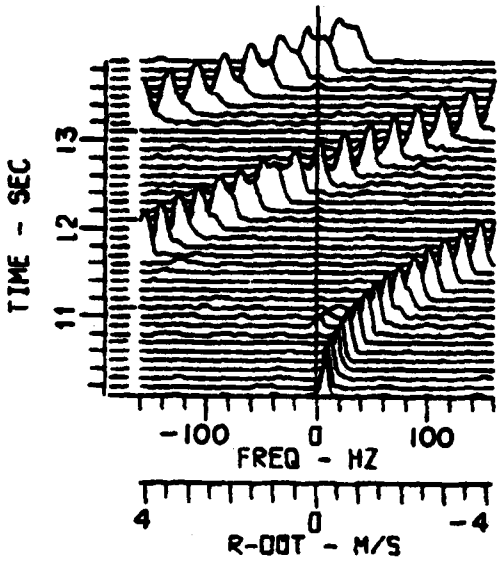


Figure 1. Raw data.

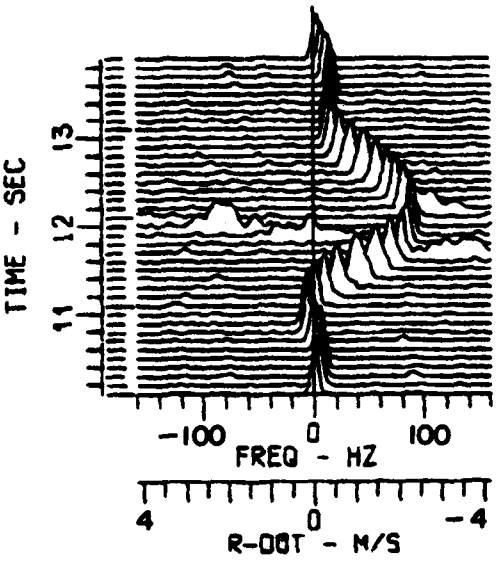


Figure 2. Range compensated data.

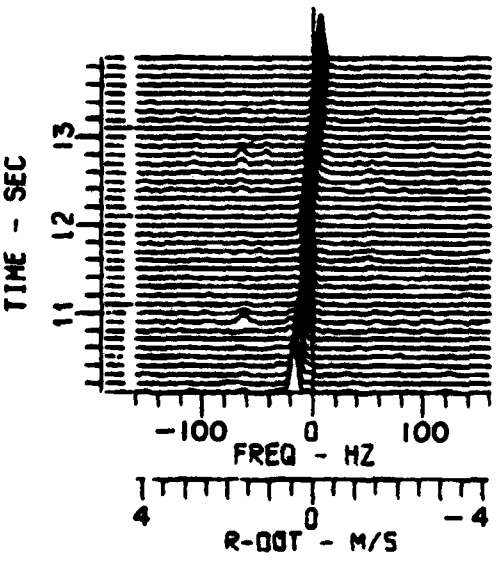


Figure 3. Translational wander compensated data.

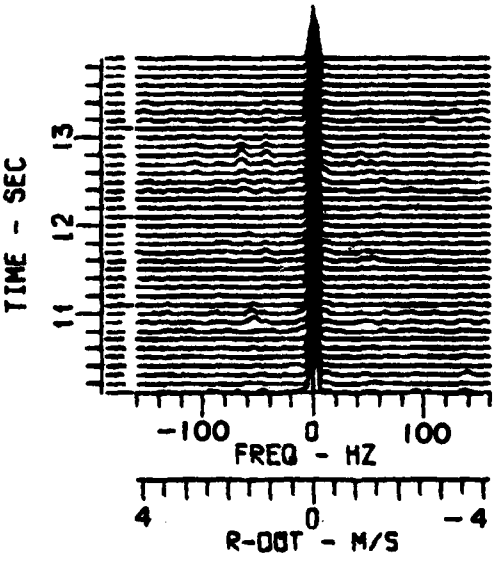


Figure 4. Precisely aligned data.

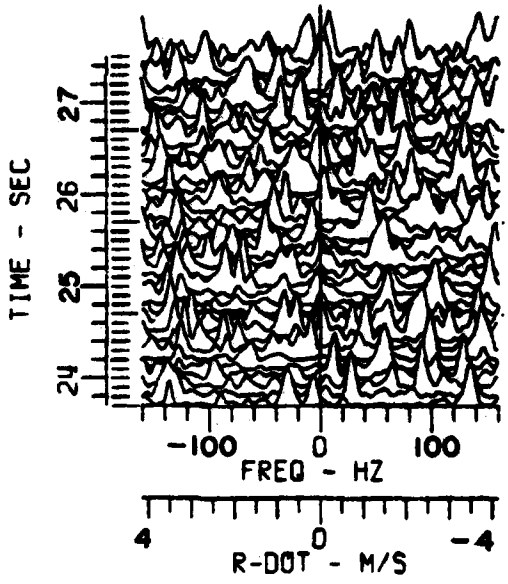


Figure 5. Clutter

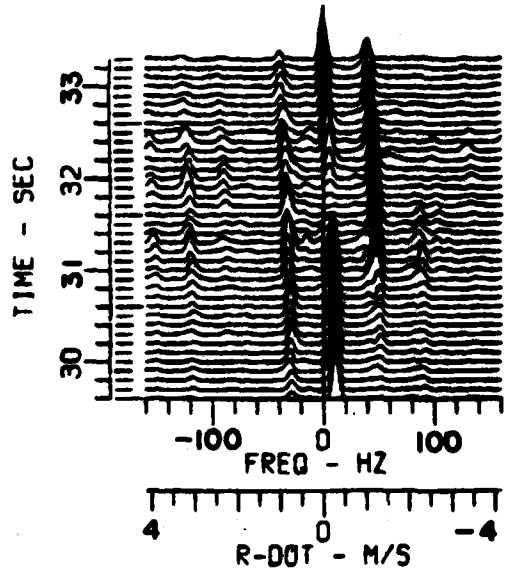


Figure 6. Desired return amplitude fading.

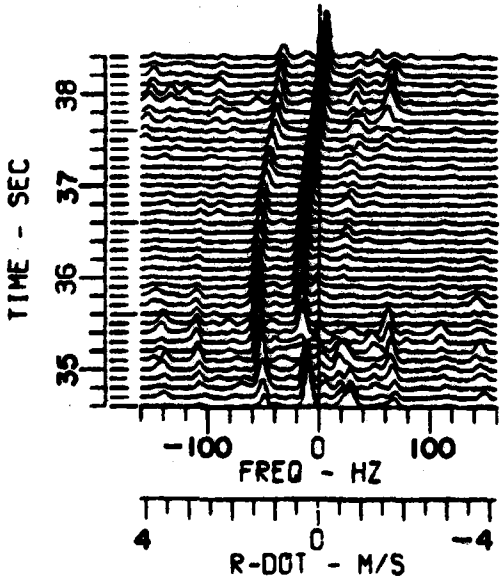


Figure 7. Large spin frequency amplitudes.

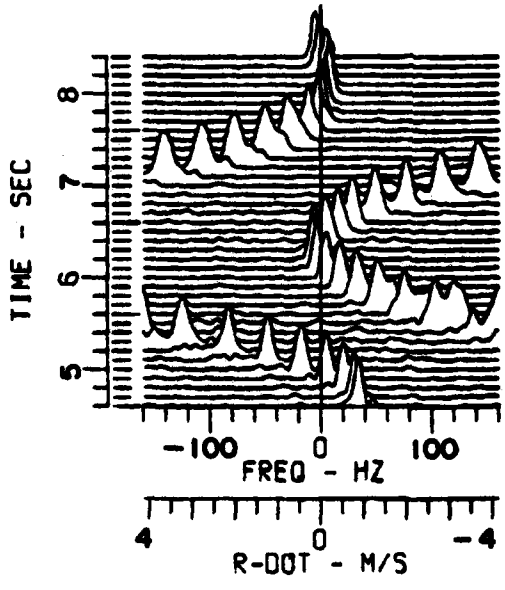


Figure 8. Velocity ambiguities.

Figure 7 shows the effect of large spin frequency returns even when the desired return does not fade. Figure 8 shows an example of velocity ambiguities, due primarily to an inadequate range compensation. Velocity ambiguities usually indicate large acceleration components, where the velocity of a particular scattering center, as indicated by its position in a doppler history plot, changes significantly as a function of time.

Early, batch mode efforts to align data with these problems were excessively long, generally taking as much as 2 days. In this process, two college level, trained analysts would manually identify and measure the doppler shifts of the desired returns and develop the parameters for spline adjustment before the computer was even used. An interactive graphics approach significantly improved the process, wherein an analyst could then align the data by interacting with plots of doppler history and the largest peak locations for the respective spectra. Figure 9 shows examples of such plots for an MLRS flight with a particularly bad range compensation, exhibiting all four classes of problems. Using the doppler history for reference, the analyst would attempt alignment by overlaying the plot of largest peaks with breakpoints of a linear spline approximation. Without the aid of the doppler history plot it would often be difficult to tell if the largest peaks were the desired ones. The option to correct for velocity ambiguity discontinuity in the plot of largest peaks was controlled by a heuristic. If a velocity ambiguity occurred, the distance between desired peaks before and after ambiguity generally differed by more than half the entire interval. Successive peak locations would then be adjusted by an entire interval to preserve continuity. Such a heuristic, however, often causes problems in that the wrong segment may be adjusted or, as seen in the following example, in that it is susceptible to large undesired returns.

In attempting to align the data of figure 9, an analyst would use the continuity option since velocity ambiguities occur at about 5.5 and 7 seconds. The doppler history would be recomputed and the plot of largest peaks redrawn, as seen in the right half of figure 10. Unfortunately, use of the option caused false ambiguity jumps at 21.3 and 25.7 seconds due to large clutter returns. A reasonable attempt to spline adjust, in spite of the false ambiguity jumps, would be indicated by the linear spline superimposed on the plot to the right in figure 10. The corresponding phase adjustment is seen in the doppler history to the left. Faced with the alignment being outside ±20 Hz, the analyst may regenerate the original plot in figure 9 and try again. The corresponding attempt, and the resulting failure, would probably be as indicated in figure 11. Again faced with inadequate alignment, the analyst may try reprocessing either or both cases, or even splitting both splines and combining the better halves of each. In any event, a proper alignment would require a trained analyst, as opposed to a data technician, and a considerable amount of time.

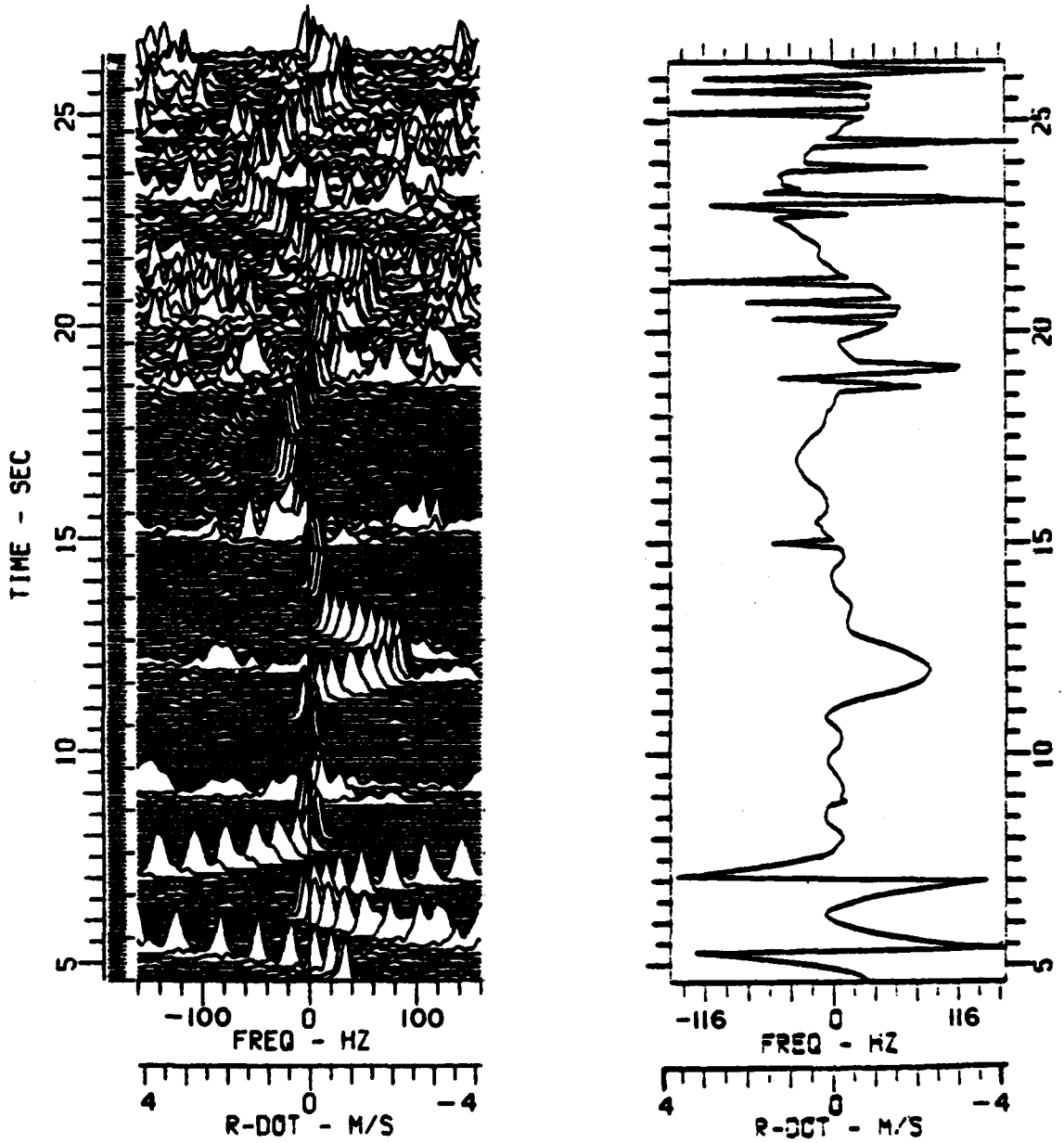


Figure 9. Unacceptable trajectory compensated data.

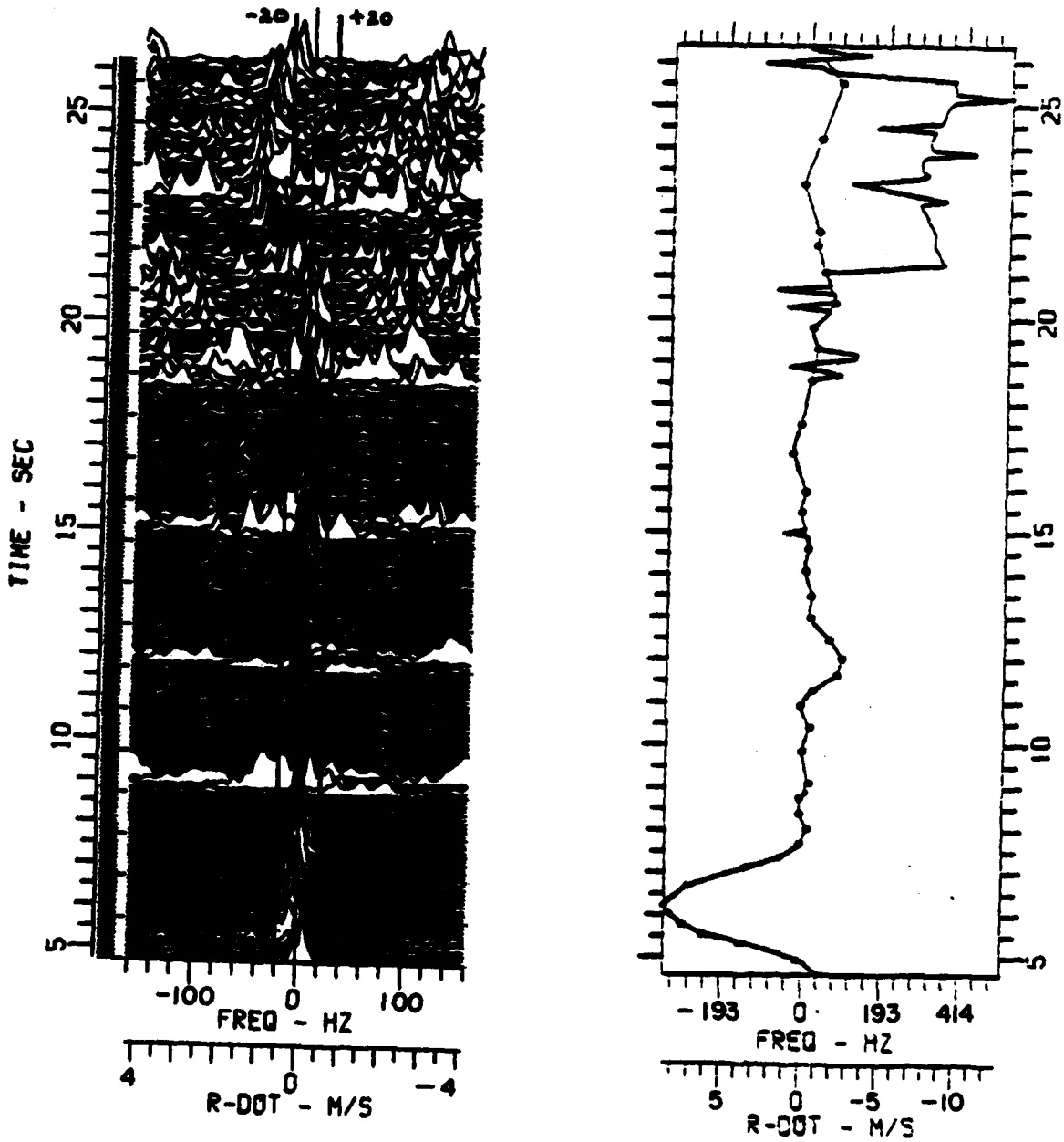


Figure 10. Example 1 of interactive graphics adjustment.

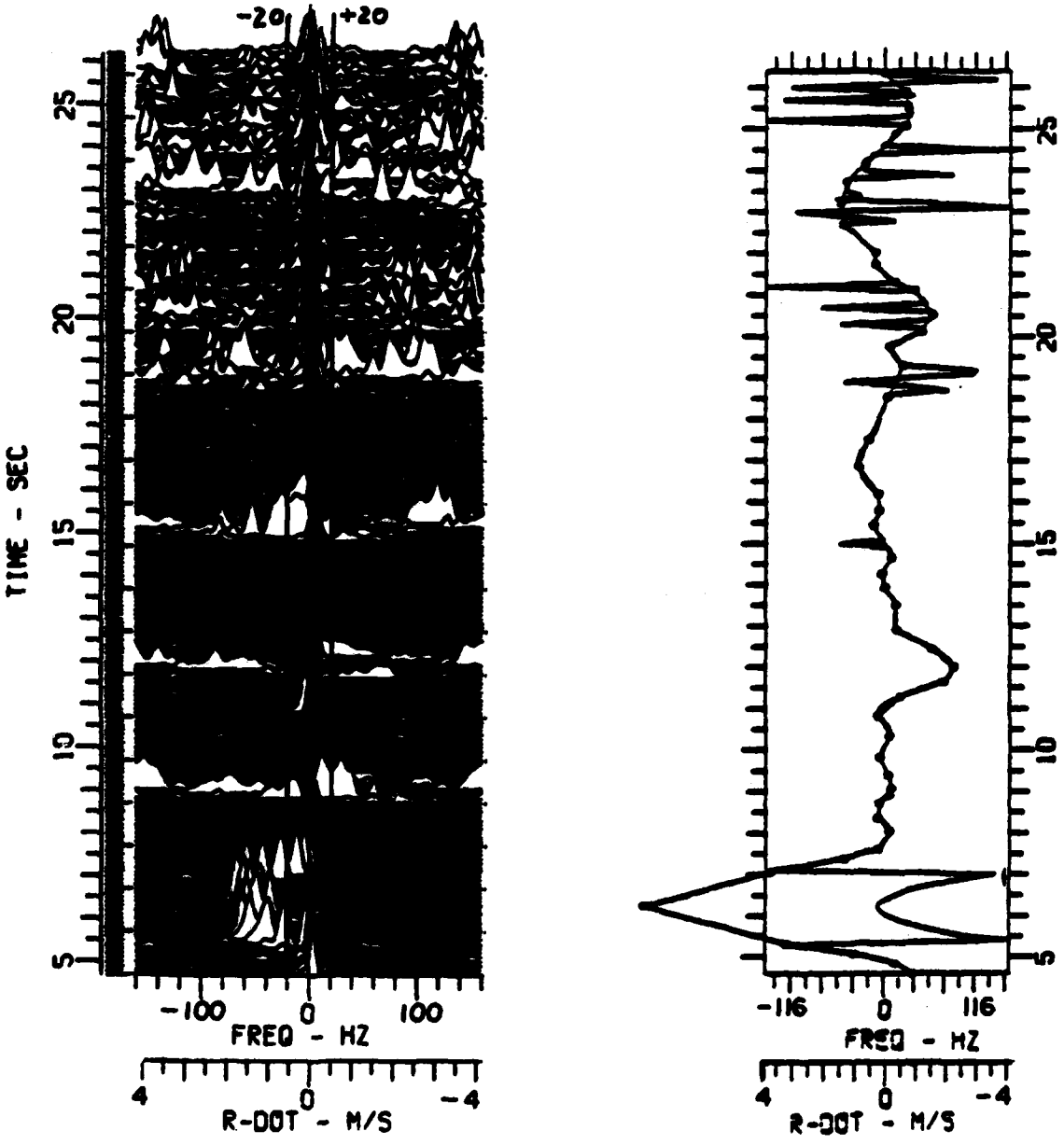


Figure 11. Example 2 of interactive graphics adjustment.

The methodology used to automate the process began with observing a well trained analyst perform the alignment. Generally, what an analyst would do is locate the desired peaks and follow their trend through the data in a windowed fashion. In observing this process, several important considerations were noted: 1) If peaks ran off one side they were sought on the other, 2) fading was controlled by following the general trend until peaks reappeared, 3) large noise returns in the area of the general trend would be identified and ignored, and 4) when large spin frequency returns were noted, care was taken to keep the search window between them. It was further noted that these considerations vary, being dependent upon technical knowledge in the areas mentioned as well as physics of the flight. The first step taken toward automating the process was reducing this variability by standardizing the parameters for the production of doppler histories. Now the problem became analogous to a state-space search (4), where the desire was systematically discovering the doppler history of a particular scatterer. With row elements acting as nodes, the analyst's heuristics could be used to form a generator function to calculate candidate successor nodes, and an evaluation function to determine their likelihoods of representing the desired scatterer's return. The search would resemble a tree search in that only the start node has no parent, and that every other node has only one parent and is a descendent of the start node. This paper will refer to goal nodes as peaks in the context of returns from the desired scatterer.

Several heuristics were used in the development of generator and evaluator functions to automate the analyst's tracking considerations. First, returns from the base of the MLRS should be used for alignment since, on the average, they are the strongest axially symmetric returns by virtue of radar orientation. Second, the short segments of peaks generally extrapolate linearly into close proximity of the next peak location. Third, most acceleration components occur near the beginning of data, where spin returns are not as noticeable. Fourth, large spin returns, especially in conjunction with faded base returns, primarily occurred later in the better radar tracked, ballistic portion of the flight. What developed was a function that generated candidate peaks by extrapolating the center of a window which widened or tightened, as acceleration components increased or decreased, respectively. Extrapolation was by a linear least squares fit of the last few peaks, where the slope of the fit controlled window width. Window parameters were also adjusted to fall within large spin frequency returns during the ballistic portion of the flight. Next, the evaluator function selected the peak as the element with the largest magnitude in the window. As the interference of large clutter returns within these windows was generally random, infrequent, and in small numbers, the analyst's ability to ignore them on the basis of general trend was implemented as a median smoothing function (5). As false ambiguity jumps were no longer a problem, since the window never spanned half an interval, all that remained was compensation of the

proper segment when a jump did occur. This was controlled by the heuristic that the ballistic portion of the flight should have the best radar track, and therefore the best range compensation. This was implemented by adjusting all previously tracked segments when velocity ambiguities occurred before mid-flight, and adjusting subsequently tracked segments when velocity ambiguities occurred after mid-flight.

Together, the heuristic functions proved quite successful, as is seen by the adjustment in figure 12 of the corresponding figure 9 data. The procedure now takes seconds rather than 1/2 to 1 hour, can be run by an ordinary data technician rather than a trained analyst, and maintains better than 95 percent reliability in routine reductions. Further, the analyst has the option, while using the procedure, to change parameters necessary to gain the required alignment on similar data with differing anomalies.

SPIN FREQUENCY TRACKING

Spin frequency tracking generally encounters the same classes of problems as base return tracking, but with different variations. Now, many peaks, or spectral nodes, can exist per line as multiples of four times the spin frequency are present. Further, these multiples can independently fade for long periods of time, or interfere with one another because of higher multiples wrapping back around onto lower multiples. Multiples of four are primarily due to the four-fold symmetry of the missile's fins, with most of the intermediate lines cancelling because of opposing doppler returns. In order to clearly display the multiples in a

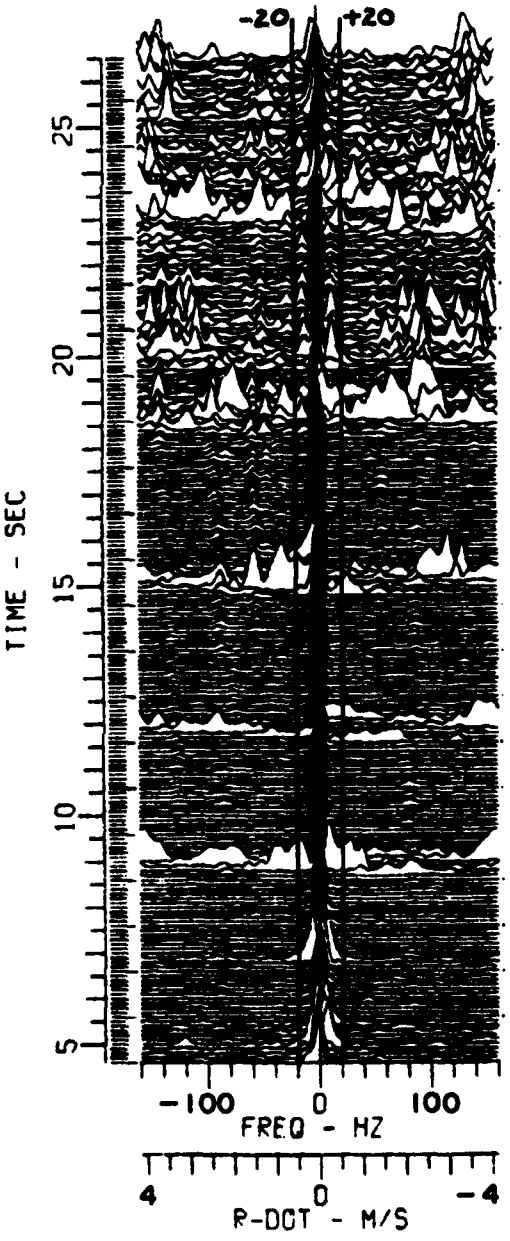


Figure 12. Main spectral peak tracker adjustment.

spin doppler plot, care must be taken to encompass several cycles of spin in the Fourier transform window so as to bring up the FM sidebands of the spin modulation (6). As descriptive examples of these problems are beyond a paper of this length, Figure 13 is instead displayed as a relatively clean example, simply to familiarize the reader with a spin doppler history. In manually processing this information, an analyst would identify and mark the unaliased spin traces present, as seen in figure 13, and then measure each of their doppler offsets, taking great care to avoid the effects of the problems discussed. Because these problems often require considerably more analysis, it was felt that a process patterned after a generalized KBS (7) would generally produce better results. The process is therefore broken down into a KBS's three primary elements: The interface, cognitive engine, and knowledge base.

The interface, which breaks down into external data, user, and expert interfaces, primarily functions as the user's two-way communication link to the expert knowledge modules and fact files which comprise the knowledge base. The external data interface is first used to create the fact files, where the current information to be processed is stored. The user interface then utilizes statistical and expert information, stored in the expert knowledge modules, to guide the user step by step through the process. These modules contain statistical averages from previously successful reductions, suggested defaults, and descriptions of what is happening at every stage of the reduction. Further, each user input is parsed and analyzed for content so

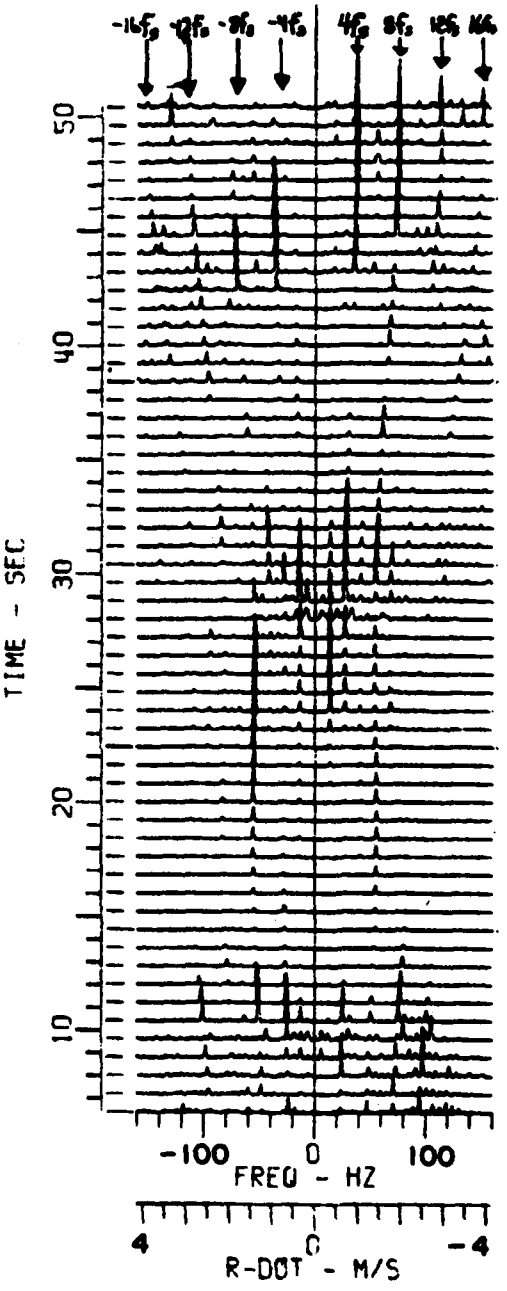


Figure 13. MLRS spin doppler history.

that part or all of this information is available even though a numeric input is requested. Statistics on the current reduction are also available to the user, and are used to update the permanent statistics if the user feels the process was successful. Finally, the expert user interface duplicates the user interface except for the capability of altering the expert knowledge modules.

The cognitive engine is the active processing element containing the generator and evaluator functions as well as the inference and reasoning algorithms that interact with the current problem state. Candidate spin returns are first generated as the largest returns in a given line, where the number generated is generally equal to the number of spin traces that do not alias. This function is based on the heuristic that most spin multiples present will probably constitute the larger returns, with most power contained in the lower, unaliased multiples. The evaluator function then works like the analyst, using a small search window to locate a spin return. The window center is first extrapolated to the expected areas of the largest multiples, on each side of the spectrum, and then to successive lower multiples if a peak is not found. Problems of clutter, aliasing, and finding no peaks are then handled interactively through the use of the inference and reasoning algorithms. For example, if no peak or more than one peak is found, tracking is halted and the user is made aware of the problem and the location. The user then has the option of allowing the process to make its best guess, based on the current state of the problem, or to enter an overriding location. Statistics are also compiled on the number and nature of such interruptions for the purpose of later advising the user in the event results are unsatisfactory. For instance, finding multiple peaks more often than not finding peaks might indicate that a smaller search window would have greater success. In the event a data technician has exhausted the resources of the process and is still unsatisfied, presentation of results and statistics to a trained analyst can usually gain an expeditious solution.

The new process contains many other heuristic based features used by trained analysts in processing spin information, including forward, backward, and segmented processing and extraneous frequency rejection. While it is too lengthy to discuss them all here, it is worth noting that collectively they have made significant advances in obtaining spin information for the MLRS. The process successfully tracks the spin frequency through fading, clutter, and the aliasing of higher spin multiples back onto lower spin multiples with better than 90 percent reliability in routine reductions. Primarily, this was achieved by placing the techniques of TMR analyst into the process and at the immediate disposal of the data technician. The analysts' capability to update the expert knowledge modules also reduces future reduction time by making data technicians even less dependent on their presence.

CONCLUSIONS

The two new processes for MLRS base and spin frequency tracking have made significant advances in the quality and speed with which data are produced for that project. Routine processing of these parameters can now be performed, using data technicians, in less than a tenth of the time formerly required by a trained analyst. While developed primarily for the MLRS, each process also has applicability to similar classes of targets, thus reducing later developmental efforts. In the future, developments in this area will continue, becoming more sophisticated and having wider applicability, helping to support the increasingly demanding test and evaluation needs of our new weapons systems.

BIBLIOGRAPHY

1. Ehling, Earnest H., Range Instrumentation, Prentice-Hall, Inc., 1967. Englewood Cliffs, New Jersey, pp. 244-281.

2. Nunn, Elwin C., "The US Army White Sands Missile Range Development of Target Motion Resolution", EASCON 1980 Record, IEEE Electronics and Aerospace Systems Convention, 1980, IEEE Publication 80CH1578-4AES, ISSN: 0531-6863.

3. Hynes, John N., "A Spectral Peak Tracker", Technical Report STEWS-ID-81-1, January 1981, STEWS-ID, White Sands Missile Range, New Mexico, DDC No. B057745L.

4. Jackson, Philip C., Introduction to Artificial Intelligence, Petrocelli Books, 1974, New York, New York, pp. 94-102.

5. Tukey, John W., Exploratory Data Analysis, Addison and Wessley, 1977, Reading, Massachusetts, pp. 205-264, 523-542.

6. Taub, Herbert and Schilling, Donald L., Principles of Communication Systems, McGraw-Hill, Inc., 1971, New York, New York, pp. 1-42, 113-155.

7. Rebane, George J., "Recommendations for Application of Knowledge Based Systems Technology to Target Motion Resolution", Report No. 308-1, June 1979, Integrated Sciences Corporation, Santa Monica, California.

Attenuation of High Intensity Reradiated Light by Photochromic Glass

Thomas V. Hynes, Ph.D.*
U.S. Army Materials and Mechanics Research Center
Watertown, Massachusetts 02172

One possible threat mechanism by which High Energy Laser Weapons may exploit the vulnerability of U.S. Army systems is through the production of large fluences of reradiated visible light when glass surfaces such as windscreens, vision blocks or lenses are struck by pulsed or cw laser radiation. At 10.6 meter laser wave lengths silicate glasses have very shallow absoprtion depths and, hence, heat to incandescence rapidly. Such blackbody surfaces at temperatures of 3000°C+ will reradiate strongly in the visible region of the spectrum. In addition, if pulsed laser conditions are such that air plasmas are formed on the target, blackbody temperatures of 17,000°C may be obtained.

In both the cw and the pulsed laser cases, the net flux obtained in the visible region will be a complicated function of fluence on target, target material, pulse shape and other parameters. Net dosage delivered to the retina and, hence, available for flashblinding will also depend on fluence rise times and blink reflex times.

This paper does not attempt to address flashblindness conditions; rather it focuses on the enhanced attenuation of the visible light in the optical materials behind the target surface.

Glasses which are prepared with certain metal salts and oxides may be photochromic, i.e., irradiation at specific absoption band wave lengths will electronically excite "color center" sites producing typically, holes and reduced metal atoms. Such activated sites will absorb radiation over a broad wavelength throughout the visible and near infrared regions of the spectrum. Visually, the materials darken under appropriate radiation. Excitation wavelengths are in the ultraviolet.

The detailed behavior of such photochromic glasses depends on both the composition and processing history of the glass. All of the revellant photochromic properties such as, activation band, absorption spectrum, rate of formation and bleaching rate vary with composition and processing. These may be chosen to emphasize one or more desired glass properties such as, optical density, darkening rate or bleaching time. Generally, the properties are not independent.

This paper reports a set of experiments to test the concept of the photochromic glass response as countermeasure against the visible reradiation obtained when transparent materials are irradiated by high energy lasers. In all of the experiments, a fused silica face plate (chose because of ultraviolet (uv) transparency) was the target and a laminate of four layers of the photochromic glass was place behind the fused silica. Attenuation of visible light traversing the photochromic glass was measured.

The photochromic glass used in the experiment was a conventional silver halide composition prepared by PPG and supplied for the experiments by Dr. Herbert D. Kivlighn of Grumman Aerospace Corporation. The glass had been one of several investigated in an unrelated program at Grumman and had been chosen for its ready availability and not for any particular photochromic property. In separate calibration experiments performed by Grumman excitation radiation in the 300 to 400 mm band, particularly between 320 and 360 mm caused photochromic darkening of the glass. Rate of darkening was dependent on excitation intensity.

AMMRC conducted laser irradiation experiments. These were performed on both cw and pulsed CO_2, 10.6 μm lasers. The pulsed laser employed was the "Humdinger" laser at Arco-Everett Research Laboratory in Everett, Massachusetts and the experiments were caried out under the auspices of the Joint Army, Navy, Air Force (JANAF) repetitively pulsed high energy laser test program. Vulnerability, effects and hardening experiments, with high energy lasers have been carried out under this program for several years. These photochromic glass experiments were carried out as part of the Army high energy laser hardening program under AMMRC auspices. The cw laser experiments were conducted at NARADCOM, Natick, Massachusetts where Natick Laboratories operates a 1.5 Kw cw CO_2 laser. AMMRC and NARADCOM have a joint program for the use of this laser.

Attenuation measurements were made by passing a HeNe laser beam through the photochromic glass during CO_2 laser irradiation and measuring the intensity of the transmitted light with a PIN diode. Care was taken to remain in the linear portion of the measurement system response Because of the restricted geometry of the pulsed laser, experimental arrangements only transverse measurements were made. That is while the laser was incident normal to the fused silica faceplate, a HeNe beam transversed the photochromic layer of interest in the plane perpendicular to the pulsed CO_2 beam. The HeNe beam was manipulated with a set of mirror reflectors and placed so it measured absorption along some convenient path through the first layer of photochromic glass.

Pulsed laser flux was focused so that plasma formation threshold was exceeded. The beam was collimated so that only 3 cm^2 of the fused silica faceplate were exposed to the air plasma. Laser pulse lengths and repetition rates were held fixed to values deemed appropriate from propagation analysis. An infrared detector monitored front face temperature. Both this device and a still camera confirmed plasma formation.

Front face damage of the fused silica faceplate confirmed the existence of high temperatures during the laser runs.

Due to program limitations, only two pulsed CO_2 laser shots were available. Timing sequence equipment failure caused the loss of motion picture coverage on one of the two shots. The high speed motion picture film of the other shot clearly shows photochromic response during the pulse train as shown in figure 1. The PIN diode output showed optical attenuation during each pulse train with more than 90% reduction of optical flux in less than one blink reflex time. Significant attenuation occurred during the first pulse.

The photographs shown in figure 1 are taken from the high speed motion picture film of the second laser shot. The darkening of the glass as seen in the film, qualitatively confirms the measurements obtained from the PIN diode attenuation experiment.

The conclusion to be drawn based on the pulsed CO_2 experiments is that photochromic glasses are a candidate passive countermeasure material for situations in which laser induced air plasmas constitute a threat to visual or optical systems.

Cw laser experiments were conducted on the same materials at U.S. Army Natick Laboratories (NARADCOM) using the 1.5 K Watt CO_2 laser system sponsored by AMMRC.

A large number of tests were performed at various fluence levels and for various irradiation times. Since the laser output could be focused it was possible to make measurements at various heating rates for the fused silica surface. Focusing enabled high temperatures to be reached. Silica deposits indicated that surface temperatures in excess of 2600°K were achieved. It is not unlikely that the vapor was heated to temperatures of 3000°K and above.

At a temperature of 3000°K, blackbody flux is between 1 and 10 $mWatt/cm^2$ the photochromic excitation band. These fluences are sufficient to trigger photochromic response under normal circumstances. Measurements made of the uv flux radiated from the front surface of the laser heated fused silica demonstrated that sufficient intensity was available to trigger the response.

HeNe measurements made in the same fashion as those described earlier for the pulsed laser case determined visible light attenuation during and after irradiation by the CO_2 laser beam. Because there was less experimental constraint in the case of these cw laser experiments, it was possible to measure attenuation of transmitted light in directions both transverse to and parallel to the CO_2 laser beam direction. The measurements were consistent with one another; there was not unexpected enhancement of attenuation that was not caused by geometry or non-uniform irradiation. The magnitudes of attenuation measured in both directions were likewise consistent.

Because of the limited power output of the cw laser available, it was not possible to irradiate the fused silica at fluences sufficient

to cause incandescence in times less than the blink reflex time of the human eye. However, the experiment did demonstrate that when uncomfortable levels of visible reradiated light were achieved sufficient photochromic response was obtained to cause substantial attenuation of the visible light. In each case measured, the attenuation of the transmit-visible flux increased at qualitatively the same rate as the visible flex transmitted through the photochromic glass remained approximately constant.

HeNe measurements were made in each of the several layers of photochromic glass to measure the diminuation of photochromic attenuation as excitation flux dropped. Results were consistent with a model which assumed the excitation radiation to fall off exponentially because of absorption in depth.

Bleaching rates as a function of temperature were measured in the cw laser experiments. These indicated that bleaching times were sufficiently short that full transmission could essentially be achieved in less than 3 seconds after irradiation had ceased. Observations during the pulsed laser experiments indicated that bleaching times were likewise short.

The overall conclusion based on both sets of experiments is that photochromic glass offers the possibility of a countermeasure for laser tactics which rely on the blinding of sensors and observers due to visual optical flux generated by heated surfaces or air plasmas.

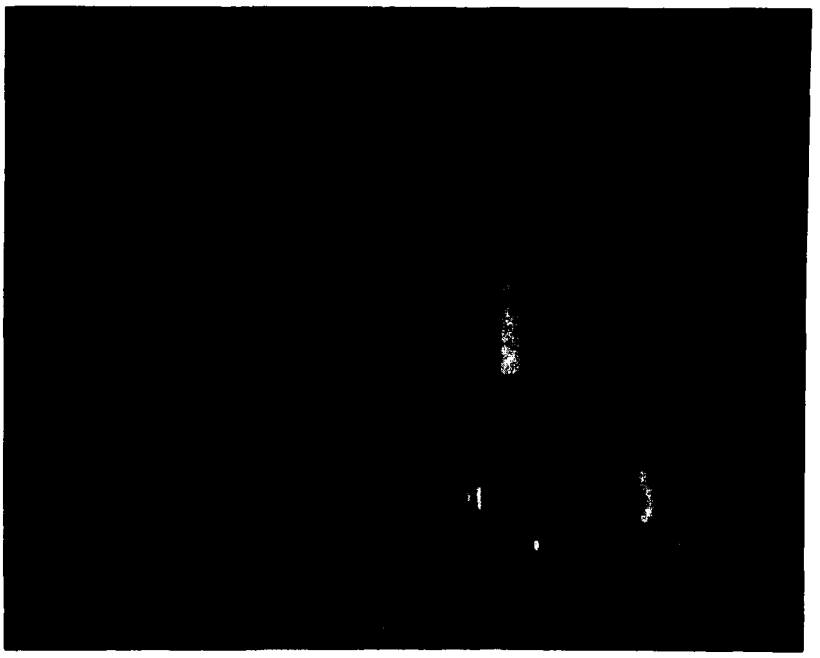


Figure 1a. Photochromic Glass Before Pulsed Laser Irradiation

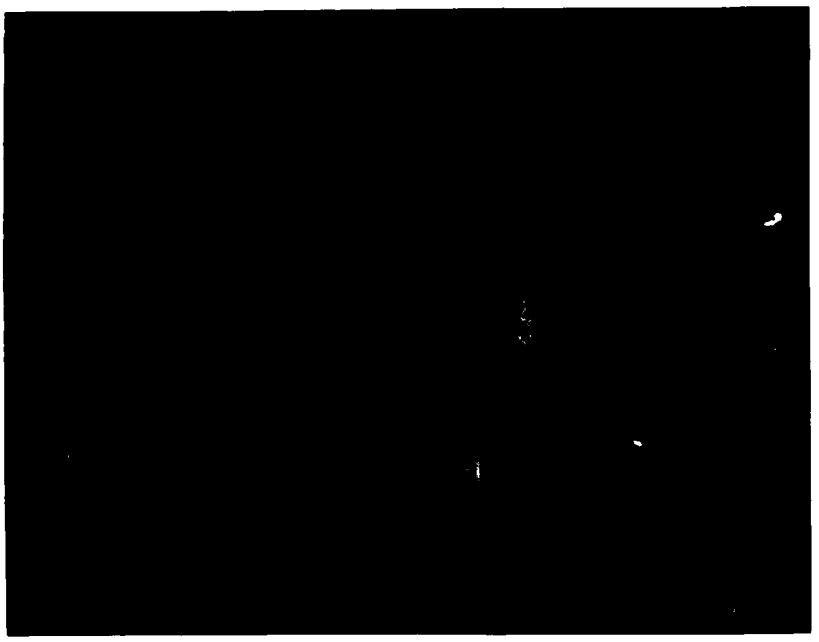


Figure 1b. Photochromic Glass After Pulsed Laser Irradiation

UTILIZATION OF QUANTUM DISTRIBUTION FUNCTIONS FOR ULTRA-SUBMICRON DEVICE TRANSPORT (U)

G.J. IAFRATE, Ph.D.
U.S. ARMY ELECTRONICS TECHNOLOGY AND DEVICES LABORATORY (ERADCOM)
FORT MONMOUTH, NEW JERSEY 07703

INTRODUCTION

As semiconductor technology continues to pursue the scaling-down of IC device dimensions into the submicron (less than ten thousand Angstroms) regime, many novel and interesting questions will emerge concerning the physics of charged particles in semiconductors. One of the more important topics to be considered is that of the appropriate transport (1) picture to be used for a given spatial and temporal regime. Moreover, from the point of view of device physics, it is most desirable to have a microscopic description of semiconductor transport which is computationally manageable or at least amenable to phenomenological treatment so that its properties can be meaningfully incorporated into device simulations. In this paper attention is focused on a useful transport methodology for the ultra-submicron regime. At present, military electronic devices with transit lengths in the ultra-submicron region are coming to fruition due to the advent of MBE (2) processing so that the need for an appropriate description of transport in this regime is indeed imperative.

Semiconductor transport in the ultra-submicron regime approaches the category of quantum transport. This is suggested by the fact that within the effective mass approximation the thermal deBroglie wavelength for electrons in semiconductors (see Fig. 1) is of the order of ultra-submicron dimensions. Whereas classical transport physics is based on the concept of a probability distribution function which is defined over the phase space of the system, in the quantum formulation of transport physics, the concept of a phase space distribution function is not possible inasmuch as the non-commutation of the position and momentum operators (the Heisenberg uncertainty principle) precludes the precise specification of a point in phase space. However, within the matrix formulation of quantum mechanics, it is possible to construct a "probability" density matrix which is often interpreted as the analog of the classical distribution function.

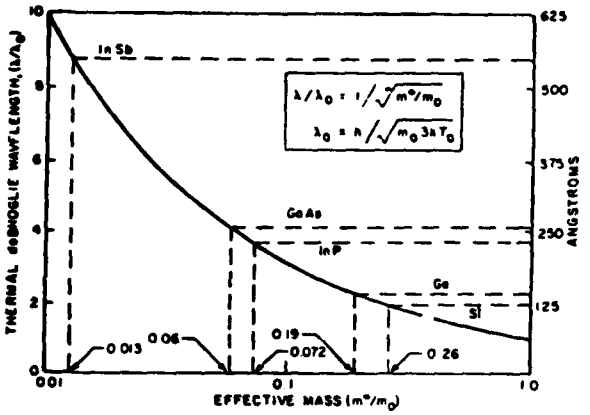


Figure 1. Thermal deBroglie wavelength vs. effective mass

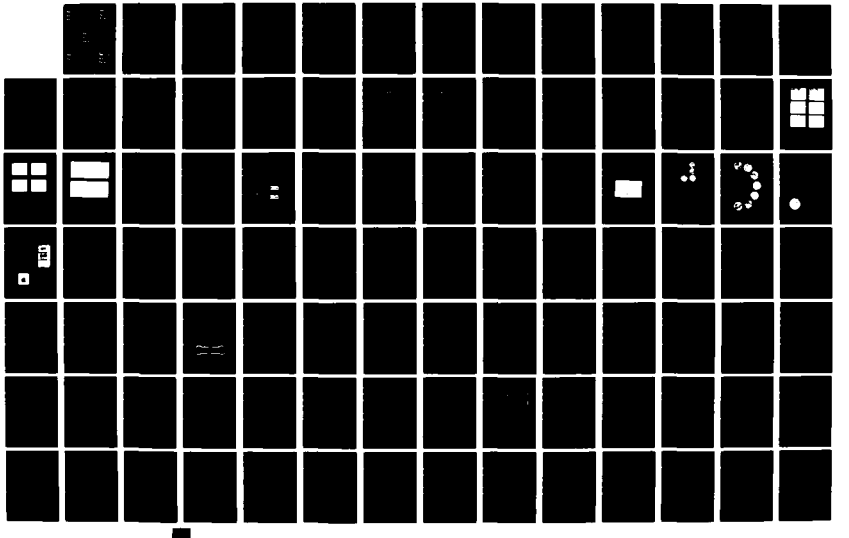
There is yet another approach to the formulation of quantum transport based on the concept of the Wigner distribution function (WDF) (3). This formalism is particularly attractive for use in ultra-submicron device transport in that it contains all of the quantum mechanical information about the state of the system yet has elements of the classical picture implicitly built in. Thus, the intent of this study is to explore the potential usefulness of the WDF as well as other possible quantum distribution functions for describing quantum (ultra-submicron) device transport in semiconductors. To this end we first review the salient features of the WDF and then discuss a new result, the derivation of the first three quantum moment equations using the WDF. It is shown that the moment equations contain quantum corrections to the classical moment equations; these quantum terms are non-negligible when the transit lengths are of the order of the carrier deBroglie wavelength.

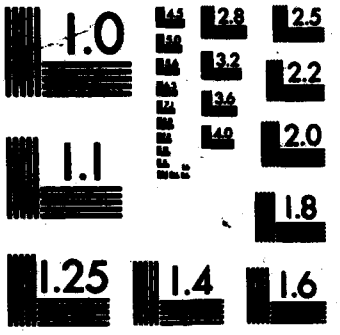
THE WIGNER DISTRIBUTION FUNCTION

The Wigner distribution function (3) is generally defined in terms of all the generalized coordinates and momenta of the system as

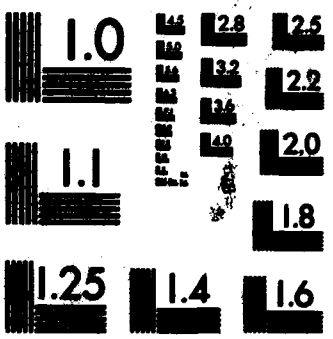
$$P_W(x_1 \ldots x_n, p_1 \ldots p_n) = \frac{1}{(2\pi\hbar)^n} \int_{-\infty}^{\infty} dy_1 \ldots dy_n \, \psi^*(x_1 + \frac{y_1}{2}, \ldots x_n + \frac{y_n}{2}) \psi(x_1 - \frac{y_1}{2}, \ldots x_n - \frac{y_2}{2}) \, e^{i(p_1 y_1 + \ldots + p_n y_n)/\hbar} . \quad (1)$$

AD-A120 812 PROCEEDINGS OF THE 1982 ARMY SCIENCE CONFERENCE HELD AT 3/5
THE UNITED STATES..(U) DEPUTY CHIEF OF STAFF FOR
RESEARCH DEVELOPMENT AND ACQUISITIO.. 18 JUN 82
UNCLASSIFIED F/G 5/2 NL

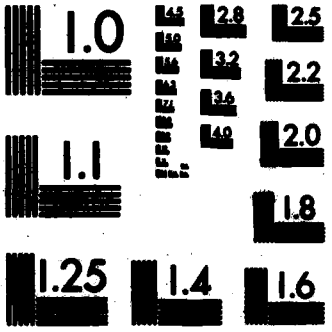




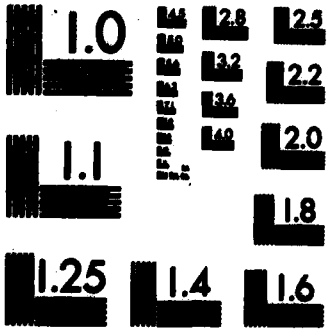
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A



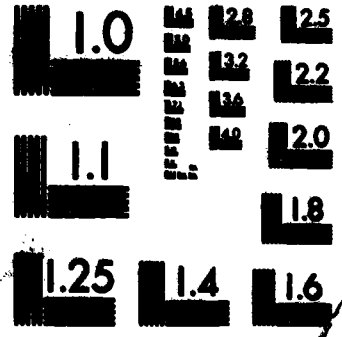
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

However, for simplicity in this paper, we discuss the properties of a single coordinate and momentum WDF:

$$P_W(x,p) = \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} dy\ \Psi^*\left(x + \frac{y}{2}\right)\Psi\left(x - \frac{y}{2}\right) e^{ipy/\hbar} \qquad (2)$$

where $\Psi(x)$ represents the state of the system in the coordinate representation. Although we will be treating the WDF for the special case of pure states, the adaptation of this formalism to include mixed states is accomplished through the generalization

$$P_W(x,p) = \sum_n P_n\left[\frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} dy\ \Psi_n^*\left(x + \frac{y}{2}\right)\ \Psi_n\left(x - \frac{y}{2}\right)\ e^{ipy/\hbar}\right]$$

where P_n is the probability to be in state "n". (For example, for a system in contact with a heat bath at constant temperature, $P_n = e^{-E_n/kT}$.)

The distribution function of Eq. (2) has interesting properties in that the integration of this function over all moments leads to the probability density in real space; conversely, the integration of this function over all coordinates leads to the probability density in momentum space. In mathematical terms,

$$\int_{-\infty}^{\infty} P_W(x,p)\ dp = \Psi^*(x)\Psi(x) \qquad (3a)$$

and

$$\int_{-\infty}^{\infty} P_W(x,p)\ dx = \phi^*(p)\phi(p), \qquad (3b)$$

where

$$\phi(p) = (2\pi\hbar)^{-1/2} \int_{-\infty}^{\infty} e^{-ipx/\hbar}\ \Psi(x)dx.$$

It follows immediately from Eq. (3) that, for an observable $W(x,p)$ which is either a function of momentum operator alone or of position operator alone, or any additive combination therein, the expectation value of the observable is given by

$$\langle W \rangle = \iint WP_W(x,p)\ dx\ dp, \qquad (4)$$

which is analogous to the classical expression for the average value. Herein lies the interesting aspect of the Wigner distribution function; the

result of Eq. (4) suggests that it is possible to transfer many of the results of classical transport theory into quantum transport theory by simply replacing the classical distribution function by the Wigner distribution function. However, unlike the density matrix, the Wigner distribution function itself cannot be viewed as the quantum analog of the classical distribution function since it is generally not positive definite and non-unique [$P_W(x,p)$ of Eq. (2) is not the only bilinear expression (3-5) in Ψ that satisfies Eq. (3)].

Further resemblance of the Wigner distribution function to the classical distribution function is apparent by examining the equation of time evolution for $P_W(x,p)$. Upon assuming that $\Psi(x)$ in Eq. (2) satisfies the Schrodinger equation for a system with hamiltonian $H = p^2/2m + V(x)$, it can be readily shown that $P_W(x,p)$ satisfies the equation

$$\partial P_W/\partial t + (p/m)\partial P_W/\partial x + \theta \cdot P_W = 0, \tag{5}$$

where

$$\theta \cdot P_W = -\frac{2}{\hbar}\sum_{n=0}^{\infty}(-1)^n \frac{(\hbar/2)^{2n+1}}{(2n+1)!}\frac{\partial^{2n+1}V(x)}{\partial x^{2n+1}}\frac{\partial^{2n+1}P_W(x,p)}{\partial p^{2n+1}} \tag{6}$$

It is evident that in the limit $\hbar \to 0$, $\theta \cdot P_W$ in Eq. (6) becomes

$$\theta \cdot P_W = -(\partial V/\partial x)(\partial P_W/\partial p) \tag{7}$$

so that Eq. (5) reduces to the classical collisionless Boltzmann equation.

The Wigner distribution function defined in Eq. (2) is derivable (8) from the Fourier inversion of the expectation value (with respect to state $\Psi(x)$) of the operator $e^{i(r\hat{p}+\theta\hat{x})}$ (here, [$\hat{x}$ and $\hat{p}$] satisfy the commutation relation $[\hat{x},\hat{p}] = i\hbar$). As such,

$$P_W(x,p) = \frac{1}{4\pi^2}\iint C_W(\tau,\theta)e^{-i(\tau p+\theta x)}d\tau\, d\theta, \tag{8a}$$

where

$$C_W(\tau,\theta) = \int \Psi^*(x)e^{i(r\hat{p}+\theta\hat{x})}\Psi(x)\, dx, \tag{8b}$$

and the interval of integration is $(-\infty,\infty)$ unless otherwise specified. In order to show that the right-hand side of Eq. (8a) is indeed the Wigner distribution function as defined in Eq. (2), note, from the Baker-Hausdorff

theorem (9), that $e^{i(\tau\hat{p}+\theta\hat{x})}$ can be written

$$e^{i(\tau\hat{p}+\theta\hat{x})} = e^{i\tau\hat{p}/2}e^{i\theta\hat{x}}e^{i\tau\hat{p}/2}, \tag{9}$$

in which case $C_W(\tau,\theta)$ of Eq. (8b) becomes

$$C_W(\tau,\theta) = \int_{-\infty}^{\infty} [e^{-i\tau\hat{p}/2}\psi(x)]^* e^{i\theta\hat{x}} [e^{-i\tau\hat{p}/2}\psi(x)]\, dx \tag{10}$$

which further reduces to

$$C_W(\tau,\theta) = \int_{-\infty}^{\infty} \psi^*(x - \tfrac{1}{2}\tau h)e^{i\theta x}\psi(x + \tfrac{1}{2}\tau h)\, dx. \tag{11}$$

Then, by inserting $C_W(\tau,\theta)$ of Eq. (11) into the right-hand side of Eq. (8a), integrating over the variable θ by using the relation

$$\int_{-\infty}^{\infty} e^{i\theta(x'-x'')}\, d\theta = 2\pi\delta(x'-x''),$$

and letting $\tau = y/\hbar$, the desired result is obtained.

The method outlined above to arrive at the Wigner distribution function is based on the notion of a characteristic function. The characteristic function of an observable, A, with respect to state $|\psi\rangle$ (here, the Dirac notation is utilized for purposes of generality) is defined as

$$C_A(\xi) = \langle\psi|e^{i\xi\hat{A}}|\psi\rangle, \tag{12}$$

where ξ is a real parameter. Assuming A to possess an eigenvalue spectrum given by $\hat{A}|A'\rangle = A'|A'\rangle$, $C_A(\xi)$ can be evaluated in the A'-representation as

$$C_A(\xi) = \int dA' \int dA'' \langle\psi|A'\rangle\langle A'|e^{i\xi\hat{A}}|A''\rangle\langle A''|\psi\rangle. \tag{13}$$

Since $\langle A'|e^{i\xi\hat{A}}|A''\rangle = e^{i\xi A'}\delta(A' - A'')$ in the A'-representation, $C_A(\xi)$ in Eq. (13) reduces to

$$C_A(\xi) = \int dA' e^{i\xi A'}|\psi_{A'}|^2, \tag{14}$$

where $|\psi_{A'}|^2 = |\langle A'|\psi\rangle|^2 \equiv P(A')$, the probability distribution function for

measuring A' while in state $|\Psi\rangle$. Hence, the characteristic function for $\hat{A}$ is the Fourier transform of the probability distribution function P(A'). Subsequent inversion of Eq. (14) above leads to

$$P(A') = \frac{1}{2\pi} \int C_A(\xi) e^{-i\xi A'} \, d\xi \,. \tag{15}$$

The Wigner distribution function was derived by taking the Fourier transform of the characteristic function for $e^{i(\tau\hat{p}+\theta\hat{x})}$. In view of the connection between the probability distribution function and the characteristic function for a given observable, this approach seems to be a natural way of obtaining a distribution function for momentum and position. Unfortunately, the noncommutative nature of the two observables destroys the convenient probability interpretation of the characteristic function implicit in Eq. (15).

In order to demonstrate this point, assume the characteristic function for two noncommuting observables, $\hat{A}$ and $\hat{B}$, to be

$$C_{AB}(\xi_1\xi_2) = \langle \Psi | e^{i(\xi_1\hat{A}+\xi_2\hat{B})} | \Psi \rangle \,. \tag{16}$$

Observables $\hat{A}$ and $\hat{B}$ are assumed to have eigenvalue spectra

$$\hat{A}|A'\rangle = A'|A'\rangle \,, \quad \hat{B}|B'\rangle = B'|B'\rangle \,. \tag{17}$$

and are chosen so that $[\hat{A}, [\hat{A},\hat{B}]] = [\hat{B}, [\hat{A},\hat{B}]] = 0$. This assumption is imposed so that the identity

$$e^{i(\xi_1\hat{A}+\xi_2\hat{B})} = e^{i\xi_1\hat{A}} \, e^{i\xi_2\hat{B}} \, e^{-\xi_1\xi_2[\hat{A},\hat{B}]/2} \tag{18}$$

may be used.

Inserting Eq. (18) into Eq. (16) while obtaining the matrix elements of $e^{i\xi_1 A}$ in the A-representation and $e^{i\xi_2 B}$ in the B-representation results in

$$C_{AB}(\xi_1\xi_2) = e^{-\xi_1\xi_2[\hat{A},\hat{B}]/2} \int dA' \int dB' \, e^{i(\xi_1 A'+\xi_2 B')} \langle \Psi|A'\rangle\langle A'|B'\rangle\langle B'|\Psi\rangle \tag{19}$$

In Eq. (19), it is assumed that [A,B] is a c-number independent of the eigenvalues A' and B'. We define F(A',B'), the generalized Wigner distribution function, to be

$$F(A',B') = \langle\Psi|A'\rangle\langle A'|B'\rangle\langle B'|\Psi\rangle , \tag{20}$$

so that

$$F(A',B') = \frac{1}{(2\pi)^2}\int d\xi_1 \int d\xi_2 \; e^{\xi_1\xi_2[\hat{A},\hat{B}]/2}\; C_{AB}(\xi_1,\xi_2)e^{-i(\xi_1 A'+\xi_2 B')} . \tag{21}$$

It is evident from Eqs. (20,21) that

$$\int F(A',B')\, dA' \equiv |\langle B'|\Psi\rangle|^2 = \frac{1}{2\pi}\int d\xi_2 \; C_{AB}(0,\xi_2)e^{-i\xi_2 B'} \tag{22a}$$

and

$$\int F(A',B')\, dB' \equiv |\langle A'|\Psi\rangle|^2 = \frac{1}{2\pi}\int d\xi_1 \; C_{AB}(\xi_1,0)e^{-i\xi_1 A'} . \tag{22b}$$

Thus, Eq. (21) establishes the relationship between the characteristic function for two arbitrary noncommuting observables and the generalized Wigner distribution function. The generalized distribution function has the essential properties of the conventional Wigner function in that an integration of the generalized function over the eigenvalue spectrum of one observable leads to the probability density in the canonically conjugate observable [Eq. (22)].

There is no simple probability interpretation of $F(A',B')$ in Eqs. (20, 21) because of the necessary overlap between the states of the noncommuting observables. However, if $\hat{A}$ and $\hat{B}$ are made to commute so that $|A'\rangle$ and $|B'\rangle$ are a common set of eigenvectors, then $F(A',B')$ reduces to the probability distribution function for $\hat{A}$ and $\hat{B}$.

Finally, it is noted that the conventional Wigner distribution function for observables $\hat{A}$ and $\hat{B}$ is obtained from

$$P_W(A',B') = \frac{1}{(2\pi)^2}\int d\xi_1 \int d\xi_2 \; C_{AB}(\xi_1,\xi_2)e^{-i(\xi_1 A'+\xi_2 B')} , \tag{23}$$

with $C_{AB}(\xi_1\xi_2)$ defined in Eq. (16), whereas the alternative distribution function, $F(A',B')$, introduced in Eqs. (20,21) differs from the Wigner function due to the presence of the phase factor $e^{\xi_1\xi_2[\hat{A},\hat{B}]/2}$ in the integrand of Eq. (21). For $\hat{A} = \hat{x}$ and $\hat{B} = \hat{p}$, $P_W(x,p)$ in Eq. (23) reduces to the Wigner function of Eq. (2), whereas $F(x,p)$ defined from Eq. (20) becomes

$$F(x,p) = \frac{1}{2\pi\hbar}\int_{-\infty}^{\infty} dy\, \Psi^*(x)\Psi(x-y)e^{ipy/\hbar} \equiv (2\pi\hbar)^{-1/2}\Psi^*(x)e^{ipx/\hbar}\,\phi(p), \tag{24}$$

where $\phi(p)$ is defined in Eq. (3b). It is evident that there is a family of functions which are bilinear in Ψ yet satisfy the sum rules of Eqs. (3a,b).

There are some interesting questions to be resolved concerning the uniqueness and positive definiteness of Wigner-type quantum distribution functions. Nevertheless, these distribution functions serve a useful purpose for calculating quantum mechanical observables in transport (7) studies and numerous solid-state (10,11) problems.

MOMENT EQUATIONS

In this section, we derive the first three moments of a "Wigner-Boltzmann"-like transport equation,

$$\frac{\partial P_w}{\partial t} + \frac{p}{m}\frac{\partial P_w}{\partial x} + \theta \cdot P_w = \left(\frac{\partial Pw}{\partial t}\right)_{coll.} \tag{25}$$

This equation was constructed to include an ad-hoc collision term which may not necessarily express the same phenomenology as that of the classical Boltzmann transport equation, since P_w is not a true probability distribution function. These problems are conceptually reduced when dealing with moments in a relaxation approximation. The moment equations are obtained by multiplying Eq. (25) by an appropriate function of momentum, $\phi(p)$, and then integrating over all momenta to obtain:

$$\frac{\partial \langle\phi\rangle}{\partial t} + \frac{1}{m}\frac{\partial}{\partial x}\langle\phi p\rangle$$

$$-\sum_{n=0}^{\infty}\left(\frac{\hbar}{2i}\right)\frac{1}{(2n+1)!}\frac{\partial^{2n+1}}{\partial x^{2n+1}}V(x)\int\phi(p)\frac{\partial^{2n+1}}{\partial p^{2n+1}}P_w\,dp = \left\langle\phi\left(\frac{\partial Pw}{\partial t}\right)_{coll.}\right\rangle \tag{26}$$

where $\langle\ \rangle$ refers to an integration over momentum. In making the assumptions that $\phi(p)$ be an analytic function of momentum and that $P_w(x,p)$ vanish at the momentum limits, it follows that

$$\int\phi(p)\left(\frac{\partial^{2n+1}}{\partial p^{2n+1}}P_w\right)dp = (-1)^{2n+1}\int\left(\frac{\partial^{2n+1}}{\partial p^{2n+1}}\phi(p)\right)P_w\,dp \tag{27}$$

and that Eq. (26) becomes

$$\frac{\partial<\phi>}{\partial t} + \frac{1}{m}\frac{\partial}{\partial x}<\phi p>$$

$$+ \sum_{n=0}^{\infty} \left(\frac{\hbar}{2i}\right)^{2n} \frac{1}{(2n+1)!}\left(\frac{\partial^{2n+1}}{\partial x^{2n+1}} V(x)\right) < \frac{\partial^{2n+1}\phi}{\partial p^{2n+1}} > = < \phi \left(\frac{\partial Pw}{\partial t}\right)_{coll.} > . \qquad (28)$$

For specific values of $\phi(p) = p^o$, p, and $\frac{p^2}{2m}$, Eq. (28) becomes

$$\frac{\partial \rho}{\partial t} + \frac{1}{m}\frac{\partial}{\partial x}<p> = 0 , \qquad (29a)$$

$$\frac{\partial<p>}{\partial t} + \frac{1}{m}\frac{\partial}{\partial x}<p^2> + \rho\frac{\partial V}{\partial x} = <p\left(\frac{\partial Pw}{\partial t}\right)_{coll.}> , \qquad (29b)$$

$$\frac{1}{2m}\frac{\partial}{\partial t}<p^2> + \frac{1}{2m^2}\frac{\partial}{\partial x}<p^3> + \frac{<p>}{m}\frac{\partial V}{\partial x} = <\frac{p^2}{2m}\left(\frac{\partial Pw}{\partial t}\right)_{coll.}> \qquad (29c)$$

where upon reduction

$$<p^n> = \left(\frac{\hbar}{2i}\right)^n \sum_{j=0}^{n} (-1)^j \frac{n!}{(n-j)!j!} \frac{\partial^j \psi^*(x)}{\partial x^j} \frac{\partial^{n-j}\psi(x)}{\partial x^{n-j}} , \qquad (30)$$

thereby showing the dependence of $<p^n>$. For $n = 0, 1, 2, 3$ we show $<p^n>$ explicitly:

$$<p^o> = \psi^*\psi , \quad <p^1> = \frac{\hbar}{2i}\left(\psi^*\frac{\partial\psi}{\partial x} - \psi\frac{\partial\psi^*}{\partial x}\right) , \qquad (31a)$$

$$<p^2> = \left(\frac{\hbar}{2i}\right)^2 \left(\psi^*\frac{\partial^2\psi}{2x^2} - 2\frac{\partial\psi^*}{\partial x}\frac{\partial\psi}{\partial x} + \frac{\partial^2\psi^*}{\partial x}\psi\right), \qquad (31b)$$

$$<p^3> = \left(\frac{\hbar}{2i}\right)^3 \left(\psi^*\frac{\partial^3\psi}{\partial x^3} - 3\frac{\partial\psi^*}{\partial x}\frac{\partial^2\psi}{\partial x^2} + 3\frac{\partial^2\psi^*}{\partial x^2}\frac{\partial\psi}{\partial x} - \frac{\partial^3\psi^*}{\partial x^3}\psi\right). \qquad (31c)$$

Note that use of Eqs. (31a) in Eq. (29a) results in the correct quantum mechanical continuity equation.

In order to see the transition between the quantum and classical regimes, we invoke a wavefunction of the form $\Psi(x,t) = A(x,t)e^{iS(x,t)/\hbar}$, and so

$$A^2(x,t) \equiv \Psi^*\Psi = \rho(x,t) \ , \quad \frac{1}{m} \frac{\partial S(x,t)}{\partial x} = v(x,t) \tag{32}$$

where $\rho(x,t)$ is the probability density and $v(x,t)$ is the ensemble velocity. Inserting the wavefunction of Eq. (32) into Eq. (31) results in

$$< p^0 > = \rho(x,t) \ , \quad < p^1 > = mv\rho \ , \tag{33a}$$

$$< p^2 > = (mv)^2 \rho - \frac{\hbar^2}{4m} \rho \frac{\partial^2}{\partial x^2} \ell n\rho \ , \tag{33b}$$

$$< p^3 > = (mv)^3 \rho - \frac{\hbar}{4} \rho \{ 3mv \frac{\partial^2}{\partial x^2} \ell n\rho + \frac{\partial^2}{\partial x^2} (mv) \} \ . \tag{33c}$$

It is evident from Eq. (33) that the terms possessing an explicit dependence on "$\hbar$" are the quantum corrections to the classical momentum-density moments. Furthermore, insertion of these momentum-moments into Eqs. (29) results in a set of moment equations which contain explicit quantum corrections as well. Of course, as $\hbar \to 0$, these moment equations reduce to the classical, zero temperature, moment equations.

The logarithmic derivative term appearing in Eqs. (33) is not uncommon; such a term appears in the real part of the Schrodinger equation when the wavefunction associated with Eq. (32) is used (9). In the limit where this term is negligible, the Schrodinger equation goes over to the Hamilton-Jacobi equation.

We have made order of magnitude estimates of the strengths of the quantum correction term appearing in Eq. (33). Using a Gaussian spatial variation for $\rho(x)$, and a thermal ensemble value for mv^2, typical of central valley GaAs electrons, we find that the quantum correction is substantial for distances of the order of 100°A.

SUMMARY

In this paper, we have put forth a formalism for treating ultra-submicron device transport. The formalism results in a useful and attractive methodology for describing "quantum" device transport in that the theory is derived from a fully quantum mechanical representation, yet implicitly contains elements of the semiclassical semiconductor transport picture. The basic three semiconductor quantum transport equations were

derived using the Wigner distribution function. These transport equations were shown to contain explicit quantum corrections; these quantum corrections are non-negligible when the transit lengths of the semiconductor device are of the order of the carrier deBroglie wavelength. Since the carrier deBroglie wavelength for carriers is of the order of hundreds of angstroms in III-V device materials of interest to the Army, the quantum description of transport as described here will play a vital role in predicting the electrical behavior of present and future generation ultra-submicron devices.

In future studies, we plan to expand our theoretical efforts to include device modeling of submicron and ultra-submicron semiconductor devices such as P-N junctions, planar doped barriers, and one- and two-dimensional superlattices; for these devices, it is clear that quantum transport will indeed be necessary to explain their semiconductor transport characteristics.

ACKNOWLEDGMENT

The author wishes to thank Kathleen Lape´ for editing and typing this manuscript.

REFERENCES

1. J.R. Barker and D.K. Ferry, Solid-State Electron. 23, 519 (1980).

2. G.J. Iafrate, D.K. Ferry, R.K. Reich, Proc. 4th Int'l Conf. Electronic Properties of Two-Dimensional Systems, Aug. 82; R.J. Malik, K. Board, L.F. Eastman, G.E.C. Wood, T.R. AuCoin and R.L. Ross, Inst. Phys. Conf. Ser. No. 59, Ch. 9, p. 697 (1981).

3. E.P. Wigner, Phys. Rev. 40, 749 (1932).

4. R.F. O'Connell and E.P. Wigner, Phys. Lett. 83A, 145 (1981).

5. G.J. Iafrate, H.L. Grubin and D.K. Ferry, Bull. Am. Phys. Soc. 26, 458 (1981); Phys. Lett. 87A (1982).

6. H.L. Grubin, D.K. Ferry, G.J. Iafrate and J.R. Barker, VLSI Electronics: microstructure science, ed. N.G. Einspruch (Academic Press, New York) to be published.

7. G.J. Iafrate, H.L. Grubin and D.K. Ferry, J. de Phys. Coll. C7, Suppl. au n°10, Tome 42, Oct 81.

8. J.E. Moyal, Proc. Cambridge Philos. Soc. 45, 99 (1949).

9. A. Messiah, Quantum Mechanics, Vol. 1 (Interscience, New York, 1961) p. 442.

10. G. Niklasson, Phys. Rev. B10, 3052 (1974).

11. F. Brosens, L.F. Lemmens and J.T. Devreese, Phys. Stat. Sol. (b)81, 551 (1977).

CO_2 LASER WAVEGUIDING IN GaAs MBE LAYERS (U)

* HOWARD A. JENKINSON, MR.
JOHN M. ZAVADA, DR.
Fire Control and Small Caliber
Weapon Systems Laboratory
ARRADCOM, Dover, NJ 07801

I. INTRODUCTION

Integrated optical circuits have vast potential for use as sensing and signal processing elements in future Army fire control systems. One of the requirements for these systems is 24 hour/all weather operation. In order to meet this requirement, the Army has been emphasizing systems which operate at longer wavelengths, in particular, at the 10.6 micron CO_2 laser line. Such systems, including CO_2 laser radars, miss-distance sensors, and sensors for smart munitions form a natural area of application for infrared integrated optical circuits. These circuits could be used to perform various functions including phased array transmission, heterodyne detection, parallel processing and optical correlation. In order to perform these operations it is first necessary to develop efficient waveguides for confining the laser light (1).

The simplest version of an optical waveguide is a planar structure with an uniform dielectric constant (2). For an electromagnetic wave to be guided in this structure, the dielectric constant must be larger than that of the adjacent media. The guided wave will then propagate in a zig-zag fashion within this structure undergoing total internal reflection at each interface.

There are two main problems with this type of optical waveguide. The first arises from surface and interfacial irregularities that scatter light out of the waveguiding region. This scattering can lead to large propagation losses and to poor signal/noise characteristics. The second is lack of phase coherence for modes of different order. This difficulty limits the use of such structures in multimode propagation and reduces the information content that can be transferred. Similar problems have been encountered with optical fibers used in communication systems and have led to the development of Graded Index (GRIN) or Self-Focusing (SELFOC) fibers (3). The refractive index of these fibers is no longer uniform but varies according to a parabolic profile. Due to this index variation, light rays

in the fiber are contained in a sinusoidal envelope. Since the light ray never touches the interface, scattering losses from surface irregularities are drastically reduced. Also, the index profile is adjusted to maintain phase coherence for a wide range of propagating modes.

While such advances have been made in the fabrication of cylindrical optical fibers, controlled index profiling has not as yet been achieved in planar optical waveguides. The reason for this lies in the considerable difficulty of introducing impurity atoms in a precise manner over a two dimensional region of several square centimers.

Molecular beam epitaxy (MBE) is a vacuum deposition processing technique (4) in which several different atoms or molecules can be deposited producing a film with the same crystalline structure as the substrate (epitaxial growth). MBE machines have been available for over a decade but the quality of the deposited films has often been erratic. Recent progress in vacuum technology has allowed the fabrication of high quality films in a reproducible manner.

With modern MBE machines, it is also possible to grow multilayer planar structures in which the optical properties of individual layers can be accurately controlled through electronic doping. Index tailoring of this type can lead to novel electro-optical devices for use in integrated optics and microelectronics. To effectively utilize the advantages of multilayer structures for guided wave applications, it is necessary to be able to determine the optical waveguiding characteristics of such structures.

In this report, experimental analyses of the electronic and optical properties of a prototype single layer MBE waveguide are presented. A polished (100) wafer of n^+-GaAs heavily doped with silicon was used as the waveguide substrate. The free carrier concentration was approximately $3x10^{18}/cm^3$. An epitaxial layer of n-GaAs lightly doped with tin was then grown on the substrate. The carrier concentration of the epitaxial layer was targeted to be $3\text{-}5x10^{15}/cm^3$.

II. DIELECTRIC MODEL

The dependence of the refractive index on the carrier concentration can be illustrated by a simple dielectric model in which the dielectric function of GaAs is composed of Lorentz oscillator terms representing the lattice contribution (5) and a Drude term representing the depression of the dielectric function due to the presence of the free carriers (6). In the spectral region between the band gap and the IR phonon resonance, the dispersion due to the lattice contribution is much less than that due to the plasma. Thus, to good approximation, the dielectric function may be

written, in terms of wavenumber σ, as

$$n^2(\sigma) = K_L - \frac{\sigma_p^2}{\sigma(\sigma + ig)} \qquad 269\ cm^{-1} \ll \sigma \ll 2x10^4\ cm^{-1} \quad (1)$$

where K_L is the high frequency dielectric constant of GaAs, taken to be 10.73. σ_p is related to the carrier concentration N through

$$\sigma_p^2 = \frac{Ne^2}{(2\pi c)^2 m^* \varepsilon_o} \quad (2)$$

where c is the speed of light, e is the electronic charge, $m^* = .08m_e$ is the effective mass of electrons in GaAs, and ε_o is the vacuum permittivity. The damping factor, g, leads to losses and is related to the mobility, μ, of the material through

$$g = \frac{e}{2\pi c \mu m^*} \quad (3)$$

Neglecting losses, the increase in the refractive index at the substrate-epilayer interface due to the carrier concentration difference is approximately

$$\Delta n \approx \frac{\sigma_{p,s}^2 (1 - N_f/N_s)}{2n\sigma^2} \quad (4)$$

where $\sigma_{p_s}^2$ is σ_p^2 evaluated using the original carrier concentration N_s of the substrate and N_f is the carrier concentration in the MBE layer.

III. INFRARED REFLECTANCE

Prior to waveguiding experiments the optical properties of the MBE sample were characterized by infrared reflectance measurements. Differential reflectance measurements, made over the spectral range of 4000 cm^{-1} to 1400 cm^{-1}, yielded interference fringes typical of a thin film/substrate structure, as illustrated in Figure 1. The interesting features of this curve are the periodicity of the fringes and their increase in amplitude as the measurement progresses to smaller wavenumbers. A first-order model has been developed to aid in deducing the carrier concentration and the thickness of the epitaxial layer from the reflectance spectrum. Basically, losses are neglected and the layer is considered to be a film of refractive

index n_f sitting atop a substrate whose index has been depressed a small amount Δn by its higher concentration of free carriers. The reference sample is considered identical to the substrate. Expanding the relevant reflectance equations in a McLaurin series expansion to first order in Δn and using equation (4) yields:

$$R_D^2(\sigma) = \frac{R^2_{sample}}{R^2_{reference}} \approx 1 + \frac{4\sigma^2_{p,s}(1-N_f/N_s)}{n_f(n_f^2-1)}\left(\frac{\sin(2\pi t n_f \sigma)}{\sigma}\right)^2 \qquad (5)$$

From this analysis, it can be shown that the thickness of the epitaxial layer can be determined by the location, σ_m, and order, m, of the fringe minima, and the relative carrier difference, X, by the fringe amplitudes:

$$t = m/(2n_f\sigma_m) \; ; \qquad X = 1 - N_f/N_s = \frac{n_f(n_f^2-1)\sigma^2}{4\sigma^2_{p,s}}(R^2_{D,max}-1) \qquad (6)$$

Table 1 presents the reflectance data for the fringe minima located between 1400 and 3000 cm^{-1} that appear in Figure 1. The order of each fringe is identified by dividing its location by the average separation of the fringes in the interval. Using a value of $3.28 = \sqrt{K_L}$ for n_f the thickness of the epitaxial layer was calculated to be 11.45 microns with a standard deviation of .03 microns. The low valve of the standard deviation is due to the uniform periodicity of the reflectance fringes which in turn implies that the epitaxial layer is optically homogeneous. The growth in fringe amplitudes shown in Figure 1 is consistent with the electronic model for the refractive index change. It can be seen from equation (5) that the fringe amplitudes are proportional to (1-Nf/Ns). As this factor reaches 90% of its maximum valve for Nf = .1 Ns, the method is not suitable for accurate determination of epilayer carrier concentrations which are more than an order of magnitude lower than in the substrate.

IV. CV PROFILING

Since reflectance measurements can not determine relatively low carrier concentrations, it is necessary to resort to electronic techniques to measure Nf. One suitable method is that of capacitance voltage (C-V) profiling. This technique is based on the principle that a Schottky barrier (metal-semiconductor interface) formed on the surface of the epitaxial layer will behave like a parallel plate capacitor. In this picture one plate of the capacitor is formed by the Schottky barrier, the other is formed by the free carriers within the epitaxial layer. As the Schottky barrier is reversed biased, the free carriers in the epilayer are repelled, causing the moveable "plate" to recess further into the epilayer and resulting in a decrease in the measured capacitance. From measurements

of the rate at which the capacitance C decreases as a function of bias voltage V, the carrier concentration as a function of depth can be determined. The equations that describe this measurement technique are (7):

$$W = \varepsilon\varepsilon_o A/C(V) \quad ; \quad N(V) = \left(-\frac{e\varepsilon\varepsilon_o A^2}{2} \frac{d}{dV} \frac{1}{C^2(V)} \right)^{-1} \qquad (7)$$

Here ε is the low frequency dielectric constant, taken to be 12.5 for GaAs, A is the area of the Schottky barrier electrode, and W is the depletion width of the junction. If N is sufficiently low, then it is possible to bias the junction such that the transition region between the epilayer and the substrate can be profiled. The results of such a measurement, made at 1MHz with a gold Schottky barrier vacuum deposited through a stainless steel mask, are shown in Figure 2. This curve indicates that the carrier concentration is uniform throughout the epitaxial layer at a level of approximately $1.6 \times 10^{15}/cm^3$. The sharp rise in this curve at a depth of 10.89 micrometers represents the transition to the more heavily doped substrate.

V. INFRARED OPTICAL WAVEGUIDING

Since the reflectance spectra and the CV profile support the approximation of a single dielectric film/substrate structure, the equations for a planar asymmetric dielectric waveguide were used as the waveguide model: (1,2)

$$kt\sqrt{n_f^2-(\beta/k)^2} - \phi_{fo} - \phi_{fs} = m\pi \qquad (8)$$

where $k=2\pi/\lambda_o$, $\beta/k=n_f \sin\theta_f$ and the phase angles are defined by:

$$\tan\phi_{fi}^{TE} = \left(\frac{(\beta/k)^2-n_i^2}{n_f^2-(\beta/k)^2} \right)^{1/2} \quad ; \quad \tan\phi_{fi}^{TM} = (n_f/n_i)^2 \tan\phi_{fi}^{TE} \qquad (9)$$

The subscript i refers to either the cover layer o or to the substrate s. Equation (8) is the familiar eigenvalue condition which must be satisfied for optical waveguiding to occur and is expressed in terms of the ratio β/k. This factor can be interpreted as the effective refractive index of the waveguide for each mode and is restricted to values between n_s and n_f. Based on the dielectric model, these parameters are expected to be 2.64 and 3.28 respectively at 10.6 microns.

Infrared optical waveguiding was achieved in this sample using a 1 watt

CO_2 laser operating at 10.6 microns. The radiation was focused through an f-10 ZnSe lens on a dove shaped germanium prism which effected coupling into and out of the waveguide. The emerging beam, totally internally reflected from the base of the prism, was imaged as a spot on a thermographic screen. Waveguiding was identified by an absorption notch in the reflected spot, indicating that radiation was removed from the incident beam. Measurements of the angles at which synchronous coupling was achieved were made for both s and p polarizations of the incident beam.

Experimentally, the TE modes were much easier to observe than the TM modes. This is due to the greater amount of energy that can be coupled into the modes. A total of eight modes, five TE and three TM, were observed and Table 2 contains the experimental results. In principal, the coupling data from any two of these modes can be used to numerically invert the mode equations to obtain the refractive index and thickness of the epitaxial layer. Since there was a larger degree of uncertainty in the TM measurements listed in Table 2, only the TE coupling data were used for this purpose. From these measurements, the refractive index of the epitaxial layer was calculated to be 3.27 with a standard deviation of .01 and the thickness to be 12.19 microns with a standard deviation of .23 microns.

VI. DISCUSSION

The results of these three independent measurements are summarized in Table 3. In the IR reflectance and CV profiling measurements, it is necessary to postulate a dielectric model for n-type GaAs in order to relate the experimental data to the optical properties of the epitaxial layer. The synchronous coupling experiments, however, provide a direct determination of n_f through the waveguiding equations. The excellent agreement of the three determinations of n_f supports the validity of the dielectric model, and indicates that the free carrier concentration is the dominant mechanism responsible for the refractive index change.

The thickness determinations are also in good agreement, but are more subject to systematic measurement errors arising in the different experimental techniques. The IR reflectance determination was calculated from a first order approximation to the exact mathematical expression for the reflectance of a thin film-substrate structure. Hence, the absolute error may be larger than indicated by the standard deviation. The C-V determination involves precise measurements of small values of capacitance and the electrode area. Experimental uncertainties in these quantities limit the overall accuracy of the thickness determination to 5%. Furthermore, the optical thickness is determined by the depth at which the carrier concentration rises from 10% to 100% of the substrate concentration. This region was not completely probed in the current experiments. The

synchronous coupling determination is again a direct measurement and may be more representative of the actual thickness of the layer in spite of the larger standard deviation. Finally, it is possible that the transition region from the MBE layer to the substrate will ultimately be shown to have a carrier concentration gradient extending over a few thousand Angstroms. This would lead to a graded index layer which would necessitate data interpretation with more complex models than used here.

While molecular beam epitaxy is still an emerging technology, the present experiments show that it can be used to form high quality CO_2 laser waveguiding structures. The epitaxial layer examined in this study displayed very uniform optical and electronic properties that can be related to an elementary Lorentz-Drude dielectric model. These results substantiate using this model to design more complicated graded index CO_2 waveguides through MBE.

ACKNOWLEDGEMENT

The authors wish to thank J. Comas of the Naval Research Laboratory for providing the MBE sample used in this investigation.

VII. REFERENCES

1. J. M. Zavada, H. A. Jenkinson, T. J. Gavanis, R. G. Hunsperger, D. C. Larson and J. Comas, "Ion Implanted Guided Wave Devices for Army Fire Control," 1980 Army Science Conference Proceedings, Vol. III, 445 (1980).

2. D. Marcuse, Theory of Dielectric Optical Waveguides, Academic Press, New York, 1974.

3. M. S. Sodha and A. K. Ghatak, Inhomogeneous Optical Waveguides, Plenum, New York, 1979.

4. C. E. C. Wood, "Progress, Problems, and Applications of Molecular-Beam Epitaxy," in Physics of Thin Films, Vol. II, Academic Press, 1980.

5. A. H. Kachare, W. G. Spitzer, J. E. Frederickson, and F. K. Euler, "Measurements of Layer Thicknesses and Refractive Indices in High-Energy Ion-Implanted GaAs and Gap." J. Appl. Phys., 47, no. 12, 5374 (1976).

6. J. M. Zavada, H. A. Jenkinson, and T. J. Gavanis, "Optical Properties of Proton Implanted n-Type GaAs," Proc. Soc. Photo-Optical Instrum. Eng., 276, 104 (1981).

7. A. Glasser and G. Subak-Sharpe, Integrated Circuits Engineering, Addison-Wesley, Reading, 1977.

Table 1. IR REFLECTANCE DATA: GaAs MBE 8126

σ_m (cm^{-1})	m	σ_m/m (cm^{-1})	(micrometers)
2925	22	132.95	11.466
2790	21	132.86	11.474
2655	20	132.75	11.483
2530	19	133.16	11.448
2390	18	132.78	11.481
2260	17	132.94	11.467
2135	16	133.44	11.424
1995	15	133.00	11.462
1865	14	133.21	11.444
1735	13	133.46	11.422
1605	12	133.75	11.397
1470	11	133.64	11.407

Table 2. SYNCHRONOUS COUPLING DATA AT 10.6 MICRONS: GaAs MBE 8126

MODE	β/k
TE_0	3.249
TM_0	-
TE_1	3.175
TM_1	3.173
TE_2	3.051
TM_2	3.038
TE_3	2.878
TM_3	2.847
TE_4	2.667
TM_4	-

Table 3. SUMMARY OF MEASUREMENT RESULTS: GaAs MBE 8126

	CARRIER CONCENTRATION ($1/cm^3$)	REFRACTIVE INDEX 10.6 um	THICKNESS (um)
Substrate (Manuf. Spec)	3.0×10^{18}	2.64	-
Film-IR Reflectance	$<1 \times 10^{17}$	3.28	$11.45 \pm .03$
C-V Profile	1.6×10^{15}	3.28	10.89
Synchronous Coupling	$<1 \times 10^{17}$	$3.27 \pm .01$	$12.19 \pm .23$

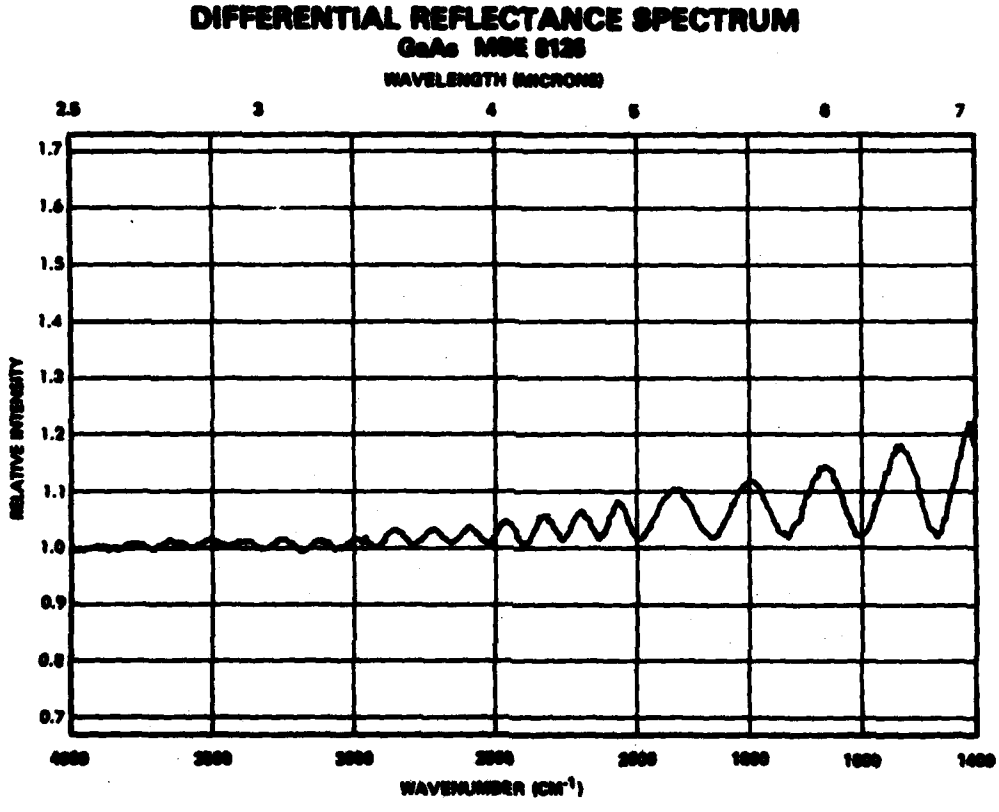

DIFFERENTIAL REFLECTANCE SPECTRUM
GaAs MBE 8126
WAVELENGTH (MICRONS)
2.5
3
4
5
6
7
1.7
1.6
1.5
1.4
1.3
1.2
1.1
1.0
0.9
0.8
0.7
RELATIVE INTENSITY
4000
3000
2500
2000
1500
1000
1400
WAVENUMBER (CM-1)

Figure 1.

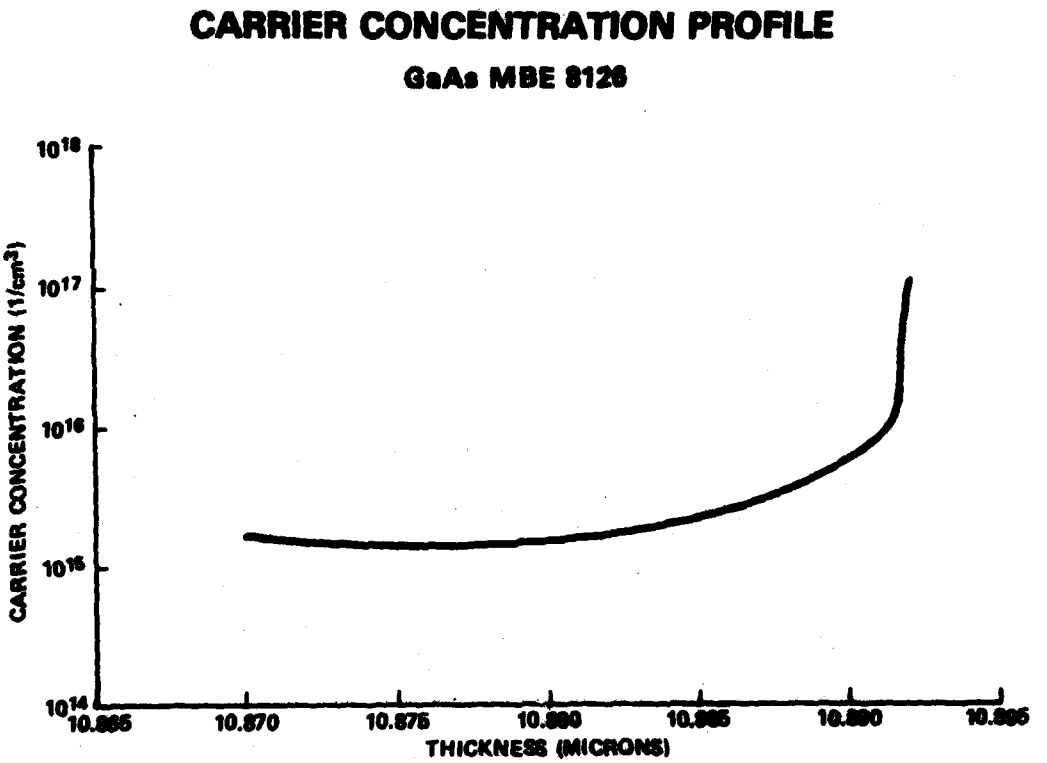

CARRIER CONCENTRATION PROFILE
GaAs MBE 8126
CARRIER CONCENTRATION (1/cm³)
10¹⁸
10¹⁷
10¹⁶
10¹⁵
10¹⁴
10.865
10.870
10.875
10.880
10.885
10.890
10.895
THICKNESS (MICRONS)

Figure 2.

ACTIVE IMAGING OF RANGE TARGETS AT 1.2 MILLIMETERS (U)

JOHN LAMAR JOHNSON, DR., KENNETH A. HERREN, MR.
ROBERT L. MORGAN, DR. AND *GEORGE A. TANTON, DR.
RESEARCH DIRECTORATE, US ARMY MISSILE LABORATORY
US ARMY MISSILE COMMAND
REDSTONE ARSENAL, ALABAMA 35898

Several years ago P. W. Kruse (1) and R. L. Hartman et. al (2) showed that for detection and tracking of "obscured" targets at ranges up to 3 km, submillimeter wavelengths in the atmospheric windows between 0.74 - 1.3 mm offer great potential advantages over wavelengths in both the optical and in the microwave regions. At submillimeter wavelengths a significant reduction over the conventional microwave antenna size needed for a given resolution can be realized. Submillimeter waves also penetrate fogs and aerosols much more effectively than do the shorter optical and IR wavelengths. We have demonstrated active target imaging through the atmosphere with a submillimeter wave laser operated at a wavelength of 1.2 mm. Copper corner cube targets were imaged over a maximum atmospheric path of 650 meters (two-way) using an 0.3 meter antenna.

The system used (3) has a CO_2 pumped SMMW gas laser which is passed through a beam expander to a 0.3 meter diameter f/5 primary mirror and then to a computer controlled two-axis scan mirror which directs the beam down a target range adjacent to the laboratory building. The return beam is collected through the same optics and sent by a beamsplitter to the detector, and the signal is sent to the computer from a lock-in amplifier.

The targets consisted of a trio of copper corner cubes at 100 meters from the primary, and a single copper corner cube at 325 meters. The cubes at 100 meters were 30 centimeters on a side, and the cube at 325 meters was 90 centimeters on a side.

The output of the SMMW system at the scan mirror is on the order of a hundred microwatts. The detectors used were Golay cells, and the chopping frequency was 13.5 Hz. The scan mirror is an Aerotech tracking gimbal which is limited to 500 steps per second, at 0.25 arcsecond per step, and thus each image scan required two hours to complete. This is not a fundamental restriction to the frame rate, as faster detectors and faster scan

mechanisms exist or can be constructed. Table 1 shows the current system parameters and some near-term improvements.

Figure 1 shows scanned images of the three corner cubes at 100 meters. The variation in relative signal strength is due to the difference in construction among the three cubes. The images are reversed left to right from the actual scene.

The images are all 40 x 40 pixels. Those in the left column are the original data, those on the right are the result of one pass with a 3 x 3 recursive median filter after the image was recorded. (a) shows that the system was detector noise limited and that only cooperative targets were detected. None of the support structure or surrounding terrain features are visible in the processed image on the right. (b) was taken with an iris in the return beam focal plane set at a 30 millimeter aperture. (c) is the same as (b) but with the iris closed to a 7 millimeter aperture.

This aperture gave a peak signal of one-half that of case (b). In (c) the signal on the right in the processed image was thresholded in the computer display so that the white center section corresponds to the FWHM. The cubes were six feet apart horizontally and vertically. In the other images the threshold was set to show low level pixels, and the apparent diameters of the targets do not correspond to the measured FWHM.

Figure 2 shows the corner cube at 325 meters. (a) is with the 7 milli meter aperture and (b) is with the 30 millimeter aperture. Note in (b) the presence of interference fringes due to unwanted radiation scattered from the components in the optical train.

Figure 3 is an isometric plot illustrating the effect of the recursive median filter after one application. Note that the strongest return is clipped as a result of the graphics display program. The recursive filter differs from the standard median filter in that the new value of a pixel is substituted before the filter mask is shifted to the next set of data points. This filter gives effective random noise reduction in a single pass. N. C. Gallagher (4) has shown that in one-dimensional signals, a single pass of a recursive median filter produces a root function that is not altered by additional passes. We observe that in two dimensions there are low-level changes in our images with repeated passes of the 3 x 3 recursive filter; but that the initial pass has by far the major smoothing effect on the images.

Two measurements of the FWHM were made, one using the iris and the other by measuring the FWHM of the scanned images. As a check, the system was centered on a corner cube, then scanned in each axis to the half-power points, and the change in location was visually observed with a sighting telescope which looked through the scan mirror. The FWHM measured in this

way was, in both axes, about 5/3 the length of the side of the corner cube, or 5 milliradians. The FWHM measured from the scanned images was 5.4 milliradians. Both of these are ± 10% measurements. The iris measurement corresponded to 4.6 milliradians, and was good to ± 8%. The Rayleigh limit for a plane-wave, using a 30 cm aperture and a wavelength of 1.2 millimeters is 1.22 λ/D = 4.88 milliradians.

Since the resolution of an active system depends only on the transmitted spot size, these numbers show that the transmitted beam was the size of a diffraction limited plane-wave Airy disc at the target. The resolution can be improved by operating with shorter wavelengths and larger antennas.

Based on previously developed criteria (5) for detection and recognition of military targets, the experimentally determined resolution of 5 mr would be sufficient to recognize a M-48 tank at a range of $\sim$ 0.5 km and to determine its orientation at a range of $\sim$ 1.5 km using a one-meter diameter antenna, and operating in the atmospheric "window" near 0.75 mm wavelength

In addition, the resolution of an active system using coherent laser radiation can be significantly improved by operating with a Gaussian beam profile, as shown in figure 4. Gaussian beams have a beam waist located beyond the primary mirror antenna as indicated in the figure, and thus the apparent angular spot size as seen from the primary is a function of the range. The solid curve shows this for the current system of 0.3 meter aperture and 1.2 mm wavelength, and the dashed curve for a 0.5 meter aperture at .75 mm wavelength. In general, this gives smaller apparent spot sizes than would be expected using the Rayleigh criteria. Since it requires neither shorter wavelengths nor larger primary apertures, this gain in resolution is one reason to prefer an active coherent imaging system over a passive and/or incoherent system.

REFERENCES

1. P. W. Kruse, "A System Enabling the Army to See Through Inclement Weather", 1974, US Army Scientific Advisory Panel Report.

2. R. L. Hartman, W. L. Gamble, B. D. Guenther, and P. W. Kruse, "Submillimeter System for Imaging Through Inclement Weather", The Optical Submillimeter Propagation Conference, Puerto Rico, 6-10 December 1976.

3. C. A. Tanton, J. F. Cummiskey, R. L. Morgan, H. C. Meyer, and J. G. Castle, Jr., "Atmospheric Propagation of Submillimeter Waves: Observed Correlations with Fog Conditions at 0.89 mm Wavelength", Technical Report RR-80-3, US Army Missile Command, February (1980).

4. N. C. Gallagher and G. L. Wise, "A Theoretical Analysis of the Properties of Median Filters", IEEE Trans. on Acoustics, Speech, and Signal Processing, Vol. ASSP-29, pp 1136-1141, December (1981).

5. J. Johnson, Image Intensifier Symposium, Ft. Belvoir, VA, 6-8 Oct. (1958).

TABLE 1.

SUBMILLIMETER WAVE IMAGING

CURRENT SYSTEM	NEAR TERM IMPROVEMENTS
RANGE SYSTEM	
Corner Cubes	Non-cooperative military targets
DETECTORS	
Golay cells, 10^{-9} W/$\sqrt{Hz}$ N.E.P., 10 Hz BW	InSb bolometer, 10^{-12} W/$\sqrt{Hz}$, up to 1 KHz BW. Ge bolometers, 10^{-14} W/$\sqrt{Hz}$, 100-500 Hz BW. Heterodyne detectors, $< 10^{-20}$ W/Hz, GHz BW.
SMMW LASER	
1.6 m cavity 100 μw at 1.2 mm output.	2.5 m cavity, CO_2 pump, gas laser, output of 10's of mw, improved optical and mechanical design, operation in 0.89 and 0.75 mm regions.
OPTICS	
f/5, 30 cm primary Two-axis full aperture scan	one-axis counter rotating mirrors, one-axis tilting mirror, focal plane scanner.
CONTROLLER	
HP 21MX Computer	Upgraded F-series mainframe
SOFTWARE	
Scanning program Median filters Simple Algebra	standardized mag tape format, expanded image manipulation, range data processing, program compatibility

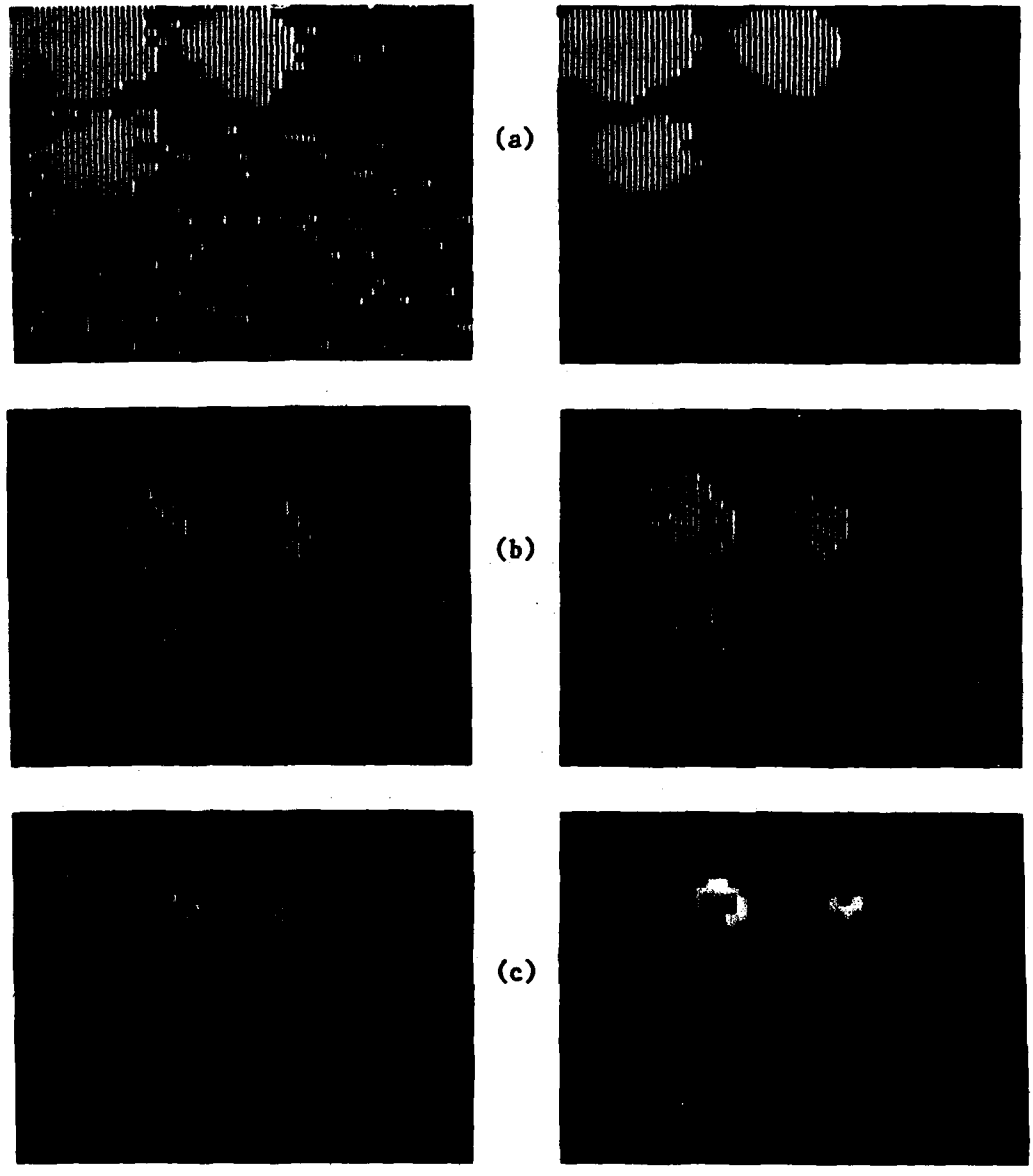


Figure 1. THREE 30 cm CORNER REFLECTORS AT 100 METERS

The images in the left column are the original data, those on the right are the result of one pass with a 3 x 3 recursive median filter after the image was recorded. (a) three corner reflectors, (b) and (c): effect of an aperture in the focal plane, (b) 30 mm aperture, (c) 7 mm aperture. This aperture gave a peak signal of approximately one-half that of case (b).

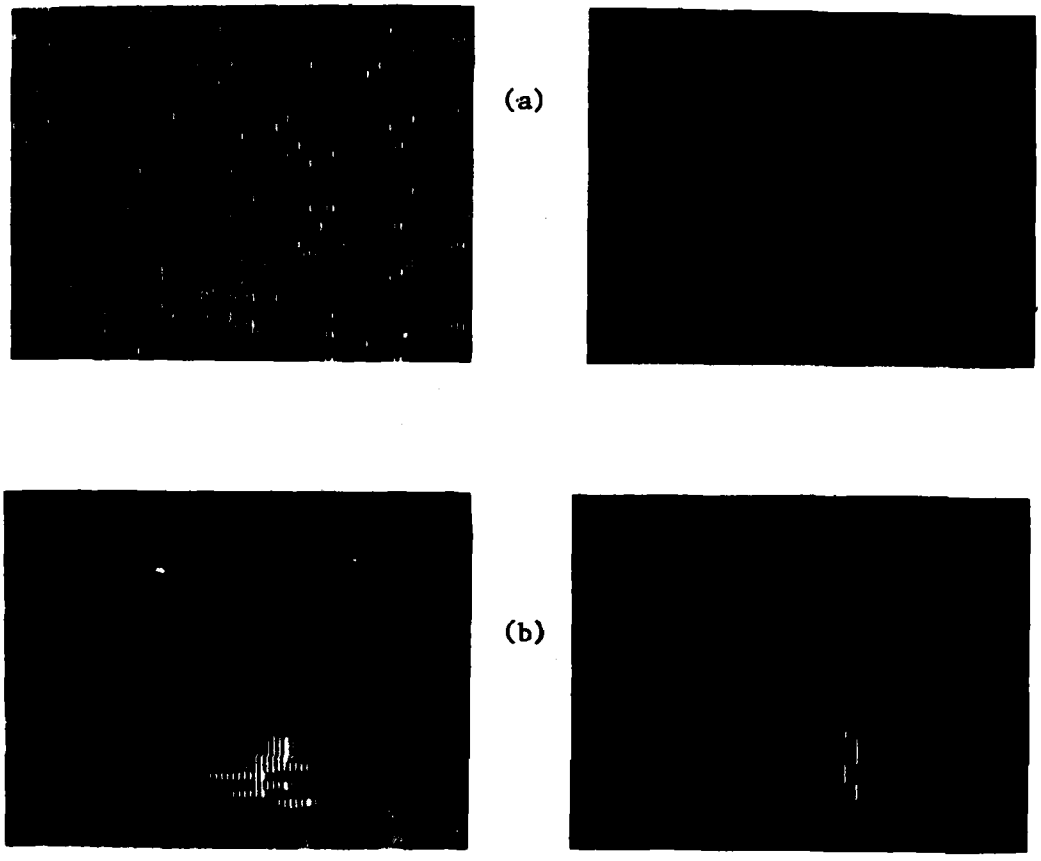
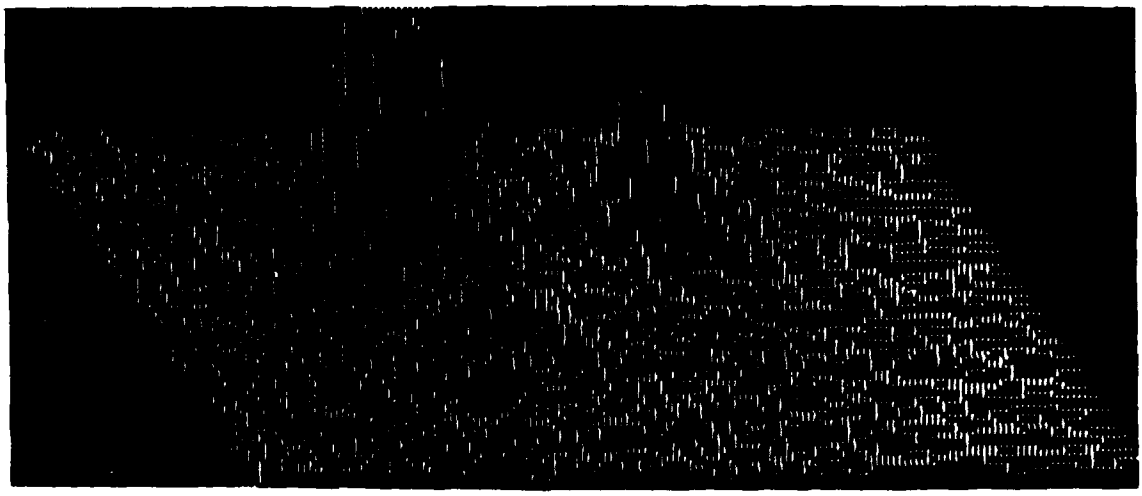


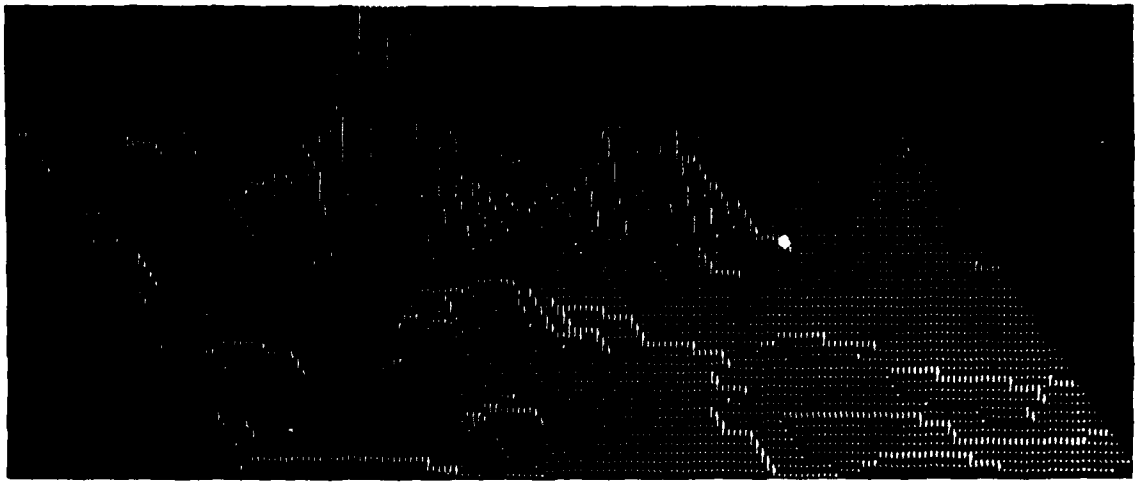
Figure 2. ONE 90 cm CORNER REFLECTOR AT 325 METERS

(a) (Left) Original data, with a 7 mm focal plane aperture stop.
(Right) Same image after application of a 3 x 3 recursive median filter.

(b) (Left) Original data, with a 30 mm focal plane aperture stop.
(Right) Data after software filtering. The interference fringes are due to the optical components in the scanning system.



(a)



(b)

Figure 3. ISOMETRIC PLOT OF RANGE IMAGE

(a) Original data, showing three corner reflectors with a 7 mm focal plane aperture stop.

(b) Two dimensional 3 x 3 recursive median filter applied to the original data.

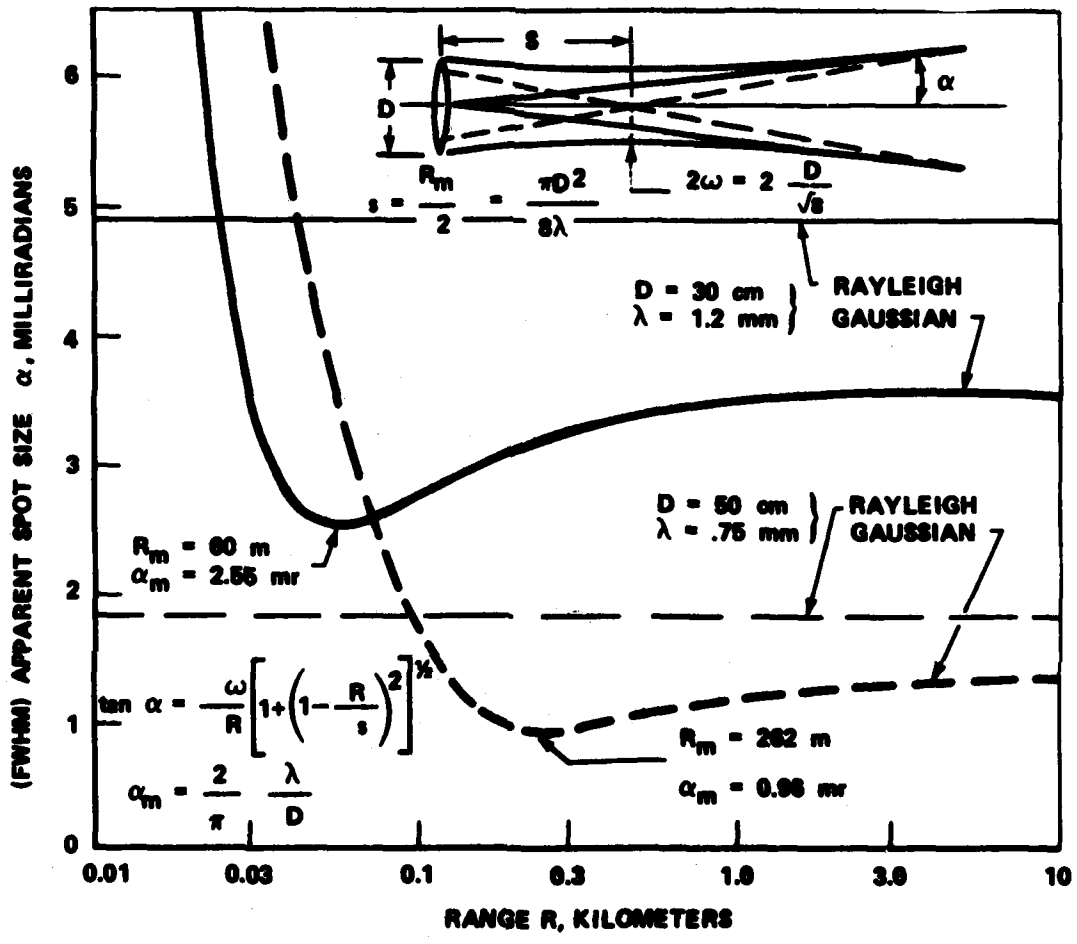


Figure 4. Apparent spot size of a Gaussian beam compared to a classical optical beam in the image plane of a lens, plotted as a function of range.

HOLOGRAPHIC INTERFEROMETRY TECHNIQUE FOR ROTARY WING AERODYNAMICS AND NOISE (U)

*JOHN K. KITTLESON
YUNG H. YU, DR.
AEROMECHANICS LABORATORY (AVRADCOM)
AMES RESEARCH CENTER, NASA
MOFFETT FIELD, CALIFORNIA 94035

INTRODUCTION

Helicopter impulsive noise, sometimes called "blade slap," is the most annoying and easily detectable sound generated by a helicopter. This extremely loud noise is generated by two different mechanisms which occur on the main and tail rotors of modern helicopters. One phenomenon, labeled high-speed impulsive noise, occurs on the advancing-blade side of the rotor disc where advancing blade Mach numbers are quite high ($M_T \approx 0.90$) and local transonic flow is inevitable. Rotor-blade thickness near the tip causes a low-frequency "thumping" and results in a nearly symmetrical acoustic signature (Figs. 1a and 1b). As the Mach number increases, local transonic events distort and amplify the time history waveform (Fig. 1c). The resulting sawtooth waveform, which radiates large amounts of acoustic energy, causes the sound to be very intense and distinctive.

The second phenomenon responsible for impulsive noise is known to occur when the rotor blades interact with previously generated tip vortices (1). Much less is known about the essential mechanisms in this case. Mach number and wake geometry are thought to play key roles in the ultimate strength of the resulting blade-vortex interaction noise. This sound is known to be a significant problem during descent and deceleration and can occur on both the advancing and retreating sides of the helicopter disc. However, it is believed that the advancing-blade interactions radiate significantly more energy.

Both of the mechanisms for loud impulsive noise have one parameter in common - high Mach number flow. For high-speed impulsive noise, local shocks on the blade surface have been shown to extend off the tip of the blade directly to the acoustic far field (Fig. 2). These shocks directly affect the amount of energy that is radiated by the helicopter. Detailed predictions of the far-field acoustics depend critically on the physical phenomena in the tip region. In particular, the three-dimensional velocity

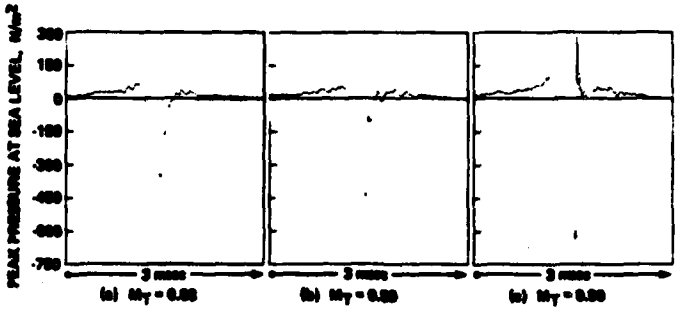


Figure 1. Waveform transition; the development of a radiating discontinuity, in-plane, r/D = 1.5 (from Ref. 2).

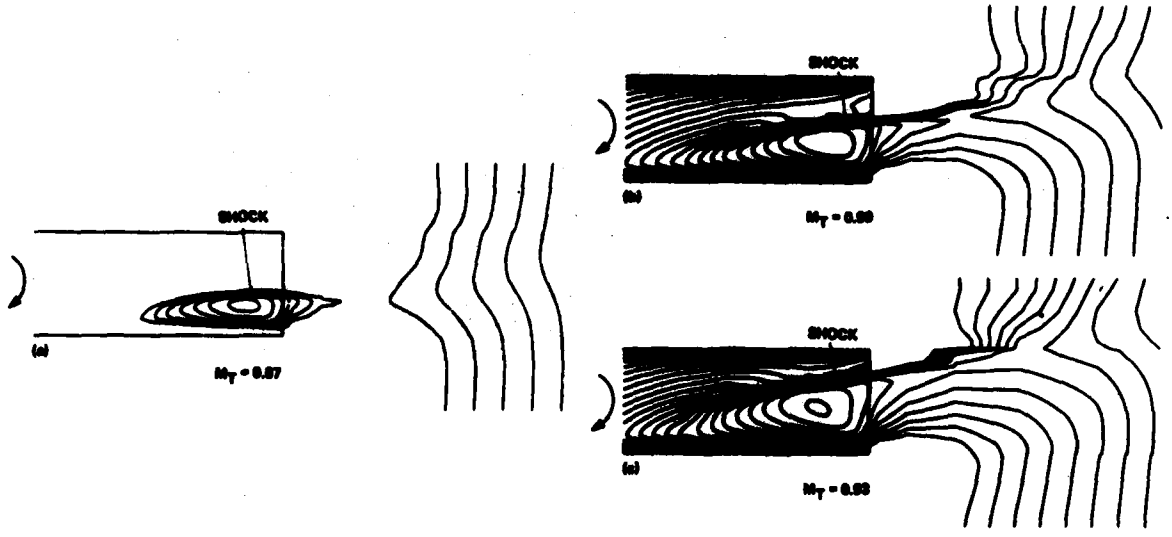


Figure 2. Velocity contours near a rotor tip in the transonic range (from Ref. 9).

distribution over a volume large enough to capture the local shock waves is required (2). The blade-vortex interaction problem also requires a vast quantity of local transonic flow information. In particular, unsteady local shocks and their resulting pressure distribution near the blade surface, as well as flow characteristics (e.g., vortex geometry and inflow), are

required. In both cases, good experimental data of the three-dimensional transonic flow field around a rotor blade and its wake system are essential for the theoretical prediction of the acoustic signature. The measurements are also required for the verification of aerodynamic transonic flow numerical codes.

Unfortunately, conventional visualizing and measuring techniques have severe limitations when applied to the large three-dimensional transonic flow field near a rotor. Hot-wire anemometry disturbs the flow and measures only one point at a time. Laser velocimetry requires seeding of the flow (which can be inaccurate because of acceleration effects) and also measures only one point at a time. Classical interferometry (such as Mach-Zehnder) divides the interfering waves spatially, requiring very expensive optical components and making large-scale experiments almost impossible. Schlieren and shadowgraph photography provide only a two-dimensional representation of a three-dimensional flow field. Clearly, another technique which overcomes these limitations must be employed.

Recent developments in the field of holographic interferometry in combination with computerized tomography proves to be the most promising technique for both visualizing and quantitatively measuring the flow field near a rotor blade. Holographic interferometry is nonintrusive and can be applied to a large-scale experiment with relatively low cost. It also records information over a large volume on a single interferogram, and can be used to visualize the flow field in a three-dimensional manner. To date, most applications of this technique have been limited to two-dimensional flow (3), axis-symmetric flow, or a simple three-dimensional flow with a small model under ideal laboratory conditions (4,5).

In this paper, the concepts of holography and holographic interferometry, as applied to the visualization and measurement of the three-dimensional flow field near a rotor tip, are previewed, and initial experimental results of investigations of local shock structures and tip vortices behind the blade are presented. An additional method to visualize the flow in a three-dimensional manner is demonstrated, and finally, a method to quantitatively measure the three-dimensional flow, which will provide the necessary information to help improve helicopter performance and reduce noise, is introduced.

BACKGROUND CONCEPTS

Holography

Holography is a two-step imaging process which records and reconstructs diffracted light waves (6,7). An interference pattern is formed on film by adding a coherent reference wave to the object wave (the wave of interest), as shown in Fig. 3a. The irradiance at the film plane is

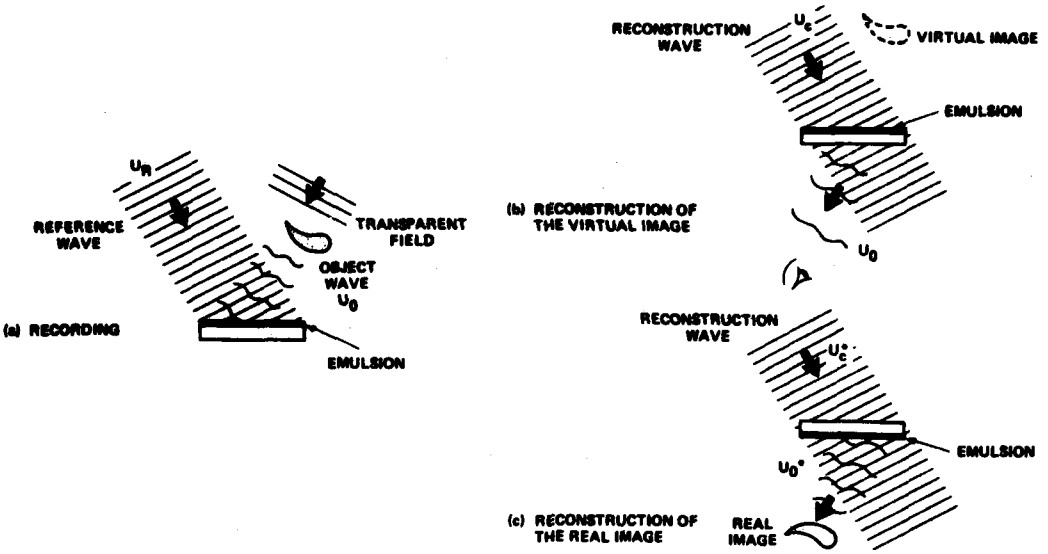


Figure 3. Optical holography recording and reconstruction.

$$I = |U_R + U_O|^2$$

$$= I_R + I_O + U_R^* U_O + U_R U_O^* \qquad (1)$$

where U_R represents the reference wave, U_O represents the object wave, I_R and I_O are amplitude squares of the reference and object waves, and the asterisk represents a conjugate wave. The last two terms in Eq. (1) represent the interference pattern recorded on the film, which contains both amplitude and phase information.

The amplitude transmittance of the developed film, called a hologram, is proportional to the irradiance

$$t = \beta(I_R + I_O) + \beta U_R^* U_O + \beta U_R U_O^*$$

where β is a property of the film. When the reconstruction wave U_C illuminates the hologram as shown in Fig. 3b, the transmitted light is

$$U_I = U_C t = \beta U_C(I_R + I_O) + \beta U_R^* U_C U_O + \beta U_R U_C U_O^* \qquad (2)$$

The first term of Eq. (2) represents an undiffracted zero-order wave. The second and third terms represent two first-order diffracted waves. The second and most important term of Eq. (2) can be written as $(\beta U_R^* U_C)U_O$. This is a replica of the original object wave U_O.

If the hologram is rotated 180° about a vertical axis so that the emulsion of the hologram is placed on the observer's side, the transmitted light with the reconstruction wave U_c^* will be

$$U_I = U_c^* t = \beta U_c^*(I_R + I_o) + \beta U_R^* U_c^* U_o + \beta U_R U_c^* U_o^* \qquad (3)$$

The last term of Eq. (3) can be written as $(\beta U_R U_c^*)U_o^*$; it represents a reconstructed wave that is conjugate to the original object wave. A real image in space is formed on the observer's side of the hologram, as shown in Fig. 3c. The real image can be photographed without the use of a lens by placing a sheet of photographic film in the real-image space.

Holographic Interferometry

Holographic interferometry is the interferometric comparison of two object waves that are recorded holographically. Although holographic interferograms and classical Mach-Zehnder interferograms are similar in many respects, there are some advantages in using holographic interferograms for this project.

Holographic interferometry has the very important property of cancellation of pathlength errors. In a classical interferometer, the interfering light waves are divided spatially. That is, the two interfering waves are recorded simultaneously but travel different paths. Any difference in pathlengths in the optical system would give rise to extraneous fringes, which could be eliminated only by using very high quality optics. This is a difficult and expensive requirement to meet when using large optical components such as were used in this experiment. However, in a holographic interferometer, the two interfering object waves are divided temporally. That is, the two interfering waves are recorded at two different times but travel the same path. Since there is no difference in the pathlength of the two interfering object waves in the optical system, only changes in the pathlength between exposures in the flow field are displayed. This makes it possible to use relatively low quality 2-ft-diameter mirrors in the optical system.

A double-exposure holographic interferogram is made by holographically recording two object waves sequentially in time. The first exposure records the undisturbed object wave (Fig. 4a), and the second exposure records the disturbed object wave (Fig. 4b). The irradiance at the film plane is

$$\begin{aligned} I &= |U_R + U_{o1}|^2 + |U_R + U_{o2}|^2 \\ &= 2I_R + I_{o1} + I_{o2} + U_R^*(U_{o1} + U_{o2}) + U_R(U_{o1} + U_{o2})^* \end{aligned} \qquad (4)$$

where U_{o1} represents the undisturbed object wave, and U_{o2} represents the disturbed object wave. As in single-exposure holography, the last two

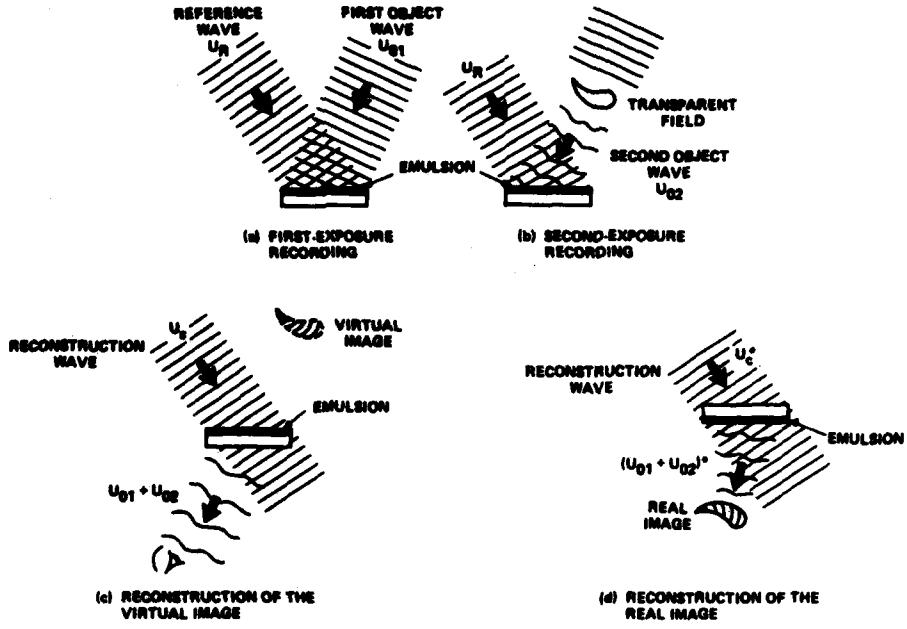


Figure 4. Double-exposure holographic interferometry.

terms of Eq. (4) represent the interference pattern recorded on the film, which contains both amplitude and phase information.

The amplitude transmittance of the developed film, called a holographic interferogram, is proportional to the irradiance

$$t = \beta(2I_R + I_{o1} + I_{o2}) + \beta U_R^*(U_{o1} + U_{o2}) + \beta U_R(U_{o1} + U_{o2})^*$$

When the reconstruction wave U_c illuminates the interferogram as shown in Fig. 4c, the transmitted light is

$$U_I = U_c t = \beta U_c(2I_R + I_{o1} + I_{o2}) + \beta U_R^* U_c(U_{o1} + U_{o2}) + \beta U_R U_c(U_{o1} + U_{o2})^* \quad (5)$$

The second term of Eq. (5), $\beta U_R^* U_c(U_{o1} + U_{o2})$, is the composite virtual image reconstruction of the two object waves $(U_{o1} + U_{o2})$. The composite real-image reconstruction can be obtained by rotating the interferogram 180° about the vertical axis (Fig. 4d).

Recall that the irradiance of the reconstructed wave is proportional to

$$I = |U_{o1} + U_{o2}|^2$$

which can be written as (6,7)

$$I = 2A^2[1 + \cos(\Delta\phi)] \tag{6}$$

where A is the amplitude of the object waves, and $\Delta\phi$ is the optical pathlength difference between the two object waves. Equation (6) represents an interferogram with a fringe pattern $2[1 + \cos(\Delta\phi)]$. Dark and bright fringes are contours of constant values of $\Delta\phi$, where $\Delta\phi$ is given by

$$\Delta\phi = \int_s k(\rho - \rho_o)ds = N\lambda \tag{7}$$

and where k is the Gladstone-Dale constant, which is a property of the gas ($k = 0.225$ m^3/kg for air); ρ and ρ_o are density values for the disturbed and undisturbed air; N is the fringe order number; and λ is the wavelength of the ruby laser ($\lambda = 0.6943$ μm). Equation (7) shows that the fringe patterns are the integrated values of local density differences $(\rho - \rho_o)$ along the path length s. An interferogram is thus a two-dimensional projection of a three-dimensional field. For an asymmetric three-dimensional flow field, several interferograms at various azimuthal angles must be analyzed to invert the integral to obtain the density value at a specific point.

EXPERIMENTAL SETUP AT THE ANECHOIC HOVER CHAMBER

The basic optical system required to record interferograms of the flow field around a rotor tip is shown in Fig. 5. The actual test setup in the Anechoic Hover Chamber at the Aeromechanics Laboratory is shown in Fig. 6. A pulsed ruby laser, with a very short pulse width, is used to "freeze" the rotor at any desired azimuthal angle. The object wave is expanded by a beam expander and collimated by a large spherical mirror to a 2-ft-diameter beam as it passes through the region of interest. The reference wave is lengthened (the difference in path lengths of the object wave and the reference wave must remain within the 7-ft coherence length of the ruby laser), expanded, and collimated. The two waves are brought together at the photographic film plate.

The first exposure is recorded by exposing the film plate while the blade is stationary. Ideally, the air in this case has no velocity and a uniform density distribution. The second exposure is recorded by exposing the same film plate while the blade is rotating. The nonhomogeneous density distribution introduces phase changes in the second object wave which give rise to an interference pattern on the film plate. The photographic plate is then removed from the Hover Chamber, developed, and reconstructed in a convenient place with a continuous-wave He-Ne laser.

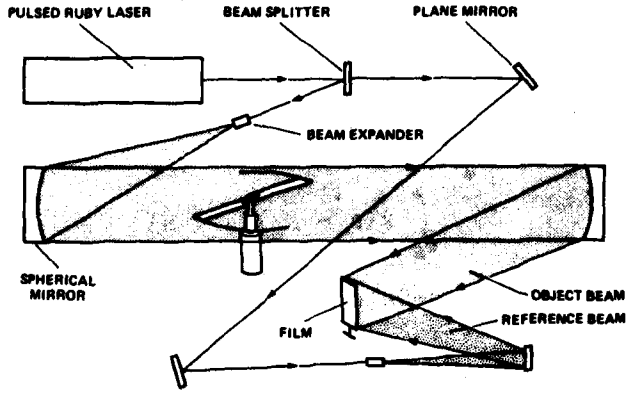


Figure 5. Basic holographic interferometry system.

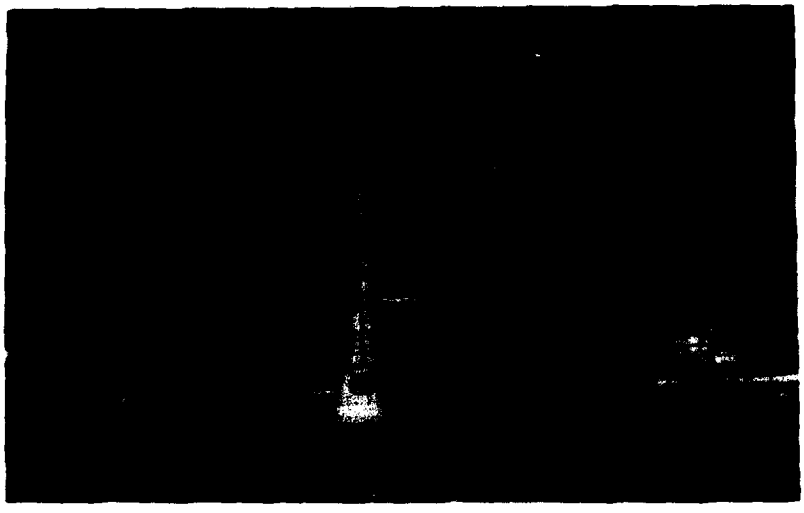


Figure 6. Actual holographic interferometry setup at the Anechoic Hover Chamber.

EXPERIMENTAL RESULTS

Effect of Tip Mach Number

Shock formation and its connection to the far field is very important in understanding the acoustic waveform transition, as shown in Fig. 1. The progress of shock formation can be seen in double exposure interferograms at various tip Mach numbers (Fig. 7). For $M_T = 0.88$ there is very little

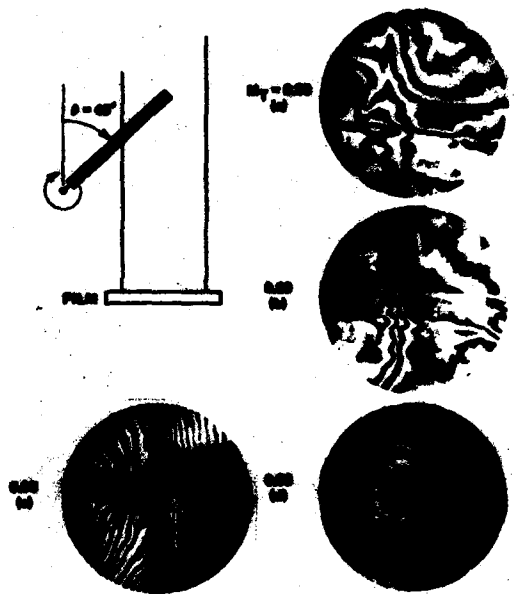


Figure 7. Interferograms at various tip Mach numbers.

shock extension beyond the blade tip (Fig. 7a). For $M_T = 0.89$ the shock is about to extend to the far field (Fig. 7b). For these conditions the local shock is confined to the neighborhood of the blade (Fig. 2a), while the far-field acoustic waveform generally maintains a symmetric shape (Figs. 1a and 1b). However, for $M_T = 0.90$ the shock has extended to the far field (Figs. 7c and 2b). The characteristics of the shock structure can now propagate to the far field, generating an abrupt change in the acoustic waveform. A discontinuity on the compression side of a sawtooth waveform results (Fig. 1c). This discontinuity produces extremely annoying high-frequency components of sound. For $M_T = 0.93$ the shock has strengthened in intensity and moved to near the trailing edge of the rotor (Figs. 7d and 2c). The acoustic waveform widens while maintaining a discontinuity on the compression side. Note that the shock appears to move inboard for increasing tip Mach numbers in Fig. 7. This is because the shock bends farther backward for higher tip speeds (Figs. 2b and 2c).

Interferograms at Various Azimuthal Angles

As previously mentioned, to determine the density value at a specific point in the flow field, Eq. (7) must be inverted. This inversion can be performed only if several interferograms at various azimuthal angles around the flow field are analyzed. Information from these interferograms will be combined with a computer tomography algorithm to determine the density value at any point of interest in the flow field. Figure 8 shows several

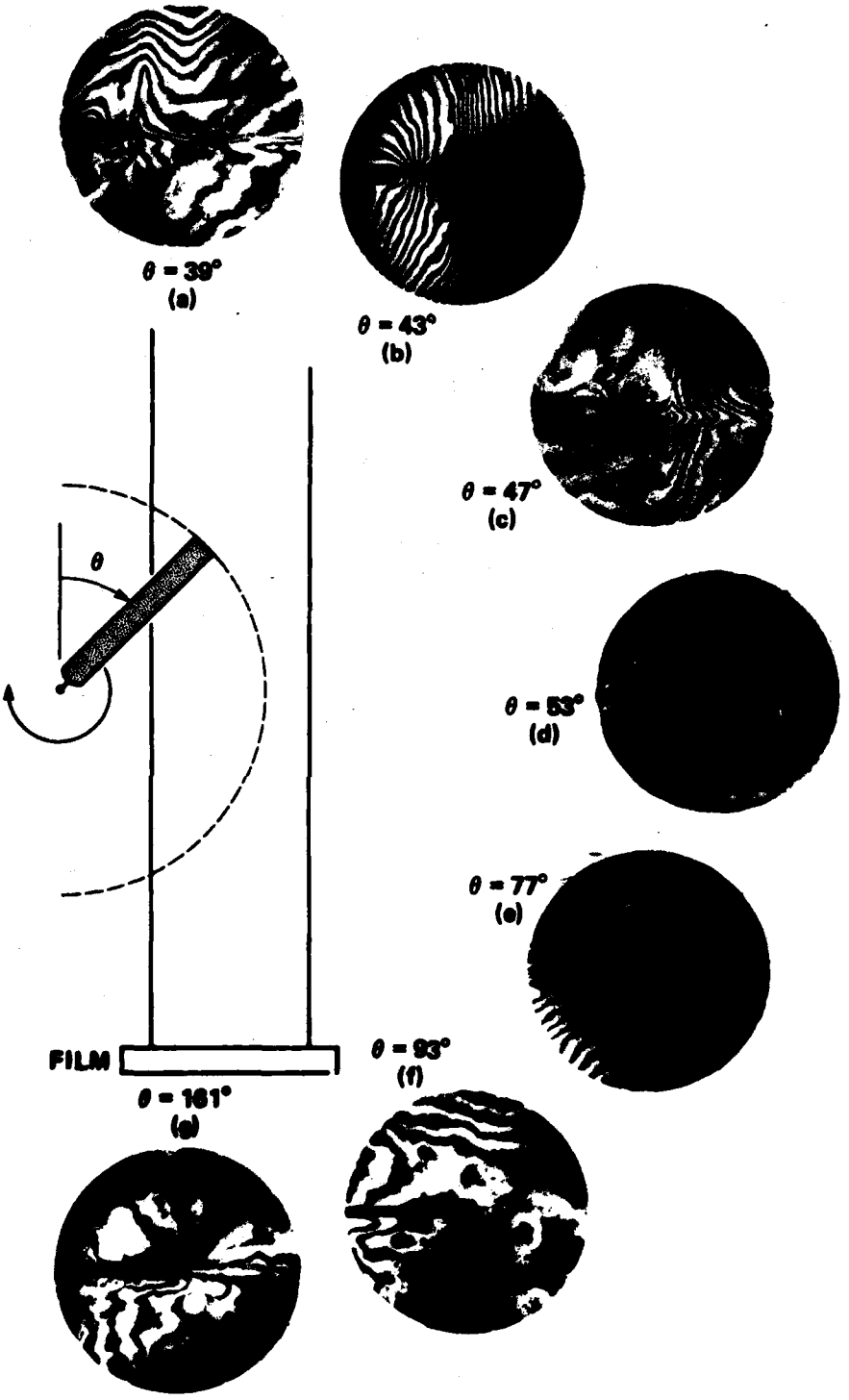


Figure 8. Interferograms at various azimuthal angles.

interferograms at various azimuthal angles for a rotor with $M_T = 0.90$. The shock extension to the far field can be seen in Figs. 8a-8e but not in Figs. 8f and 8g. This is basically a result of part of the shock being parallel to the optical path (creating a large phase change) in Figs. 8a-8e, while the shock is generally perpendicular to the optical path (creating very little phase change) in Figs. 8f and 8g.

Wake System and Tip-Vortex Trajectory

The tip-vortex geometrical position and strength play the dominant role in studies of blade-vortex interactions. This interaction generates unsteady pressure fluctuations over the blade, contributing to the impulsive nature of the acoustic wave and to the severe vibration felt by helicopter passengers. Figure 9 shows the details of tip vortices generated by two rotor blades. This is an interferogram where $M_T = 0.4$ and the collective pitch angle is 8°. The downwash and contraction of the tip vortices obtained from this interferogram agree reasonably well with hot-wire measurements (8) (Fig. 10).

Three-Dimensional Visualization

Holographic interferometry provides the unique ability to visualize a three-dimensional field with a single interferogram. Although no quantitative information will be obtained, this visualization can aid in understanding the geometry of a field. To illustrate this point, a simple test was carried out as shown in Fig. 11. The model consists of three objects: (1) a double post near the film plate, (2) a heating element and model rotor blade slightly farther from the film plate, and (3) a single post farther still from the film plate. A diffuser was inserted into the object wave and a double-exposure interferogram was recorded. The first exposure was made with the heating element at room temperature, simulating a stationary blade. The second exposure was made with the heating element at a



Figure 9. Double-exposure interferogram of tip vortices at $M_T = 0.4$ and collective pitch angle of 8° (NACA 0012, aspect ratio = 13.7).

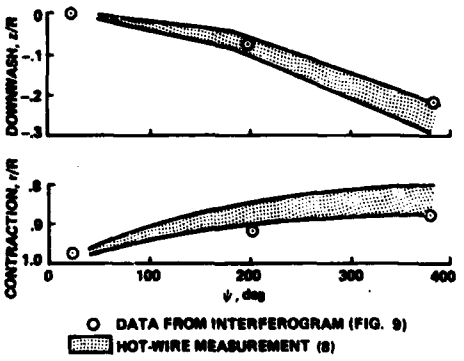


Figure 10. Comparison of hot-wire and holographic interferometry measurement of tip vortices position.

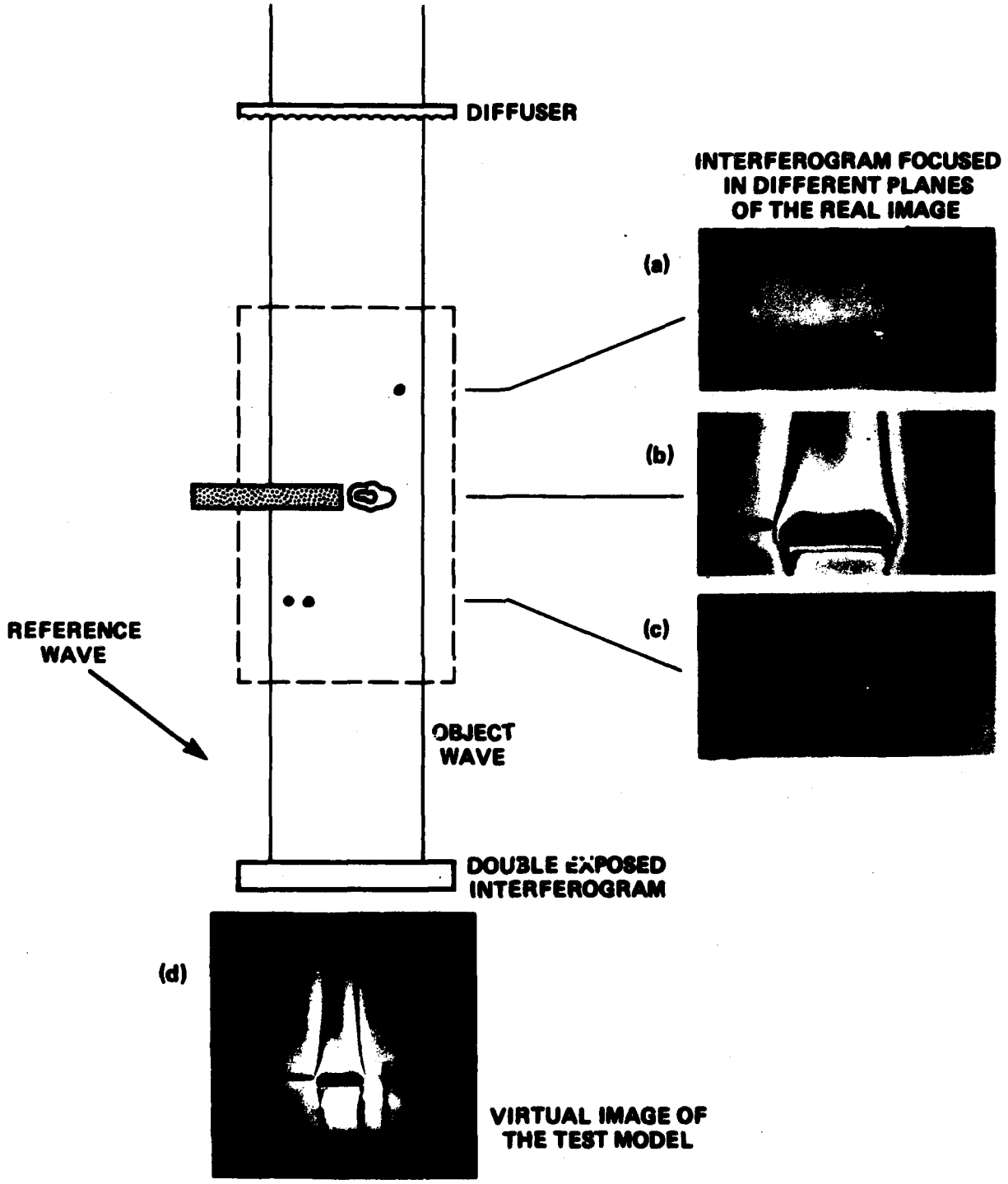


Figure 11. Three-dimensional visualization.

high temperature, creating a nonhomogeneous density distribution in the air and thus simulating a rotating blade. The three-dimensional virtual image was reconstructed and photographed (Fig. 11d). The real image was reconstructed and photographed at three planes in the real-image space (Figs. 11a-11c). It is possible to determine the distribution of density change by noting where the fringes come into focus.

CONCLUDING REMARKS

A holographic interferometry technique has been demonstrated as the most promising method for both visualizing and quantitatively measuring the three-dimensional transonic flow field near a rotor blade. This technique overcomes many limitations of conventional experimental techniques and has yielded some initial results.

Holographic interferograms show the behavior of shock structures near a high-speed rotor tip. Interferograms recorded at various tip Mach numbers (M_T = 0.88 to 0.93) verify the formation of a local shock and its extension to the far field at high tip Mach numbers. Interferograms recorded at various azimuthal angles show the general position and shape of an extended shock structure. The position and shape of such a shock recorded in these interferograms is in general agreement with numerical calculations.

Holographic interferograms also show the geometrical position of tip vortices behind a rotor blade. The characteristic rotor contracted slip stream is clearly visible. The contraction and downwash position of the tip vortices recorded in this interferogram agrees reasonably well with hot-wire anemometry measurements.

In the future, an additional method to visualize the flow in a three-dimensional manner will be employed. Double-exposure holographic interferograms will be recorded with a diffuser added to the optical system. The reconstructed real image can be photographed in any desired plane to identify the position of the shock.

To quantitatively measure the three-dimensional flow field near a high-speed rotor, several interferograms must be analyzed. This is because an interferogram records integrated information and is therefore a two-dimensional projection of the three-dimensional flow field. Interferograms recorded of various azimuthal angles will be processed with a computer tomography algorithm to determine the velocity at a specific point. This experimental quantitative measurement of the three-dimensional transonic velocity field near a rotor blade, with detailed information of the shock structure and wake system, will help the researcher of both acoustics and aerodynamics to design a better blade at high tip speeds.

ACKNOWLEDGMENTS

We thank several of our colleagues at the Aeromechanics Laboratory for their unselfish contributions to this effort, especially Mr. H. Andrew Morse and Dr. Fredric Schmitz for their consistent and untiring support. We also owe a special thanks to Professor Charles Vest of the University of Michigan for many fruitful discussions and guidance.

REFERENCES

1. Boxwell, D.A.; and Schmitz, F. H.: Full-Scale Measurements of Blade-Vortex Interaction Noise. Presented at the 36th Annual Forum of the American Helicopter Society, paper No. 61, May 1980.

2. Schmitz, F. H.; and Yu, Y. H.: Transonic Rotor Noise - Theoretical and Experimental Comparisons. Vertica, vol. 5, 1981, pp. 55-74.

3. Johnson, D. A.; and Bachelo, W. D.: Transonic Flow about a Two-Dimensional Airfoil - Inviscid and Turbulent Flow Properties. AIAA Paper 78-1117, July 1978.

4. Matulka, R. D.; and Collins, D. J.: Determination of Three-Dimensional Density Fields from Holographic Interferometry. J. Appl. Phys., vol. 42, 1971, pp. 1109-1119.

5. Sweeney, D. W.; and Vest, C. M.: Reconstruction of Three-Dimensional Refractive Index Field from Multidirectional Interferometric Data. Appl. Opt., vol. 12, 1973, pp. 2649-2664.

6. Vest, C. M.: Holographic Interferometry. Wiley-Interscience, 1979.

7. Collier, R. J.; Burckhardt, C. B.; and Lin, L. H.: Optical Holography. Academic Press, 1971.

8. Tung, C.; Pucci, S. L.; Caradonna, F. X.; and Morse, H. A.: The Structure of Trailing Vortices Generated by Model Rotor Blades. NASA TM-81316, 1981.

9. Caradonna, F. X.: The Transonic Flow on a Helicopter Rotor. Ph.D. Dissertation, Stanford U., Stanford, Calif., Mar. 1978.

Fluoropolymer Barriers to Stress Corrosion in Optical Fibers

Liliana Klinger
USA CECOM, Ft. Monmouth, NJ 07703

James R. Griffith
Naval Research Laboratory
Washington, DC 20375

Organic polymer coatings play a critical role in determining the overall performance of optical fibers. They protect the very fine silica glass core from abrasive damage and environmental effects such as stress induced corrosion. In addition they act as a buffer in reducing the sensitivity of the optical fiber to microbending loss[1]. Military requirements for optical fiber coatings are more stringent then those for coatings used in commercial applications. No more than two or three coatings may be applied to the fiber so that it remains light enough for possible use in helicopter payout and missile tracking systems. Application of the coating must be done in line with the drawing of the glass fiber; thus low viscosities (10^3-10^4 cps), rapid solidification and fast crosslinking reactions are required. Ultraviolet cures are preferable, but rapid thermal cures are possible[2].

Stress induced corrosion by H_2O is associated with stress concentrations at surface flaws under deformation. The only way to protect the glass fiber against stress corrosion is to prevent the buildup of H_2O and OH at the fiber surface. Polymer coatings for optical fibers must thus be designed to reduce permeation of H_2O through the polymer matrix. Moisture may permeate through the polymer coating(s) while it is being stored under tension (wound on drums) or during deployment on hydrous terrain. The process of permeation of H_2O through a polymer film involves, wetting of the polymer surface[2], sorption of H_2O into the polymer matrix, diffusion through the film along a concentration gradient and desorption of H_2O from the polymer surface[3,4]. Incorporation of fluorine into the polymer chain has been found to reduce surface wettability and sorption of H_2O[3]. In addition, halogenated polymers have been found to have amongst the lowest moisture vapor permeabilities of all polymers (see Table 1)[3].

Two fluorinated polymers were investigated as possible coatings for optical fibers; a heat cured fluoroepoxy (see Fig. 1) and a photocured fluoroacrylate (see Fig. 2). In addition to barrier properties such factors as flexibility, strength and modulus must be considered in designing a suitable coating.

Fluoroepoxy resin is synthesized by mixing the diglycidyl ether with bis-γ-aminopropyl (tetramethyl disiloxane) at 50°C until both reactants become mutually soluble. Films were then solvent cast using Freon TF solvent and cured at 80°C.

Fluoropolyolacrylate is a long chain monomer containing aromatic and oxerane rings; photopolymerization occurs through pendant ester groups. Films were solvent cast using Freon TF, in a nitrogen atmosphere and subsequently exposed to ultraviolet radiation.

In both systems fluorine is incorporated into the polymer chain as fluoromethyl groups. The C_6 fluoroepoxy has an additional fluorinated hexane attached to the aromatic ring. The long chain adds flexibility whereas the aromatic ring gives strength and resiliency.

Determining the effects of chemical and physical structure on moisture vapor permeability of these films requires balancing several conflicting factors; these include effects of polar groups, fluorinated substituents, symmetry and segmental mobility. Polar groups which improve adhesion to glass and prevent buildup of H_2O at the interface also increase moisture vapor permeability. This is due to strong localized interactions between the small polar H_2O molecule and polar groups within the polymer[4]. Although the silicone amine curing agent reduces the fluorine content and thus barrier properties of the epoxy film, it acts in a manner similar to silane coupling agents in improving adhesion to glass. Both low moisture vapor permeability and good adhesion are important factors in preventing stress corrosion of the optical fiber.

Wettability studies are reported in terms of the contact angle, θ_{H_2O}, made by triply distilled water on fluoroepoxy + fluoropolyol-acrylate films. Contact angles were measured at ambient temperature on an NRL contact-angle goniometer. Critical surface tensions of wetting γ_c were determined from contact angles of a series of n-alkanes.[5] The cosines of the contact angles of the n-alkanes were plotted against the surface tensions γ_{LV} of these liquids and the intercept of the resultant straight line at $\cos\theta = 1$ ($\theta = 0^o$) is denoted as γ_c (see Fig. 3).

Moisture vapor permeability was determined using the cup method (Gardner-Park Permeability Cup - ASTM Standard D1653-72) maintained in a relatively anhydrous environment (15% ± 1 humidity) at 22°C ± 1. A thin polymer film was placed over an aluminum cup containing 7 ml of distilled H_2O; the entire system was weighed and placed in a dessicator. It was subsequently weighed every 24 hours to determine how much H_2O had permeated the polymer sample. Weight was plotted against time; when the system reaches a steady state all points lie on a straight line (see Fig. 4).

Preliminary results (see Table 2) indicate that despite the difference in contact angle and surface tension, moisture vapor permeability of the two films remain about the same. As mentioned earlier, permeability is a composite quantity governed by two independent parameters, diffusion and solubility of H_2O in the polymer. Sorption studies on various fluoroacrylate and fluoroepoxy systems indicate that the acrylates absorb less H_2O than the epoxies[6]. Diffusion properties have as yet not been determined. This implies that the low moisture vapor permeability of the polyfluoropolyolacrylate film is in part due to low H_2O absorption given the superior surface properties of the fluoroepoxy film.

Polymer coatings can be designed to reduce stress concentrations at surface flaws as well as sensitivity to microbending loss. (Small external forces due to an irregular surface are sufficient to cause lateral deformations, mode coupling and optical loss.) The elastic or relaxation modulus E of the polymer is a good indicator of how effective the coating is in protecting the clad fiber from these effects. In order to determine their elastic modulus E and glass transition temperature T_g, dynamic mechanical tests were performed on the fluoroepoxy (see Fig. 5) and fluoropolyolacrylate polymer films (see Fig. 6) using a Rheo-Vibron DDV-II. It has been observed that coatings with relaxation modulus values less than 10^8 dynes/cm^2 are effective in protecting the glass fiber from stress concentrations and reducing microbending loss.

Both fluoroepoxy and fluoropolyolacrylate polymers respond in a similar manner to changes in temperature under the applied oscillatory force of the Rheo-Vibron. This is probably due to similarities in chemical structure. At ambient temperature (22°C) the elastic modulus of the fluoroepoxy polymer is 2×10^{10} dynes/cm^2, while that of the polyfluoropolyolacrylate is 4×10^{10} dynes/cm^2. The glass transition or drop in modulus of both systems, to about 10^8 dynes/cm^2, occurs at about 50°C. It is possible to reduce the temperature of the glass transition and the modulus below Tg by reducing the crosslink density of the polymer network[7].

In order to reduce the crosslink density of the polyfluoropolyolacrylate, it would be necessary to react some of the pendant OH groups with nonreactive chemical groups (such as dimethyl sulfate or pentafluorophenylisocyanate) before attaching the ester group ($-O-\overset{O}{\overset{\|}{C}}-C=C$) that participates in the crosslinking reaction. Since this procedure would also reduce the modulus below T_g, we would obtain a polymer with a modulus less than 10^8 dynes/cm^2 at ambient temperatures, as required.

In designing a coating system for optical fibers it is necessary to strike a balance between two factors: a soft compliant enclosure that acts as a buffer in masking the nonuniform surface of the glass and a hard enclosure to act as a stiffener and make the fiber resistant to the environmental surface. These requirements are best suited by using a photocured polyfluoropolyolacrylate of reduced crosslink density as a primary coating and the high modulus fluoroepoxy resin as a secondary coating. Although it is possible to lower the glass transition and modulus of the fluoroepoxy system, the longer cure times required for heat curing systems, make them less attractive as a primary coating. Also the superior surface properties of the fluoroepoxy suggest themselves for use as an external coating.

References

1. Blyler, Jr., L. L., Eichenbaum, B. R. and Schonhorn, H., Coatings and Jackets. In "Optical Fiber Telecommunications" (S.E. Miller and A. G. Chynoweth, eds), Chapter 10, Academic Press (1979), New York.

2. Blyler, Jr., L. L., Polymer News 8 (1), 6 (1981).

3. Salame, M., Steingiser, S., Polym-Plast. Technol. Eng., 8(2), 155-175 (1977).

4. Barrie, J. A., Water in Polymers. In "Diffusion in Polymers" (J. Crank and G. S. Park, eds), Chapter 8, Academic Press (1968), London.

5. Hunston, D. L., Griffith, J. R., Bowers, R. C., Ind. Eng. Chem. Prod. Res. Dev., 17 (1), 10 (1978).

6. Griffith, J. R., Bultman, J. D., Ind. Eng. Chem. Prod. Res. Dev., 17 (1), 8 (1978).

7. Ferry, J. D., "Viscoelastic Properties of Polymers", Chapter 10, John Wiley & Sons, Inc., New York.

Table 1

Permeability Properties of High Barrier Polymers

Polymer	Polymer Class	Water[b]
Polyvinylidene chloride	Halogen	0.2
Lopac[a]	Nitrile	5.0
Barex	Nitrile	6.1
Cycopac	Nitrile	5.0
Saran Wrap	Halogen	0.5
Epoxy (bisphenol A/amine cure)	Thermoset	4.0
Kel-F (PCTFE)	Halogen	0.3
Trogamid T	Polyamide	5.2
Kynar (PVDF)	Halogen	1.0
Nylon 69-nylon 610	Polyamide	6.0
Polyethylene terephthalate	Polyester	2.0
Phenoxy	Condensate	4.5
PVC (rigid)	Halogen	2-3[c]

[a]Acrylonitrile-styrene copolymer used for manufacturing Monsanto Cycle-Safe containers (70% AN).

[b]g/24 hr/100 in.^{2}/mil at 100°F (100% RH).

[c]Depends on exact compound formulation.

Table 2

Surface Properties and Permeability

Film (3 mil)	H_2O	c (dyne/cm)	MVP g/(25 cm^2-24 hrs)
Fluoroepoxy	81°	18.3	.033
Fluoropolyolacrylate	71°	25.4	.031

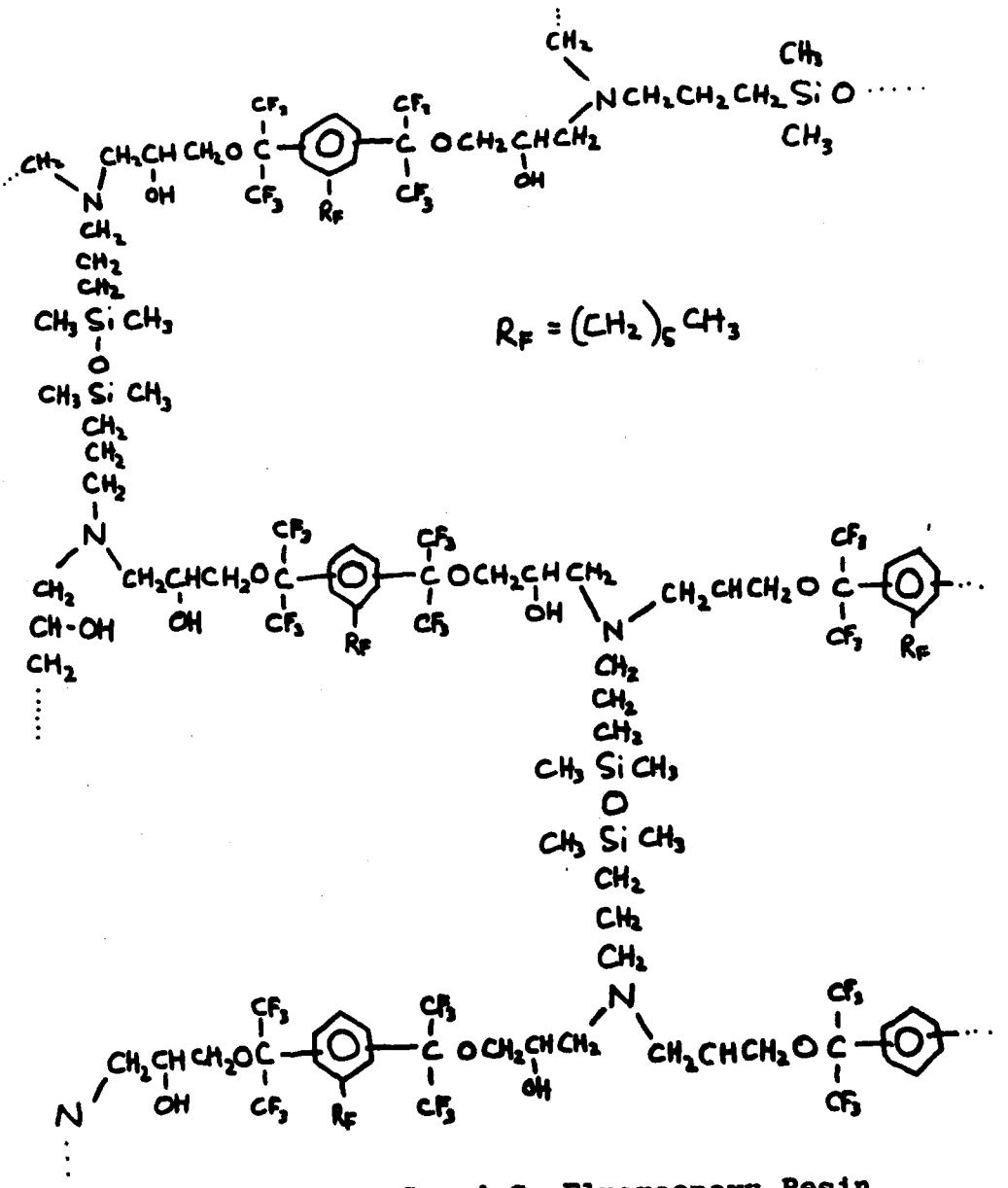


Fig. 1: Silicone Amine Cured C_6 Fluoroepoxy Resin

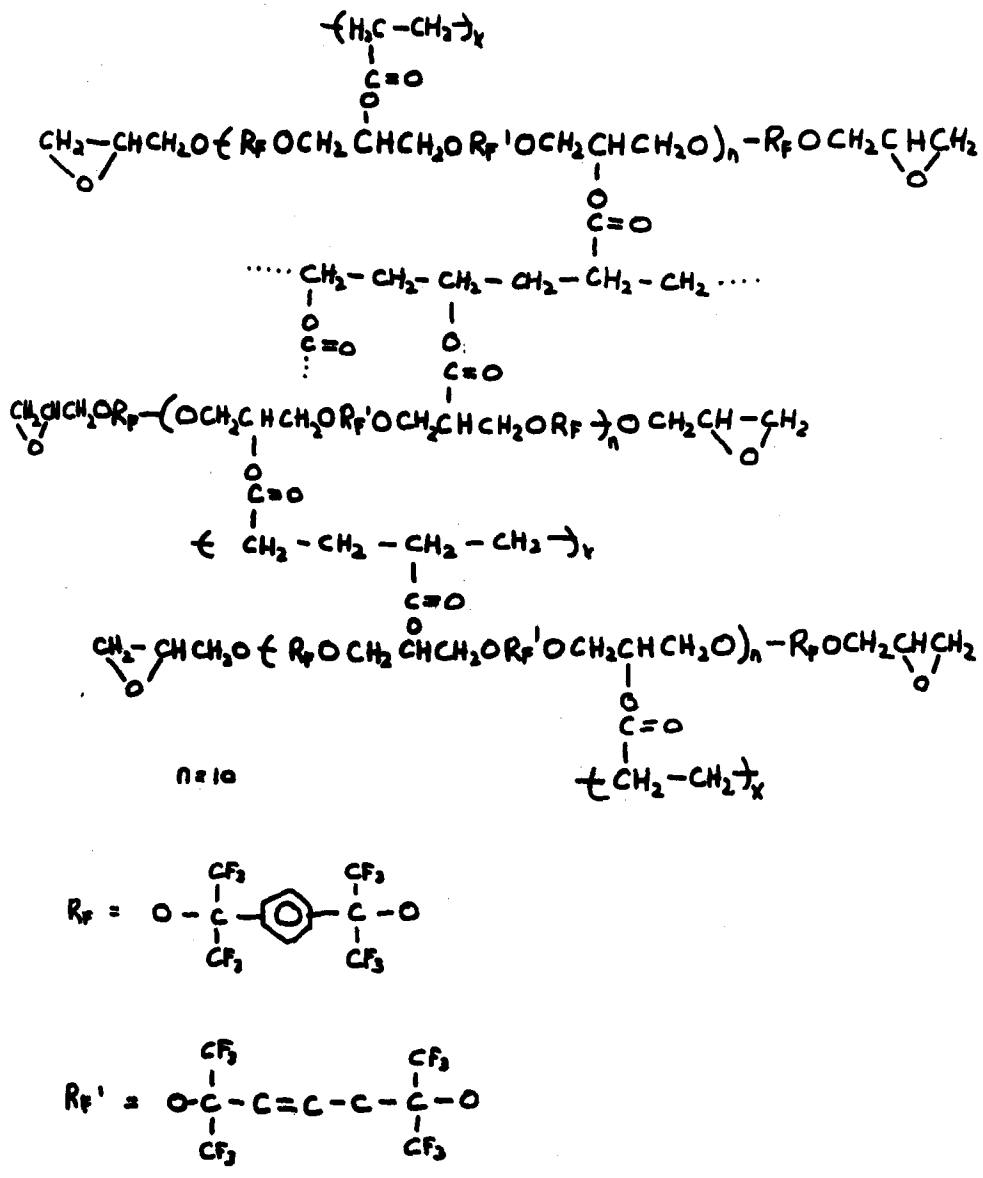


Fig. 2: Poly(fluoropolyolacrylate)

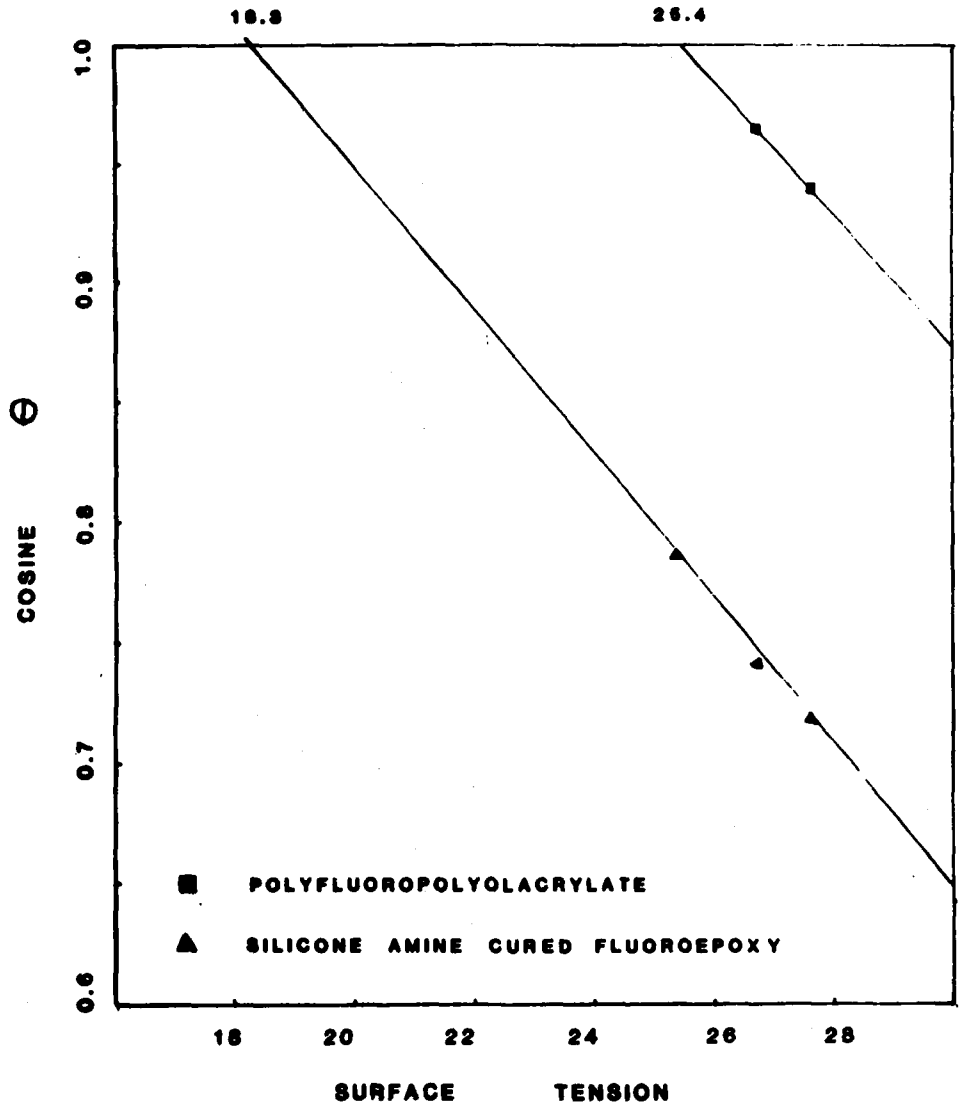


Fig. 3: Plot of Cosine θ vs. Surface Tension for N-Alkanes on Fluoroepoxy and Fluoroacrylate Polymer Films.

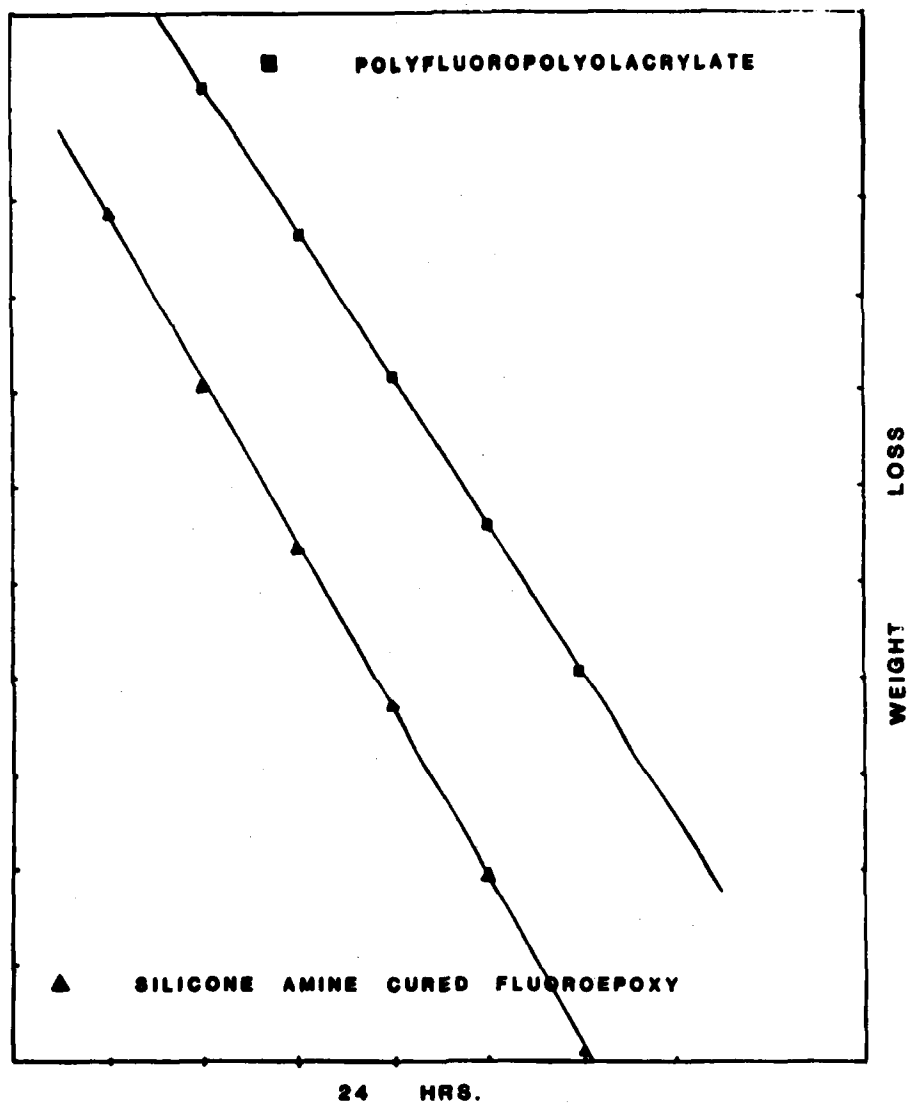


Fig. 4: Plot of Weight Loss Due to H_2O Permeation vs. 24 hr. Interval.

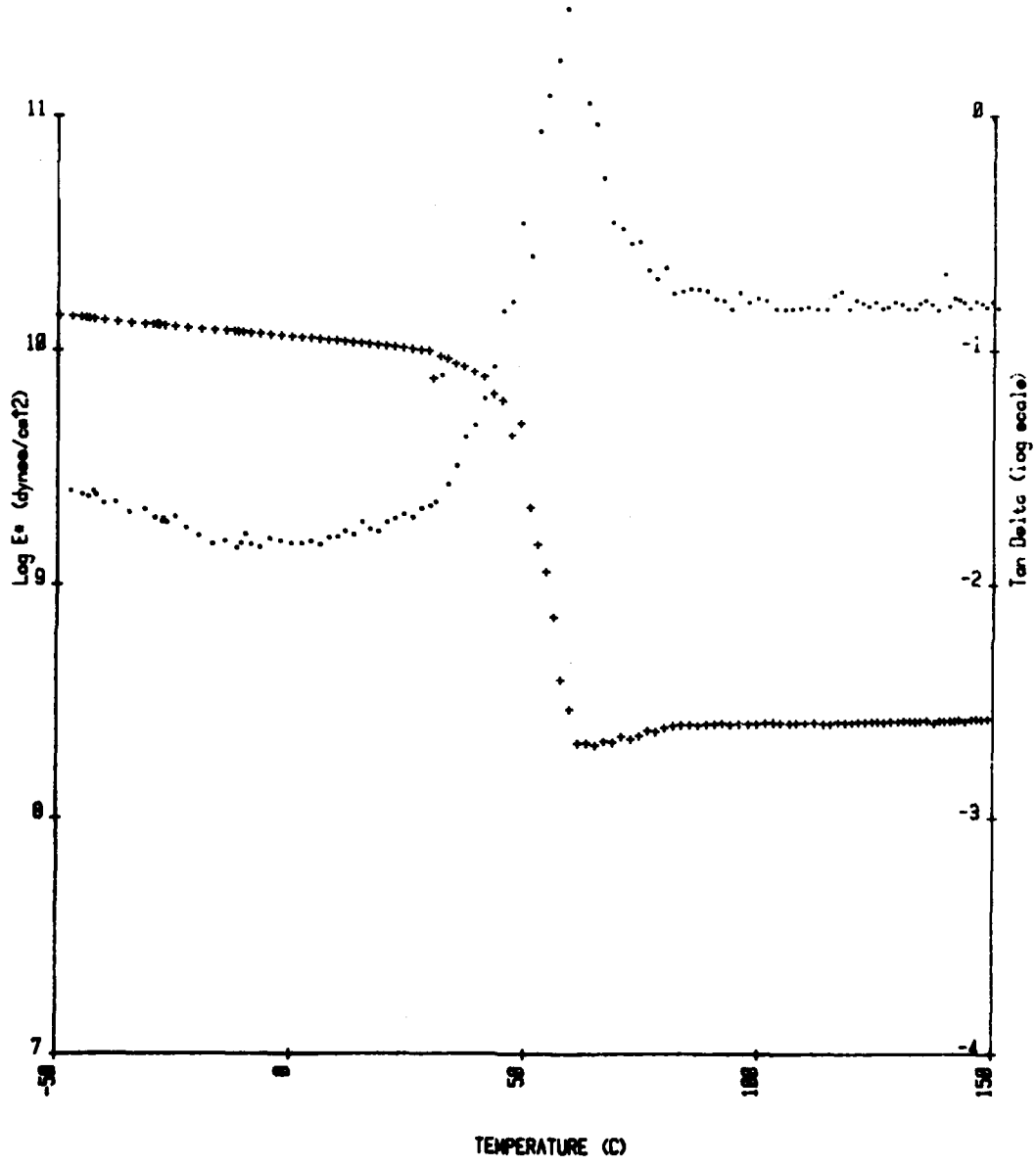


Fig. 5: The Dynamic Mechanical Properties as a Function of Temperature of Silicone Amine Cured C_6 Fluoroepoxy.

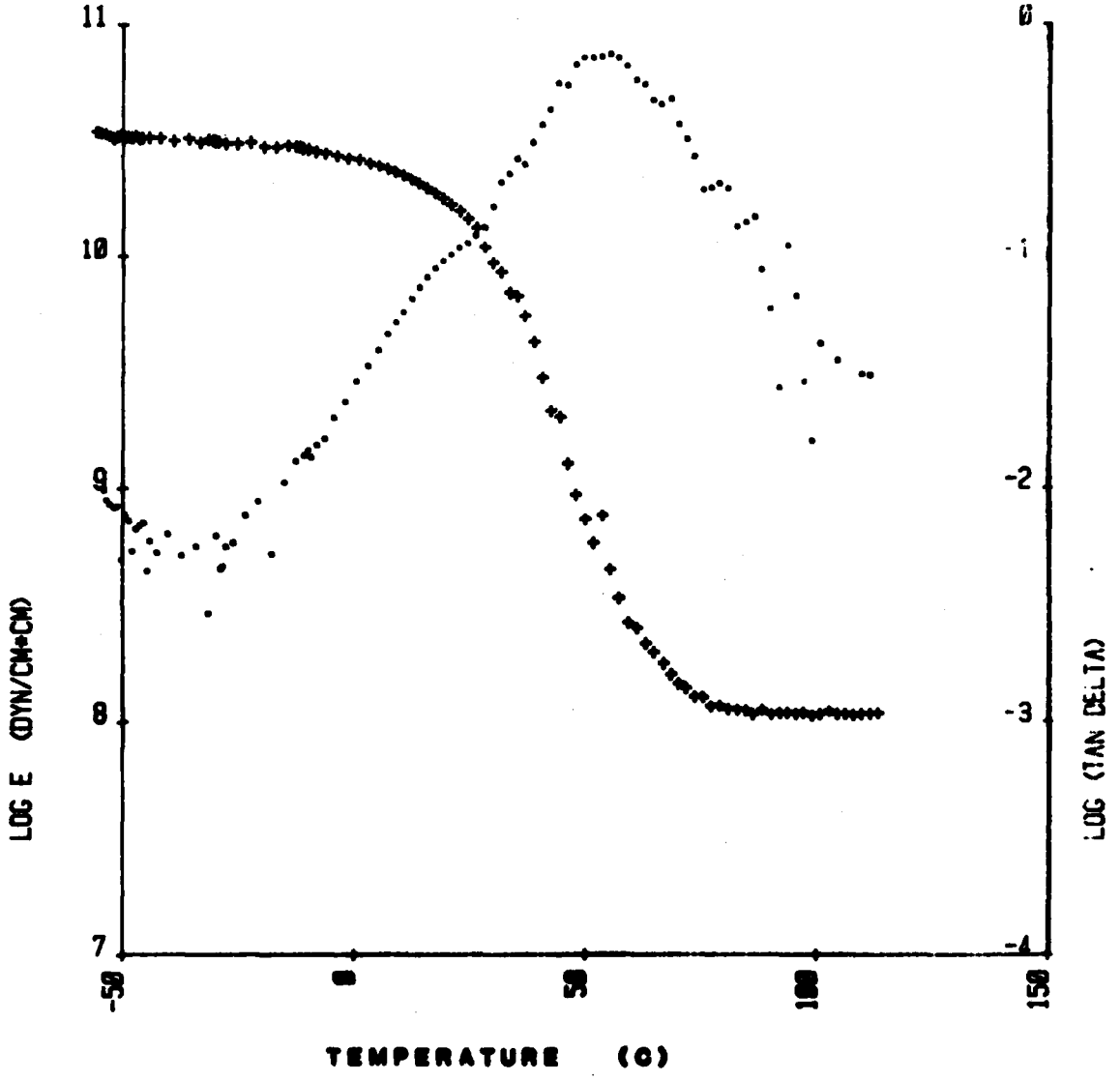


Fig. 6: The Dynamic Mechanical Properties as a Function of Temperature of Polyfluoropolyolacrylate.

TACTICAL GAMMA AND FAST NEUTRON DOSIMETRY WITH LEUKO DYE OPTICAL WAVEGUIDES (U)

STANLEY KRONENBERG, Ph.D.
U.S. ARMY ELECTRONICS TECHNOLOGY AND DEVICES LABORATORY (ERADCOM)
FORT MONMOUTH, NEW JERSEY 07703

INTRODUCTION

The change of color induced by nuclear radiation in radiochromic leuko dyes can be used to measure radiation doses (1,2). To significantly increase the sensitivity of such dosimeters, the layer of the dye-containing solution must be made very thick. To solve this problem in a practical manner we constructed optical waveguides (fiber optics) whose core is a solution of radiochromic leuko dye and the outside layer a thin teflon or FEP tubing. Proper choice of solvents with a refractive index greater than the refractive index of the wall results in total reflection of the light traveling through the core, thus satisfying the waveguide condition. This approach is analogous to our earlier approach to increase resistivity of glass dosimeters by using fiber optics.(3).

RADIATION EFFECTS ON THE INDEX OF REFRACTION OF THE DYE SOLUTION

While experimenting with the radiochromic waveguide dosimeters we observed an effect which makes the leuko dye dosimetry in waveguides operationally different from leuko dye dosimetry with bulk solution and which can be exploited for construction of dosimeters with a very wide dynamic range.

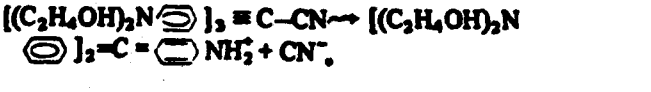
For example, an irradiated solution of hexahydroxyethyl pararosaniline cyanide appears blue in a test tube but red after being incorporated in a waveguide and observed through its end. This effect can be understood from the theory of anomalous dispersion (4) illustrated qualitatively in Fig. 1. Exposure of the dye solution to radiation induces an absorption band. This absorption band spans the region between the red and the blue, and results in an increase in the value of the refractive index (n) of the solution at the longer wavelengths and a decrease in n at shorter wavelengths. At shorter wavelengths, where n of the irradiated core solution drops below the value of n of the cladding, the waveguide condition no longer applies

and the device ceases to be a propagator of light through its core. In that wavelength region the light is not channeled through the waveguide, but either escapes axially or is absorbed by the cladding.

The result is that irradiation of the waveguide makes it virtually opaque for a certain optical frequency band whose width increases with increasing dose. The observed resulting dynamic range of such dosimeters covers at least six orders of magnitude.

MATERIALS AND CONSTRUCTION OF DOSIMETERS

We have chosen hexahydroxyethyl pararosanile cyanide (molecular weight 578.7) from among commercially available triphenylmethane radiochromic leuko dyes.



This leuko dye changes to deep blue by exposure to ionizing radiation. It is also sensitive to ultraviolet light and must be shielded from it. In slightly acidic polar solvents containing oxygen there is no back reaction, thus ensuring stability (no fading) of the radiation-induced color change.

Polar solvents which can be used are dimethyl sulfoxide ($(CH_3)_2SO$), triethyl phosphate ($(C_2H_5)_3PO_4$) or N,N-dimethyl formanide ($HCON(CH_3)_2$). The later two solvents are liquid between -56.4 C and +153 C; the indices of refraction are 1.48, 1.41 and 1.43, respectively. A 5-20% solution of the dye is drawn into a tube of "FEP", a teflon-like flexible thermoplastic whose refractive index is 1.34. We used tubing with 3.0 mm outer diameter and 2.3 mm inner diameter.

Two approaches were used to terminate the ends of such waveguides. One approach by means of commercially available glass beads is shown in Fig. 2. The beads provide a hermetic seal as well as a lens effect for injecting and extracting light from the waveguide. The second approach uses no explicit termination. Mixing the dissolved dye with vinyl pyrrolidone (with the inhibitor removed) and letting the mixture polymerize results in a stable jelly-like substance which is sufficiently solid to prevent leakage through the ends of the waveguide. This very convenient approach makes it possible to fill a long piece of FEP tubing with the liquid mixture and, after partial polymerization takes place, to cut it up in pieces of desired length.

READING THE DOSE

Dosimetry is accomplished by comparing ratios of transmissivity of the waveguide at two different wavelengths. The same calibration applies for all dosimeters having the same dimensional and chemical parameters.

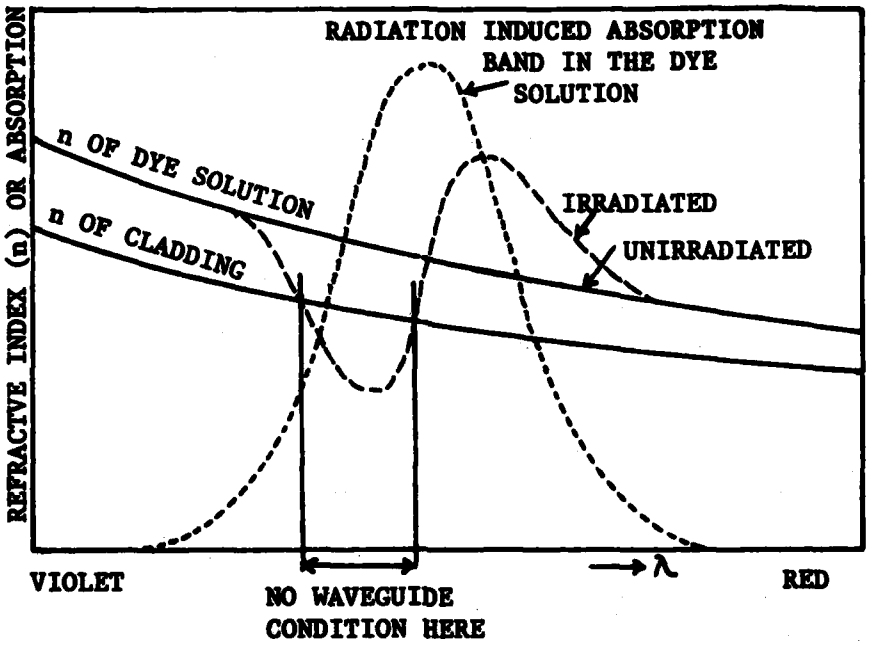


Figure 1. Illustration of radiation-induced anomalous dispersion and its effect on the optical waveguide.

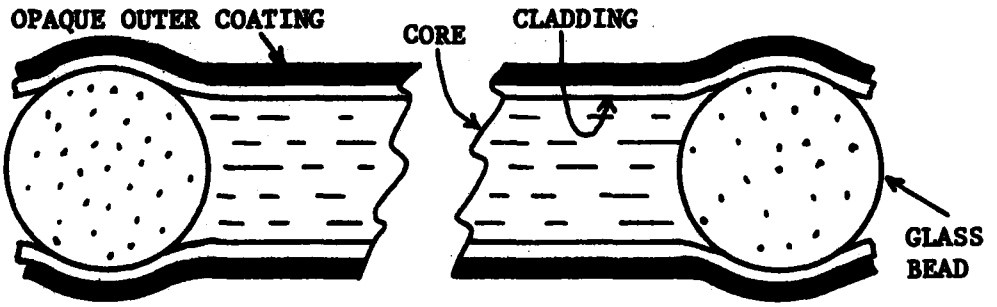


Figure 2. Construction of an optical waveguide whose core is a solution of radiochromic dye.

Figure 3 illustrates a laboratory dosimeter reader. Because the critical measurement is a ratio of optical transmittance values, reproducibility of the waveguide mounting and stability of the instrumentation are not essential. The two parameters which are required to be fixed are the spectral distribution of the light source and the response characteristics of the light detector.

A different approach to dose reading is the use of different spectral bands instead of monochromatic light sources. Figure 4 shows the schematics of such a dosimeter reader which fits in a box 10 x 15 x 5 cm. The spectral output of the three light-emitting diodes used in this device is shown in Fig. 5. To read the dose to which the 11 cm long waveguide was exposed, place it between one of the three LED's and the photo diode and then compare the digital readouts with the light on and the light off. Unfolding the difference between the two readings on a calibration curve yields the dose. The choice of which one of the three LED's should be used depends on the dose to which the dosimeter was exposed.

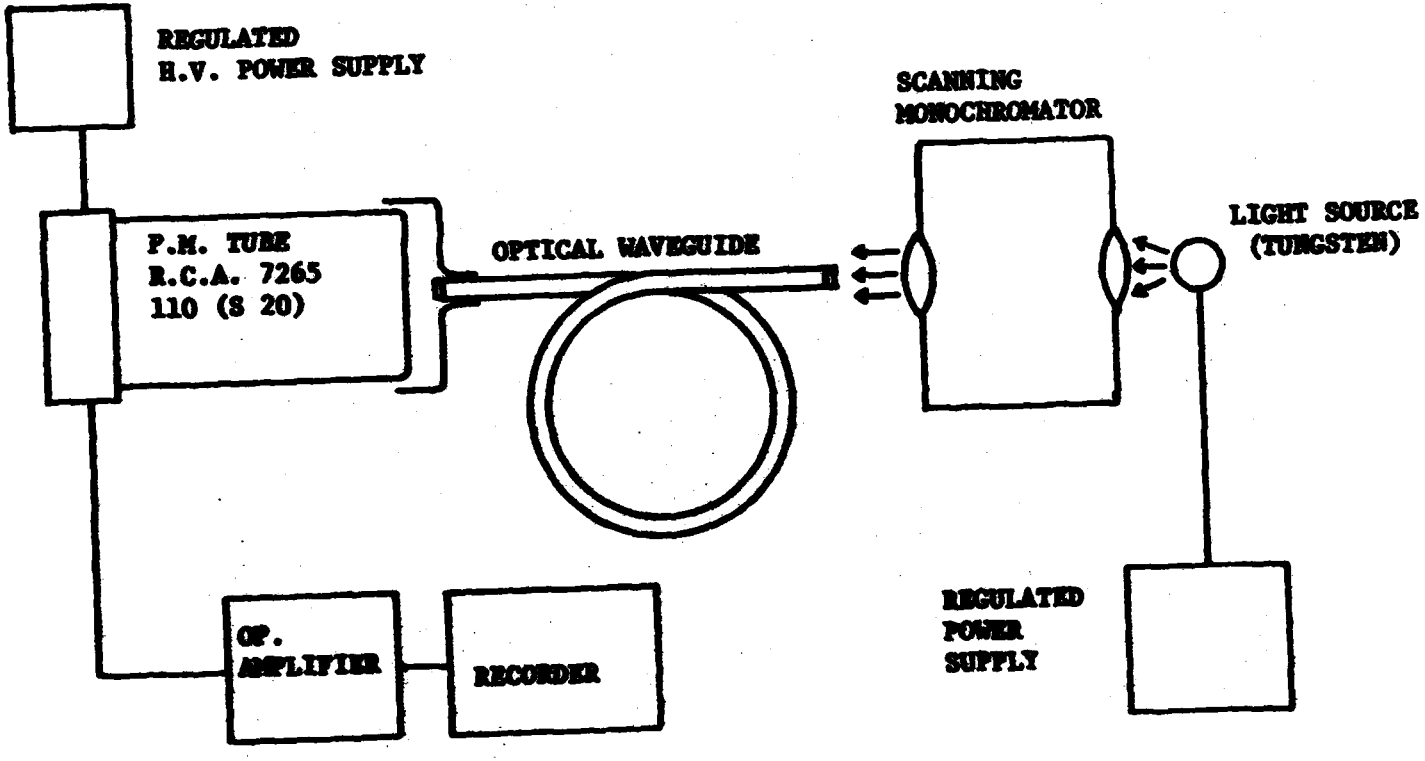
GAMMA RAY DOSE MEASUREMENTS AND GAMMA ENERGY DEPENDENCE

The absorbed dose is measured most accurately by taking the ratios of transmitted intensities before and after exposure at two different wavelengths, calculating the ratio of these ratios and reading the dose from calibration curves or tables. As an example, Fig. 6 shows such a calibration curve for a 14.5 cm long waveguide containing hexahydroxyethyl pararosaniline cyanide in N,N-dimethylformamide (concentration by weight: 0.09651). Electron equilibrium was obtained by surrounding the sample during the irradiation with tissue-equivalent material so that 1 roentgen exposure corresponds to an absorbed dose of 1.54 Rad (tissue). The photocathode was 110 (S-20) and the light source of the scanning monochromator was a quartz halogen lamp.

The dose received by any dosimeter of the same construction can be read by means of this calibration over a dose range of six orders of magnitude by selecting appropriate wavelengths for analysis. All data shown in Fig. 6 was obtained from a single waveguide by measuring its transmittance after consecutive exposures to gamma rays. The effect of bleaching by radiation, which is visible in Figure 6 above 2 megaroentgen, exists only in liquid dye solutions; it does not occur in solid solutions.

The quantum energy response to gamma rays of these dosimeters is determined mainly by the composition of the solvent and of the cladding. For example, if the dosimeter is kept in a container made of tissue-equivalent material during exposure to gamma radiation under conditions approximating electron equilibrium, the response will be close to that of tissue over a broad range of photon energies.

To obtain rad readings in materials other than tissue it is practicable to expose the solution not in a waveguide but in a container made of the material in question. The solution should be placed in a layer sufficiently thin to assure the applicable electron equilibrium according to



Figur Apparatus for recording relative transparency of waveguides versus wavelength for visible and near infrared light.

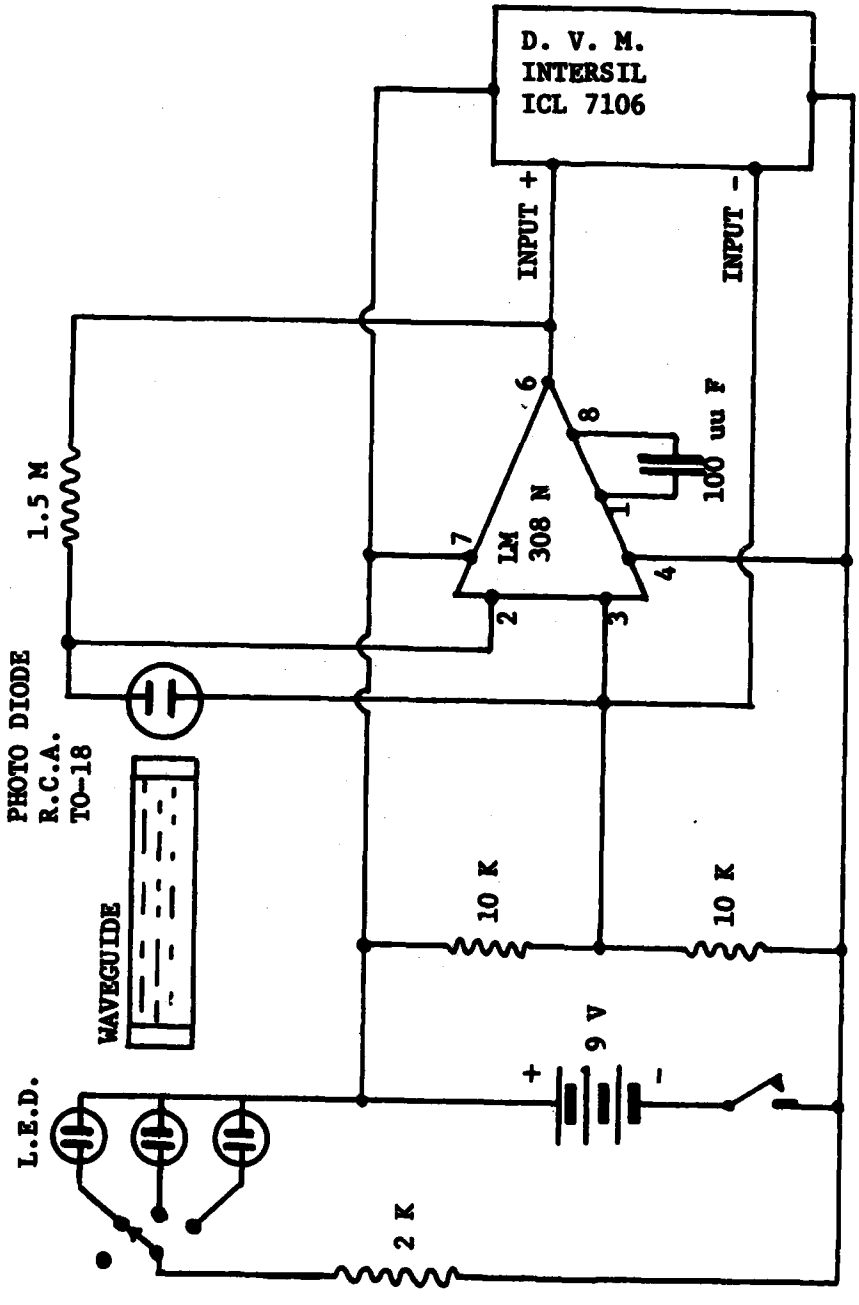


Figure 4. Waveguide dosimeter reader.

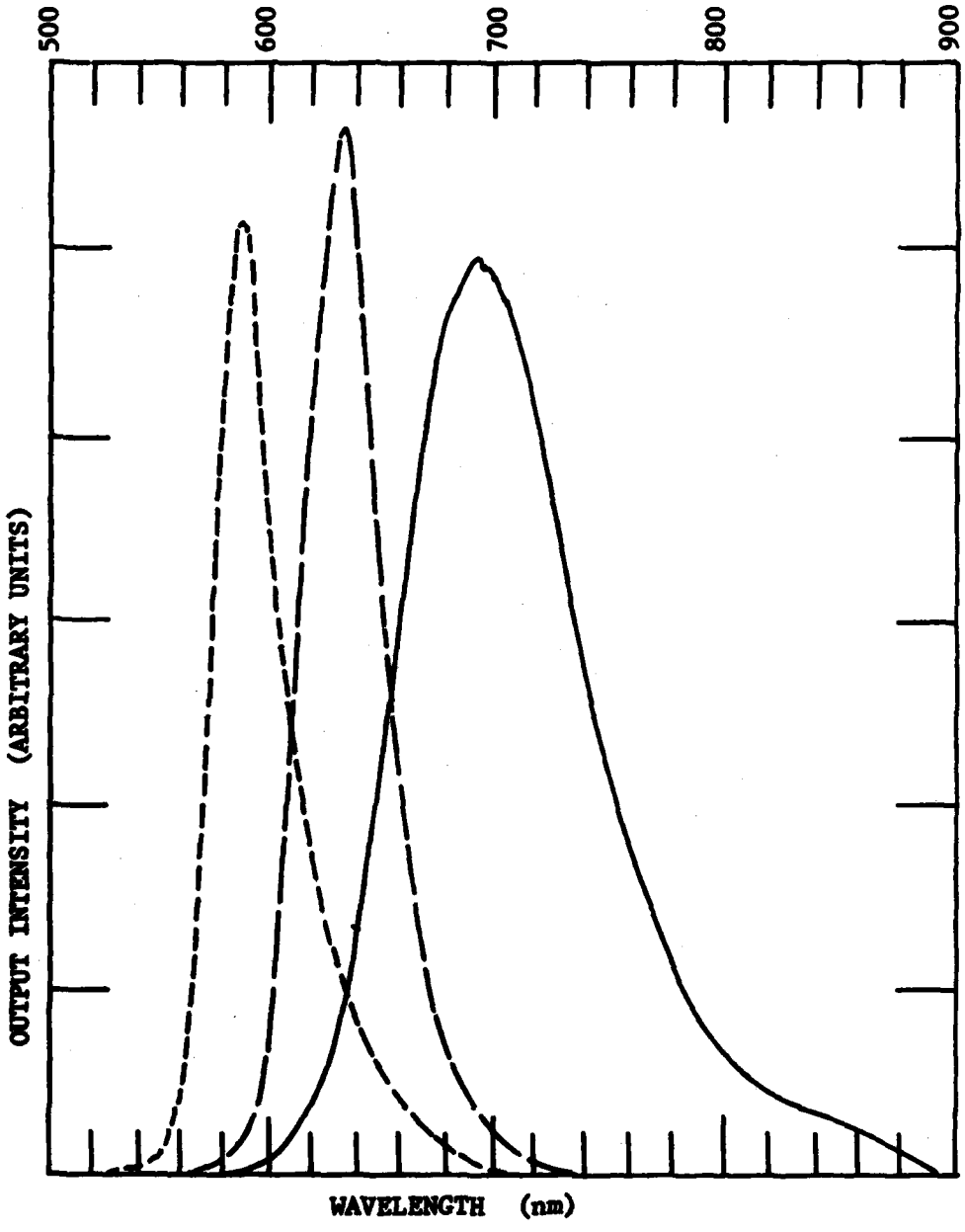


Figure 5. Spectra of the outputs of the three LEDs used in the dosimeter reader illustrated in figure 4.

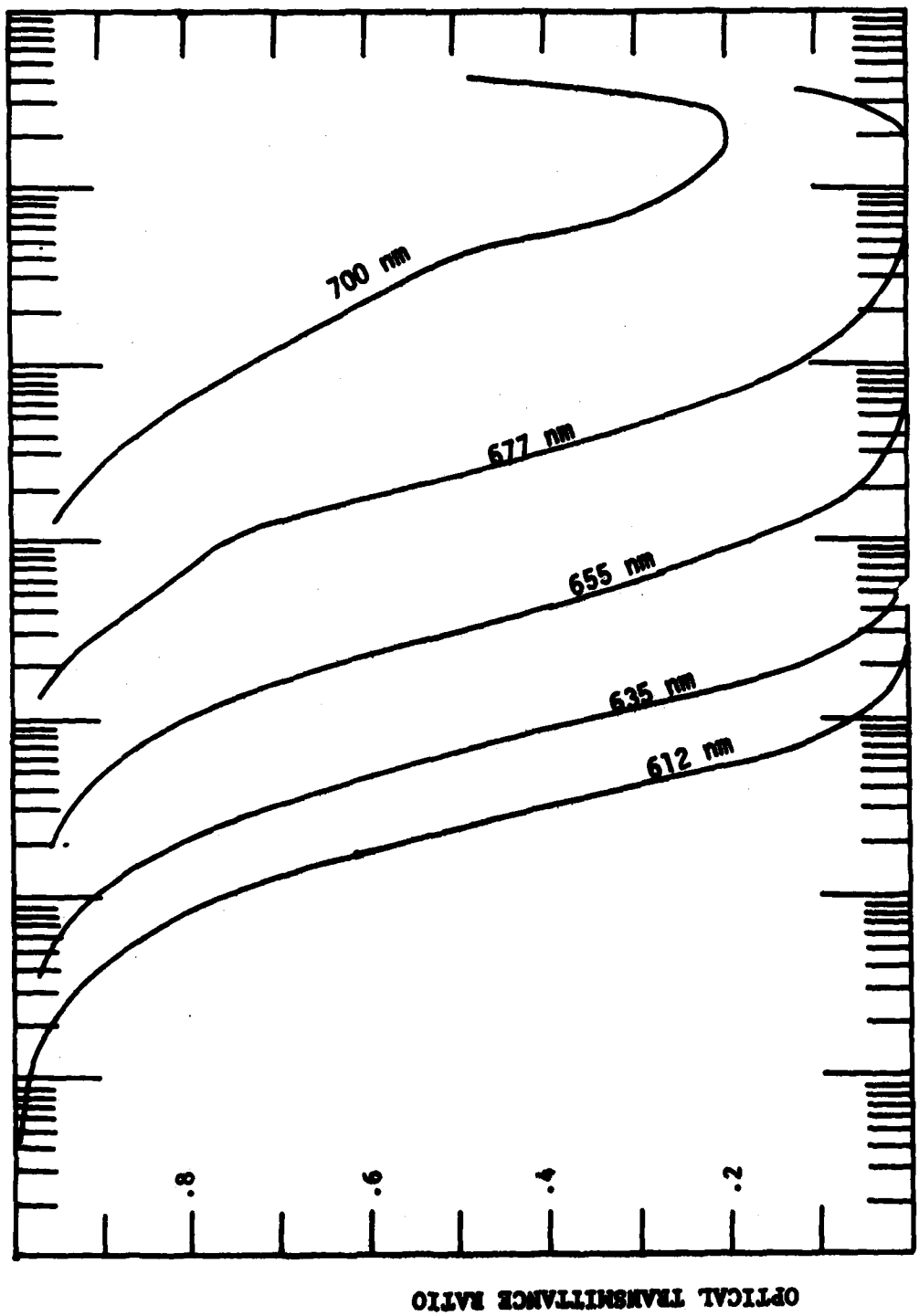


Figure 6. Optical transmittance ratio of irradiated to unirradiated waveguide vs exposure at several different wavelengths.

the Brag-Gray principle. After exposure the solution should be filled into a FEP tubing to form a waveguide and to be read.

FAST NEUTRON DOSE MEASUREMENTS AND NEUTRON ENERGY DEPENDENCE

For readout wavelengths below 620 nm, fast neutrons and gammas produce virtually the same effects in radiochromic waveguide dosimeters. Above λ = 620 nm, fast neutrons produce an additional effect illustrated in Fig. 7. This effect may be exploitable for differentiation between gamma and neutron doses and thus may lead to a future mixed radiation, rem dosimetry. It is currently being investigated and is not a subject of this paper whose scope for mixed dose measurements is limited to readouts at wavelengths below 620 nm.

The response of these dosimeters to fast neutrons is defined mainly by the abundance of fast neutron-generated ionizing recoil particles, particularly the abundance of recoil protons, in the solvent. As the teflon-like cladding contains negligible recoil-producing hydrogen, it contributes little to the neutron dose reading.

So far it has not been possible to find a suitable cladding material which contains hydrogen. Therefore, all currently available waveguide dosimeters have a variable fast neutron response which depends on the fast neutron spectrum and on the diameter of the waveguide; a waveguide with an inside diameter much larger than the range of an average energy recoil proton being more sensitive to fast neutrons than one with a small inside diameter. The proton contribution to the dose was calculated theoretically. Figure 8 shows the specific response of a waveguide dosimeter to fast neutrons of the energy E_n versus the waveguide diameter. This curve applies to an isotropic distribution of fast neutrons. The diameter of the waveguide core (ϕ) is expressed in range of maximum energy of recoil protons ($E_{pr} = E_n$) in the waveguide core material. Figure 9 shows ranges of recoil protons (in mg cm^{-2}) versus proton energy for interpretation of Fig. 8.

Analogous to what was said above, in the case of gamma rays the neutron response problem can be simplified by exposing the dye solution not in a waveguide but in a container whose walls are made of a hydrogenous material and by transferring it later into the waveguide to read the dose.

UTILIZATION IN ARMY TACTICAL DOSIMETRY

The purpose of tactical dosimetry (as opposed to administrative dosimetry) is to provide the troop commander with current data on the radiological status of his troops. The very low-cost leuko dye optical waveguide dosimeter is an eligible candidate to replace the current IM 185, which is a reliable tactical dosimeter but costly and not easily maintained. The leuko dye optical waveguide dosimeter may also be used as the integrated dose component of the Army Miniature Multipurpose Radiac (MMR).

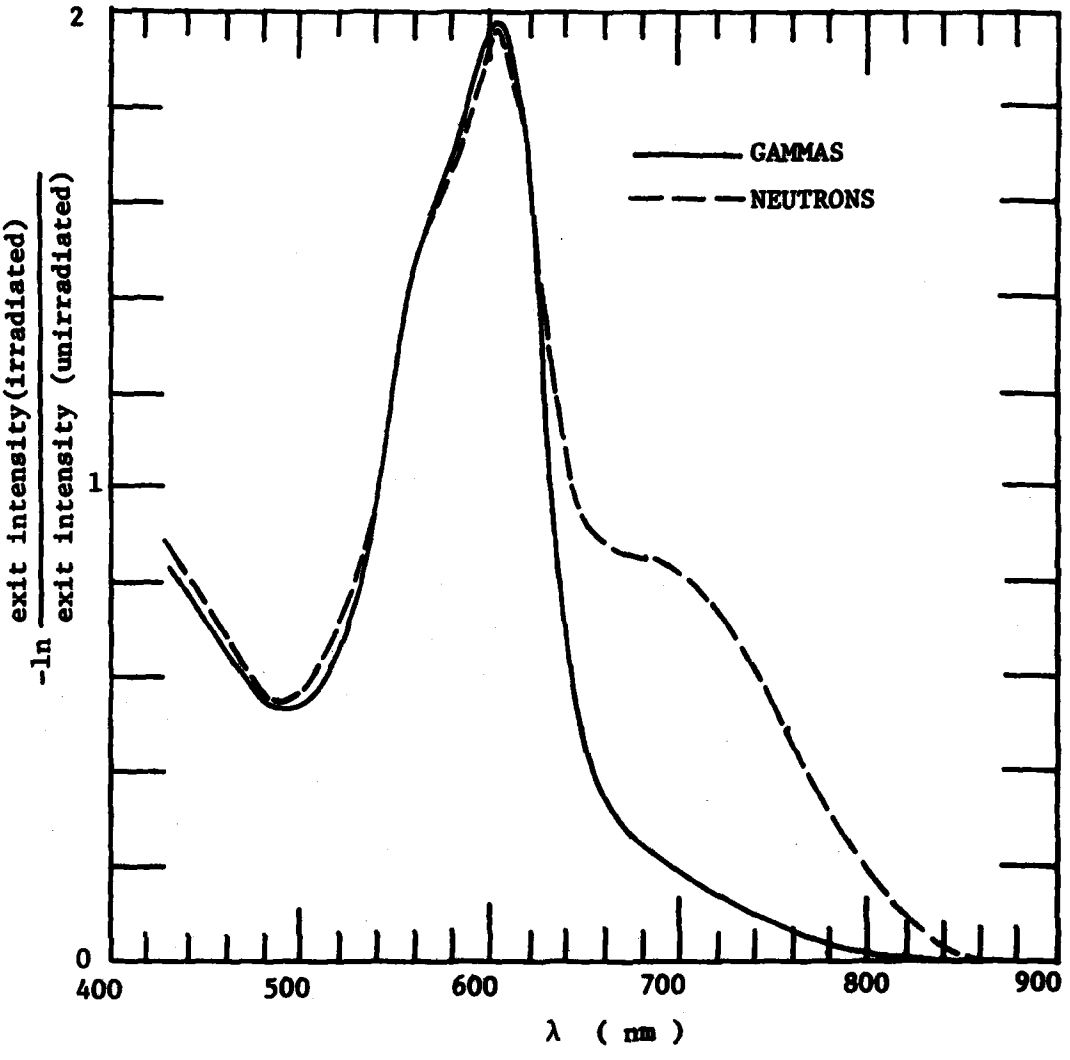


Figure 7. Negative natural logarithm of exit intensity ratio versus wavelength for identically constructed waveguides, 103 mm long, 2.3 mm inner diameter. The core solution (6.7 % dye in DMF) was exposed in two polyethylene capsules, one to 1 k rad (tissue) 14 MeV neutrons + 70 rad (tissue) associated gammas, the other capsule to 1 k rad (tissue) ^{60}Co gammas.

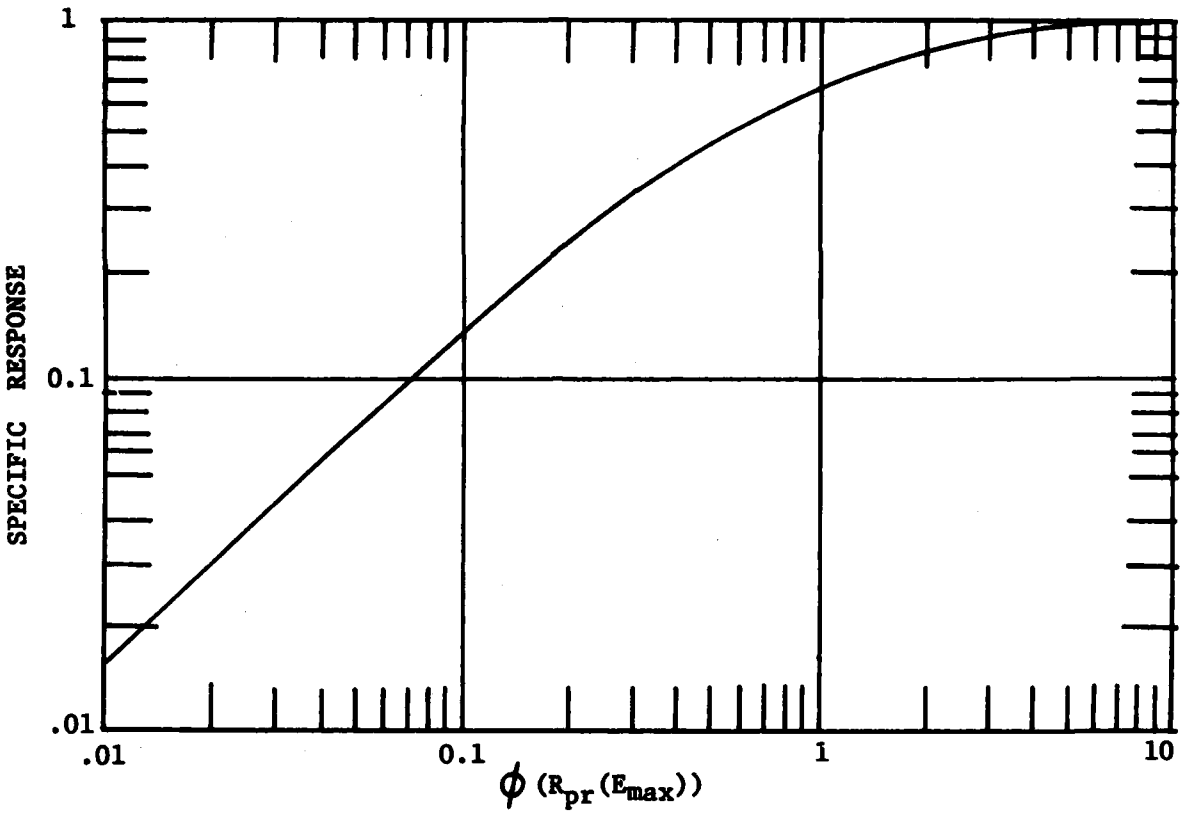


Figure 8. Calculated specific response of a waveguide dosimeter to fast neutrons of the energy E_n vs the waveguide diameter which is expressed in range of recoil protons of maximum energy ($E_{pr} = E_n$) in the waveguide core material.

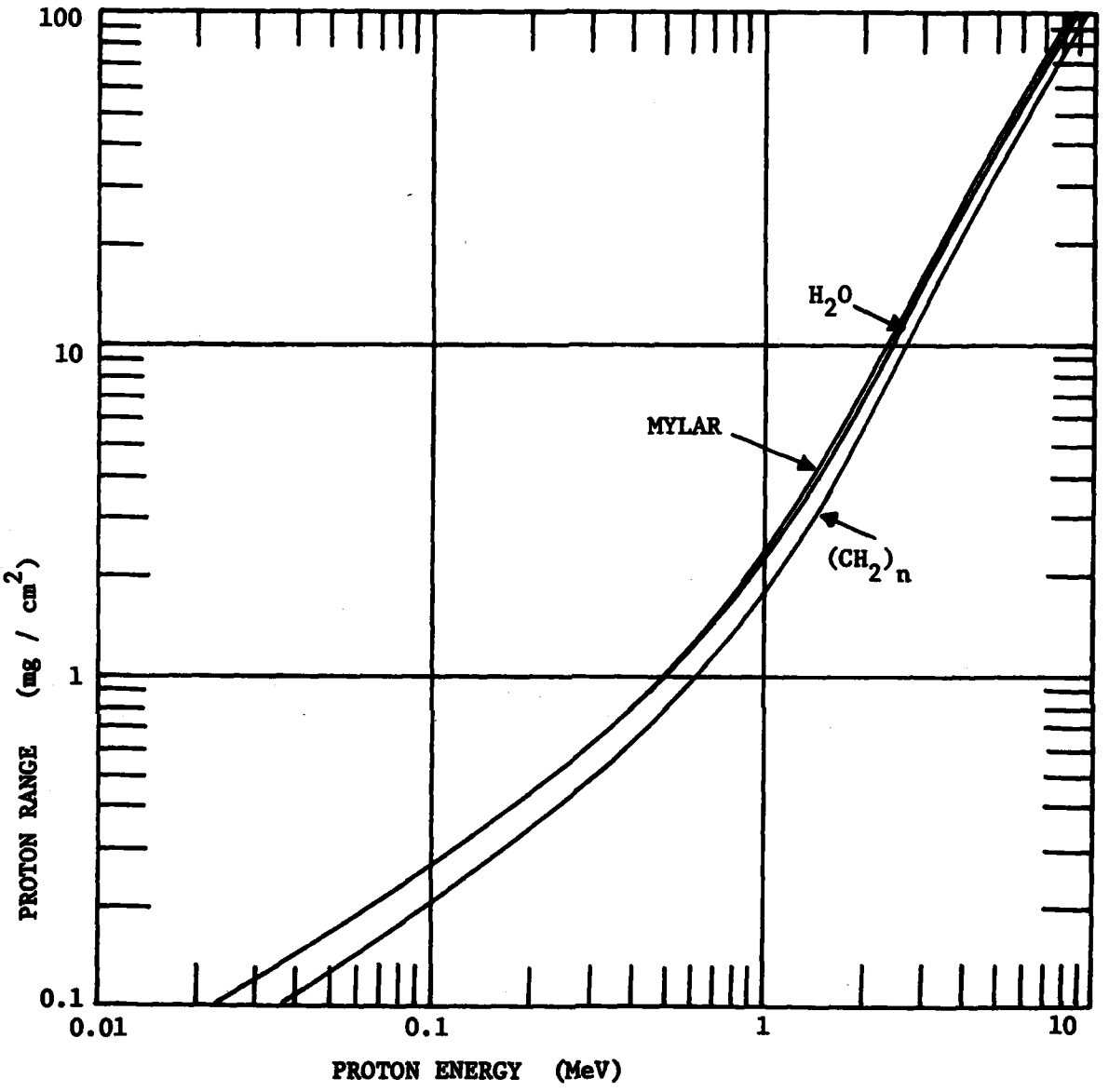


Figure 9. Proton ranges versus proton energy needed for the interpretation of figure 7.

The dosimeter does satisfy the main criteria required for a tactical dosimeter:

a. Radiochromic dyes are dose rate independent up to 10^{12} rad/sec and are thus capable of responding to the prompt initial radiation from an exploding nuclear weapon as well as to the low dose rate fallout radiation (5).

b. The prompt initial radiation of a nuclear weapon consists of gammas and fast neutrons, the fallout radiation of gammas only. This dosimeter can be constructed to give mixed radiation readings as described above. The neutron radiobiological effectiveness (RBE) for Army tactical applications is defined as 1 (1 rad = 1 rem).

c. The radiation-induced changes remain fixed within required limits (e.g., lack of fading) provided that the solution is prepared correctly. A waveguide dosimeter was exposed to 1 K rad (tissue) ^{60}Co gamma rays and its transmittance displayed at room temperature on the reader shown in Fig. 4 for 18 days. The observed output remained constant within 2% which was the readout accuracy of this experiment.

d. The sensitivity and the dynamic range of this system can be adjusted for the required 10 - 1000 rad (tissue) using a wavelength band which peaks at approximately 620 nm for the readout.

e. This system can be made compact by using waveguides made of available FEP tubing 0.8 mm O.D. and 0.25 mm I.D. which can be coiled in a 3 cm diameter spool.

In summary, ionizing radiation-induced changes in the refractive index of radiochromic dye solutions results in a novel dosimetry system with a very wide dynamic range. This approach is adaptable to personnel dosimetry and to Army tactical dosimetry.

ACKNOWLEDGMENTS

The author thanks the following scientists (listed alphabetically) for participating in this research: Dr. Harry Berkowitz, ET&DL; Mr. Harry Guetzlaff, ET&DL; Mr. Milton Harmatz, ET&DL; Dr. Robert Lux, ET&DL; Dr. William McLaughlin, NBS; Dr. Robert Pfeffer, ET&DL and Mr. Carl Siebentritt, FEMA.

REFERENCES

1. W.L. McLaughlin and M.M. Kosanic´, Int. J. Appl. Radiat. Isotopes 25, 249 (1974).

2. M. Wankerl, M. Wacks and L.A. Harrah, Trans. Am. Nucl. Soc. 12(1), 60 (1969).

3. S. Kronenberg and C. Siebentritt, Nucl. Instrum. Methods 175, 109 (1980).

4. M. Born and E. Wolf, Principles of Optics, 2nd ed. (McMillan, New York, NY, 1964) pp. 90-98.

5. M.M. Kosanic´, M.T. Menadovic´, V.B. Radak, V.M. Markovic´, W.L. McLaughlin, Liquid Radiochromic Dye Dosimetry for Continuous and Pulsed Radiation Field Over a Wide Range of Flux Density, Int. J. Appl. Radiat. Isotopes, Vol 28, pp. 313-321 (1977).

ISOLATION OF HEPATITIS A VIRUS FROM THE NEW WORLD OWL MONKEY: A NEW ANIMAL MODEL FOR HEPATITIS A INFECTIONS (U)

*STANLEY M. LEMON, MAJ(P) MC, JAMES LEDUC, MAJ MSC,
LEONARD N. BINN, Ph.D.
WALTER REED ARMY INSTITUTE OF RESEARCH
WASHINGTON, D.C. 20012

Infectious hepatitis has historically been a medical problem of major proportions during large scale military campaigns. Although specific serologic tests capable of defining infection with hepatitis A virus did not become available until the early 1970s, it is very likely that the great majority of cases of viral hepatitis recorded among American soldiers during World War II and the Korean and Vietnam conflicts were due to this virus. Hepatitis A virus was probably the single "greatest cause of disabling illness among United States forces in the Mediterranean (North African) Theater of Operations" (1) during 1943, a period during which 15,865 cases of hepatitis were hospitalized out of a mean strength of 423,114 soldiers (2). Since then, it is highly likely that the overall prevalence of antibody to hepatitis A virus (anti-HAV) has fallen significantly among incoming recruits, resulting in a military force which has a greater susceptibility to the virus than that which existed forty years ago. In fact, during 1981, a survey of 233 members of the 82nd Airborne Division, a combat-ready unit maintained at the highest level of readiness, revealed that only 35, or 15 per cent, had pre-existing antibody and could be considered to be immune to hepatitis A infection (3). Thus, hepatitis A virus represents a significant and well proven threat to American forces following any future deployment under combat conditions and should receive the highest possible priority for the development of effective control measures.

The successful replication of hepatitis A virus in cell culture during 1979 (4) provided the first real hope that a hepatitis A vaccine might eventually be developed. Nonetheless, continued efforts directed at the control of this agent have been greatly hampered because there are no genetic markers for attenuation of this virus, because available animal models (chimpanzees and several species of marmosets) (5) have been of limited availability, and because it has thus far not been possible to obtain sufficient replication of the virus in vitro to permit development of an inactivated vaccine. Studies carried out recently at the Walter Reed Army Institute of Research, and described in this report, are

therefore of considerable interest in that they demonstrate that the New World owl monkey (Aotus trivirgatus)*, now used as a model of malaria infections, is also susceptible to infection with hepatitis A virus and is a valid model of human hepatitis A. Furthermore, a new strain of hepatitis A virus, identified in the feces and liver of naturally infected Aotus, has been isolated in tissue culture and shown capable of replication to high titers with free release of virus into supernatant fluids even during early passage.

MATERIALS AND METHODS

Aotus trivirgatus monkeys. Aotus is a small, nocturnal and arboreal primate which is widely distributed throughout Central and South America. Fully grown adults weigh 0.7 to 1.2 kg, and a number of different karyotypes have been described. Several successful breeding colonies have supplied monkeys for the study of new antimalarial agents, and both wild-caught and colony-bred monkeys are available. Aotus held at the Walter Reed Army Institute of Research (WRAIR) in Washington, D.C. were bled and studied for the presence of anti-HAV. Both colony-bred and procured (wild-caught) monkeys were included.

Newly captured Aotus were bled at the Gorgas Memorial Laboratory (GML) in Panama City, Panama. These monkeys had been captured by local trappers in the forests which surround Panama City, most often within a radius of 100 km from the city. Monkeys were held in a quarantine facility for not less than 90 days and were then integrated into the main colony. When it became apparent that newly captured monkeys were becoming infected shortly after arrival at the colony, Aotus admitted to the colony during the interval from 25 September 1980 to 31 December 1980 were studied prospectively. Feces were collected twice weekly and liver was taken at necropsy from monkeys dying during their first three months of captivity. Both liver and fecal specimens were stored at -20 C. until tested for viral antigen. A portion of each liver was fixed in 10% neutral buffered formalin, embedded in paraffin, and later stained with hemotoxylin and eosin for histopathologic studies.

Serologic methods. Anti-HAV was determined by commercial radioimmunoassay (Havab, Abbott Laboratories, N. Chicago, IL). IgM antibody to hepatitis A virus (IgM anti-HAV) was detected by radioimmunoassay (6), and antibody to hepatitis B virus core antigen (anti-HBc) by blocking radioimmunoassay (Corab, Abbott Laboratories).

Detection of hepatitis A virus. Viral antigen was detected by a

*In conducting the research described in this report, the investigators adhered to the "Guide for the Care and Use of Laboratory Animals", as promulgated by the Committee on Care and Use of Laboratory Animals of the Institute of Laboratory Animal Resources, National Research Council.

modification of the solid-phase radioimmunoassay method described by Purcell et al. (7). The central 60 wells of a flexible microtiter "U" plate were coated for 4 hrs. at 30 C. with 100 ul of human convalescent hepatitis A serum diluted 1:200 in 50 mM sodium carbonate buffer, pH 9.6. The plates were washed five times with phosphate buffered saline, pH 7.4, (PBS) and each well filled with 1% bovine serum albumin (Calbiochem, La Jolla, CA) in PBS. After one hour at 30 C., plates were washed with PBS containing 0.05% Tween-20 (PBS-Tween) (Fisher Scientific Co., Fairlawn, NJ) and 50 ul of a sample to be tested was placed into each well. The plates were incubated at 4 C. for 18 hrs., washed with PBS-Tween, and 40 ul of ^{125}I-labelled human anti-HAV (6) was added to each well. After four hrs. at 4 C., plates were washed again with PBS-Tween, the wells cut apart, and bound ^{125}I determined by counting in a Rackgamma-II gamma counter (LKB-Wallac, Turku, Finland). Test specimens were compared with reference chimpanzee fecal specimens which either did or did not contain virus. A negative mean cpm value (N) was determined for each microtiter plate based on the results obtained from two wells loaded with virus-free chimpanzee material. The mean cpm value obtained from two wells containing a test sample (P) was divided by N to obtain a test specimen/negative mean value ratio (P/N). Samples generating a P/N ratio of 2.1 or greater were considered positive. The specificity of each positive reaction was confirmed by blocking the reaction with hepatitis A convalescent sera; for this purpose, paired chimpanzee sera were obtained from the Research Resources Branch, National Institute of Allergy and Infectious Diseases, Bethesda, MD, and from previous experiments carried out at WRAIR. Twenty ul of a 1:100 dilution of blocking sera was added to each microtiter well fifteen minutes prior to the addition of the ^{125}I-labelled antibody. A positive radioimmunoassay reaction for viral antigen was considered confirmed if a greater than 50% reduction in cpm was noted in wells blocked with post-infection chimpanzee serum compared with wells blocked with homologous pre-infection serum.

Fecal specimens were tested as 20% suspensions in PBS and were centrifuged for 30 minutes at 8000 x G prior to testing. Previously frozen liver tissue was minced into fine pieces, subjected to 15 strokes with a Ten Broeck homogenizer as a 20% suspension in 0.05 M Tris-HCl containing 0.04 M KCl and 0.005 M sodium acetate and clarified at 8000 x G for 30 minutes prior to testing.

Alternatively, hepatitis A virus antigen was detected in acetone-fixed, infected cell monolayer cultures by direct immunofluorescence, using fluorescein-isothiocyanate conjugated with the IgG fraction of a convalescent human serum specimen containing anti-HAV. Fixation was carried out at room temperature for one minute.

Electron microscopy. Studies were carried out by Dr. Stephen Feinstone of the National Institute of Allergy and Infectious Diseases, Bethesda, MD., as previously described (8).

Comparison of Aotus and human hepatitis A viruses. To determine the

degree of antigenic relatedness between virus recovered from Aotus and human virus, a parallel radioimmunoassay was developed using only Aotus-derived reagents. The assay was performed as described above, with the exception that Aotus convalescent serum having an anti-HAV titer of greater than 1:6400 was used as the coating antibody. A liver suspension (5%), prepared as above from an infected monkey, served as viral antigen. Aotus IgG was obtained by precipitation of 1 ml convalescent Aotus serum (anti-HAV titer > 1:6400) with 45% ammonium sulfate, followed by centrifugation at 8000 x G for 30 minutes. The pellet was resuspended in 1 ml distilled water and dialyzed extensively against 0.0175 M potassium phosphate buffer, pH 7.2. The dialyzed material was then passed through a 10 cm x 1.5 cm DEAE-Sephacel (Pharmacia Fine Chemicals, Piscataway, NJ) column and eluted with the same buffer. Fractions comprising the first protein peak were pooled, filter concentrated using an Amicon XM300 filter (Amicon Corporation, Lexington, MA) to approximately 1 mg protein/ml, and radiolabelled with ^{125}I as described (6). A human convalescent serum collected during an outbreak of hepatitis A in Alaska, and a convalescent Aotus serum obtained at the GML, were diluted serially and the blocking activity of each serum was determined in simultaneously performed parallel radioimmunoassays using Aotus or human reagents. Serum dilutions were made in PBS containing a fixed 1:100 dilution of anti-HAV-negative chimpanzee serum.

Experimental monkey infections. Seronegative, colony-bred Aotus were selected and placed in quarantine at the United States Army Medical Research Institute of Infectious Diseases, Fort Detrick, Maryland. Baseline evaluations included weekly determinations of serum alanine and aspartate aminotransferase activities, serum bilirubin, and collection of fecal specimens. In addition, an open liver biopsy was performed on each monkey prior to intravenous inoculation with 1 ml a 0.1% filtrate of Aotus feces. The monkeys were tested twice weekly for elevations of aminotransferase activities or fecal shedding of viral antigen as detected by radioimmunoassay. Each monkey underwent a second liver biopsy at the earliest chemical evidence of hepatocellular disease. Tissues were processed as described above.

Cell cultures. Primary African green monkey kidney cells (AGMK), certified for use in vaccine production, were obtained from Lederle Laboratories, Pearl River, NJ. Continuous green monkey kidney cells (BSC-1) were the gift of H.E. Hopps of the Bureau of Biologics, Food and Drug Administration, Rockville, MD. Fetal rhesus lung (FRhL-2) (9) and kidney (FRhK-6) (10) cells were as described. MA-104, continuous Rhesus monkey kidney cells, were purchased from Microbiological Associates, Walkersville, MD. Cells were grown in Eagle's minimal essential medium variously supplemented with non-essential and essential amino acids and vitamins and 10% fetal bovine serum. Prior to inoculation with viral suspensions, cell monolayers were washed twice with Hanks' balanced salt solution (HBSS). After a period of viral adsorbtion for one to two hours

at 35 C., maintenance media containing 2% fetal bovine serum was added. Cells were subsequently refed every five to seven days. Virus was harvested by washing infected cell monolayers once with HBSS, scraping the cells into a reduced volume (20%) of HBSS and sonicating for one minute at 100W (Heat Systems-Ultrasonic, Model W185, Plainview, NY) followed by centrifugation at 8000 X g for 30 minutes.

Human hepatitis A virus strains. HM-175 strain hepatitis A virus, passed six times in marmosets, was obtained as a 10% liver suspension from S.M. Feinstone of the National Institute of Allergy and Infectious Diseases, Bethesda, MD. MS-1 strain virus was obtained from previously infected human volunteers (11) and infected chimpanzee feces as described (6).

RESULTS

Serologic surveys. An initial survey of 170 Aotus held at WRAIR demonstrated a high prevalence of anti-HAV (60%) among procured monkeys, but few sero-positive monkeys among those which were colony-bred (3%). These preliminary results prompted a study of sera collected from 50 Aotus monkeys which had been held at the GML colony in Panama for various lengths of time following their capture. Thirty-eight of these sera (76%), which were collected in November 1979, contained anti-HAV antibody. The highest antibody prevalence rates were found among those animals held in captivity longest. None of 10 animals held less than 30 days, but all of 35 animals held for longer than 100 days were positive. Furthermore, whereas immunoglobulin M (IgM) anti-HAV antibody was uniformly present in seropositive animals held for less than 300 days, it was not detected in any animal which had been held for over 600 days. None of these 50 monkeys had antibody to hepatitis B core antigen. These observations suggested that hepatitis A virus was being transmitted to monkeys following their capture, probably within the colony at the GML.

To confirm the absence of hepatitis A infection among wild Aotus, 145 newly captured monkeys arriving at the GML between 25 September 1980 and 31 May 1981 were bled shortly after arrival and tested for the presence of anti-HAV. Only two of these 145 monkeys (1.4%) were sero-positive. Thus, exposure to hepatitis A virus appears to have been a very rare event in these monkeys prior to their capture.

Prospective study of newly captured A. trivirgatus. To more fully document the transmission of virus within the GML colony, all 41 monkeys arriving between 25 September and 31 December 1980 were followed prospectively with periodic collection of serum and fecal samples. Thirty-one of these animals (76%) had been held by trappers less than eight days prior to delivery to the GML, and only one had been held longer than 15 days. All lacked anti-HAV antibody upon arrival at the colony. Of the 41 monkeys in the original cohort, 11 died relatively soon after capture and were not bled later than 25 days after arrival at the

colony. None of these monkeys showed serologic evidence of infection with hepatitis A virus and none of eight tested had viral antigen detected in liver taken at necropsy. However, seroconversion to the virus was documented by radioimmunoassay in 21 of the remaining 30 monkeys.

To evaluate fecal shedding of virus by infected monkeys, fecal specimens (collected at 2-4 day intervals during the period between the last antibody-negative and first antibody-positive serum samples) were examined for the presence of antigen by radioimmunoassay. Viral antigen was detected in feces collected from 12 of 15 cages, containing a total of 23 monkeys (Table 1). The quantity of antigen present, as estimated by the amount of ^{125}I-label bound, approximated that of feces from acutely infected chimpanzees. Two fecal specimens which gave positive results by radioimmunoassay were examined by immune electron microscopy. The presence of hepatitis A virus in these specimens was confirmed by the finding of 27nM picornavirus-like particles which were aggregated in the presence of anti-HAV (data not shown).

Table 1. Fecal Shedding of hepatitis A virus by Aotus trivirgatus*.

Cage #	Monkey(s)	Days in colony	P/N	HAV Radioimmunoassay (cpm/cpm blocked)**
254	10512	62	25.1	(427/85)
262	10538	34	9.1	(128/26)
263	10523	36	146.0	(2047/357)
264	10529	55	10.2	(326/40)
266	10542	32	55.6	(2434/598)
267	10515,16	57	1.8	(51/81)
269	10513	63	4.5	(77/32)
270	10539,40,41	49	1.3	(48/54)
271	10524,25	35	23.6	(532/78)
272	10526,27,28	32	12.5	(283/29)
275	10519,20	50	1.3	(38/29)
277	10537	36	159.0	(2704/299)
278	10532	42	26.6	(599/67)
279	10533	28	92.5	(2083/200)
280	10535,36	36	44.8	(1010/86)
Neg control	Chimp 273	-	1.0	(14-44/Not Done)
Pos control	Chimp 273	-	94.4	(2077/175)

* Results are for the specimen yielding the highest P/N value.
** (cpm blocked with pre-infection chimpanzee serum/cpm blocked by post-infection chimpanzee serum).

Of the nine remaining monkeys which survived over 25 days in the colony and which did not develop anti-HAV, eight died within 60 days of their arrival. Liver tissue, taken at necropsy, was available from four of these eight monkeys, and in each case contained viral antigen detectable by radioimmunoassay. Antigen was also present in fecal specimens collected from three of these four monkeys just prior to their death; feces were not available from the fourth monkey. Three of the four liver specimens containing viral antigen were examined histologically, but none had changes suggestive of acute viral hepatitis in man. Thus, although these four monkeys died during the acute phase of their infection, it appeared that death was not directly due to liver disease.

Comparison of human and Aotus hepatitis A virus strains. To examine more fully the degree of antigenic relatedness between the viral antigen recovered from the monkeys and human hepatitis A virus (MS-1 strain), parallel radioimmunoassays were developed employing in one case antigen and antibody derived from human infections, and in the other, material obtained from Aotus at the GML. Two sera, one collected from a convalescent human and the other from a convalescent monkey, were tested in each radioimmunoassay for blocking activity. In replicate experiments, both monkey and human sera demonstrated a similar titer (1:25,600) of blocking activity when tested against the Aotus antigen. Although the Aotus serum appeared to have a somewhat lower titer (1:12,800) of blocking activity against MS-1 antigen when compared with the human serum (1:25,600), this difference was only one tube dilution. These results indicate that these two antigens were indistinguishable by this technique.

Experimental infection of colony-bred Aotus. Six sero-negative, colony-bred Aotus were inoculated with 1 ml of a 0.1% filtrate of stool (PA-33) collected from an infected monkey (#10533, Table 1) at the GML colony and shown to contain hepatitis A virus by electron microscopy and radioimmunoassay. Results of this experiment are summarized in Table 2,

Table 2. Experimental infection of Aotus with hepatitis A virus.

Monkey	Days with fecal virus	Max. P/N	Baseline* ALT (s.d.)	Max. ALT	Days with elev. ALT	Converted anti-HAV
WR-68	6-15	21.9	27.3 (6.3)	210	14-21	+
WR-80	10-13	7.1	40.9 (17.3)	316	14-21	+
WR-130	8-20	54.4	68.5 (11.4)	372	14-28	+
WR-148	17-22	20.2	35.4 (10.8)	233	21-28	+
WR-149	7-15	26.6	73.3 (30.6)	272	14-28	+
WR-153	6-16	45.9	49.3 (6.8)	189	9-28	+

*ALT = alanine aminotransferase (s.d. = standard deviation)

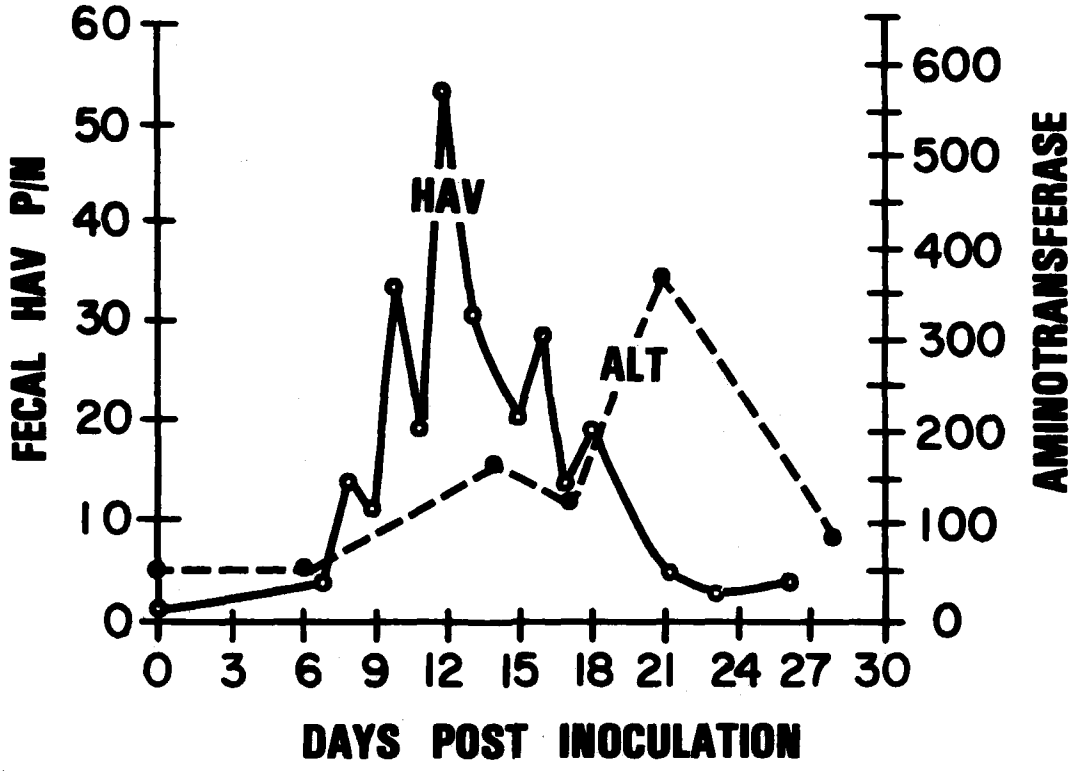


Figure 1. Experimental infection of Aotus WR-130 with PA-33 virus.
HAV = hepatitis A virus ALT = alanine aminotransferase

with the course of infection in a single monkey (WR-130) depicted in Figure 1. Following an incubation period of six to 17 days, hepatitis A virus antigen was detected in feces from each monkey by radioimmunoassay. Fecal shedding of virus persisted for four to 13 days. Four to eight days after the appearance of viral antigen in feces, both serum alanine and aspartate aminotransferase activities became significantly elevated in each monkey. Liver biopsies taken at this time revealed moderate histologic changes in five monkeys, and mild changes in the sixth. These changes included peri-portal mononuclear and polymorphonuclear infiltration and patchy areas of hepatocellular necrosis, changes very similar to those seen in human infections. Aminotransferase activities returned to normal by one month after inoculation, by which time each monkey had sero-converted to the virus.

These findings confirm the utility of Aotus as an animal model of human hepatitis A.

In vitro replication of hepatitis A virus. A variety of cell substrates were inoculated with a 10% filtrate of the same Aotus stool (PA-33) used for the experimental monkey infections, a 1% filtered suspension of Aotus liver (PA-21) taken from a monkey which died in the GML colony and was shown by radioimmunoassay to have viral antigen in the liver at necropsy, or additional material containing either HM-175 (0.5% marmoset liver) or MS-1 (10% human stool filtrate) strains of hepatitis A virus. Results are summarized in Table 3. Hepatitis A antigen was detected in cell extracts by radioimmunoassay and in cell monolayers by direct immunofluorescence. The specificity of the immunofluoresence test was confirmed by blocking studies employing pre- and post-infection chimpanzee sera. Fluorescence was discreet, granular and confined to the cytoplasm, as has been described by others (4,12). Hepatitis A antigen was detected in AGMK cells inoculated with both PA-33 and HM-175 virus strains, as well as in FRhL-2, BSC-1, and MA-104 cells inoculated with either PA-21 or PA-33. Antigen was detected by immunofluoresence as

Table 3. Initial passage of hepatitis A viruses in cell culture.

Virus Inoculum	Cell culture	Viral Antigen Detected RIA*	FA**	Day tested
HM-175	AGMK	2.4	Not Done	(59)
		5.0	+	(92)
PA-33	AGMK	24.3	Not Done	(30)
		6.2	+	(50)
		12.2	+	(92)
	FRhL-2	1.9	Not Done	(30)
		4.4	+	(59)
		3.8	+	(92)
PA-21	BSC-1	0	+	(42)
	MA-104	0	+	(42)
	FRhK-4	0	0	(28)
	FRhL-2	0	0	(28)
MS-1	AGMK	0	0	(92)
	FRhL-2	0	0	(92)

* RIA = radioimmunoassay, results are P/N values.
** FA = direct immunofluorescence.

early as seven days after inoculation of AGMK with PA-33 stool and could be detected by radioimmunoassay in cell extracts at 14 days. No antigen was detected in MS-1 inoculated cultures. As has been reported by other investigators (4,12,13), there was no apparent cytopathic effect.

Of special interest, viral antigen was identified by radioimmunoassay in supernatant fluids of AGMK cultures as early as 30-50 days after inoculation with PA-33 stool. This finding is in contrast with other reports of hepatitis A virus propagation in vitro (12,13), which have emphasized the cell-associated nature of this virus infection. The presence of infectious virus in supernatant fluids of infected AGMK cells was confirmed by subsequent passage.

The PA-21 and PA-33 Aotus hepatitis A virus isolates were passed a total of four to 10 times in cell culture (Table 5). With subsequent passage, there was a shortening of the lag time between inoculation of cells and detection of viral antigen. Viral titers, determined by immunofluoresence 28 days post-inoculation, have ranged from 5.5 to >7.5 $\log_{10}$ tissue culture infectious units/ml. After the third to the fifth passage, viral antigen was readily detected by radioimmunoassay in supernatant fluids of infected cultures.

DISCUSSION

These data provide strong evidence that the New World owl monkey, Aotus trivirgatus, is an accurate and potentially very useful animal model of human hepatitis A. There is little doubt that transmission of

Table 4. Propagation of hepatitis A viruses in cell culture.

Virus	Cells	Virus Antigen			Virus Titer	
		Passage	RIA*	FA**	Passage	$\log^{10}$/ml
HM-175	AGMK	7	7.0	4+	-	-
PA 33	AGMK	10	10.8	4+	5	6.0
	FRhL-2	5	2.9	3+	5	5.5
PA 21	BSC-1	5	8.4	2+	3	7.0
	FRhK-6	4	8.8	2+	3	>7.5

* RIA = radioimmunoassay P/N ratio
** FA = direct immunofluoresence.

hepatitis A virus occurred within the Aotus colony at the GML. Since initiating serological testing, all but two of 145 newly captured monkeys have arrived at the colony lacking antibody to the virus. On the other hand, 25 of 41 monkeys followed prospectively became infected with hepatitis A virus following their arrival at the colony. Only one monkey survived over 60 days and was not infected. Of special interest was the demonstration of virus, both by radioimmunoassay and by subsequent viral isolation, in the feces of these monkeys. To our knowledge, this is the first identification of hepatitis A virus in fecal material from non-human primates which had not been experimentally infected.

The course of hepatitis A virus infection observed in Aotus appears remarkably similar to that of hepatitis A in man (14,15). Viral antigen was present in liver tissue at the peak of infection, when antigen was also present in feces. It appeared to be shed in feces for about one week, and the first detectable serologic response, consisting at least in part of IgM anti-HAV, coincided with cessation of fecal antigen shedding. Fecal shedding of virus was followed by elevations of serum aminotransferase activities (Figure 1). Histologic changes, occurring coincidently with elevation of aminotransferase activities, were similar to those seen in hepatitis A in humans. However, the incubation period of the disease, from virus inoculation to onset of hepatocellular disease (about 14 days), was somewhat shorter than that seen in man (about 28 days).

The strains of hepatitis A virus isolated from Aotus (PA-33 and PA-21) were readily propagated in vitro and in this respect resemble human hepatitis A strains passed multiple times in marmosets (4,12). The yield of PA-21 virus in BSC-1 or FRhK-6 cell cultures was greater than the minimal titer of poliovirus, a related picornavirus, required for production of inactivated poliovirus vaccine ($10^{6.5}$ infectious units/ml) (16). Further efforts are in progress to improve the yield of virus and simplify methods of viral titration. In addition, cell culture derived virus will replace virus produced in experimentally infected primates as a source of antigen for both serologic assays and studies aimed at further characterizing the hepatitis A virion.

The question remains as to whether the virus isolated from Aotus was indeed human hepatitis A virus or instead a closely related virus of Aotus. The virus reacted well in conventional virus assays, and antigen detected was specifically blocked in the radioimmunoassay by reference anti-HAV sera. Furthermore, cross-blocking experiments in parallel radioimmunoassays employing either human or Aotus antigens did not reveal any significant differences between these strains. A significant biologic difference between the virus found in the GML colony and human hepatitis A virus is, however, suggested by the fact that monkeys held at the GML colony appeared to be uniformly susceptible to infection with the GML strain of virus. In contrast, neither of two Aotus inoculated orally with infectious human virus (MS-1 strain) developed anti-HAV in a previous

study (17). Although this apparent difference in susceptibility might possibly be due to differences in inoculum size, it seems likely that the virus found at the GML has become adapted to Aotus by multiple passage. Adaptation of human hepatitis A virus to a non-human primate species, as evidenced by a shortening of the incubation period, an increase in the amount of viral antigen found in liver, and an enhanced ability to infect members of that species, has been observed following repeated passage of hepatitis A virus in the marmoset Saguinas mystax (18).

Whether or not the virus described here was ultimately of human origin, our observations indicate that Aotus will be a useful animal model for further investigation of the pathogenesis of human hepatitis A. Aotus, in addition to its current availability, has several apparent advantages over existing marmoset models of hepatitis A. Aotus is a somewhat larger primate, and therefore may be more suitable for some pathogenesis studies. Perhaps more importantly, hepatocellular damage due to hepatitis A virus in Aotus may be monitored biochemically with the same enzyme assays which are useful in hepatitis A infections in man (alanine aminotransferase and aspartate aminotransferase), whereas hepatitis in marmosets generally requires the monitoring of special enzymes (isocitrate dehydrogenase (5,17)) and therefore appears to be less analogous to disease in man. While the chimpanzee may be considered to remain the animal model of choice for hepatitis A infections, these large primates are very expensive and remain heavily endangered. Aotus should prove exceptionally valuable as an alternative model for human hepatitis A and should play a significant role in the eventual development of a vaccine for the protection of American soldiers against an epidemic disease of major military importance.

SUMMARY

1. Hepatitis A virus infections occurred in naturally exposed Aotus trivirgatus. Monkeys newly admitted to the Aotus colony in Panama became infected during the quarantine period, shed hepatitis A virus in feces, and had viral antigen demonstrable in the liver. Infected monkeys developed a serologic response to the virus, including transient IgM anti-HAV.

2. Experimental infection of colony-bred Aotus monkeys with PA-33 strain hepatitis A virus confirmed the field observations of hepatitis A infections in Aotus. Viral shedding was followed by chemical evidence of hepatocellular damage, and histologic changes in the liver were similar to viral hepatitis in man. These studies clearly define the Aotus as a useful and accurate model of human hepatitis A infections.

3. Two strains of hepatitis A virus (PA-21 and PA-33) were isolated from Aotus in cell cultures suitable for use in human vaccine production. The titer of virus yields achieved in cellular extracts approached or exceeded those required for production of inactivated

picornavirus vaccines. Preparation of a tissue culture-derived, inactivated hepatitis A vaccine is now a feasible goal.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the assistance of C. Keenan, R. Graham, N. Gates, R. Marchwicki, S. Feinstone, and J. Burdick.

REFERENCES

1. Gauld RL (1946). Epidemiologic field studies of infectious hepatitis in the Mediterranean Theater of Operations. American Journal of Hygiene 43: 248-254.
2. Havens WP (1968). Viral hepatitis. In, Anderson RS, Havens WP [eds]. "Internal Medicine in World War II", Volume III. "Infectious Diseases and General Medicine". Office of the Surgeon General, Washington, D.C., pp 331-384.
3. Prior R, Lemon SM. (1982). Unpublished observations.
4. Provost PJ, Hilleman MR (1979). Propagation of human hepatitis A virus in cell culture in vitro. Proceedings of the Society for Experimental Biology and Medicine 160: 213-221.
5. Deinhardt F (1976). Hepatitis in primates. Advances in Virus Research, 20: 113-157.
6. Lemon SM, Brown CD, Brooks DS, Simms TE, Bancroft WH (1980). Specific immunoglobulin M response to hepatitis A virus determined by solid-phase radioimmunoassay. Infection and Immunity 28: 927-936.
7. Purcell RH, Wong DC, Moritsugu Y, Dienstag JL, Routenberg JA, Boggs JD (1976). A microtiter solid-phase radioimmunoassay for hepatitis A antigen and antibody. Journal of Immunology 116: 349-356.
8. Feinstone SM, Barker LF, Purcell RH (1979). Hepatitis A and B. In Lennette EH, Schmidt NJ [eds]. "Diagnostic Procedures for Viral, Rickettsial and Chlamydial Infections", Fifth Edition. American Public Health Association, Washington, D.C., pp 879-925.
9. Wallace RE, Vasington PJ, Petricciani JC, Hopps HE, Lorenz DE, Kadanka Z (1973). Development of a diploid cell line from fetal rhesus monkey lung for virus production. In vitro 8: 323-332.
10. Wallace RE, Vasington PJ, Petricciani JC, Hopps HE, Lorenz DE, Kadanka Z (1973). Development and characterization of cell lines from subhuman primates. In vitro 8: 333-341.
11. Boggs JD, Melnick JL, Conrad ME, Felsher BF (1970). Viral hepatitis. Clinical and tissue culture studies. Journal of the American Medical Association 214: 1041-1046.
12. Daemer RJ, Feinstone SM, Gust ID, Purcell RH (1981). Propagation of human hepatitis A virus in African green monkey kidney cell culture:

Primary isolation and serial passage. Infection and Immunity 32: 388-393.
13. Frosner GG, Deinhardt F, Scheid R, Gauss-Muller V, Holmes N, Messelberger V, Siegl G, Alexander JJ (1979). Propagation of human hepatitis A virus in a hepatoma cell line. Infection 7: 303-305.
14. Dienstag JL, Feinstone SM, Kapikian AZ, Purcell RH, Boggs JD, Conrad ME (1975). Fecal shedding of hepatitis A antigen. Lancet, 1: 765-767.
15. Decker RH, Overby LR, Ling CM, Frosner G, Deinhardt F, Boggs J (1979). Serologic studies of transmission of hepatitis A in humans. Journal of Infectious Diseases 139: 74-82.
16. Code of Federal Regulations, 21 (1977). Chapter 1 "Food and Drug Administration, Department of Health Education and Welfare", Subchapter F "Biologics", Part 630 "Additional standards for viral vaccines". U.S. Government Printing Office, Washington, §630.2(c).
17. Purcell RH, Dienstag JL (1978). Experimental hepatitis A virus infection. In Oda T [ed]. "Hepatitis Viruses." University Park Press, Baltimore, pp 3-12.
18. Provost PJ, Villarejos VM, Hilleman MR (1977). Suitability of the rufiventor marmoset as a host animal for human hepatitis A virus. Proceedings of the Society for Experimental Biology and Medicine 155: 283-286.

THE MOBILITY OF WATER IN FROZEN SOILS (U)

*VIRGIL J. LUNARDINI, DR., RICHARD BERG, DR., RICHARD MCGAW,
THOMAS JENKINS, YOSHISUKE NAKANO, DR., JOSEPH OLIPHANT, DR.,
KEVIN O'NEILL, DR., AND ALLAN TICE
U.S. ARMY COLD REGIONS RESEARCH AND ENGINEERING LABORATORY
HANOVER, NEW HAMPSHIRE 03755

INTRODUCTION

The movement of water in soil systems has long been of interest for applications such as road, dam, and reservoir design; agronomy; and ground thermal storage. Recent proposals for buried, chilled pipelines raise the possibility of water movement in freezing soils over very long periods of time. The need to apply soil water and ice relations to practical problems such as frost heave and thaw weakening has prompted an expansion of research, since rules of thumb have repeatedly failed. The physics of the processes must be studied in increasing depth to rationalize mathematical models.

The relation of theoretical microscopic soil water physics to observed macroscopic parameters has been evaluated here, emphasizing frost heave phenomena. Experimental work was carried out to determine how water moves in isothermal frozen clays and to measure simultaneous pressures and temperatures within the freezing zone of silts.

SOIL WATER AND ICE

Taber [1] showed that the major cause of volume increase when soil freezes is the formation of segregated ice: ice formed from water in excess of that originally in the soil. A complete, quantitative explanation for the movement of excess soil water in response to a freezing condition is still sought. Beskow [2] demonstrated that a pressure gradient, or a gradient of free energy, developed within the water of fine-grained soil during freezing. Soil water will move under appropriate potential gradients and an understanding of these fluxes in frozen soils is vital to explain frost heaving.

Most descriptions view the soil as a macroscopic continuum with reference made to microscopic phenomena at the length scale of a single pore. Thus basic laws are cast in terms of local averages which allow equations to be written in continuous variables despite the microscopic discontinuities. Local microscopic mechanical and thermodynamical equilibria are usually assumed. Macroscopic nonequilibria, such as temperature variation, may be viewed as driving potentials, while equilibrium relations are applied locally. This approach is valid in soil physics since events occur slowly and phases are in contact over large specific surfaces. Soil hysteresis can greatly complicate computations, but no frost heave theories as yet include it. Latent heat effects are large and may be computationally troublesome, especially as they enter via non-linearities; references [3, 4] give useful reviews. If phase transition is concentrated within a macroscopic discontinuity, then a variety of methods recently devised for Stefan-type problems may be used [5-7]. If the scale of analysis entails a spatially extended phase change zone, a non-step freezing zone representation follows naturally [8, 9].

Conservation principles of momentum, energy, and mass enter into a complete frost heave model. Experimental evidence shows that an increase in overburden pressure diminishes the heave produced by forces acting in the freezing zone [10]; overburden pressure is defined macroscopically as the net force per unit area over all phases and constituents. The pressure in any phase will normally differ from the overburden pressure, but may equal it in special circumstances. If the pore water gage pressure in a saturated, coarse-grained, unfrozen soil is zero, then a weight upon the soil must be supported entirely by an "effective stress" transmitted through mineral intergranular contacts. Microscopically the effective stresses are complex and vary abruptly, but macroscopically they manifest themselves simply as a net intergranular force per unit area of soil.

The overburden pressure is not balanced by an effective stress reaction alone but also by the pore content stress. The macroscopic partitioning of pore stress, or "neutral stress," by the pore constituents results from a microscopic interplay of constituent stresses, interfacial energies, and geometry. Miller's [11] suggestion that the use of a stress partitioning factor to quantify the relative participation of macroscopic pore constituent stresses is experimentally supported by [12]. Using thermodynamics, equations were derived relating overburden pressure, ice pressure, liquid water pressure and volume ratios of ice and liquid [13, 14]. Agreement, however, has not yet been reached on the quantitative distribution of stress within the phase change regions.

One may question the validity of viewing stresses as having pressure-like manifestations macroscopically, especially when the stresses refer to an ensemble of grains, or act within extremely thin adsorbed water layers. It has been suggested that film water can pull nearby water into it while

simultaneously pushing away the ice, if the water is stressed anisotropically [15, 16]. However, explanations of the surface layer forces which separate a particle from nearby ice can be advanced without recourse to peculiar water or stresses [17-19]. Data on the film water may not be necessary, if it is locally in equilibrium with the experimentally accessible pore water. Osmotic pressure also enters into the film water - pore water - ice pressure relationships [20]. However, no frost heave model takes account of the effects of solutes nor the effects due to ions in clays.

Ice can maintain a complex state of stress and develop stress concentrations, although it tends to "relax" and relieve such conditions. The idea of simple ice pressure assumes that events happen slowly, allowing micro-variations of stress to be relieved. Since the unfrozen water attains a locally uniform state and the water coats the pore ice which forms in equilibrium with it, the ice will also tend to uniformity. When the soil matrix and unfrozen water react to an imposed load, they may move in any direction, equalizing the apparent state of stress in all constituents. The pore ice may also be considered to be isotropically stressed, behaving as a liquid-like component, with any macroscopic anisotropy entering through the stress-strain characteristics of the granular skeleton.

PHASE COMPOSITION

The water in a soil does not all freeze at a unique temperature; rather, the unfrozen water content decreases with temperature [21], and some temperature drop below 0°C is required before pore ice can exist. The finer grained a soil is, the greater the freezing point depression, and the more unfrozen water it will retain under a given set of conditions. Williams and Burt [22] measured the hydraulic conductivity of frozen soil using a permeameter with an aqueous solution of lactose. However, the lactose complicates the interpretation of the data. To overcome this difficulty Horiguchi and Miller [23] measured the hydraulic conductivity of frozen soil in the temperature range 0°C to -0.15°C using pure, supercooled water. These experiments quantify the moisture flux in frozen soils.

The partially frozen nature of heaving soil accounts for the important unfrozen water flow occurring within the warmest, least frozen portion of the frozen soil. Evidence shows ice segregation occurring within the frozen zone, behind a "frozen fringe" [24, 25]. This fringe may be defined as the zone over which volumetric ice content increases from zero at the freezing front (0°C isotherm) to 100% at the warmest lens. High overburden pressures and low temperature gradients tend to increase the fringe thickness.

An equilibrium relation which appears applicable to the ice-liquid mixture in frozen soil is the generalized Clapeyron equation [26]:

$$(p_w - \pi)v_w - p_i v_i = \frac{L(T-T_o)}{T_o} \qquad (1)$$

Applied microscopically, this equation can be used to explain equilibrium between pore ice and unfrozen water. Macroscopically, p_w and π in this equation presumably pertain to measured pore water values. The Clapeyron equation has been linked with another relation which is usually presented in terms of soil capillarity. When different water phases co-exist with curved phase interfaces (as in soil pores), there is a pressure jump between the two phases due to interfacial effects,

$$p_j - p_w = \sigma_{jw} F(\theta) \qquad (2)$$

where j refers to ice in saturated, frozen soil, or air in unsaturated, unfrozen soil. $F(\theta)$ is a macroscopic function of unfrozen water content, which corresponds to the mean curvature of the microscopic phase interface. Koopmans and Miller [27] hypothesized that $F(\theta)$, at a fixed water content, was the same for an unsaturated, unfrozen soil and a saturated, frozen soil. If p_w is zero, then from Eq. (2)

$$p_i/p_a = \sigma_{iw}/\sigma_{aw} \qquad (3)$$

For the same granular soil, Koopmans and Miller [27] found the pressure ratio to be about 0.45, indicating that capillary effects dominate in the same way for both freezing and drying. For a clay the ratio was unity.

When a particular unfrozen, unsaturated soil is considered, the curves of Eq. (2) are often called soil moisture characteristic curves. Typically, the curves differ somewhat for drying or wetting processes (hysteresis). At atmospheric air pressure, neglecting hysteresis, one obtains a curve relating moisture content and liquid pressure. With a frozen, saturated soil, one may construct freezing characteristic curves by relating θ to p_w and T using Eqs. (1) and (2). For a fixed value of ρ_w, a relation is found between θ and T. Thus unfrozen water relations, discussed earlier, are special cases of freezing characteristic curves at constant pressure. One cannot obtain liquid pressure in a freezing soil directly from an unfrozen moisture characteristic curve. $F(\theta)$ may be the same in both cases, but the ice-water surface tension σ_{iw} must be used in place of σ_{aw} and the product $\sigma_{iw} F(\theta)$ only provides values for the difference between ice and liquid pressures. The ice pressure need not be zero when the overburden pressure is small because, in general, any number of combin-

ations of liquid and ice pressure may hold for given values of θ and overburden pressure. Data do not support application of the same characteristic curve data to both frozen and unfrozen soils, if both soils are unsaturated [20]. In the freezing of unsaturated soil, three pore phases are present, and a more complex relation between pore contents must hold than that in Eq. (2); nevertheless, characteristic curves from unfrozen, unsaturated soil are generally used in frost heave models of unsaturated, freezing soil.

The movement of moisture in soil systems is usually described by some form of Darcy's law. This expresses the conservation of momentum when inertial effects are negligible. The gradient of free energy, or thermodynamic potential, represents a net force balanced by the drag of other components on the liquid, which is proportional to the flow rate. Thus the moisture flux can be written

$$\bar{q} = -k \nabla f \tag{4}$$

The difference in free energy of the soil moisture is the work required to change the water from one state to another. This gives the clearest physical picture of the driving potential for soil water movement since it contains all of the component potentials usually encountered, e.g. pressure, gravity electrostatic forces, osmosis. The free energy can be expressed as a function of external force fields and contact forces. Thus it is common to write the motive potential directly in terms of pressure and gravity fields:

$$\bar{q} = -k \nabla \left(\frac{\psi}{\rho_w g} + z\right) \tag{5}$$

The soil suction ψ is a combination of hydrostatic pressure, surface tension and adsorption forces of the soil skeleton. The soil suction ψ can be expressed as a function of water content and temperature, as noted already. Then the extended Darcy's law is

$$\frac{\bar{q}}{\rho_w} = -D_\theta \nabla \theta - D_T \nabla T - k \bar{e}_z \tag{6}$$

Eq. (6) may be substituted into the conservation of mass equation to arrive at

$$\frac{\partial \theta}{\partial t} = \nabla[D_\theta \nabla\theta + D_T \nabla T + k \bar{e}_z] \tag{7}$$

No frost heave models have used the mass flux relation as a function of both pressure and temperature gradients, exemplified by Eq. (6). If one-dimensional flow is considered, with negligible temperature effects, then

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial x}\left(D \frac{\partial \theta}{\partial x}\right) \tag{8}$$

In a number of frost heave models, soil moisture diffusivity is used in the same manner for frozen and unfrozen soil, evidently by analogy between drying of unsaturated soil and freezing of saturated soil. However, the analogy may not hold. In fact, the D_T term can be dominant in saturated, granular soil. Unlike p_a, the ice pressure will not be constant in space, but should vary from some small value where ice content is low to the overburden pressure where segregation is occurring. According to Eq. (2), the liquid pressure gradient is then

$$\frac{\partial p_w}{\partial x} = \frac{\partial p_i}{\partial x} - \sigma_{iw}\left[\left(\frac{\partial F}{\partial \theta}\right)_T \frac{\partial \theta}{\partial x} + \left(\frac{\partial F}{\partial T}\right)_\theta \frac{\partial T}{\partial x}\right] \tag{9}$$

Thus the liquid pressure gradient depends upon ice pressure gradient, and gradients in θ and T. No simple relation for diffusivity exists unless the ice pressure and thermal gradients are negligible.

Pressure or temperature gradients may cause relative motion between ice and small particles completely enveloped in the ice [18, 28]. The processes involved are not completely understood but the ice must melt on one side of a particle, with liquid being transported around the particle to a location where it refreezes. The process, called regelation, may operate in heaving soils, so that movement of ice as well as unfrozen water may occur [11]. The extent of such activity in soil has not been established, but the possibility of pore ice movement exists.

ISOTHERMAL TRANSPORT OF WATER IN SUBFREEZING SOIL WITHOUT ICE

Frozen soils need not contain ice if the moisture content is less than the equilibrium value for the temperature of interest. To examine the movement of water under such conditions, tests were run whereby wet and dry soil columns were joined and the water was observed to move from the wet, ice-free part of the frozen soil column to the dry region [29].

Assuming the extended Darcy law is valid, the movement of the soil water is governed by Eq. (8), with the following conditions:

$$\theta(x,0) = \begin{cases} \theta_o & x < 0 \\ \theta_c & x \geq 0 \end{cases} \tag{8a}$$

$$\lim_{x \to -\infty} \theta(x,t) = \theta_o$$
$$\lim_{x \to +\infty} \theta(x,t) = \theta_c \qquad t > 0 \qquad (8b)$$

Nakano [30] showed that a weak solution of Eq. (8-8b) converges asymptotically to a similarity solution of Eq. (8), assuming that the soil diffusivity is given by

$$D(\theta) = D_o (\theta - \theta_c)^{\beta} \qquad (10)$$

An experimental asymptote can be obtained at some finite time, chosen so that the moisture profile no longer varies with respect to a similarity parameter. With a best fit of the experimental and theoretical profiles the values of D_o and β can be found. This method is also applicable to the determination of the diffusivity for unfrozen porous media and is a useful new technique.

The soil was Morin clay with a specific surface area of 60 m^2/g-soil. The measured unfrozen water content (with 21% total water content) versus temperature showed that about 10% of the water remained unfrozen at -1.0°C.

Graduated plastic pipettes were used to enclose soil columns 0.80 cm in diameter by 20.33 cm in length. Two soil columns with the same soil dry density of 1.3 g/cm^3 were required per test, one uniformly dry and the other uniformly wet at 8.0%. The thoroughly mixed water and soil was allowed to set for at least three days to attain moisture uniformity.

An experiment began by joining pipettes with dry and wet soil, both held at -1.0°C. The surfaces of contact were smoothed to avoid contact resistance. The joined pipettes were then placed in a constant temperature bath containing an ethylene glycol/water mixture held at -1.0°C and maintained to within ±0.03°C. After a specified time, the compound pipette was removed from the constant temperature bath and quickly sectioned into 64 equal segments for water content determination.

Tests were carried out with time durations of 2, 4, 8 and 16 days and also at 4, 8, 16 and 32 days, to examine the reproducibility of the data. The water content is plotted versus the distance in Fig. 1. The two sets of profiles, after 8 days, indicate the difficulty in reproducing the same profile. There was about 0.5% water content difference in the wetter end of a soil column. The water content θ is plotted versus a similarity parameter ξ (Fig. 2), which indicates that the experimental profiles tend to stabilize as if they were converging upon an asymptotic profile reasonably well after 8 days. Using the least-squares method the parameters in Eq. (10) were determined as

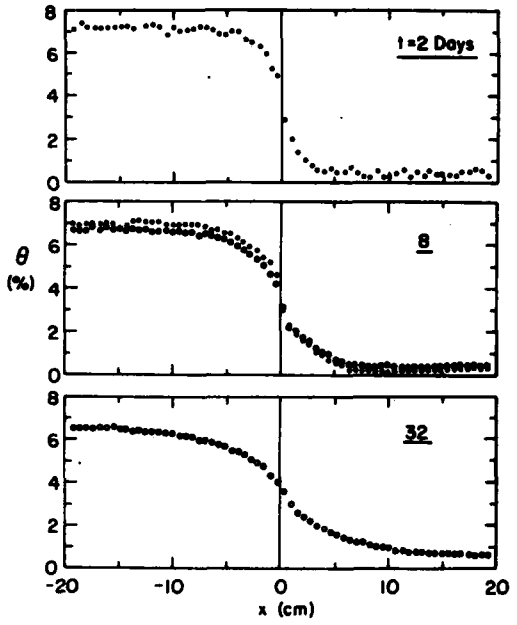


Figure 1. Water movement in ice-free Morin clay. θ_o = 8.0%; T = -1.0°C.

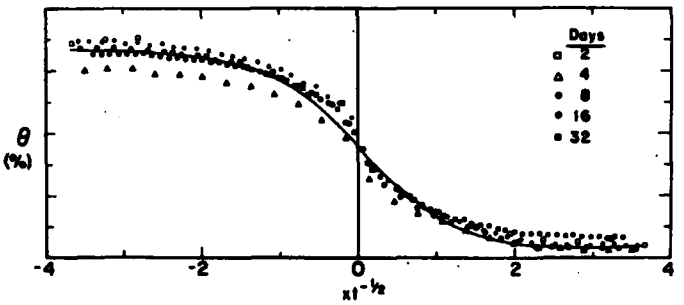


Figure 2. Similarity correlation of water content in ice-free Morin clay. θ_o = 8.0%; T = -1.0°C.

$$D(\theta) = 1.21\ (\theta - .0038)^{.272} \qquad (11)$$

where $D(\theta)$ is in cm^2/day and θ is in g water/g dry soil. The curve in Fig. 2 is the computed θ profile using Eq. (11) and is a reasonably good approximation, but is not quite a close fit. If $D(\theta)$ is evaluated directly from the experimental profile at 8 days, using the same mathematical basis, it appears that there is a slight peak in $D(\theta)$ near θ = 1%. The validity of this calculation is uncertain, due to the error in the numerical evaluation of derivatives, but it is possible that the fit is somewhat erroneous because the actual $D(\theta)$ may not be a monotonically increasing function of θ.

ISOTHERMAL TRANSPORT OF WATER IN SOIL WITH ICE

Tests were run using the same technique and apparatus as described in the previous section except that the water content was adjusted so that ice was present in the wet section [31]. The total water content was 20.0% and the dry density of the clay was 1.42 Mg/m^3. The unfrozen water content was evaluated by a pulsed nuclear magnetic resonance (NMR) technique to yield the following unfrozen water relation for a warming soil:

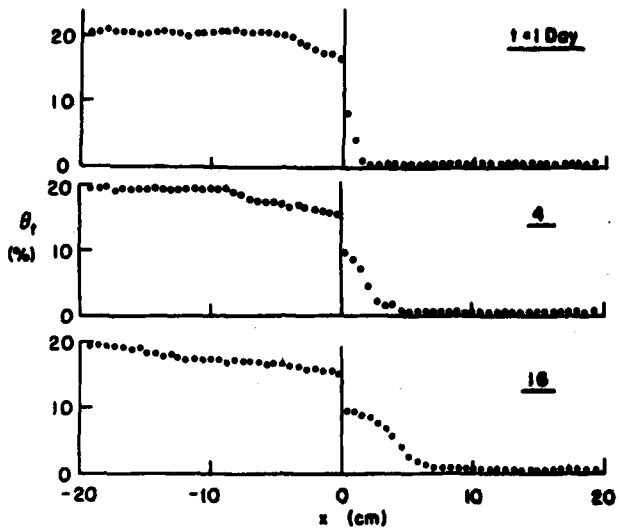


Figure 3. Water movement in Morin clay with ice present. θ_o = 20.0%; T = -1.0°C.

$$\theta = .1269\ (-T)^{-.5080} \qquad (12)$$

Since the tests were run at -1°C, the initial water and ice contents were 12.7% and 7.3% respectively. Some test results are shown in Fig. 3, which indicate that water moved slowly but steadily from the wet part of the soil column to the dry part.

Supplemental tests were carried out to determine if the phase equilibrium of the unfrozen water was maintained. The total water content of the soil was measured gravimetrically while the unfrozen water content was determined by NMR. As the total water content decreased the unfrozen water content remained at the initial value until the total water content dropped to about 14%. The unfrozen water content of the wet column remained constant at the equilibrium unfrozen water content at -1°C and any net water removal from the wet soil column was at the expense of the ice. The ice which melted in the wet column absorbed energy from the surroundings but the system temperature did not vary more than ±.01°C. Thus the ice melted so slowly that latent energy could be transferred with a very minor temperature gradient.

If Darcy's law is valid for water flow with ice then

$$q = -\rho_w D(\theta_t) \frac{\partial \theta_t}{\partial x} \qquad (13)$$

Using Eq. (13) it was possible to calculate the diffusivity using the data for θ_t and q as

$$D(\theta_t) = 49.9\ (\theta_{to} - \theta_t)^{.678} \qquad (14)$$

By solving the water content equation with a diffusivity given by Eq. (14), the cumulative water flux was predicted as a function of time and position. Since these values compared quite well to the data, a Darcy type law may govern the flow of film water in icy soils under conditions similar to those of the experiment.

FREEZING ZONE TEMPERATURE AND PRESSURE

Although moisture tensions are known to exist below an advancing freezing front, and are suspected to exist within the freezing zone itself, the cause of these tensions has not been adequately explored. Critical assumptions in frost heave theories are not adequately supported by data on temperatures and pressures that exist within the freezing zone. Tests were designed to provide concurrent measurements of temperature and pressure in the moisture phase of freezing silt soils. Graves sandy silt and Northwest silt were frozen unidirectionally from the upper surface with the lateral surfaces insulated. Selected physical properties for these soils are listed in reference [31].

A special apparatus was designed and fabricated to contain the 300-mm-long soil specimen during freezing and to support the measuring devices for temperature and moisture tension [31]. Top and bottom temperatures were maintained to a tolerance of 0.01°C by separate fluid baths; the upper temperature was adjusted daily to advance the freezing front into the specimen at a generally decreasing rate between 20 and 5 mm/day over a 24-day period. De-aired water was supplied continuously to the bottom of the soil through a porous stone connected to an external 4°C water supply held level with the bottom of the specimen.

Temperatures along the length of the soil column were measured using calibrated thermistors mounted within the wall of the containing vessel in direct contact with the soil. These registered soil temperatures within 0.05°C. Moisture tensions were measured in the unfrozen and frozen portions of the soil column by means of special tensiometers designed at CRREL for temperatures above and below 0°C [32]. A 10% solution of ethylene glycol and water was used within the porous cup which remains unfrozen to a temperature of approximately -4°C and does not diffuse into the soil water. Within experimental error the soil moisture tension is found to be the same for the 10% solution as for water.

Figure 4 is an example of the data obtained for the Graves sandy silt. The Northwest soil, a true silt, developed similar features. Tension was registered continuously by the tensiometers and differed significantly from atmospheric pressure only when the soil temperature was below

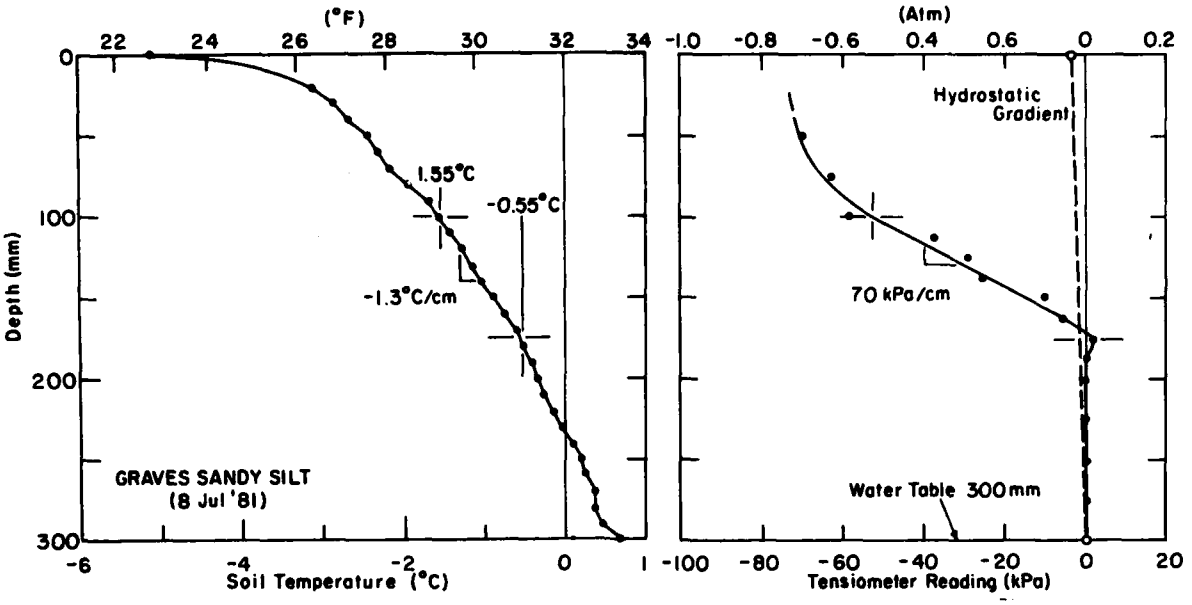


Figure 4. Temperature and apparent moisture tension in freezing zone of Graves sandy silt.

-0.4°C. The value of moisture tension increased rapidly as temperatures became lower, finally reaching an apparent maximum of approximately 70 kPa, after which it remained constant or fell off slightly at a temperature of -1.5°C for the two soils tested. Water flowed through the unfrozen soil at a barely detectable moisture tension gradient, consistent with the measured water intake and the hydraulic conductivity of the soils tested. Moisture tension gradients in the freezing zone reached a maximum of approximately 2 kPa/mm.

Figure 5a is a graph of the moisture tensions and temperatures measured in the freezing zone of the Graves sandy silt plotted for various depths. Figure 5b is a graph of unfrozen water content for this soil, as a function of below-freezing temperature, measured with the NMR procedure.

Combining the data of Fig. 5a and 5b leads to an important new step in the assessment of frost heaving: a moisture-tension characteristic for the unfrozen water behind the freezing front (Fig. 6). The combining of pressure and temperature data assumes that the pressure measured by the tensiometers is that of the unfrozen water itself, i.e. the state of the water is the same in the active freezing zone as it is in the unfrozen water tests. An observed correspondence in the two types of test provides strong support for this assumption: the temperature of nucleation found by ice seeding in the unfrozen water tests (approximately -0.5°C) is the same temperature at which moisture tension first exceeds the hydrostatic level in

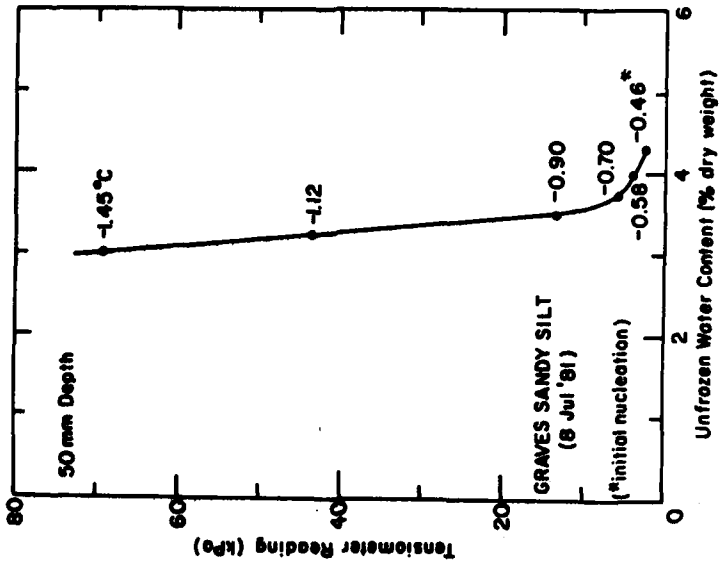


Figure 6. Freezing-zone moisture characteristic; Graves sandy silt.

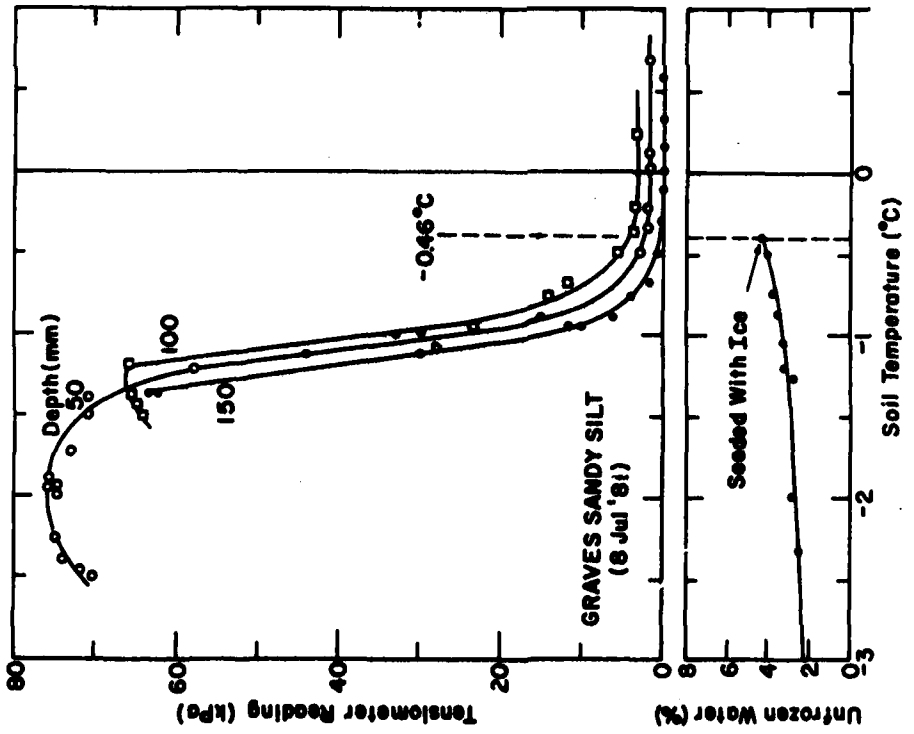


Figure 5. Moisture tension (a) and unfrozen water content (b) versus temperature in freezing zone of Graves sandy silt.

the advancing freezing front. Moreover, a steep gradient of tension develops at temperatures immediately lower than -0.5°C.

Unlike a moisture characteristic for an unfrozen soil, the freezing-zone moisture characteristic is not independent of temperature; each intersection of moisture tension and unfrozen water content on such a plot corresponds to a narow range of temperatures below 0°C. The range is most likely the result of local variations in gradation, density, or water content in the vicinity of a tensiometer. A normal moisture characteristic corresponds to an air/water interfacial system, while the freezing-zone moisture characteristic apparently conforms to a water/ice interfacial system having a lower surface energy.

CONCLUSIONS

New relations for the flow of water in partially frozen soils are now being incorporated into frost heave models. The thermodynamic equilibrium equations for freezing soil water are well established but the partition of the soil system stresses within freezing zones is poorly understood.

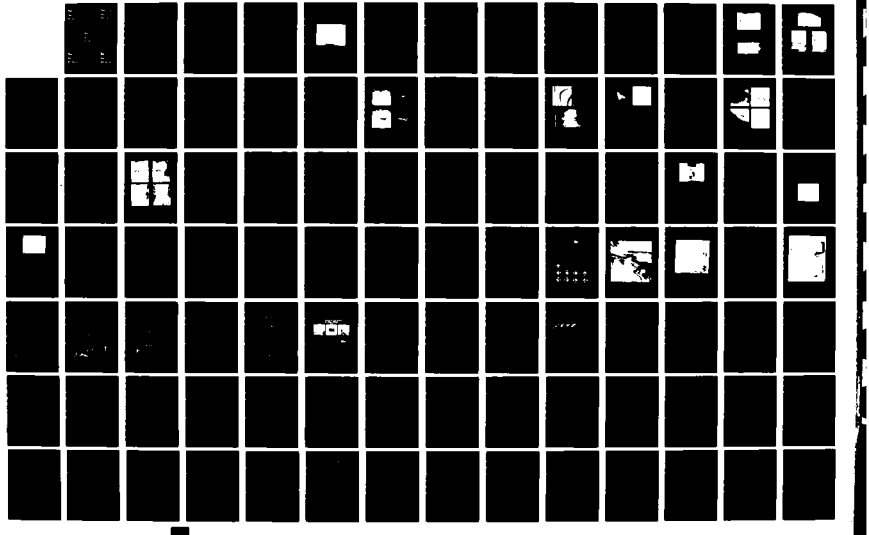
It has been demonstrated that the hydraulic conductivity of water in frozen soils is small but non-zero. Over time spans of several years, which are appropriate for new large scale engineering projects in cold climates, water flow in frozen soils will be significant. The experimental work has shown that Darcy's law is valid for film water flow in ice-free frozen soils and is a reasonable approximation for frozen soil with ice if the ice pressure gradient is negligible.

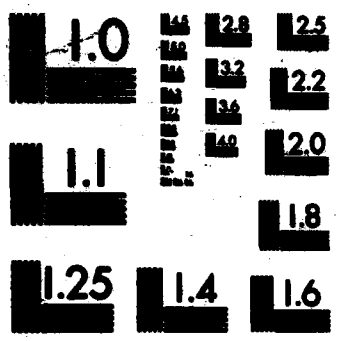
The capability of deriving the freezing-zone moisture characteristic for a frost-susceptible soil, demonstrated here, is important in the theoretical prediction of the amount of segregated ice in a freezing soil. The empirical correspondence between temperature, moisture tension, and unfrozen water content may lead to physically realistic computer algorithms for the freezing zone. Quantitative predictions from frost heave models will rely on new information of this kind.

REFERENCES

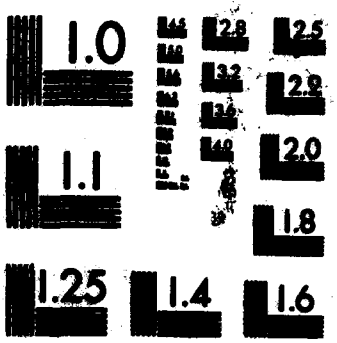
1. Beskow, G., "Soil Freezing and Frost Heaving, with Special Reference to Highways and Railroads." Statens Vaginstitut, Meddelande 48, Swedish Geological Society, Series C, No. 375, Stockholm, 1935, 242 pp.
2. Taber, S., "Surface Heaving Caused by Segregation of Water Forming Ice Crystals." Eng. News-Rec., Vol. 81, 1918, pp. 683-684.
3. Crank, J., The Mathematics of Diffusion, Oxford at Clarendon Press, 1975.

AD-A120 812 PROCEEDINGS OF THE 1982 ARMY SCIENCE CONFERENCE HELD AT 4/5
THE UNITED STATES..(U) DEPUTY CHIEF OF STAFF FOR
RESEARCH DEVELOPMENT AND ACQUISITIO.. 18 JUN 82
UNCLASSIFIED F/G 5/2 NL

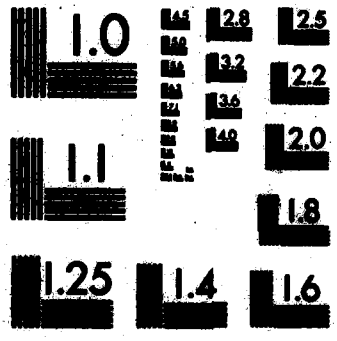




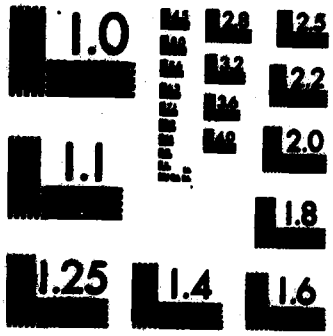
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A



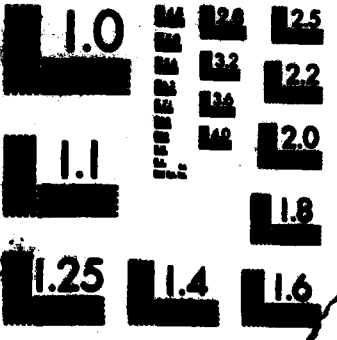
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

4. Lunardini, V., Heat Transfer in Cold Climates, Chap. 9. Van Nostrand Reinhold, 1981, New York.
5. Crank, J. and R.S. Gupta, "Isotherm Migration Method in Two Dimensions." Int. J. Heat Mass Trans., Vol. 18, 1975, pp. 1101.
6. Bonnerot, R. and P. Jamet, "A Third Order Accurate Discontinuous FEM for the One-Dimensional Stefan Problem." J. Comp. Phys., Vol. 32, 1979, pp. 145-167.
7. O'Neill, K. and D.R. Lynch, "A Finite Element Solution for Porous Medium Freezing Using Hermite Basis Functions and a Continuously Deforming Coordinate System." Proceedings First International Conference on Numerical Methods in Thermal Problems, 1979, Pineridge Press, Swansea.
8. Luthin, J. and G. Taylor, "A Model for Coupled Heat and Moisture Transfer During Soil Freezing." Can. Geo. J., Vol. 15, 1978, pp. 548-555.
9. O'Neill, K. and R.D. Miller, "Numerical Solutions for Rigid Ice Model of Secondary Frost Heave." Second International Symposium on Ground Freezing, Norwegian Institute of Technology, June 24-26, 1980, pp. 656-669.
10. Penner, E. and T. Ueda, "Effects of Temperature and Pressure on Frost Heaving." Eng. Geol., Vol. 13 (1-4), 1979, pp. 29-39.
11. Miller, R., " Frost Heaving in Non-Colloidal Soils. Third International Conference on Permafrost, Edmonton, Alberta, 1978, pp. 707-713.
12. Snyder, V.A., "Theoretical Aspects and Measurements of Tensile Strength in Unsaturated Soils." PhD. Thesis, 1980, Cornell U., Ithaca, N.Y.
13. Groenvelt, P.H. and B.D. Kay, "Water and Ice Potentials in Frozen Soils." Water Res. Res., Vol. 13, No. 2, 1977, pp. 445-449.
14. Groenvelt, P.H. and B.D. Kay, "Pressure Distribution and Effective Stress in Frozen Soils." Second International Symposium on Ground Freezing, Norwegian Institute of Technology, Trondheim, Norway, June 24-26, 1980, pp. 597-610.
15. Vignes, M. and K. Dijkema, "A Model for the Freezing of Water in a Dispersed Medium." J. Coll. Int. Sci., Vol. 49, 1974, pp. 165-172.
16. Takagi, S., "Summary of the Adsorption Force of Frost Heaving." Cold Reg. Sci. Tech., Vol. 3, No. 2-3, 1980, pp. 233-35.
17. Derjaginn, B.V., and N.V. Churaev, " The Theory of Frost Heaving." J. Coll. Int. Sci., Vol. 67, 1978, pp. 391-396.
18. Gilpin, R.R., "A Model of the 'Liquid-like' Layer Between Ice and a Substrate with Applications to Wire Regelation and Particle Migration." J. Coll. Int. Sci., Vol. 68, No. 2, 1979, pp. 235-251.
19. Gilpin, R.R., "A Model for Prediction of Ice Lensing and Frost Heave in Soils." J. Water Res. Res., Vol. 6, No. 5, 1980, pp. 918-930.
20. Miller, R.D., "Freezing Phenomena in Soils." Chap. 11 in Fundamentals of Soil Physics, D. Hillel, Ed., Academic Press, 1980.
21. Tice, A.R., C.M. Burrous, and D.M. Anderson, "Phase Composition Measurements on Soils at Very High Water Contents by the Pulsed Nuclear Magnetic Resonance Technique." Trans. Res. Rec., No. 675, 1978, p. 11-14.

22. Williams, P.J. and T.P. Burt, "Measurement of Hydraulic Conductivity of Frozen Soils." Can. Geotech. J., Vol. 11, 1974, p. 674-680.
23. Horiguchi, K. and R.D. Miller, "Experimental Studies with Frozen Soil in an 'Ice Sandwich' Permeameter." Cold Reg. Sci. Tech., Vol. 3, 1980, p. 177-183.
24. Hoekstra, P. "Water Movement and Freezing Pressures." J. Soil Sci. Soc. Am., Vol. 33, 1969, pp. 512-518.
25. Loch, J. and B.D. Kay, "Water Redistribution in Partially Frozen, Saturated Silt Under Several Temperature Gradients and Overburden Loads." J. Soil Sci. Soc. Am., Vol. 42, 1978, pp. 400-406.
26. Loch, J., "Thermodynamic Equilibrium Between Ice and Water in Porous Media." Soil Sci., Vol. 126, 1978, pp. 77-80.
27. Koopmans, R.W.R. and R. Miller," Soil Freezing and Soil Water Characteristic Curves." J. Soil Sci. Soc. Am., Vol. 30, 1966, pp. 680-685.
28. Romkens, M.J.M. and R. Miller, "Migration of Mineral Particles in Ice with a Temperature Gradient." J. Coll. Int. Sci., Vol. 42, 1973, pp. 103-111.
29. Nakano, Y., A. Tice, J. Oliphant and T. Jenkins, "Transport of Water in Frozen Soils: I. Experimental Determination of Soil-Water Diffusivity Under Isothermal Conditions." Adv. in Water Res., in press.
30. Nakano, Y., "Asymptotic Behavior of Solutions to the Problem of Wetting Fronts in One Dimensional, Horizontal and Infinite Porous Media." Adv. in Water Res., in press.
31. Lunardini, V.J., Ed., "Water Movement and Properties in Frozen and Freezing Soils - Report on 1981 ILIR Project." CRREL Internal Report 748, Feb. 1982.
32. Ingersoll, J.E., "Laboratory and Field Use of Soil Tensiometers Above and Below 0°C." U.S. Army CRREL Special Report 81-7, 1981.

NOMENCLATURE

D	diffusivity
D_θ	$k/\rho_w g\left(\frac{\partial\psi}{\partial\theta}\right)_T$, moisture diffusivity
D_T	$k/\rho_w g\left(\frac{\partial\psi}{\partial T}\right)_\theta$, temperature diffusivity
$\bar{e}_z$	unit vector in z-direction
f	free energy
F	surface tension function
g	gravitational constant
k	hydraulic conductivity
L	latent heat of fusion
p	pressure
q	liquid flux
t	time
T	temperature
T_o	freezing temperature of bulk water
v	specific volume
x	coordinate in flow direction
z	coordinate in gravity direction
θ	liquid water content
θ_o	initial water content
θ_t	total water content
ξ	$x\ t^{-\frac{1}{2}}$
π	osmotic pressure
ρ	density
σ	surface tension
ψ	matric potential
∇	nabla operator

Subscripts

a,i,w	air, ice, water, respectively
aw,iw	air/water, ice/water respectively

*LUND, BEATRICE, SCHUSCHEREBA

BIOEFFECTS DATA CONCERNING THE SAFE USE OF GaAs LASER TRAINING DEVICES (U)

DAVID J. LUND
EDWIN S. BEATRICE, COL MC
STEVEN T. SCHUSCHEREBA

Modern weaponry poses a vexing problem to the military training community: how can realistic battlefield games be conducted which mimic the undeniable ability of weaponry to produce lethal action at a distance? A child's pointed "BANG, you're dead!" suffers in that it has an unverifiable effect, is limited in range, and exposes the originator to immediate retaliation by the target. It does, however, contain the essential ingredients of a weapons simulation system; the ability to engage a target and transmit a message which can be interpreted by the target in terms of its subsequent ability to function. A key aspect of the message is that it implies a fatal hit but does not possess that element of real lethality.

Systems have been devised which simulate the firing of live munitions for training purposes. The MILES system, developed by PM TRADE, is an example. A gallium arsenide (GaAs) laser transmitter is mounted on, and boresighted with, each weapon, and all potential targets are equipped with detectors sensitive to GaAs laser radiation. When a weapon is triggered, no projectile is fired; rather a signal is transmitted from the laser and directed at the intended target. The success of the round is scored upon receipt of the signal at the target. The system is effective; however it does present a new problem. The MILES system transmits laser beams at personnel; the probability of their eyes being [illegible] is high. It is essential that the consequences of such ocular exposure understood to insure that the signal used for simulation does not carry potential harm.

Within the Army, safety restrictions on the use of all lasers are governed by the provisions of laser safety standards, AR 40-46 and TBMED 279 (1,2). In a quest for effectiveness, the designers of MILES have pushed the emitted power to the maximum permissible exposure (MPE) allowed by the standards. An understanding of the potential hazard of the MILES system can be gained by testing the accuracy of the provisions of the safety standards.

TBMED 279 dictates the MPE for laser viewing as a function of several laser parameters including wavelength, exposure duration, effective irradiance diameter and repetitive pulse factors. The relative values of the MPE were derived from bioeffects data which related the potential for tissue alteration to the operational characteristics of the laser. The eye is the most vulnerable part of the body to visible and near infrared laser radiation. This is true because within the eye a light absorbing layer of tissue (retina) lies at the focus of the eye lens system. Just as the rays of the sun can be concentrated by a lens to burn wood, so is a laser beam concentrated onto the retina where the concentrated energy can induce thermal, mechanical, and chemical processes which alter the retinal tissue.

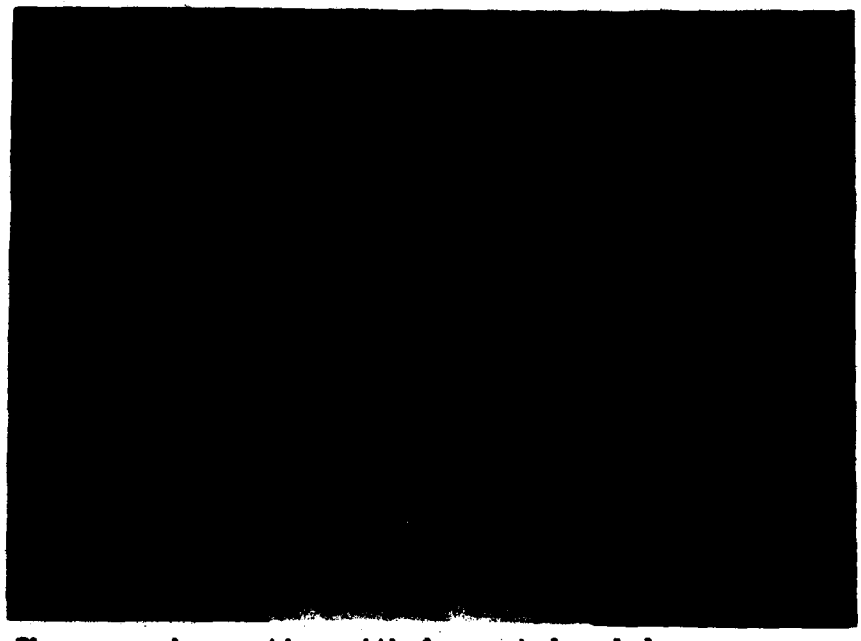


FIGURE 1. Rhesus monkey retina with laser-induced damage.

The MPE has, for the most part, been based upon the ED_{50} for visible retinal alteration in rhesus monkeys under a given set of exposure conditions. What does that mean? Figure 1 is a photograph of the retina of a live rhesus monkey. The retina is the thin layer of tissue at the back of the eye which contains the visual photodetectors. Damage to the retina can diminish vision. The relatively dark circle near the center is the macula, the area of central vision, and in the center of the macula lies the fovea, the area of most acute vision. Around the periphery of the macula in this photograph are a series of small white spots. These are alterations to the tissue caused by laser exposure. The criterion for laser induced damage, in studies upon which MPEs are based, is the appearance of such a visible alteration. The retina is not uniform in

appearance but exhibits large and small scale variations in pigmentation. Because of this variation, the proportion of an incident laser beam absorbed by the tissue will not be the same for all exposure sites. If a series of retinal sites are subjected to laser exposure all at the same incident energy, not all will exhibit the same effect. Some sites will show visible alteration; some will not. For a given incident energy level, there will be a probability of alteration, computed by dividing the number of exposures producing alteration by the total number of exposures. When such an experiment is performed for a number of exposure levels, a curve is derived which relates the probability of alteration to the exposure level. The probability is low for low level exposures and high for higher level exposures. The data relating the probability of damage to the exposure energy can be processed by the statistical technique of probit analysis (3) to determine the incident energy having a fifty percent probability of producing alteration. This incident energy, or dose, is called the ED_{50}. The ED_{50} is not a damage threshold but rather a statistical point which has greater confidence than any other point on the dose-response curve. The MPE is the maximum permissible exposure for safe viewing of laser radiation. No alteration should ever occur upon exposure at the MPE, which therefore must be lower than the ED_{50}. Based on a number of considerations, the MPE has been chosen to be a factor of 10 to 100 below the ED_{50} upon which it is based. The variation of safety factor results from simplification of the dependence of MPE upon exposure parameters.

The most recent version of TBMED 279 was issued in 1975. At that time essentially no laser bioeffects data existed for the spectral range between 694.3 nm (ruby laser) and 1060 nm (neodymium laser). Faced with the absence of data, the writers of the standards chose to compute the MPE in this range by interpolation between the MPEs at 694.3 nm and 1060 nm. The interpolation was not arbitrary but was based on the transmission and absorption properties of ocular tissue. Figure 2 shows the relationship between the MPE and wavelength for the visible and near infrared. The ED_{50}s are also shown for some specific laser wavelengths, valid for ocular exposure to single short duration pulses. The MPE and ED_{50} are given in terms of total intraocular energy (TIE), that is, the total energy entering the eye. If the MPE is to be an accurate derivation of the ED_{50} for wavelengths between 694.3 nm and 1064 nm, the ED_{50} should closely follow a straight line between these wavelengths. The ED_{50} for 850 nm (erbium laser), recently obtained at LAIR (4), lies significantly below the expected value. Thus we have doubts about assumptions underlying the interpolated MPE in this spectral region. In light of this evidence, it became urgent to determine the wavelength dependence of ED_{50} near the GaAs wavelength of 900 nm.

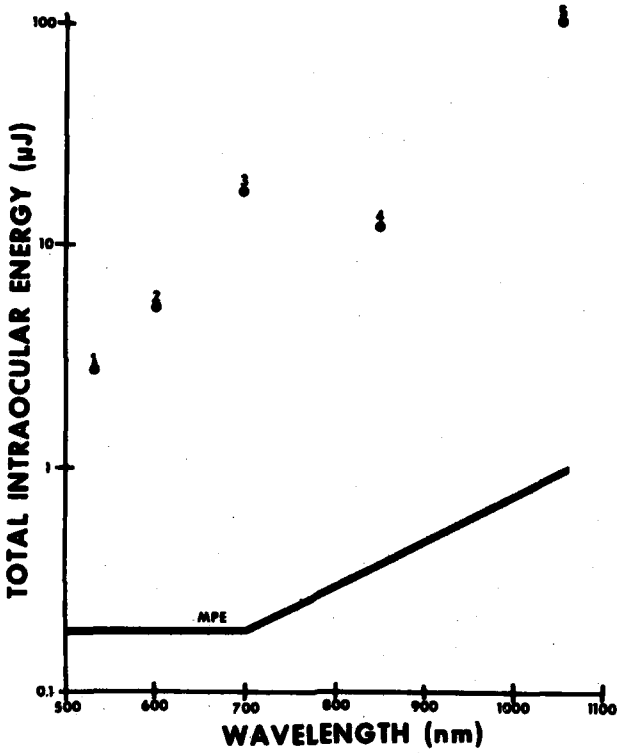


FIGURE 2. Wavelength dependence of the maximum permissible exposure (MPE) and the ED_{50} for retinal damage in rhesus monkey. Given ED_{50}s are for: 1 - 532 nm, 140 ns; 2 - 600 nm ,400 ns; 3 - 694.3 nm, 15 ns; 4 - 850 nm, 180 ns; 5 - 1064 nm, 180 ns

Advances in dye laser technology have made this study possible. The lasing medium of such a laser is a fluorescent dye carried in a suitable liquid solvent. When optically pumped with intense radiation from a flashlamp or laser, the dye can be caused to emit laser radiation at any wavelength within its fluorescent spectrum. The fluorescent bandwidth of most dyes is nanometers wide, and dyes are available which allow selection of any wavelength from the ultraviolet to greater than 900 nm.

The experiments and results of our studies with eyes of rhesus monkeys are reported in this document.

MATERIALS AND METHODS

The source of laser radiation in this study was a Molectron DL-18 dye laser coupled to a Molectron MY33 Nd:YAG laser. The neodymium laser

emitted 15 ns duration pulses at a repetition rate of 10 Hz and was provided with internal second and third harmonic generators which could be positioned so that the laser output was any of three wavelengths: 1064 nm (fundamental), 532 nm (second harmonic), or 355 nm (third harmonic). The output energy of the neodymium laser served as an excitation source for the dye laser which consisted of a side-pumped cell through which the dye was circulated, a diffraction grating which served as a wavelength tunable resonator mirror, and an output resonator mirror. The output wavelength of the dye laser was determined by the grating with the restriction that it be tuned to a wavelength within the fluorescent spectrum of the dye. The fluorescent spectrum of the dye in turn was dependent on the specific dye chosen, the dye solvent and concentration, and the excitation wavelength.

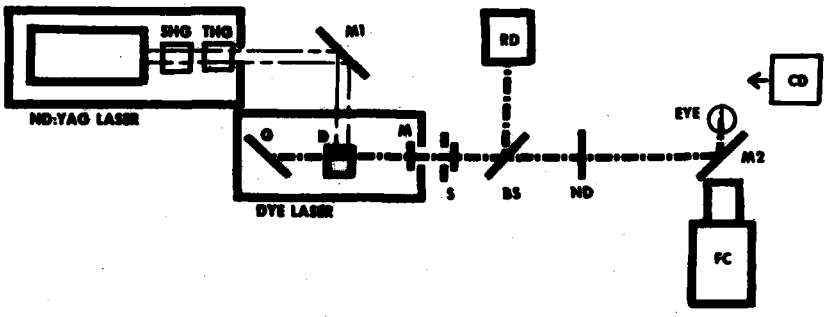


FIGURE 3. Exposure configuration.
BS - beamsplitter CD - calibrated detector for dosimetry
D - laser dye cell FC - fundus camera
G - diffraction grating for dye laser tuning
M - dye laser output mirror
M1 - redirecting mirror for Nd:YAG beam
M2 - dichroic mirror ND - neutral density attenuating filter
RD - reference detector S - shutter
SHG - second harmonic generator THG- third harmonic generator

Figure 3 is a schematic of the exposure system. The laser emitted a continuous train of pulses at 10 Hz; a shutter allowed selection of a single pulse for exposure. A beam splitter deflected a constant proportion of the pulse energy into a reference detector while the remainder of the energy passed through attenuators and onto a dichroic mirror. The mirror had high reflectivity at wavelengths longer than 700 nm but was transparent in the visible. A fundus camera, looking through the mirror, permitted observation of the retina to be exposed. The fundus camera, mirror, and laser beam were aligned so that the laser energy reflected by the mirror passed through the center of the ocular pupil and struck the retina at the site corresponding to the crosshairs of the fundus camera viewing optics. Before the rhesus monkey was positioned, a calibrated detector, which

directly read the incidence pulse energy, was placed so that it would receive all the energy that would normally enter the eye. The ratio of the energy at this position to the energy at the reference detector was obtained with the attenuator removed. Subsequently, when the eye was exposed, the energy entering the eye for each exposure was determined by multiplying the energy at the reference detector by the ratio previously determined and by the transmission of the attenuating filter chosen to give the desired energy. The laser wavelength, beam divergence, and pulse duration were determined for each wavelength. A Jarrell-Ash 1/2 meter spectrometer was used to measure the wavelength. The wavelength scale of the spectrometer was calibrated against a mercury spectral source; the subsequent laser wavelengths error was less than 0.1 nm. The beam divergence was measured by a linear detector array at the focal plane of a one meter lens.

Rhesus monkeys were used in this study. Each animal was sedated and anesthetized for exposure, its pupils were dilated, and the eye to be exposed was held open by a lid speculum. While the eye was open, the cornea was periodically irrigated with physiological saline solution to maintain clarity. For each test, an animal was positioned and 30 exposures were placed in an array in the extramacular retina. The initial exposures in each sequence were at a dose high enough to produce an immediate visible tissue response. Subsequent exposures were at successfully lower doses so that the range of doses in the array varied by about a factor of ten. The retina was photographed and the exposure sites marked on the photograph for subsequent identification. The exposure sites were examined by ophthalmoscope one hour after exposure and the presence or absence of visible alteration noted for each site. The response at each site was correlated to the dose at that site. For each wavelength, the data obtained by exposure of four to six eyes were statistically evaluated to determine the ED_{50} and associated 95% confidence limits. One animal, exposed to 900 nm radiation, was sacrificed one hour after exposure and the retinas prepared for histological evaluation.

RESULTS

The ED_{50} for single Q-switched exposure was determined for six laser wavelengths obtained from the dye laser. The wavelengths and exposure conditions are listed in Table 1. The solvent for all the laser dyes was DMSO. The laser linewidth at 912 nm was 0.4 nm. For the other wavelengths the laser linewidth was less than 0.08 nm, the resolution limit of the monochrometer used for wavelength measurements. The laser beam was nearly gaussian in profile. The beam divergence was measured at the diameter where the intensity fell to 1/e times the peak value.

*LUND, BEATRICE, SCHUSCHEREBA

TABLE 1
DYE LASER PARAMETERS

WAVELENGTH (nm)	PULSE DURATION (nsec)	BEAM DIVERGENCE (mrad)	DYE	CONCENTRATION (molar)	EXCITATION WAVELENGTH (nm)
850.2	11	1.4	DTTC + HIDC	1.5×10^{-3} 1.1×10^{-3}	532
859.6	10	1.6	IR144	6×10^{-4}	532
867.0	7	1.6	IR144	6×10^{-4}	532
880.0	14	1.4	IR125 + HIDC	1.8×10^{-4} 1.8×10^{-3}	532
899.7	6	1.6	IR125	2×10^{-3}	355
912.0	7	1.6	IR140	3×10^{-3}	532

The ED_{50} for visible retinal alteration at one hour in rhesus monkey is given in Table 2 for each of these wavelengths. Also given are the 95% confidence limits about the ED_{50} and the slope of the regression line, defined as ED_{84}/ED_{50}. The data are for extramacular exposure.

TABLE 2
RETINAL ED_{50} FOR Q-SWITCH DYE LASER EXPOSURE IN RHESUS MONKEY

WAVELENGTH (nm)	ED_{50} (μJ)	95% LIMITS (μJ)	SLOPE
850.2	9.1	7.8-10.7	1.71
859.6	6.7	5.6-8.0	1.77
867.0	5.2	4.4-6.0	1.69
880.0	6.3	5.2-7.7	1.80
899.7	4.3	3.4-5.3	2.38
912.0	5.5	4.6-6.7	1.87

Histologic evaluation

Retinal tissue from two eyes exposed to 900 nm radiation was processed and sectioned for light microscopy. Figure 4 is a retinal photograph of one eye taken just prior to sacrifice of the animal. Sites marked 1, 2, 3 and 5 were each exposed to a train of 100 pulses at 10 Hz. The energy per pulse in the train was 17 μJ. Between sites 1 and 2 and between sites 3 and 5 were placed 6 exposures, each consisting of a single pulse having an energy of 17 μJ. Figure 5 shows a section through one of the extramacular sites exposed to 100 pulses. The sensory retina is a complex tissue within which have been defined a number of layers. The retinal pigment epithelium (RPE) is a single layer of cells which contain the pigment melanin. Melanin is the strongest optical absorber in the retina; thus for most laser wavelengths, the RPE is the center of damage. Each photoreceptor of the retina extends through four layers, the outer segment layer (OS), the inner segment layer (IS), the outer nuclear layer (ON), and the outer plexiform layer (OP). The outer segment of the photoreceptor contains the photochemicals which convert the optical signal to a bioelectric signal, the inner segment and nucleus contain the life support system of the cell, and a nerve process extends into the outer plexiform layer where the bioelectrical signal is passed to other nerve cells which convey the information to the brain. Figure 5 shows that, although the damage to the RPE is slight, the photoreceptors at the exposure site have been damaged throughout their length.

The retina contains two types of photoreceptors; the rods which respond to dim light, and the cones which respond to high ambient light and provide color vision. The two types of photoreceptors are not uniformly distributed in the retina: cones are more common in the macula and rods are more common in the extramacular retina. Figure 6 shows a lesion near the edge of the macula where both rods and cones are found. The dose producing this lesion was the same as that for Figure 5. The RPE is more extensively damaged in this lesion, and again the photoreceptors are damaged throughout their length. Figures 7 and 8 are magnified views of the lesions of Figure 5. It can be seen that although the rods are extensively damaged, the cones are relatively unaltered.

DISCUSSION

When the ED_{50} is plotted as a function of wavelength (Figure 9), a minimum is seen near 900 nm. This is difficult to explain on the basis of the known optical qualities of the rhesus eye. Incident radiation must be absorbed by tissue to produce an alteration. A laser beam entering the eye passes through the cornea, aqueous, lens, and vitreous before reaching the retina. These transparent media absorb a fraction of the incident radiation. Of the radiation reaching the retinal surface, part is reflected and part transmitted through the retina. The remainder is

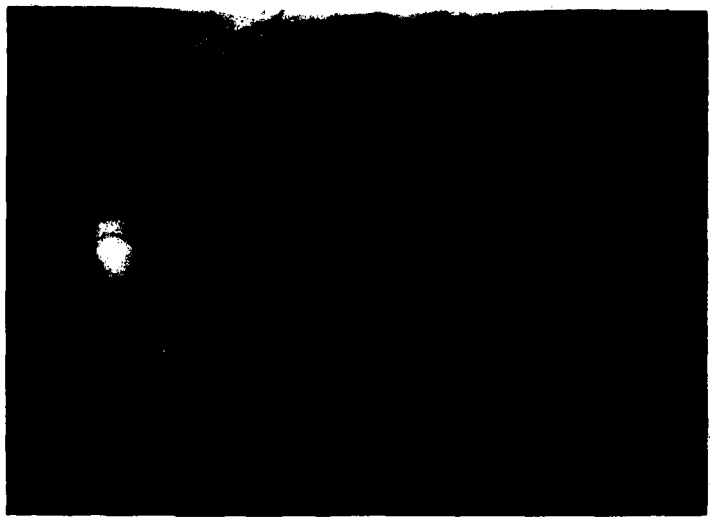


FIGURE 4. Photograph of rhesus monkey retina. The large white area is the optic disc and the dark area the macula. Tiny white spots (arrows) are focal 900 nm dye laser lesions. There are six barely visible lesions between arrows 1 and 2 and between arrows 3 and 5. Arrow 4 points to one of the lower level lesions.



FIGURE 5. Light micrograph of lesion 2 of figure 1. The section shows the following layers of the retina: inner nuclear layer (IN); outer plexiform layer (OP); outer nuclear layer (ON); photoreceptor inner segments (IS); photorecptor outer segments (OS); retinal pigment epithelium (RPE). Small arrows indicate vacuolization. Dark stained nuclei are present in the RPE and ON. Some of the outer segments above the lesion are highly swollen. BAR = 100 µm.

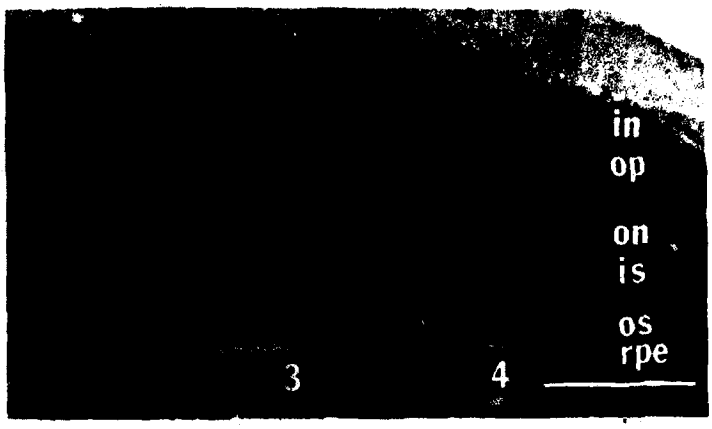


FIGURE 6. Light micrograph of lesions 3 and 4. Large vacuoles are present in the basal region of the RPE in both lesions and small vacuoles are present in the ON. Dark staining nuclei appear in the RPE and the ON. BAR = 100 µm

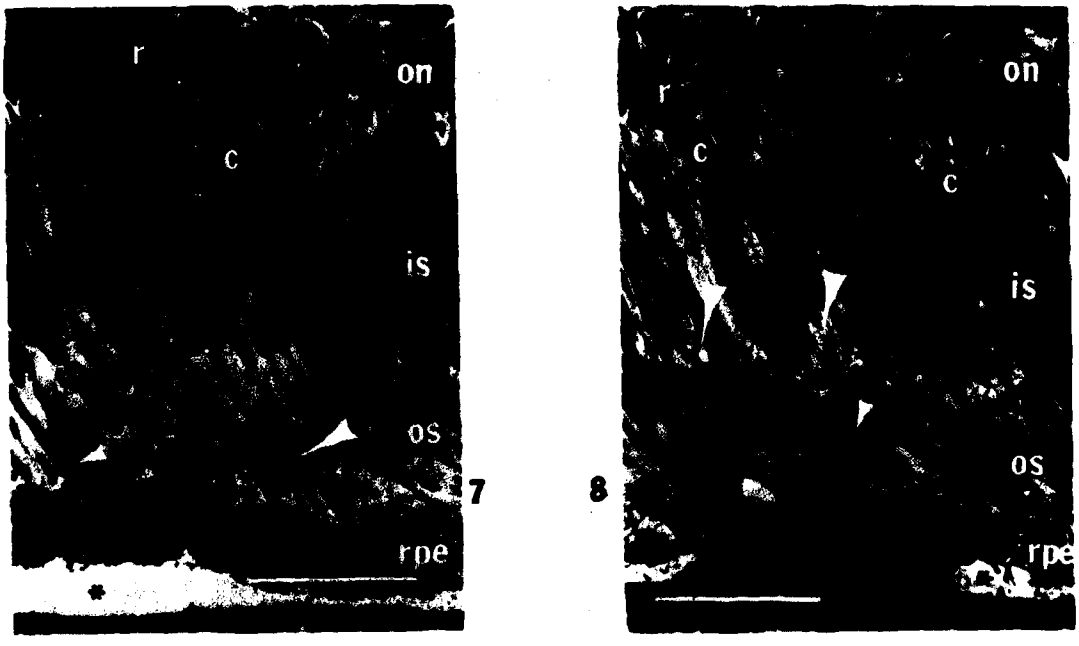


FIGURE 7. Light micrograph of lesion 3. A large vacuole (*) is seen in the basal region of the RPE. Melanin granules are clumped and disarrayed in this layer. Dark staining nuclei are present in the RPE as well as in the ON. Dark stained nuclei belong to rods (r), while cone nuclei (c) stain normally. The slender inner segments of rods are swollen and vacuolated while the larger cone inner segments are normal. BAR = 50 µm

FIGURE 8. Light micrograph of lesion 4. A vacuole (*) is present in the basal region of the RPE. Dark stained nuclei are present in the basal region of the RPE and one dark nucleus is present in the ON. The dark nucleus belongs to a rod (r) that has its outer segments near the lesion in the RPE. Inner segments of rods are highly swollen and vacuolated. The most severe vacuolation occurs at the junction of IS and OS (large arrow) adjacent to the high energy lesion. The cone (c) outer segments show whorl formation about midway in their lengths. BAR = 50 µm

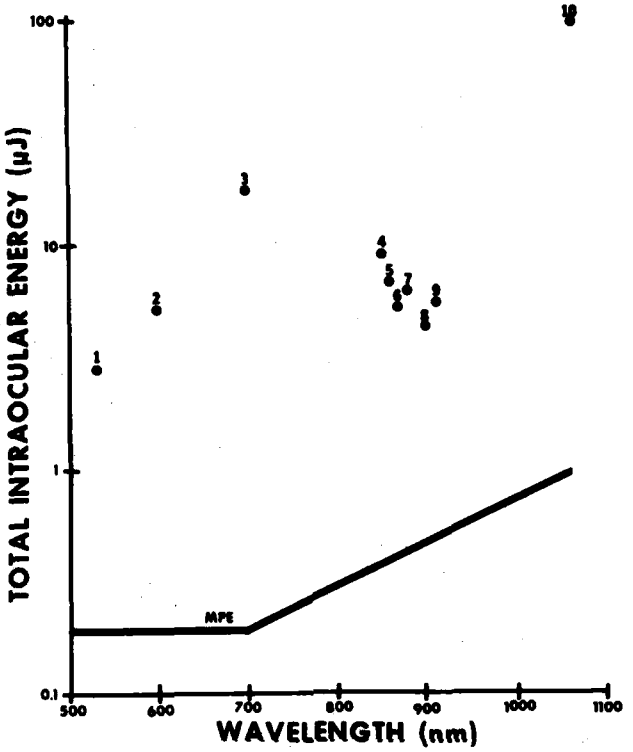


FIGURE 9. Wavelength dependence of the MPE and the ED_{50} for retinal damage in rhesus monkey. Given Ed_{50}s are for:

1 - 532 nm, 140 ns	6 - 867 nm, 7 ns
2 - 600 nm, 400 ns	7 - 880 nm, 14 ns
3 - 694.3 nm, 15 ns	8 - 899.7 nm, 6 ns
4 - 850.2 nm, 11 ns	9 - 912 nm, 7 ns
5 - 859.6 nm, 10 ns	10 - 1064 nm, 180 ns

absorbed in the retinal tissue. Absorption and reflection by the ocular media are wavelength dependent quantities which have been measured by several investigators (5-7). From these data, the energy absorbed in the retina at each wavelength can be computed as a fraction of the energy incident at the cornea. A transformation of this relationship provides the relative energy at the cornea as a function of wavelength to produce constant absorbed energy in the retina (Figure 10). Comparing this result to the bioeffects data for 850 nm to 900 nm, we are driven to consider two possibilities: either there is a flaw in the ocular absorption measurements; or laser radiation in this spectral range produces retinal alteration with lower absorbed energy. The optical absorption measurements of the rhesus eye were performed in vitro with low level illumination.

These measurements would not detect absorption mechanisms which operate only in living tissue nor would they detect power dependent interactions. Photochemical interactions, normally associated with shorter wavelengths and longer exposure durations, require lower energy for initiation than do the thermal interactions normally associated with near infrared retinal alteration. The retina is a beehive of photochemical and biochemical processes which might be altered by specific wavelengths. The histopathology suggests that 900 nm radiation is absorbed by the rod photoreceptors to the extent that the underlying RPE is relatively spared in the rod-rich areas of the retina. The data currently available are not sufficient to explain the reduced ED_{50} near 900 nm.

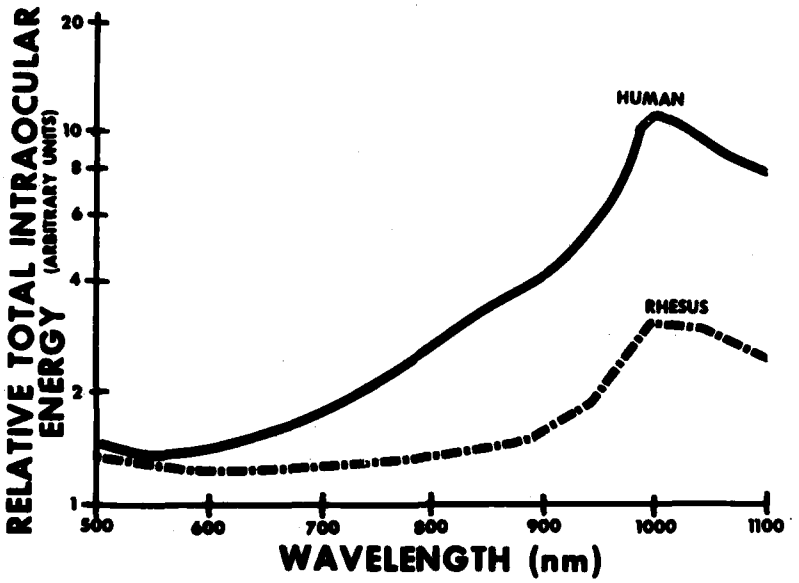


FIGURE 10. Wavelength dependence of the total intraocular energy required to produce a constant level of absorbed energy in the retina of human and rhesus monkey eyes.

The determination of ED_{50}s is not an exact science; the shape of the ED_{50} versus wavelength relationship from 700 nm to 1060 nm may not be exactly represented by the data of Figure 9. However, the data are internally consistent in that all the data were collected at LAIR by the authors and have a commonality of dosimetry and determination of results. All the ED_{50}s of Figure 9 are for single short pulse exposure to the extramacular retina in rhesus monkey. There is a real decrease in the ED_{50} between 850nm and 900 nm as compared to the ED_{50}s at 700 nm and 1060 nm. The MPE as provided in TBMED 279 is not consistent with these new bioeffects data. When the MPE at 900 nm was compared to the data of this experiment, we found a safety factor of 10. The current MPE should provide

adequate protection for exposure to single pulses from the GaAs laser.

CONCLUSIONS

The MPEs for wavelengths affecting GaAs laser training devices were derived based on assumptions concerning interaction of laser radiation with retinal tissue. The data of this study show that those assumptions were not adequate, and that retinal damage occurs at lower doses than expected. The results of this study do not necessarily have an adverse affect on the use of GaAs laser training devices - the safety standards still provide an adequate safety margin. The bioeffects data base in this spectral region is still incomplete; additional data now being collected will serve to define the wavelength dependence of ED_{50} between 700 nm and 1060 nm. It is possible that, when the data base is complete, a recommendation will be made to change the operating wavelength of training devices to a wavelength with increased safety margin. The data currently available do not warrant a recommendation for modification of the safety standards.

REFERENCES

1. AR 40-46. Control of Health Hazards from Lasers and Other High Intensity Optical Sources. Headquarters, Department of the Army, Washington, DC, 15 November 1978

2. TBMED 279. Control of Hazards to Health from Laser Radiation. Headquarters, Department of the Army, Washington, DC, 30 May 1975

3. FINNEY, D.J. Probit Analysis. New York: Cambridge University Press, 1952

4. LUND, D.J., B.E. STUCK, and E.S. BEATRICE. Biological research in support of project MILES. Report No. 96. Presidio of San Francisco, California: Letterman Army Institute of Research, July 1981

5. GEERAETS, W.J. and E.R. BERRY. Ocular spectral characteristics as related to hazards from lasers and other light sources. Am J Ophthalmol 66: 15-20, 1968

6. BOETTNER, E.A. and J.R WOLTER. Transmission of the ocular media. Invest Ophthalomol 1: 776-783, 1962

7. MAHER, E.J. Transmission and absorption coefficients for ocular media of the Rhesus monkey. Report SAM-TR-78-32. San Antonio, Texas: Brooks Air Force Base, April 1978

In conducting the research described in this report, the investigators adhered to the "Guide for the Care and Use of Laboratory Animals" as promulgated by the Committee on Revision of the Guide for Laboratory Animal Facilities and Care, Institute of Laboritory Animal Resources, National Academy of Sciences, National Research.

Citation of trade names in this report does not constitute an official endorsement or approval of the use of such items.

COMPUTER DETECTION OF LOW CONTRAST TARGETS

RAUL G. MACHUCA, PHD
US ARMY WHITE SANDS MISSILE RANGE
WHITE SANDS MISSILE RANGE, NEW MEXICO 88002

Introduction. At White Sands Missile Range we are interested in detecting manmade objects (targets) usually against a natural background such as the sky or a mountain. A widely used method of detecting a target is to find the edges in a picture using some operator whose value is high when there are many changes in gray levels. An object which contrasts with the background because of step and ramp edges gives rise to high values of the edge detector. After the detection of the step and ramp edge points in a scene, the problem of detecting a target is reduced to that of determining which of the points classified as edges belong to the target. If the target is the object of greatest contrast then its location can be found by simply thresholding for the highest value of the edge detector. However, if the target is not the object of highest contrast other methods must be used to detect a target in a scene. Thus, we must use properties of the target that are not shared by other edge points in a scene. The targets that we track, planes and rockets, have as one of their main geometric features several points of very high curvature. Thus, to detect edges which belong to a target a reasonable procedure would be to look for edge points which come from objects which have high curvature. Another property which makes manmade objects different from natural objects is the strong interior edges (extremal edges) which are present. This suggests that another procedure for detecting objects such as planes and rockets is to measure in some way the extremal edges present in a scene and to look for the target amongst those points of higher values. It may be possible that the information present in the gray level picture does not have any significant difference in gray level between the target and the background but the target and the background may be of different colors. In this case, we must look a the color information and extract the edges from the color data before we begin to look for a target. The problems of finding extremal edges, finding curvatures along the boundary of objects, and detecting color edges, at first sight appear unrelated. The formulation of these problems in the framework of vector fields and the application of

this theoretical model to data both from the range and from the laboratory are the subjects of this paper.

The paper is divided into three parts. In the first section we discuss extremal edges and their detection, section 2 discusses curvature and its measure and section 3 formulates the detection of color edges in vector field language and uses the theoretical results of sections 1 and 2 to detect color edges. Each section follows the same paradigm, namely:

1. The problems is formulated as a problem in the theory of vector fields.

2. The detection of a particular property is formulated as the computation of an integral over a vector field.

3. The theoretical results are applied to data obtained from the field or in the lab.

An outline of the mathematical methods are mostly contained in section 1 and are based on classical vector analysis. Details of the mathematics used appear in [8] and [9], for supporting material on differential geometry, see for example [1] or [2]. What we call the rotation degree of a vector field is also called the topological degree of the corresponding mapping from the unit circle to itself. Two books treating topological degree theory are [3] and [4].

Section I. VECTOR FIELDS AND EXTREMAL EDGES

The two planes of figure 1 are typical of the tracking situations that occur at WSMR. To set these we started with the original video frames and digitized the analog signal at the rate of 512 points per line. The value assigned to each point is between 0 and 255 and is proportional to the strength to the analog signal. In these two pictures, that are of very low contrast, we have been able to segment the target by finding the extremal edges arising from the shape of the fuselage and wings. In terms of a function of two variables an extremal edge is a connected set of extrema very similar to the top points of a roof. The classical development of extrema via partial derivatives does not give a method for finding these extrema. The standard treatment of critical points looks at the quadratic form $f_{11}f_{22}-f_{12}^2$ and at f_{11} using the Hessian

$$Hf = \begin{bmatrix} D_{11}f & D_{12}f \\ D_{21}f & D_{22}f \end{bmatrix}$$

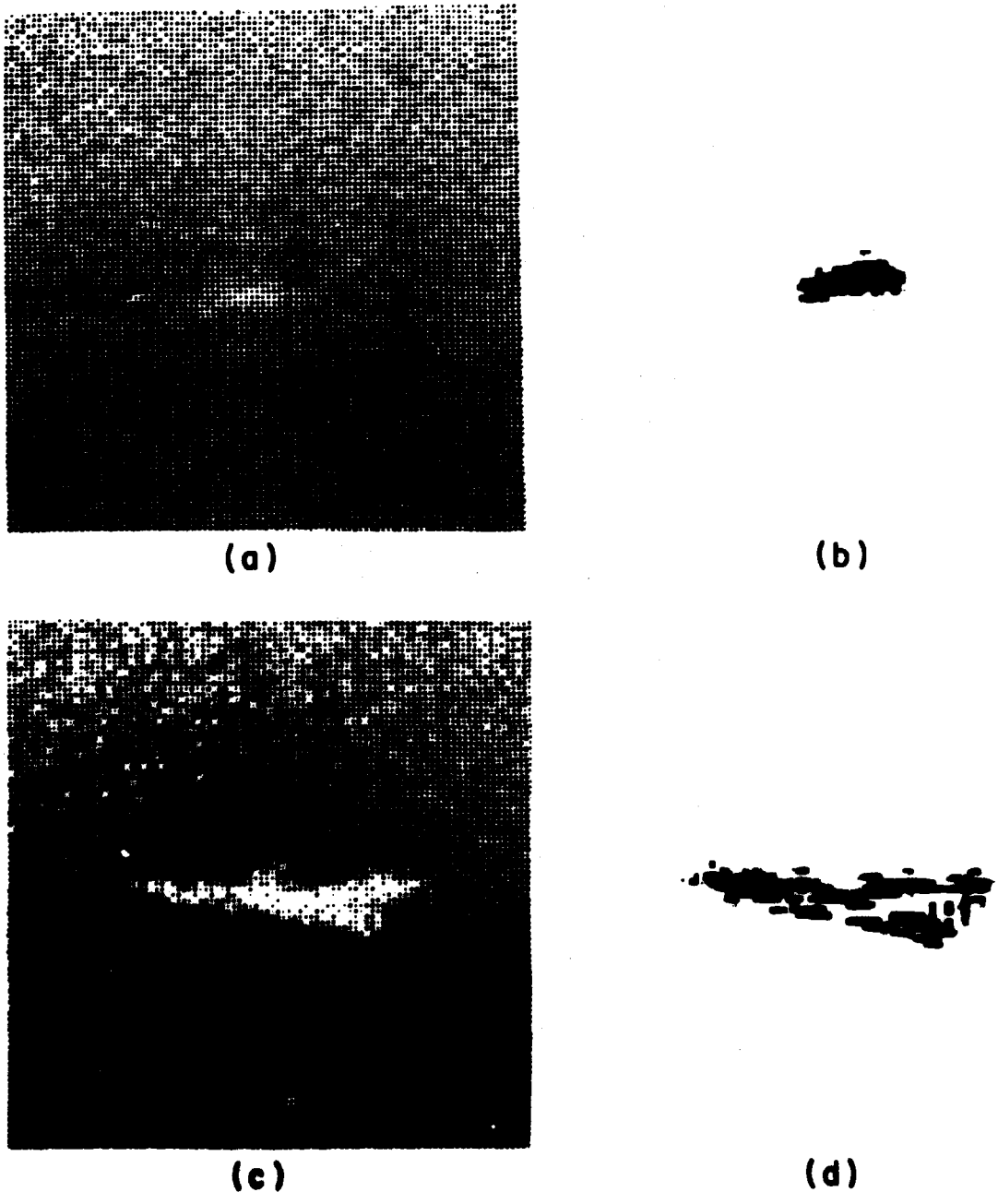


Figure 1. (a) (c) Two planes as seen through a telescope. (b) (d) extremal edges of (a) and (c) segment the planes from the background.

the conditions for an extrema are:

max: determinant (Hf) >0 and $f_{11}<0$
min: determinant (Hf) >0 and $f_{11}>0$
non-extrema: determinant (Hf) <0
cannot say: determinant (Hf) = 0

An example of a function about which nothing can be said from this test is $Z = X^4 + Y^4$. It has a Hessian which vanishes at 0 and thus from this test no conclusion can be made about the type of critical point at (0, 0) even though it is obviously a minimum. The gradient field shows that the function is increasing at every direction away from (0, 0) and so the information as to the nature of the extrema is in the gradient field.

We can extract this information from the vector field by calculating the integral for $z=x^4 + Y^4$

$$\eta = \frac{1}{2\pi} \int_\gamma d\theta \; \frac{1}{2\pi} \int_\gamma K = 1$$

where

(a) $$K(\tilde{x}) = \frac{Hf(\tilde{x}, \tilde{\nabla} f(\tilde{x}))}{|f(\tilde{x})|^2}$$

(b) $$Hf(\tilde{x}, \tilde{\nabla} f(\tilde{x})) = \begin{bmatrix} D_{11}f & D_{12}f \\ D_{21} & D_{22}f \end{bmatrix} \begin{pmatrix} -D_2 f \\ D_1 f \end{pmatrix}$$

(c) $$\tilde{\nabla} f(\tilde{x}) = (-D_2 f, D_1 f)$$

and $d\theta$ is the $d\theta$ of (figure 2).

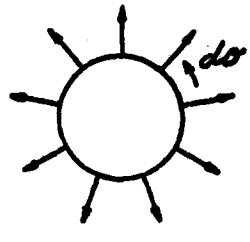


Figure 2. A portion of a vector field.

The fact that Z is an extrema at (0,0) is reflected by its nonzero value of η. But we are mainly interested in detecting extrema which are not isolated as in the case $f=X^2$. This function has the line x=0 as the set of its extrema. If we look at the Hessian it gives no information which can be used to detect the extremal set: As before we must look at the gradient field to find the extrema. The gradient field of the function as before reveals the behavior of the function at x=0, namely it is increasing as we move away from x=0, thus there is an extremal edge along the line x=0.

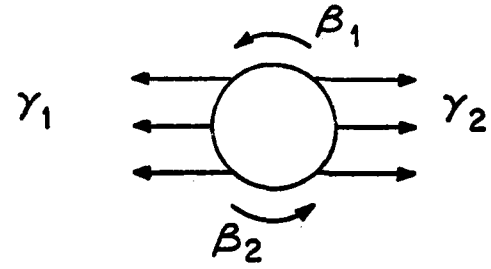


Figure 3. A small neighborhood of the gradient field for the function $z=x^2$.

Because of the ambiguous $d\theta$ at β^1 and β^2 (figure 3) the analytical detector of extrema edge points becomes somewhat complicated. The analogy of isolated extrema is followed and we define the detector as an integral along the places where it can be defined and where there is an ambiguity we take $d\theta=\pi$. Using the quantities defined at (a) and (b) the analytical detector is

$$\eta = \int_{\gamma 1} K dr + \int_{\gamma 2} K dr + \beta_1 + \beta_2$$

This gives a function which is equal to zero at nonextremal edges and 1 at extremal edges.

Once we have identified the extermal edges we want to distinguish between weak and strong extremal edges. In the one dimensional case (figure 4) the quantity $M(e) = |f'(a)| + |f'(b)|$ will distinguish between sharply sloped roofs and weakly sloped roofs. The two dimensional analogue of this quantity is $\int_\gamma |\nabla f|$.



Figure 4. Weak and strong extremal edges.

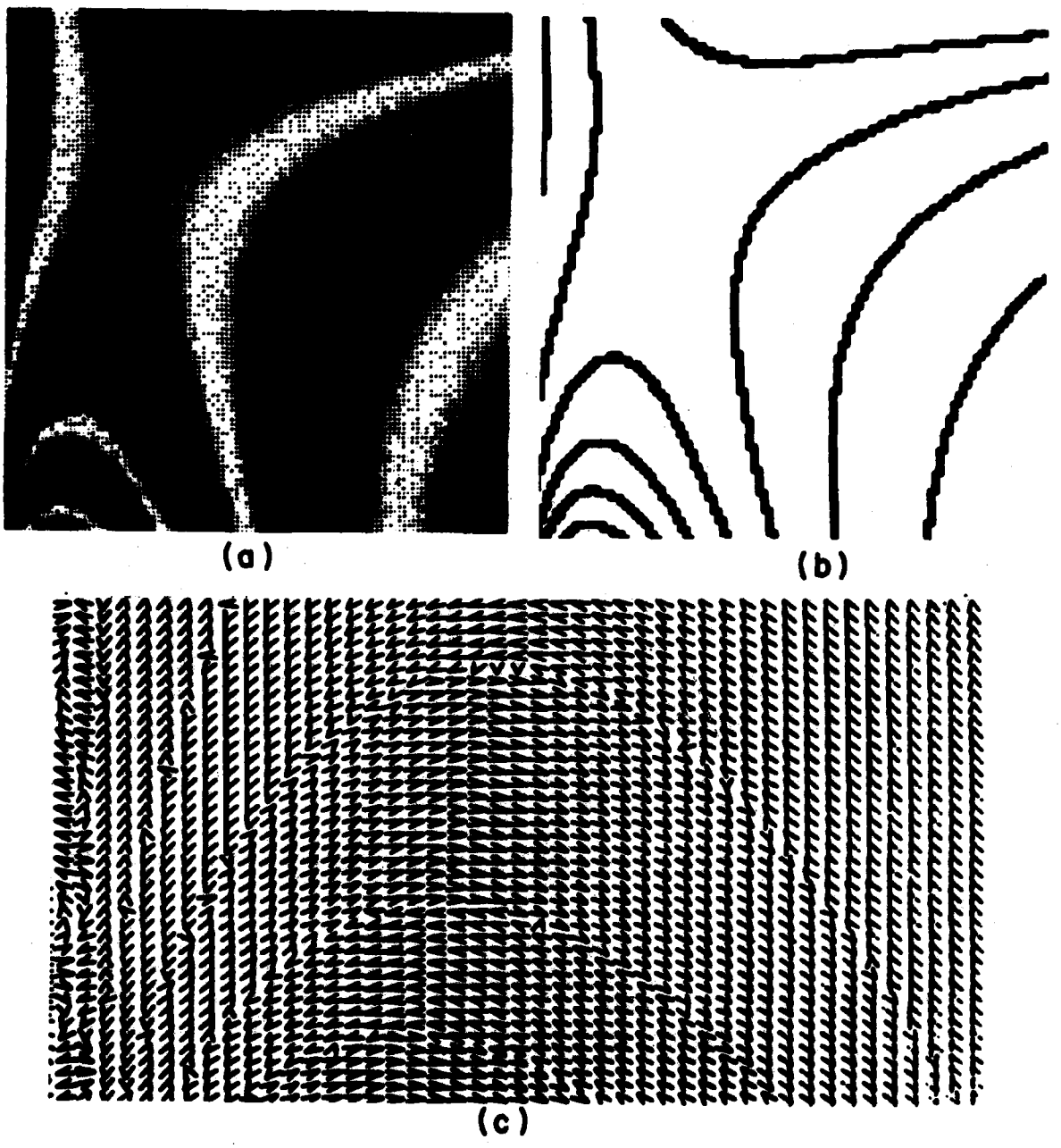


Figure 5. (a) Computer generated example, (b) edge points detected, and (c) partial gradient field of (a) rotated 90 degrees.

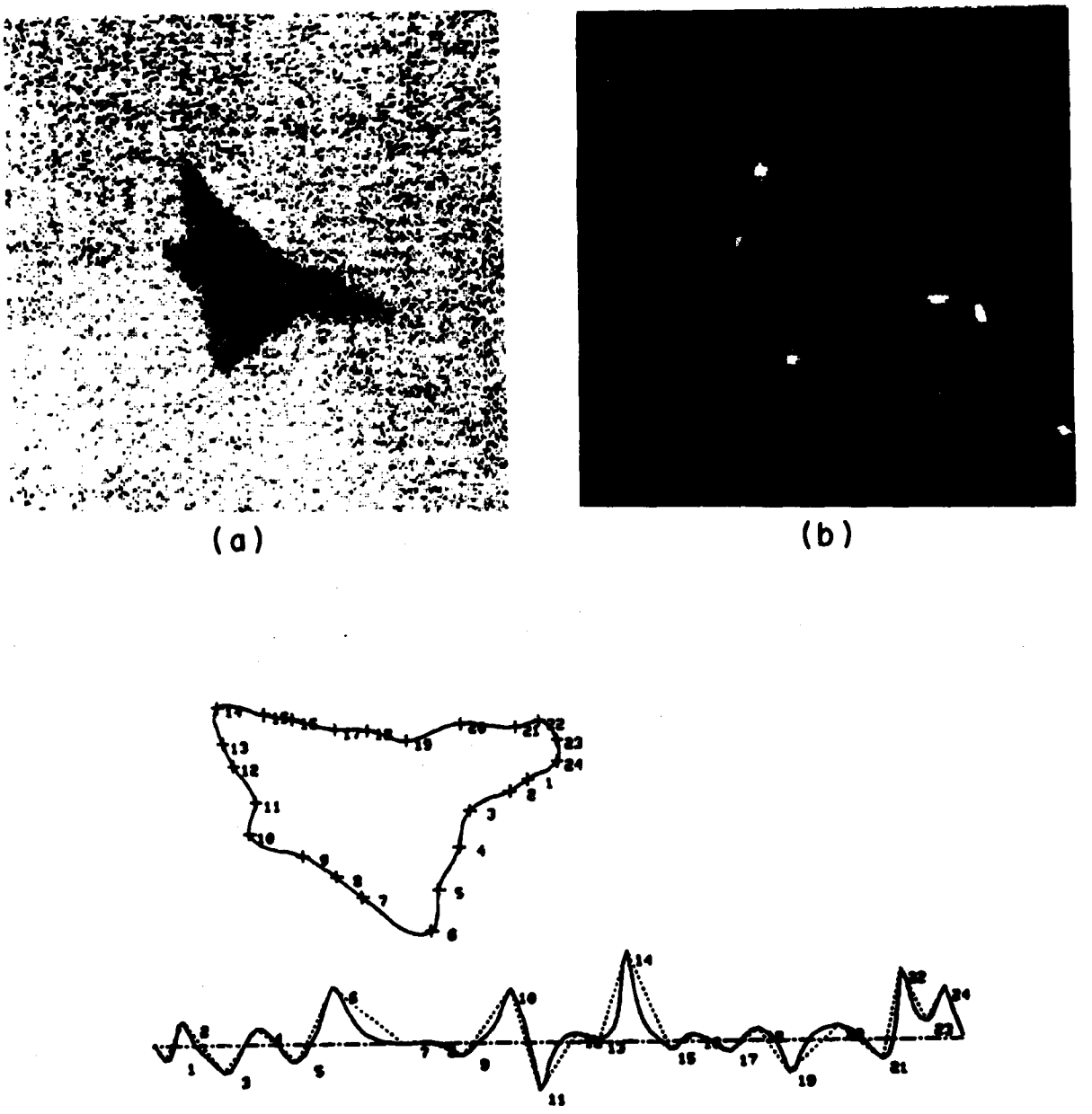


Figure 6. (a) Digitized video frame of an F102, (b) areas of high curvature, and (c) contour of (a) and its curvature.

The computer application of these methods is done by first approximating the partials in a 3x3 neighborhood of a point and from this computing the quantities $d\theta$ and ∇f. A new file is created which contains $\int_\gamma |\nabla f|$ when n=1 and 0 when n = 0.

In figure 5 we have several extremal edges of the type $z=x^2$ and $z=-x^2$. The darkest parts are like the extreme points $z=+x^2$ while the brightest points are like the extrema $z=-x^2$. The edge points exhibited in figure 5 have been found using the procedure described above. Figure 1 contains pictures of a plane taken through long distance optics. There is very little contrast in both pictures between the plane and the background. However, the extremal edges which result from the wings and the fuselage are very easily detected by the methods described above.

Section II. CURVATURE

This concept of studying the gradient field to find properties of a scene can be used to find other features of a target. For example in figure 6, the points of the original in which the contour has high curvature have been found by looking at the gradient vector field.

The analytical tools used to detect curvature are the same as those used in detecting extremal edges. With the definitions the same as (a), (b), and (c) above, the pointwise curvature can be computed from the Hessian and the gradient and is given by the formula

$$W(\bar{x}) = - \frac{< Hf(\bar{x}, \tilde{\nabla} f(\bar{x}), \tilde{\nabla} f(\bar{x})) >}{|\nabla f|^3}$$

Because of the amount of noise present in these scenes a better quantity to compute is the average curvature:

$$A^*(K,\beta) = \frac{1}{\beta} \int \beta \, W \, dr$$

To see that the quantity is a measure of curvature we look at the function $f(x,y) = x^2 + y^2$ and compute A* for this function. If A* is computed over a small circle centered at n as in figure 7,

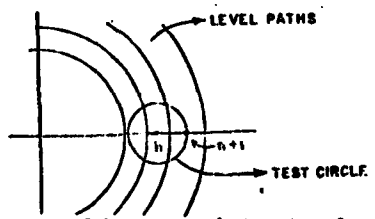


Figure 7. Contour lines and test circle.

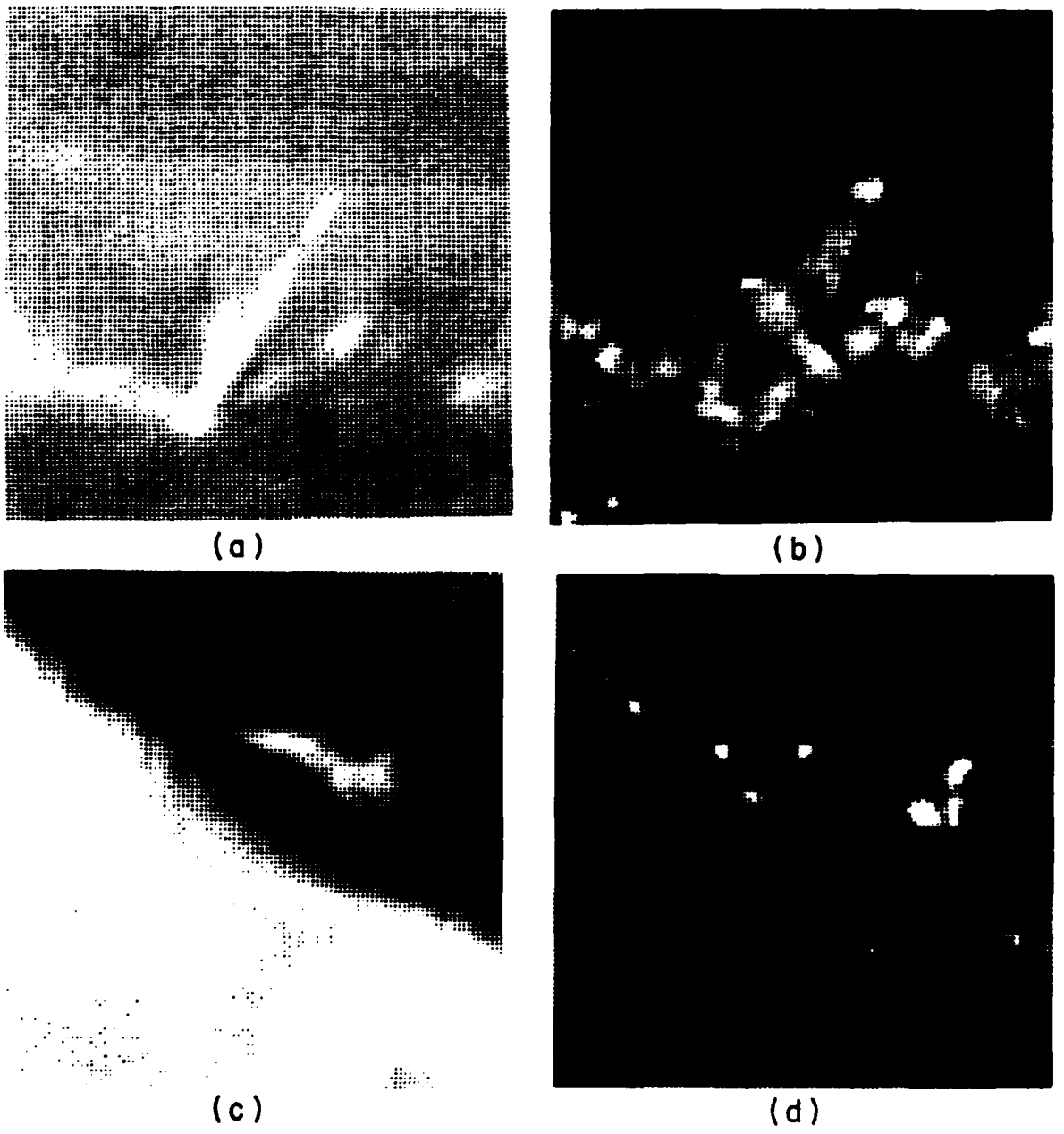


Figure 8. Missile (a) and plane (c) with area of high curvature (b) and (d).

then if A* is a good measure of curvature it should be small when n is large and large when n is small. Using polar coordinates the pointwise curvature is 1/r and with the circles parametrized by $(n,0) + e^{it}$, A* becomes the integral

$$\int_{-\Pi}^{\Pi} \frac{dt}{(n2 + I + 2n \cos t)\ 1/2}$$

and this integral is a maximum for n=1 and decreases as n increases, exactly what a good measure of curvature should do. Figure 8 contains two scenes obtained from range missions. Figure 8 (a) is a rocket at takeoff. The nose of the rocket has been identified as one of the points of high curvature and is the brightest point of (b) closest to the top. The points of high curvature generated by the plume are also indicated. In figure 8 (c) there is a plane flying by a mountain. The points of high curvature at the nose, wing, and tail have been detected. There are also some others from the mountain edge. Thus, the gradient field reflects many properties of the original scene. We can also look at other vector fields generated by scenes and find other properties such as color edges.

Section III. COLOR EDGES

We are mainly interested in the analysis of color data in the case where there is no gray level difference between the target and the background, i.e., in the cases where the above methods fail. That is we are looking for a method for finding edges in the case where the target cannot be separated from the background by looking at the luminance component of a scene but can be detected by using the chrominance information. A widely used color representation that breaks up a color into luminance and chrominance information is the (Y,I,Q) color representation of standard video. In this representation Y is the luminance component, which we ignore since our principal assumption is that the target and background are similar in luminance. Thus we must look at the I and Q components to detect the target from the background, and in particular we are interested in edges caused by differences in hue between target and background.

For any color(Y,I,Q) the vector v = (I,Q) has the property that the angle it makes with the I - axis, θ(v), specifies its hue (see figure 9) while its absolute value gives the saturation.

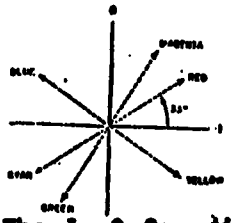


Figure 9. The I, Q Coordinate System

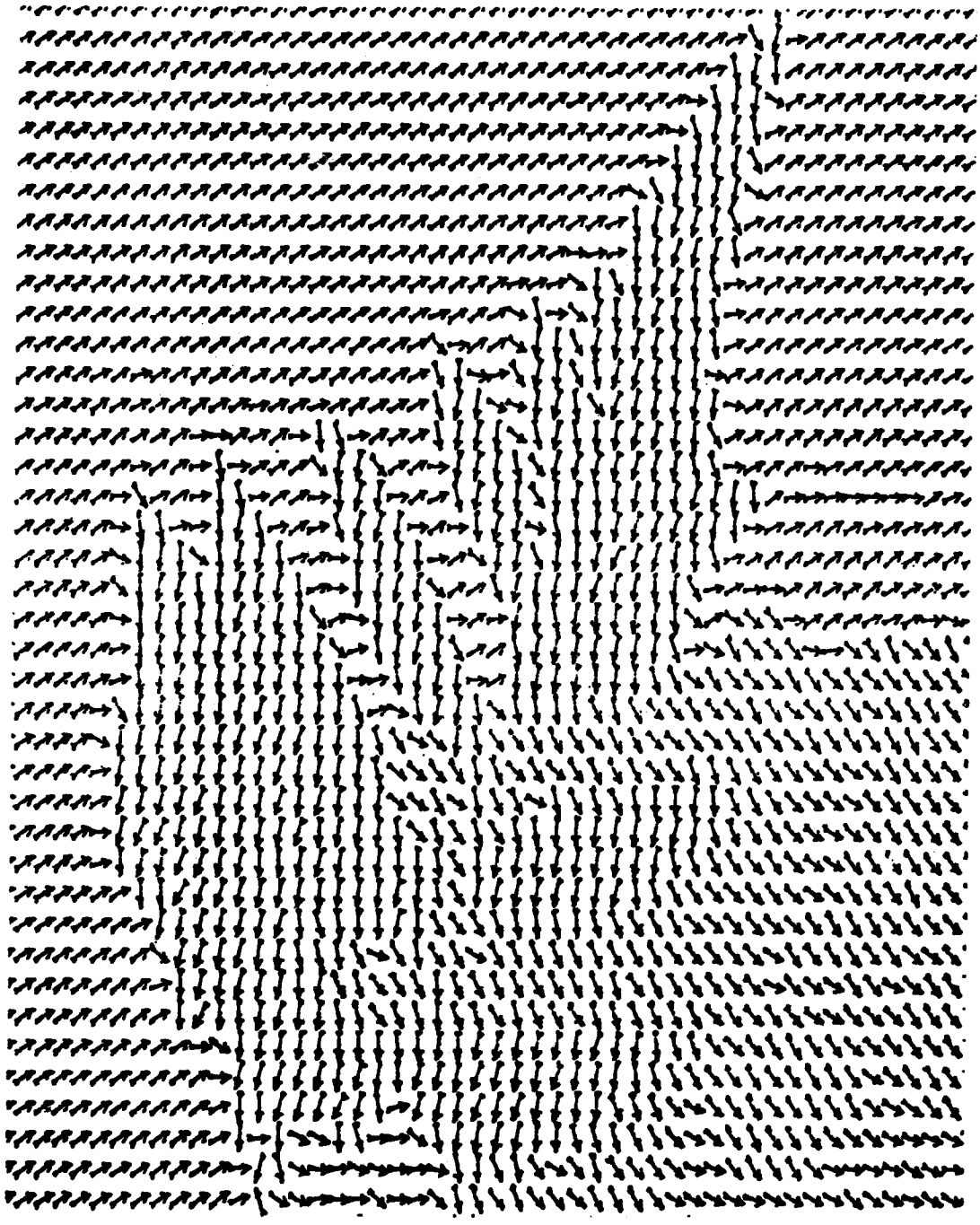


Figure 10. The (I,Q) vector field in the neighborhood of a leaf.

If we assume that the saturation is the same for every color of an image (this is the case for the scenes we study) then the color information available to detect the target is all contained in $\theta(v)$. In order to detect the edges from this component, it is useful to consider the vector field $T = v/|v|$ and use the vector field formulation of section I to detect edges due to hue.

An example of a vector field T so generated appears in figure 10. A portion of the vector field on a test circle, selected from near the center of figure 10 is shown in figure 11.

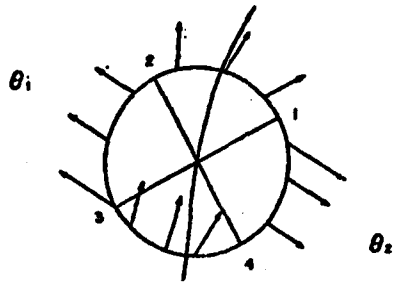


Figure 11. A portion of a color edge.

In figure 11, we detect an edge between the two colors, θ_1(blue) and θ_2(yellow), simply because the vectors point in opposite directions. Note that T rotates counterclockwise from 1 to 2 to 3, clockwise from 3 to 4 to 1. There are convexity changes at approximately "1" and "3". If the circle γ is parametrized on [0,1] by $\gamma(t) = e^{2\pi it}$, then the argument function $\theta(t)$ is simply $\theta = \mathrm{Arg}(v)$ at the point $e^{i2\pi t}$ on the circle:

(i) $$T(t) = (\cos(\theta(t)), \sin(\theta(t))).$$

Clearly, because if we use the integral as the color edge detector,

(ii) $$\int_0^1 |\theta(t)|\,dt = \int_\gamma |d\theta|$$

we ignore the change in convexity and measure net rotation. From figure 11 it is clear that

(iii) $$\int_\gamma |d\theta| = 2\,|\theta_1 - \theta_2| \ .$$

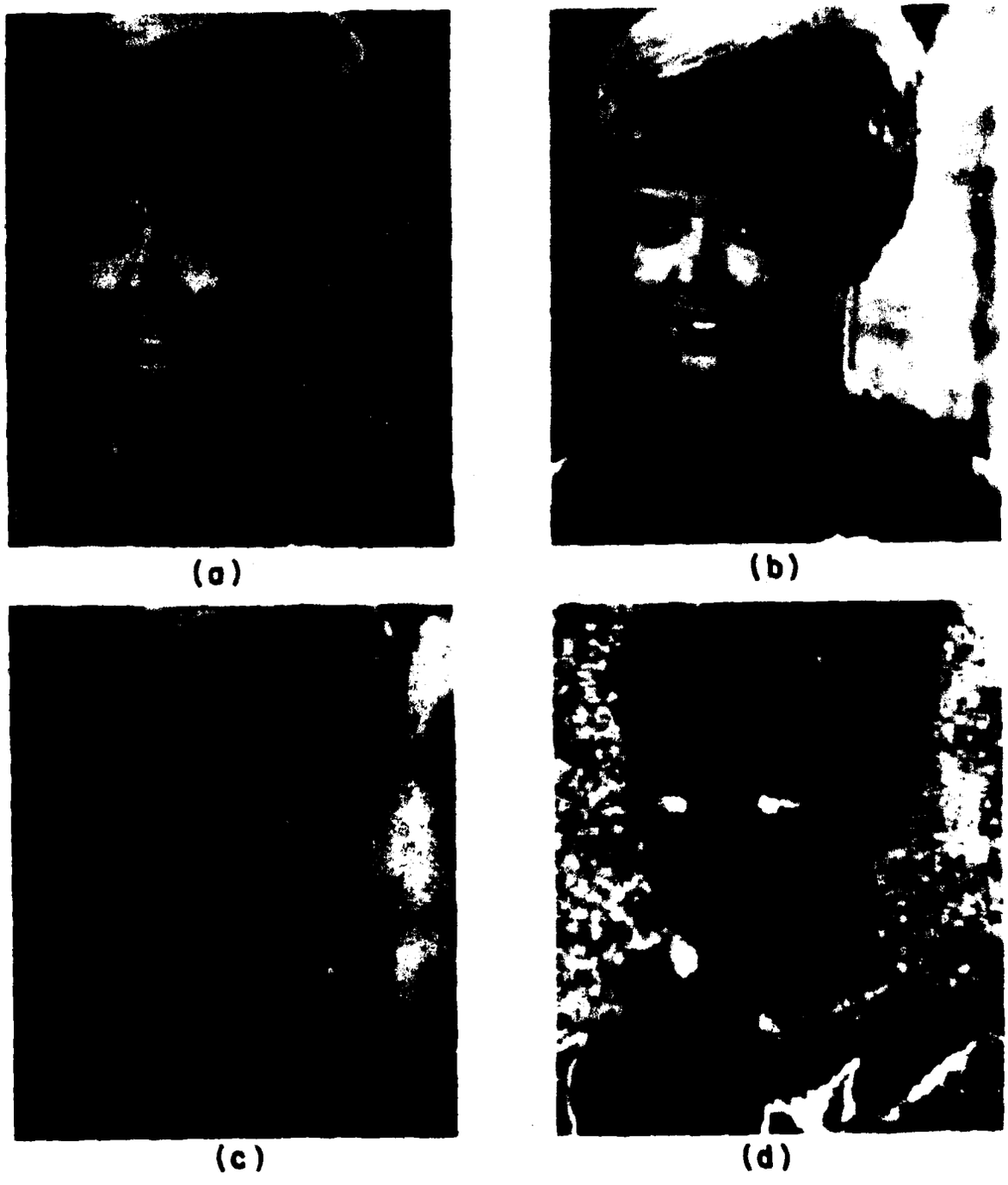


Figure 12. The R, G, B components of "girl picture" (a), (b) (c) and the hue edgepoints (d).

The integral $\int_\gamma d\theta$ in (ii) can be interpreted as the average curvature of the vector field T; e.g., as in (iii) for the curve γ. Hence, measuring net rotation corresponds to measuring the average curvature of T on γ. We have found $\int_\gamma d\theta$ to be a good color edge detector. In fact the vector field shown in figure 10 is generated by the leaf in the R,G,B components shown in figure 12 a-b-c. The result of the average curvature method - $\int_\gamma d\theta$ - is shown in figure 12.

An alternate approach to detecting color edges would be to follow the classical development for gray level edges. That is compute abs(grad(R)), abs(grad(G)), and abs(grad(B)) and combine these in some way to find the color edges. This process requires that two convolutions take place for each of the gradients (one for each partial derivative), and after the gradients are computed there remains the problem of how to combine them into one output.

The edge detection method described in this section is particularly provocative because of its implications with regard to hardware implementations of color edge detectors. The phase of the (I,Q) vector is easily available from analog hardware and so the data input to this algorithm is available in real time. The integration that takes place for the edge detection is a type of "convolution" easily performed by state-of-the-art real time hardware. Thus, with this method an implementation can be constructed with components that are practically available off the shelf, and such a machine is being built at White Sands Missile Range.

Conclusion. We have shown that a variety of image processing problems can be stated in terms of vector field problems and that quantities easily calculated in real time over the associated vector fields can be used to detect image features. Extremal edges can be found by computing the rotation number of a curve α via a simple process. Points of high average curvature are found directly from the vector field, without approximate surfaces or statistics. It was shown that the detection of hue edges could be done by computing the rotation associated with the vector field (I,Q), a number that can easily be computed in real time by existing hardware components.

REFERENCES

1. Klingenberg, Wilhelm, A Course in Differential Geometry, Springer-Verlag, 1978.

2. Spivak, Michael, Differential Geometry, vol. II, 1970.

3. Milnor, John, Topology From the Differentiable Viewpoint, University of Virginia Press, 1965.

4. Lloyd N. G., Degree Theory, Cambridge University Press, 1978.

5. Machuca, R. and Gilbert, A., "Finding Edges in Noisy Scenes", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 3, No. 1, pp. 102-111, January 1981.

6. Barnard, S. C., and Thomson, W. B., "Disparity Analysis of Images", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 2, No. 4, pp. 333-340, July 1980.

7. Kitchen, L., and Rosenfeld, A., "Gray Level Corner Detection", TR-887, Computer Sciences Center, University of Maryland, College Park, MD, April 1980.

8. Machuca, R., and Phillips, K., "Applications of Vector Field Theory to Extremal Edge Detection", Proceedings of the IEEE SOCON, Orlando, FL, March 1982.

9. Machuca, R., and Phillips, K., "Applications of Vector Fields to Image Processing", To appear in IEEE Transactions on Pattern Analysis and Machine Intelligence.

This work was supported by the Department of the Army, Office of the Assistant Secretary for Research and Development under ILIR tasks DA OM1510 and DA OM1832.

THE PLANAR DOPED BARRIER: A NEW CLASS OF ELECTRONIC DEVICES (U)

*ROGER J. MALIK, DR., THOMAS R. AU COIN, MR.,
RAYMOND L. ROSS, MR., AND ROBERT O. SAVAGE, MR.
US ARMY ELECTRONICS TECHNOLOGY AND DEVICES LABORATORY (ERADCOM)
FORT MONMOUTH, NEW JERSEY 07703

INTRODUCTION

The requirement for improved performance components in high resolution, lightweight radars and secure communication systems at millimeter-wave frequencies has led to an investigation of alternate approaches to metal-semiconductor junctions. Up to the present time, Schottky barriers have been used almost exclusively for very high frequency switching, mixing, and rectifying applications. However, they have inherent limitations with respect to barrier heights, metallurgy, interface states, and capacitance. A revolutionary new semiconductor device structure, the planar doped barrier (PDB), has been invented and demonstrated at ET&DL in gallium arsenide grown by molecular beam epitaxy (MBE). The PDB offers significant advantages over conventional Schottky barriers with respect to cutoff frequency, metallization, I-V characteristics, barrier height, and capacitance. The unique features of this structure are twofold: (1) the degree of symmetry/asymmetry in the rectifying I-V characteristics is completely designable, and (2) for the first time, the zero-bias barrier height can be varied from approximately zero to the bandgap of the semiconductor.

Entirely new classes of electronic and opto-electronic devices are rapidly evolving based upon the planar doped barrier concept. PDB diodes with symmetric I-V characteristics have been successfully tested in subharmonically-pumped stripline mixer circuits. A single PDB diode was used to replace two precisely matched Schottky diodes, resulting in the lowest local oscillator pump power (4 dbm) reported for a subharmonically-pumped mixer - of major significance to projected 94 GHz and higher frequency radar systems. Photo-diodes with subnanosecond response times and high-speed switching devices utilizing PDB's have been demonstrated. In addition, a new PDB transistor has been invented for three-terminal amplification at MMW frequencies. This new generation of devices incorporating planar doped barriers will have a significant and far-reaching impact upon future Army and DOD microwave, millimeter-wave, and near infrared systems for secure communications, ISTA, fire control, terminal homing, and guidance.

PDB Diode Model

Although Schottky barriers are used almost exclusively for very high frequency switching, mixing, and rectifying applications, they suffer from several inherent limitations. For a particular metal-semiconductor system, barrier heights are virtually constant and operational stability is related to the metallurgy of the contact system. Interface states and interfacial layers often play a dominant role in determining the Schottky barrier transport properties and can lead to hysteresis effects, especially in metal-GaAs structures.

In this paper, a revolutionary new type of majority carrier rectifying structure which is designated as a "planar-doped-barrier" (PDB), is reported. The PDB structure is practically demonstrated in GaAs grown by molecular beam epitaxy. The key features of the PDB are that the barrier height and the degree of asymmetry in the I-V characteristics can be continually and independently controlled. The barrier height can be readily varied from zero to approximately the bandgap of the semiconductor. In addition, the structure can be designed to give a capacitance which is essentially independent of bias voltage between forward and reverse conduction. Current transport is by well-behaved majority carrier thermionic emission over the barrier in contrast to a tunneling mechanism employed in Schottky barrier lowering.[4-5] Although the PDB technique is here demonstrated in GaAs, it can, in general, be used in all types of semiconductors. In addition, since the PDB structure is formed in a single semiconductor, the problems of charge trapping, surface recombination, and thermal and lattice mismatch associated with interfaces between dissimilar materials are avoided. This technique allows great design versatility in the potential barrier heights and I-V characteristics which can be used to optimize the performance of specific barrier devices.

Recently, new-designable rectifying diode structures were reported. Shannon[1] demonstrated rectification in a "camel" diode formed by shallow implantation of n^+ and p^+ regions in n-type silicon. Chandra and Eastman[2] have demonstrated rectification in an n-n GaAs:(Ga,Al)As heterojunction grown by liquid phase epitaxy where the ternary was low-doped to 1×10^{15} cm^{-3}. Allyn et al.[3] have also reported a unipolar rectifying structure where a triangular potential barrier was induced by a graded layer of $Al_xGa_{1-x}As$ between layers of n-type GaAs grown by molecular beam epitaxy.

The basic profile of the PDB has an n^+-i-p^+-i-n^+ configuration, as shown in Figure 1. The structure uses a narrow plane of acceptor atoms positioned inside a nominally undoped region bounded by two heavily doped donor regions (1) and (2). The undoped region is assumed to have a total

number of impurities which is small compared to the p^+ -plane. The p^+ plane is, within certain constraints, fully depleted and, at zero-bias, narrow space charge regions are induced in the two n^+ regions in order to satisfy charge neutrality. The proportions of the space charge in regions (1) and (2) depend upon the position of the p^+ -plane within the undoped region.

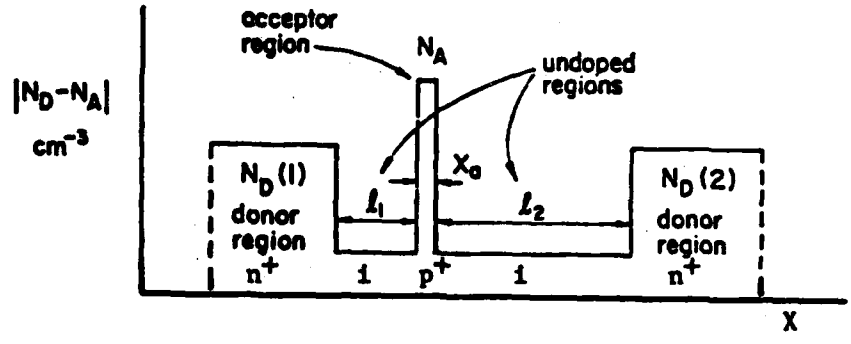


Figure 1. Doping profile in a planar diode barrier (PDB).

The resultant charge profile is that indicated in Figure 2(a). A solution of Poisson's equation yields the electric field and energy-band diagrams shown in Figures 2(b) and 2(c), respectively. In the limit, as the space charge widths approach single atomic planes and the undoped region becomes truly intrinsic, the electric field and band-energy profiles become exactly rectangular and triangular in shape, respectively.

The response of the barrier heights in a typical PDB to positive and negative bias is shown in Figures 3(a) and 3(b), respectively. Under positive bias (layer (1) positive), and assuming the plane to be much closer to layer (1) than (2), ϕ_{B1} increases slightly and ϕ_{B2} is reduced by an amount nearly equal to the bias voltage. Forward current flow occurs by emission of electrons from layer (2) over the barrier into layer (1). Under negative bias, ϕ_{B1} is reduced only slightly and ϕ_{B2} increases by almost the amount of the applied voltage. Therefore, significant reverse current flow occurs at a much higher voltage, this time by the emission of electrons from layer (1) over the barrier into layer (2). This should be compared to the reverse characteristics of Schottky barriers and p-n junctions in which avalanche breakdown is responsible for current flow. The I-V characteristics can be made asymmetric or symmetric, depending upon the position of the p^+ plane within the undoped region. Thus, the PDB offers much greater control over the I-V characteristics in comparison with the "camel" diode.[1]

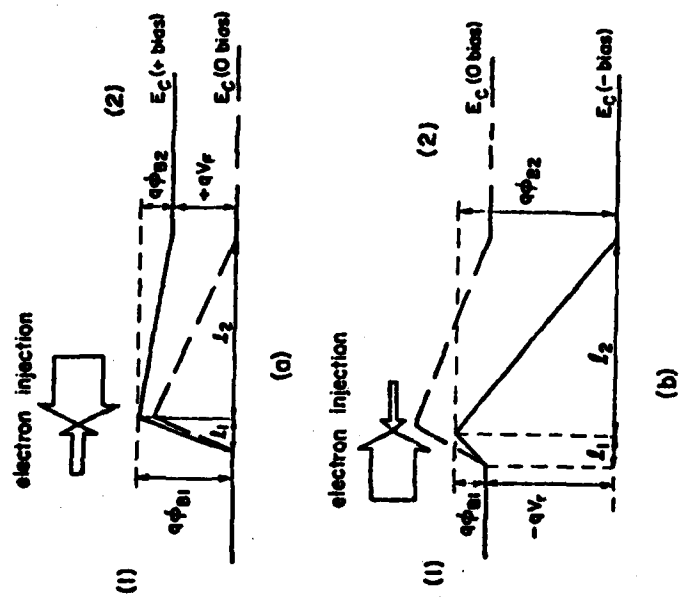


Figure 3. Barrier height response in a PDB under (a) positive bias and (b) negative bias.

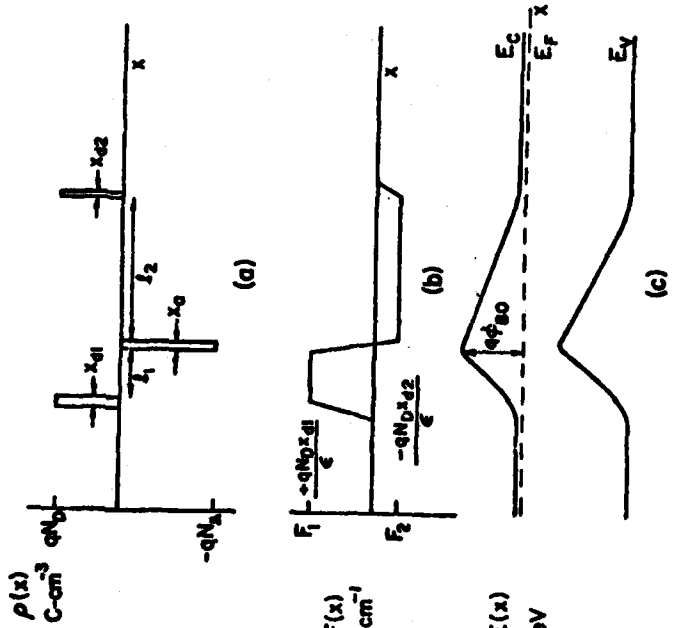


Figure 2. (a) Space charge, (b) electric field, and (c) band-energy diagrams of the PDB at zero-bias.

Theory

The solutions of Poisson's equation were derived using a simplified atomic plane charge density model where the area acceptor charge density, $-Q_A$, is assumed to be fixed. Under these conditions, the zero-bias barrier height is given by

$$\phi_{BO} = \frac{(\ell_1 \ell_2)}{(\ell_1+\ell_2)} \cdot \frac{Q_A}{\varepsilon} \tag{1}$$

where ℓ_1, ℓ_2 are the respective distances separating the acceptor plane from the donor planes of layers (1) and (2), and ε is the dielectric permittivity of the semiconductor.

The barrier near the surface (layer 1) has a height given by:

$$\phi_{B1} = \phi_{BO} + \frac{\ell_1}{(\ell_1+\ell_2)} V \tag{2}$$

and the barrier near the substrate (layer 2) has a height given by:

$$\phi_{B2} = \phi_{BO} + \frac{\ell_2}{(\ell_1+\ell_2)} V \tag{3}$$

In Equations (2) and (3), the sign of V is determined by the polarity of the surface ohmic contact.

From Equation (1), the zero-bias barrier height is continually variable and can be designed for a particular device application through appropriate choice of Q_A, ℓ_1 and ℓ_2. In addition, from Equations (2) and (3) it is clear that under bias one barrier increases while the other decreases and the relative change in ϕ is controlled by ℓ_1 and ℓ_2. Thus, the choice of ℓ_1 and ℓ_2 influences the symmetry of the I-V curves while the selection of all three parameters determines the actual I-V characteristics.

If the distances ℓ_1 and ℓ_2 and the magnitude of the barrier height are such that the potential change over a mean-free path is greater than kT/q[6], then current flows in both directions by thermionic emission over the barrier. The resultant total current can be expressed approximately by:

$$I = aA^*T^2 e^{\frac{-q\phi_{BO}}{kT}} \left(e^{\frac{q\alpha_2 V}{kT}} - e^{\frac{-q\alpha_1 V}{kT}}\right) \tag{4}$$

In Equation (4), a is the device area, A^* is the effective Richardson constant for electrons, and $\alpha_1=\ell_1/(\ell_1+\ell_2)$, $\alpha_2 = \ell_2/(\ell_1 + \ell_2)$. The capacitance C of the structure is simply given by

$$C = \frac{\varepsilon}{(\ell_1+\ell_2)} \tag{5}$$

and should be constant for all bias voltages between forward and reverse conduction.

Experimental - MBE

The GaAs PDB layers were grown in a Varian MBE-360 system using Ge^7 and Be^8 as the n- and p-type dopants, respectively. A photograph of the system, as presently configured in ET&DL's Materials Research Center, is shown in Figure 4 and a schematic diagram is shown in Figure 5. The system is comprised of a main vacuum chamber which houses the effusion furnaces, substrate heater stations and the analytical equipment. A loadlock assembly is attached to the main chamber which allows for rapid interchange of samples and is essential in maintaining low background impurity levels in the films. The UHV in the main chamber is obtained from a variety of vacuum pumps. The primary pump is a 400 ℓ_s^{-1} sputter-ion pump which is separated from the main chamber by an isolation valve. Additional pumping of CO, H_2O, and other reactive gases is achieved by a Ti sublimation pump. A further reduction in the H_2O pressure is achieved through large areas of liquid nitrogen cryopumping.

The source furnaces are arranged in a "wagon wheel" vane assembly which is liquid nitrogen cooled for thermal isolation and has provisions for eight effusion cells. The source crucibles inserted in the furnaces are radiantly heated by Ta windings on alumina insulators. Ta heat shielding around the furnaces is used to maintain temperature stability. High purity elemental sources of Ga, As, Al, Ge, and Be are evaporated from BN crucibles. The molecular fluxes emanating from the furnace cells are each individually shuttered by an electro-pneumatically driven Ta plate placed in the path between the cell and the substrate. The shutters are automatically controlled by an HP 85 microprocessor to generate any number of layers of a desired dopant concentration.

A UTI-1000 quadrapole mass spectrometer (QMS) is used to monitor the background gases and relative fluxes of the effusion cells. The QMS is located just below the substrate growth position and can indirectly measure the flux density of evaporated molecular species. The QMS effectively discriminates between the mass to charge (m/q) ratio which can be displayed on a storage oscilloscope.



Figure 4. Varian-360 Molecular Beam Epitaxy System.

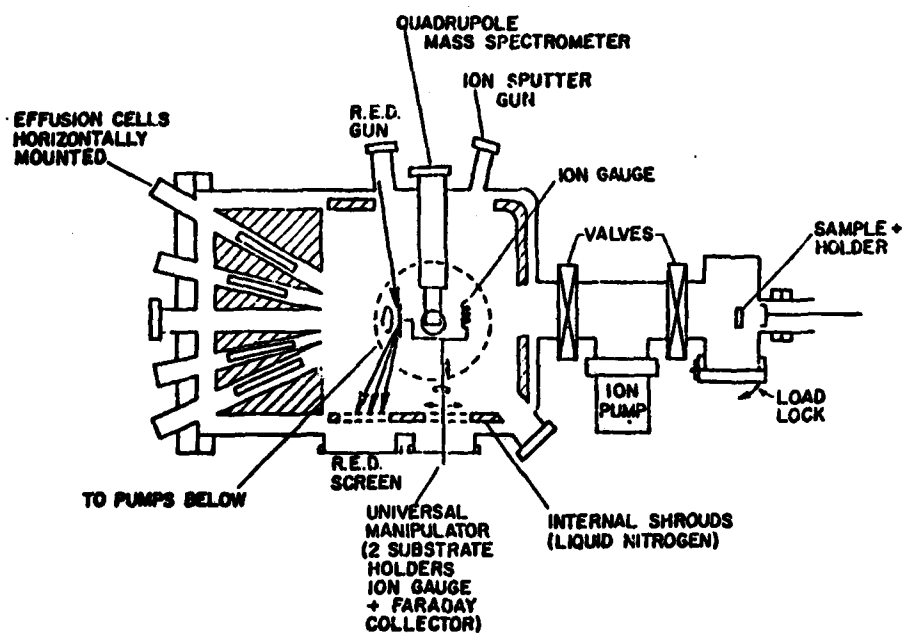


Figure 5. Schematic diagram of the Varian-360 MBE system.

Reflection electron diffraction (RED) is another important analytical tool included in the MBE system. In this technique, electrons of several keV which are elastically scattered off the substrate are displayed on a fluorescent phosphor screen. The surface structure of the substrate can then be determined from the electron diffraction pattern. The arrangement of the surface atoms is normally different than that in the bulk crystal. The growth of GaAs is normally done under an excess As flux to ensure a stoichiometric crystal. This leads to an As stabilized surface in which the As atoms on the surface are arranged in an array of which the period is a multiple of the bulk lattice and depends upon the crystallographic orientation.

Another feature of RED is its use in determining whether the surface of the film is atomically clean to begin MBE growth. A substrate in the chamber normally has a thin native oxide on the surface after cleaning and etching. Upon heating to $\sim$ 550°C under an As_4 flux, this native oxide is thermally desorbed and results in a characteristic reconstruction pattern readily observed by RED.

Experimental - PDB

Typical PDB layers contained dopant concentrations (n^+, p^+) of $\sim 2X10^{18}$ cm^{-3}, an ionized acceptor region width of < 100Å, and undoped region distances varying from 250 to 2000Å. The basic PDB diode structure is shown in Figure 6. Arrays of 100 micron diameter mesa diodes were fabricated using electron-beam lithography-generated masks, photolithography, and chemical etching. Au/Ge was evaporated and alloyed for ohmic contacts.

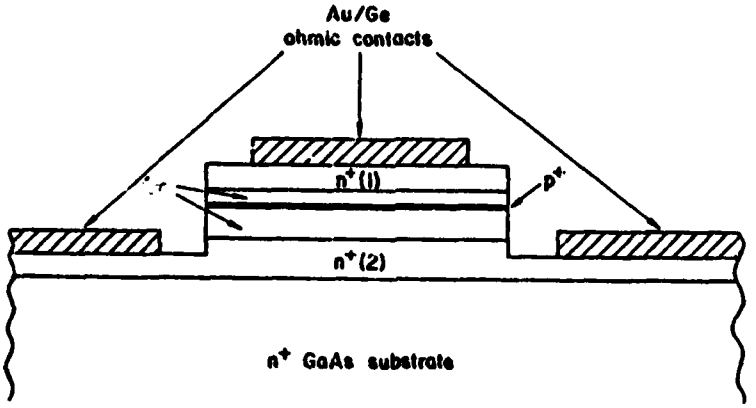


Figure 6. Schematic cross-section of a PDB diode.

DC measurements were made on the PDB diodes to examine their characteristics and to compare them with theoretical predictions. Room temperature I-V curves on two representative PDB diodes, A and B, are shown in Figures 7 and 8. The great flexibility in the design of the diodes is clearly illustrated and, in principle, the characteristics are infinitely variable. Diodes A and B both exhibit forward conduction at positive bias and show the effects of varying the ℓ_1/ℓ_2 ratio. The measured barrier heights of diodes A and B were 0.9 and 0.5 V, respectively, while their ℓ_1/ℓ_2 ratios were 8.0 and 1.0. A symmetrical I-V characteristic is demonstrated in Diode B by simply making ℓ_1 equivalent to ℓ_2. This is the first semiconductor structure in which the I-V characteristics are completely designable.

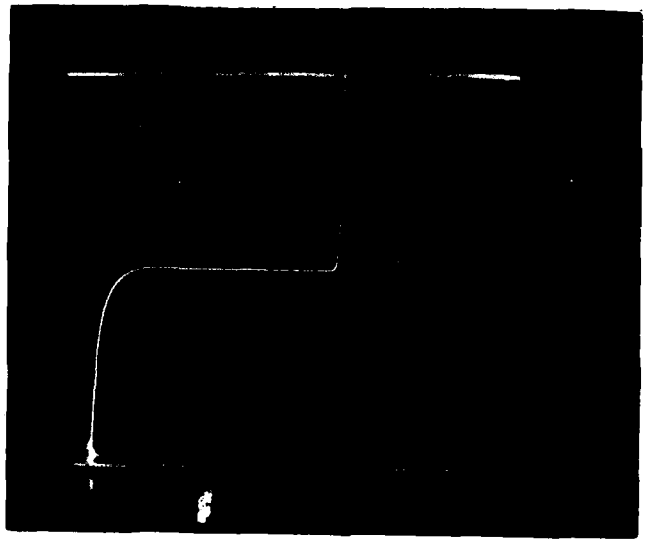


Figure 7. Room temperature I-V characteristics of assymetric PDB diode A, where ℓ_1 = 250 Å, ℓ_2 = 2000 Å and p^+ layer thickness = 100 Å.

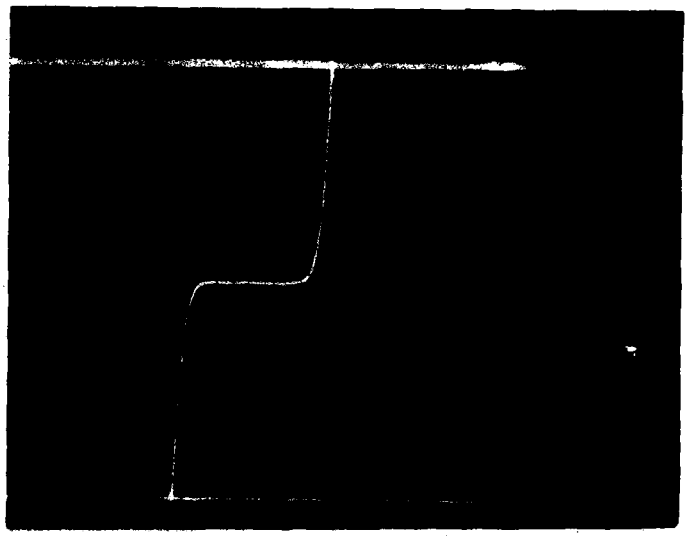


Figure 8. Room temperature I-V characteristics of symmetric PDB diode B, where ℓ_1 = 2000 Å, ℓ_2 = 2000 Å, and p^+ layer thickness = 40 Å.

PDB Transistor Model

During the course of the development of PDB diodes, it soon became apparent that a new type of transistor could be formed using the planar doped barrier concept. The basic doping profile of a planar doped barrier transistor is shown in Figure 9.

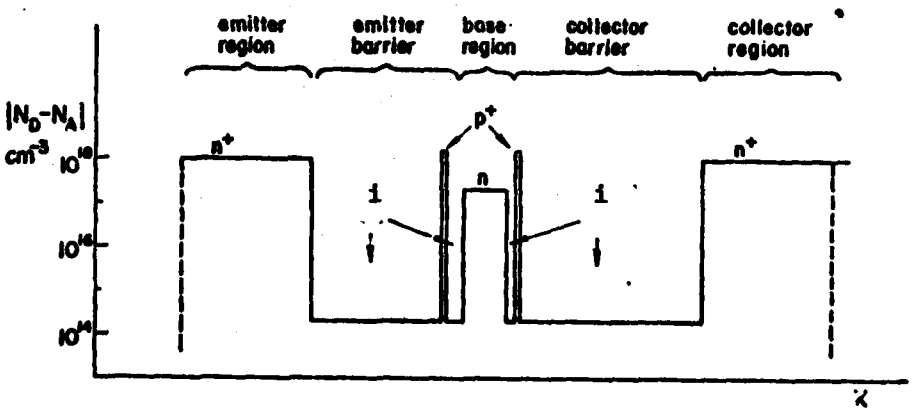


Figure 9. Doping profile in a PDB transistor.

The PDBT structure consists of two planar doped barriers joined together by a common n-region which serves as the base of the transistor. The emitter barrier is used to rapidly accelerate and inject electrons at high velocities into the base region. If the base is thin enough, then the majority of these injected electrons will retain sufficient kinetic energy and momentum to overcome the collector barrier peak and thus contribute to the collector current. A small collector barrier must be maintained at all bias conditions to restrain thermal electrons in the base from being emitted over the barrier into the collector. This serves to isolate the base from the collector load which ensures a high impedance between the base and collector. This requirement is necessary for power gain in these devices.

The electron transport theory can be described with the aid of the conduction band potential diagram of the PDBT under zero-bias and normal operating bias conditions, as illustrated in Figure 10. At zero-bias, the emitter barrier height, $\phi_{EB}(0)$, is larger than the collector barrier height, $\phi_{BC}(0)$. Under operating bias conditions, the base is biased positively with respect to the emitter and the collector is biased positively with respect to the base. Due to the barrier lowering of ϕ_{EB} on the emitter side, enhanced emission of thermal electrons occurs from the emitter region to the emitter barrier peak. Electrons which reach the barrier peak are then rapidly accelerated on the steep downhill slope of the barrier. Electron velocities approaching the crystal limited group velocity of 1×10^8 cm-s^{-1} in the (100) direction in GaAs can be easily obtained for electrons with energies, 0.3 eV < E < 0.35 eV.

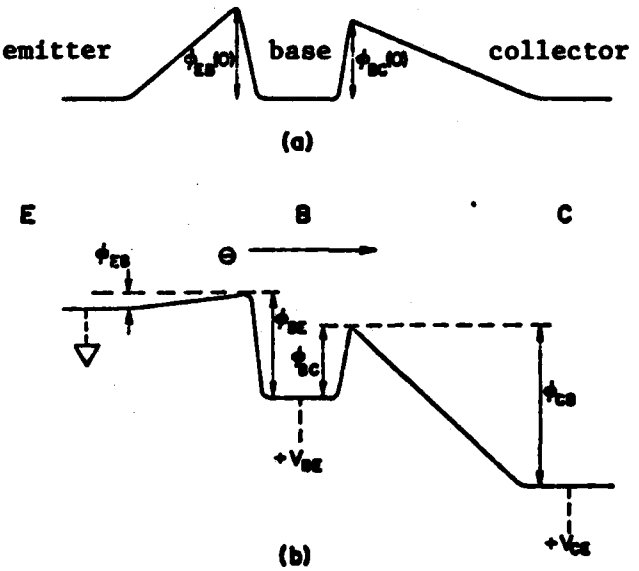


Figure 10. Conduction band potential in a PDB transistor under (a) zero bias and (b) normal operating bias conditions.

The electrons which are accelerated by the emitter barrier are then injected at high velocities into the base region of the PDBT. Electron scattering in the base region cannot be neglected due to the free electron concentration and thickness of the base. Typical values for these parameters are $n \sim 5 \times 10^{17}$ cm^{-3} and $W_b \sim 1000$ Å, respectively, and are required to reduce the base resistance and facilitate fabrication. The dominant scattering mechanisms are believed to be due to polar optical phonons and plasmons. However, the effect of scattering in the base is somewhat mitigated by the fact that the injected electrons have relatively high kinetic energies in comparison to the equilibrium thermal energy of electrons in the base. Electrons may suffer a few collisions in the base without appreciable energy loss or momentum redirection. In addition, computer calculations have shown that quantum reflections at the collector barrier should not be significant. Therefore, high base transport factors approaching one are predicted and have been measured experimentally. Prototype PDB transistors have been fabricated in GaAs epilayers grown by molecular beam epitaxy. Standard photolithography, chemical etching and alloyed eutectic Au/Ge ohmic contacts were employed to yield the transistor structure shown in Figure 11. Preliminary measurements have shown base transport factors very close to 1 at room temperature and power gains of 3dB at 77°K. These measurements have lead to design criteria for high frequency PDB transistors which are currently being developed.

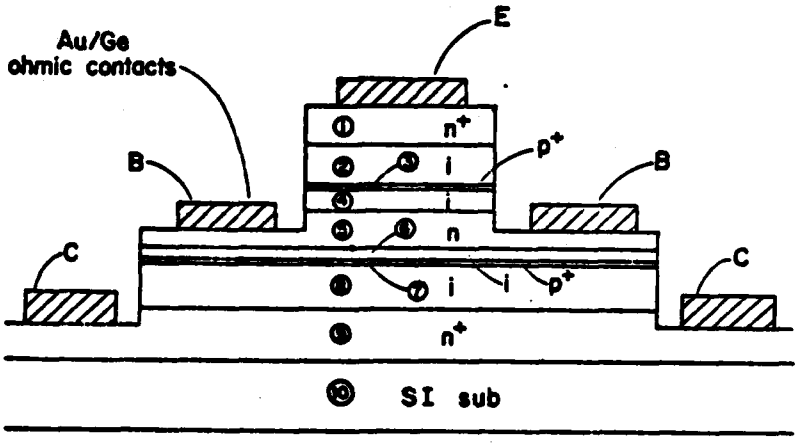


Figure 11. Schematic cross-section of a prototype PDB transistor.

PDB Device Applications

A whole new class of electronic and opto-electronic devices is beginning to evolve based upon the planar doped barrier concept. A variety of novel devices has already been proposed and demonstrated which utilize this PDB technique.

Planar doped barrier diodes are presently being developed in-house at ET&DL, and externally[9] for use at MMW frequencies. Low barrier height (0.2 - 0.5 eV) PDB mixers are being fabricated which will reduce local oscillator (LO) power requirements and conversion losses. PDB's with symmetric I-V characteristics have been tested in subharmonically-pumped mixer circuits, with outstanding results. The diodes, mounted on gold-plated aluminum oxide striplines, have required the lowest reported LO pump power (4 dbm) for a subharmonically-pumped mixer. In this configuration, a single PDB diode replaced the two precisely-matched Schottky barrier diodes used in conventional circuits. This significant accomplishment is of major importance to 94 GH_z and higher frequency radars, where LO pump power is extremely expensive and generally not available. Indeed, the ability to easily fabricate PDB's by MBE may lead to their general use at lower frequencies (X- and K-bands) for improved AM noise rejection and decreased power consumption.

A PDB photodiode has recently been demonstrated which exhibits extremely fast response times and large optical gain.[10] The structure utilizes a reverse-biased PDB diode in which minority carriers created by photon absorption are used to efficiently lower the barrier height, resulting in enhanced thermionic emission over the barrier. Rise and fall times of 100 - 400 ps and DC optical gains approaching 1000 have been measured experimentally.

A novel PDB switching device has also been demonstrated which promises very high switching speeds that may eventually be used for logic gates.[11] The structure uses a PDB in conjunction with a minority carrier injector (p-n junction or Schottky barrier). Minority carrier injection with positive feedback effectively turns the PDB barrier "on" for two stable operating states. Voltage and optically controlled switching have both been demonstrated in this device.

An analysis of the equivalent circuit model of the PDB transistor results in theoretical maximum oscillating frequencies of the order of 300 GHz without submicron lateral geometries. Therefore, the PDB transistor is one of the few device structures expected to achieve three-terminal amplification in the millimeter-wave frequency range (30 GHz <f< 300 GHz).[12]

Conclusions

A new majority carrier rectifying device known as a "planar doped barrier" (PDB) has been successfully demonstrated in GaAs. The structure has an n^+-i-p^+-i-n^+ configuration in which an extremely narrow p^+-planar region is grown in an undoped region by MBE. The barrier height can be readily varied from zero to approximately the bandgap of the semiconductor. Also, parasitic resistance and capacitance in the PDB can be independently minimized, which leads to theoretical cutoff frequencies even higher than those obtained in the best Schottky barrier mixer diodes. In addition, the capacitance of the devices is essentially constant with bias voltage over their whole operating range. The dominant transport mechanism is by majority carrier thermionic emission over the barrier for both positive and negative bias. Although the PDB concept has been experimentally demonstrated in GaAs, it is applicable to all semiconductors in general. Also, since the PDB structure is formed in a single semiconductor, the problems associated with interfaces between dissimilar materials are avoided.

A variety of new device structures has already been realized through the unique characteristics of the planar doped barrier. The PDB concept is expected to have significant applications in high frequency and millimeter-wave devices, including mixer, Gunn, and Baritt diodes, photodiodes, switches, FET's, and MMW transistors.

Acknowledgments

The authors express their gratitude to C. G. Thornton, J. A. Kohn, and S. Dixon, Jr. for program support and technical discussions, and to C. F. Cook, V. E. Rible, J. H. Kwiatkowski, W. Goodreau, M. J. Wade, and K. D. Renz for assistance in processing planar doped barrier structures.

References

1. J. M. Shannon, "A Majority-Carrier Camel Diode," Appl. Phys. Lett. 35, 63 (1979).

2. A. Chandra and L. F. Eastman, "Rectification at n-n GaAs:(Ga,Al) as Heterojunctions," Electron. Lett., 15, 91 (1979).

3. C. L. Allyn, A. C. Gossard, and W. Wiegmann, "New Rectifying Semiconductor Structures by Molecular Beam Epitaxy," Appl. Phys. Lett., 36, 373 (1980).

4. J. M. Shannon, "Control of Schottky Barrier Height Using Highly Doped Surface Layers," Solid State Electron., 19, 537 (1976).

5. R. S. Popovic, "Metal N-Type Semiconductor Ohmic Contact with a Shallow N^+ Layer," Solid State Electron., 21, 1133 (1978).

6. H. A. Bethe, "Theory of the Boundary Layer of Crystal Rectifiers," MIT Radiation Laboratory, 43-12 (1942).

7. C. E. C. Wood, J. Woodcock, and J. J. Harris, "Low Compensation N-Type and Flat Surface P-Type Ge Doped GaAs by Molecular Beam Epitaxy," Inst. Phys. Conf. Ser., 45, 28 (1978).

8. M. Ilegems, "Beryllium Doping and Diffusion in Molecular Beam Epitaxy of GaAs and $Al_xGa_{1-x}As$," J. Appl. Phys., 48, 1278 (1977).

9. J. Paul, Hughes Research Laboratories, private communication.

10. C. Y. Chen, A. Y. Cho, P. Garbinski, and C. G. Bethea, "A High Sensitivity Majority Carrier Photodetector," presented at the Device Research Conference, U. of Cal. at Santa Barbara, (1981).

11. K. Board, K. Singer, R. J. Malik, and C. E. C. Wood, "A Planar-Doped Barrier Switching Device," Eighth Biennial Conference on Active Microwave Semiconductor Devices and Circuits," Cornell U., (1981).

12. R. J. Malik, M. A. Hollis, L. F. Eastman, D. W. Woodard, C. E. C. Wood, and T. R. AuCoin, "GaAs Planar Doped Barrier Transistors Grown by Molecular Beam Epitaxy," Eighth Biennial Conference on Active Microwave Semiconductor Devices Circuits, Cornell U., (1981).

IR Algorithm Development for Fire and Forget Projectiles (U)

*MARCHESE, VINCENT P. MR.
US Army Armament Research & Development Command
Dover, New Jersey 07801

1. INTRODUCTION: ARRADCOM has under contract, on the Advanced Indirect Fire System (AIFS) Program, several contractors who are attempting to solve the problem of autonomous detection of military targets, primarily armored vehicles. The contractors are prepared to analyse data from the background as well as targets in their seeker's field of view to develop algorithms designed to distinguish the targets from the background. Several of the sensing approaches collect and analyse reflected MMW (Radar) energy, while several other approaches depend upon receiving the targets' radiated and/or solar reflected infrared (IR) energy. This paper addresses algorithms currently being used to operate on this IR energy in the 3-5u and 8-12u passbands. It should be recognized that many of the IR detection algorithm techniques are also applicable to other forms of energy.

The Precision Munitions Branch of the Munitions Systems Division of the Large Caliber Weapon Systems Laboratory has undertaken its own independent analysis of the contractors' algorithms by programming them on its in-house computer. These algorithms are then run against stored IR images of actual foreign tanks to determine the capabilities and limitations of each algorithm. Since most contractors have several versions of their proprietary algorithms which they are evaluating to determine which are most effective we have had to evaluate a considerable number of variations. Additionally we have programmed algorithms from other sources evaluating them against the same IR data set. Finally we have indpendently modified and combined various algorithms to attempt to improve their performance.

The in-depth knowledge of the contractors' algorithms gained by evaluating hundreds of IR scenes has allowed us to devise and conduct tower and captive flight tests which have verified the algorithm's performance under conditions approaching its useful limits.

2. ALGORITHMS USED FOR ARMORED TARGET DETECTION: An algorithm is a set of logical rules or mathematical instructions used to solve complex problems. The problem of autonomous target detection most complex because no theoretical model of the background clutter has been found to predict the actual observed data over the range of environmental conditions typically found to occur. Additionally the various target conditions i.e. operating history, viewing aspects, and target types combined with the atmospheric variations produce an unmanageably large set of data.

The most common discriminats used for autonomous detection of targets include: the target's radiant intensity, the variability of the intensity and the spatial extent of the energy. Algorithms should, in addition to finding military targets, reject other objects (including active countermeasures) which are: too hot (a flare or already burning targets), too small (rocks, sheds, flares), too large (buildings, ponds, trees) and the wrong shape (roads, utility lines, natural clutter).

The various algorithms being implemented by the several contractors are quite unique in their details. However several general features can be used to distinguish one from another. Typical operators which algorithms use to process the thermal signature are: Matrix Multiplication, First Moment (M1), Second Moment (M2), Cross Exceedance, Spatial Size Discrimination, Segmentation or Variability, and Edge and/or Shape Finders. It should be noted that several of these operators, although different in their methods, obtain measures of the targets which are really quite similar. For example, several of the matrix operators as well as the cross operator are used solely to obtain spatial (size) discrimination. Not suprisingly, therefore, most seekers collect IR data in a manner which preserves the spatial properties of the target and its background. The simplest method (in theory) of collecting this data is with a large focal plane array (FPA) of detector elements (pixels) since all spatial information is present simultaneously. Typically 32 X 32 or 64 X 64 pixels of information are processed for each "snapshot" or picture. Figure 1 shows a subset of a larger array of pixels, along with a matrix multiplication spatial algorithm suggested by Parenti at MIT (1). The target's thermal energy is assumed to be located at the center of a 9 pixel area of the FPA. The thermal intensity of each background pixel is multiplied by 1/16 while the single target pixel is multiplied by 1. All other pixels are ignored or multiplied by zero. The sum of all 17 pixels is used as a merit function which can be seen to generate a zero for a uniform scene and some number larger than zero for a target. This operator is used once for each different target pixel location.

An AIFS contractor with a focal plane array sensor is using a much more complex scene analysis to distinguish targets from the

background. Sizing this system for use with higher resolution optics, this algorithm calculates and compares the average and standard deviation of the background radiant energy. Figure 2 defines those target and background areas. Each target pixel is then compared to the background, using a first moment type calculation. This process is repeated for all possible target locations, just as the matrix multiplication does. The merit function thus derived is used to locate probable target locations. These locations are handed over to a second algorithm which segments that portion of the scene into areas of like intensity. The number of pixels belonging to segments lying solely within a target sized area is used as the final discriminant. This final segmentation technique may also provide a measure of the target's variability.

A non-imaging seeker, operating on a scanning linear array of 12 pixels can be used with a shift register memory to obtain enough spatial data to find targets. A second moment (M2) operator is used in this algorithm is similar in form to a first moment except that the pixel values are squared. In each nine pixel area (three adjacent pixels of the array sampled three times while the array rotates) an M2 value is computed. An M2 or variance map of the scene is thus constructed in the shift register memory. In this M2 domain a target sized area is compared to like sized areas. When the M2 values in the target area are larger by some threshold than the other areas a target is detected.

A cross exceedance algorithm, although subsequently found to be inefficient at finding targets when used alone, may be used as a pre-screening algorithm. Figure 3 demonstrates how the cross exceedance is used as a discriminant. The exceedance algorithm is satisfied if the center pixel is larger (by some threshold) than each of the pixels indicated by being blackened in any of the nine cross arrays shown.

Spatial algorithms, when used alone, lead to poor overall performance due primarily to false targets. These algorithms search for areas which have the same thermal shape as a target. The primary difficulty with this approach is that the shape of the IR image of a target is not a constant nor is it similar to the visual image except under special circumstances. This is illustrated in the following section.

3. IR IMAGES IN THE DATA SET: Several sources of IR imagery exist in the community. They include NVL, MIT, ERIM and TABILS. Much of the TABILS data was taken at Eglin Air Force Base, Florida with a foreign tank as the target. A visible spectrum photograph is available for each digitized IR image scene. A photograph of a typical scene is shown in Figure 4. An IR image constructed from the TABILS data of the identical scene is shown in Figure 5. One can see from this image that

MATRIX MODIFICATION **ALGORITHM** FROM M.I.T.

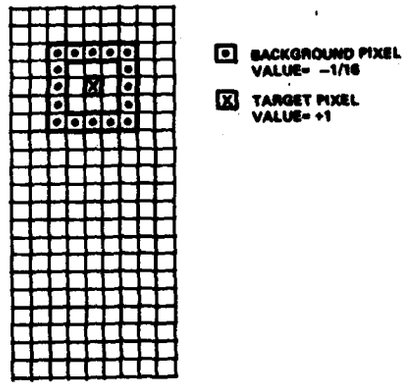


FIG 1

TARGET AND BACKGROUND DEFINITIONS AND **IMAGING SEEKER ALGORITHM**

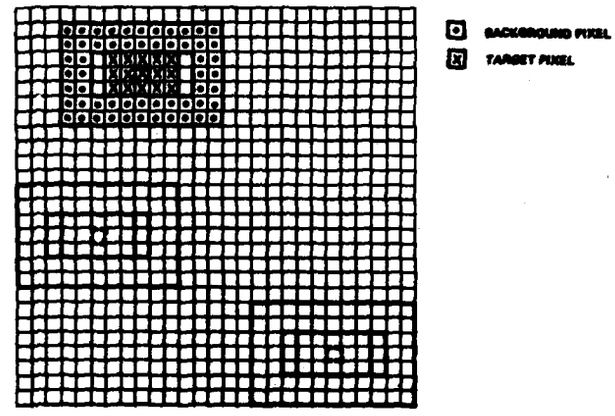


FIG 2

FIGURE 3 **CROSS EXCEEDANCE ALGORITHM**

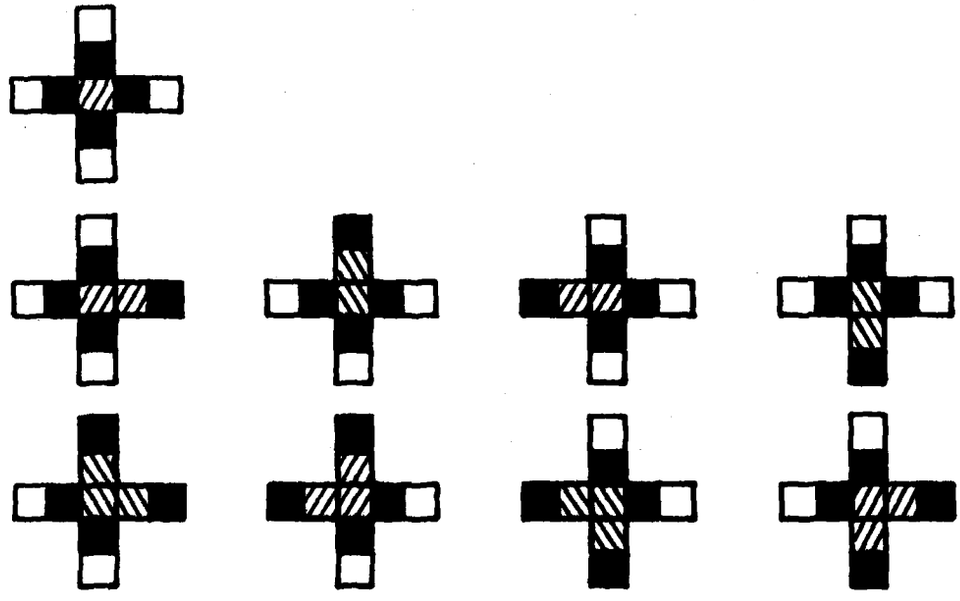




FIGURE 4 - PHOTOGRAPH OF TABILS SCENE



FIGURE 5 - IR IMAGE OF TABILS SCENE

the IR shape is not at all similar to the visual shape. (Note that registration lights - heat lamps - are placed in the scene to aid in locating the target). Another illustration of this thermal shape difference can be seen in Figure 6. These photos show a top view of the same target under two operating conditions (top exercised, bottom stationery).

4. EXAMPLES OF ALGORITHMS APPLIED TO IR DATA: The IR image shown as shades of grey in Figure 5 is shown in Figure 7 where different characters represent different intensities. The scale is such that punctuation marks and symbols are the lower values followed by the numbers (in order) with the letters (in order) being the higher intensity. The asterisk is the highest value. This scene operated upon by the M2 algorithm is shown in Figure 8. Here one can notice that the small heat lamps become larger and are of the highest level. Figure 9 shows the result of applying the cross exceedance pre-screener and a size discriminant to the M2. Note that many of the false alarms have been eliminated, and if a threshold at the "C or D" level were chosen it would result in a scene which consists of just the target and the two heat lamps. One can then easily eliminate the heat lamps by using a T-max or color ratio technique.

The MIT algorithm operating on this same scene is shown in Figure 10. It does not make the heat lamps bloom as the M2 operator does, but it has problems with separating the target from the background. A variant of this algorithm, using the MIT matrix to operate on a M2 map significantly reduces the false alarms.

The first moment algorithm "found" three potential targets in this sample scene shown in Figure 11. Figure 12 illustrates the results of the segmentation algorithm. The shaded pixels indicate those segments which lie totally within a target sized window. This is the target area.

5. RESULTS OF COMPUTER ANALYSIS AND/OR CAPTIVE FLIGHT TESTS OF SELECTED ALGORITHMS: The most successful sets of algorithms are those demonstrated in Figure 10. Not only did this combination of algorithms successfully find the target in almost all IR data scenes it also performed remarkably well during actual captive flight tests. Problem areas which surfaced during flight tests were primarily due to false targets which had not been seen in the IR data sets.

Computer analysis of the algorithms using the first moment with spatial processing and segmentation indicated several drawbacks. Many targets were missed while other background areas were selected as potential targets. Also the target areas chosen often would include only part of the target while including a considerable portion of the

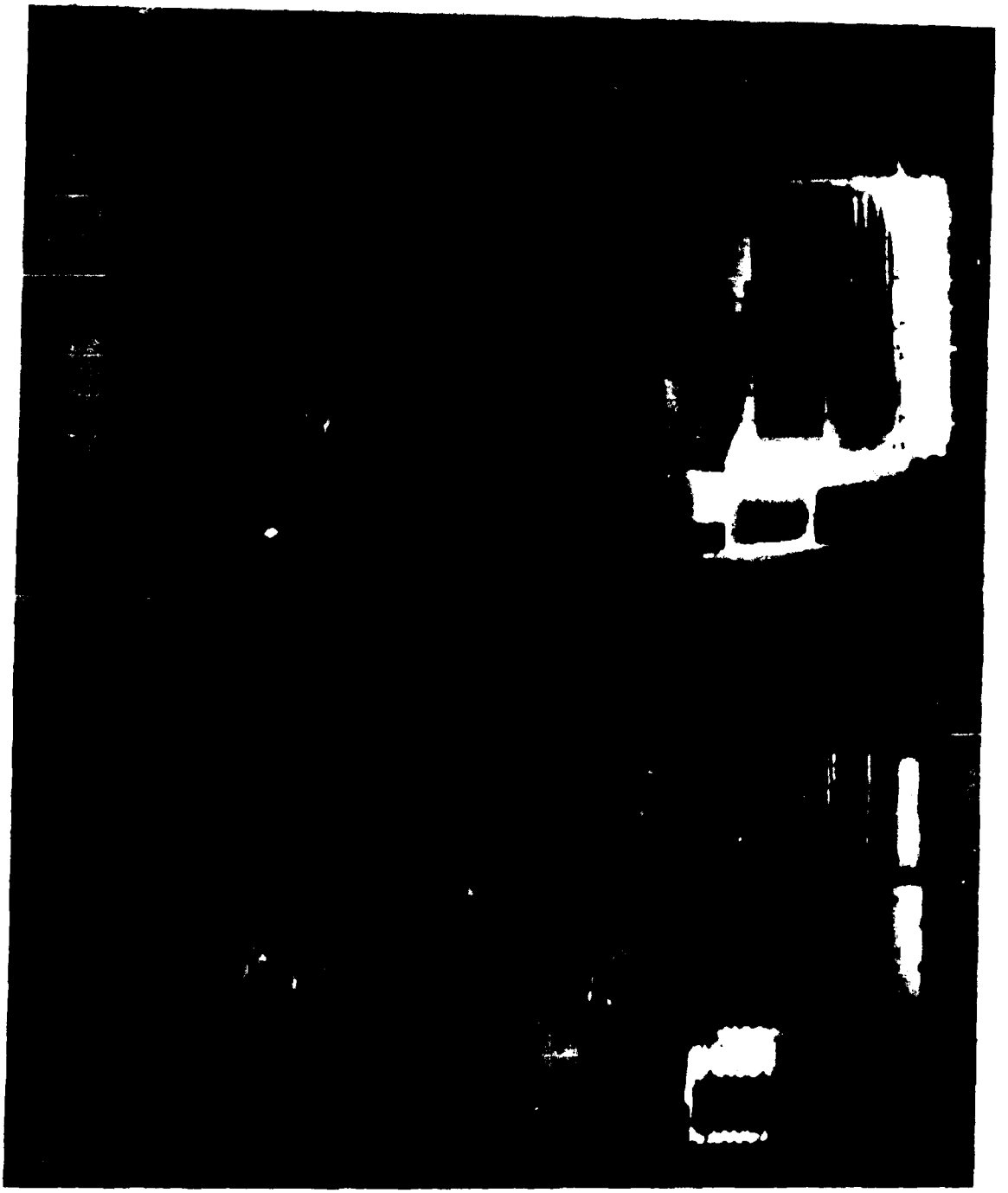


FIGURE 6 - THERMAL IMA[illegible]Y OF A TYPICAL TANK, (TOP-EXERCISED, BOTTOM-IDLE)

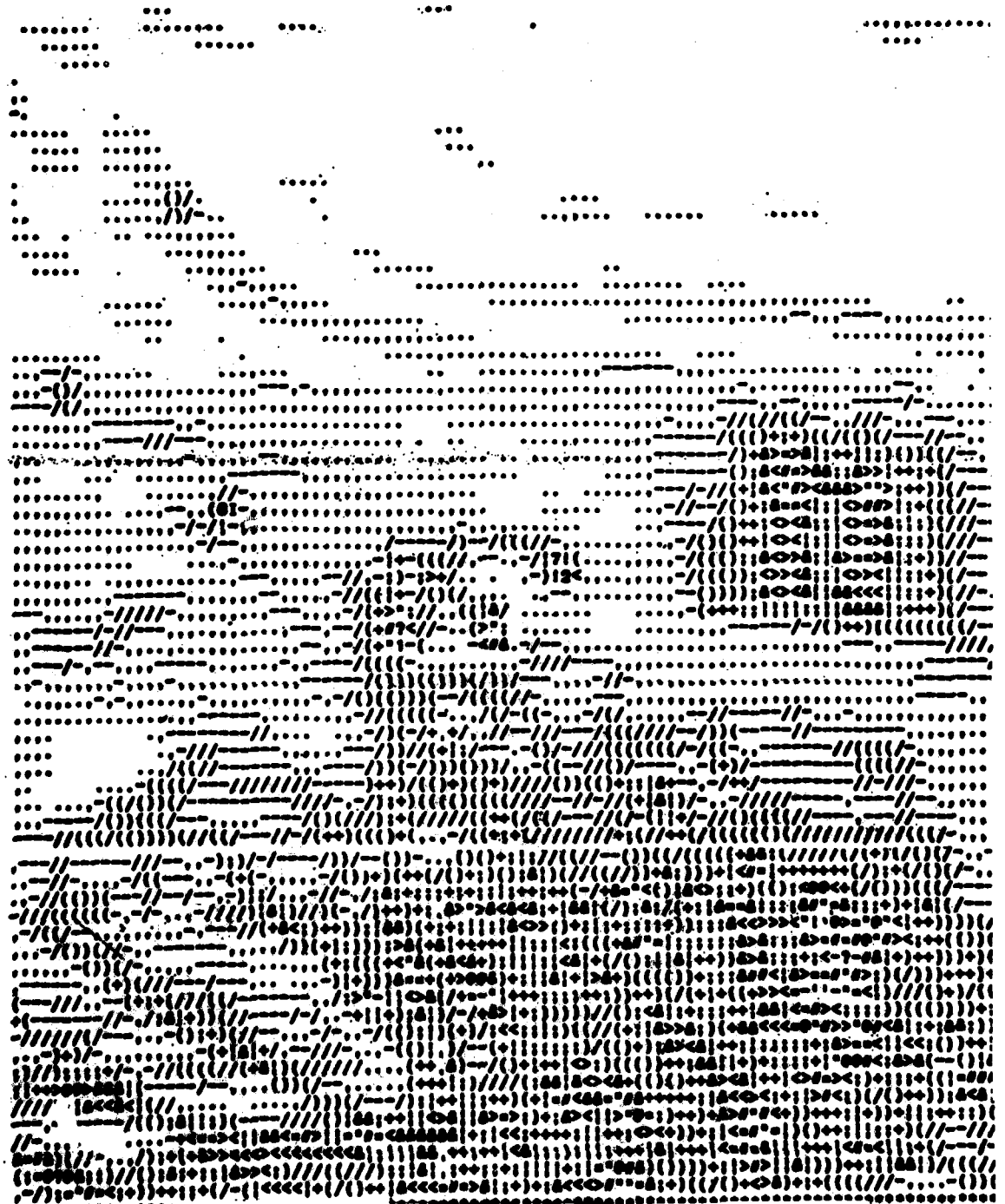
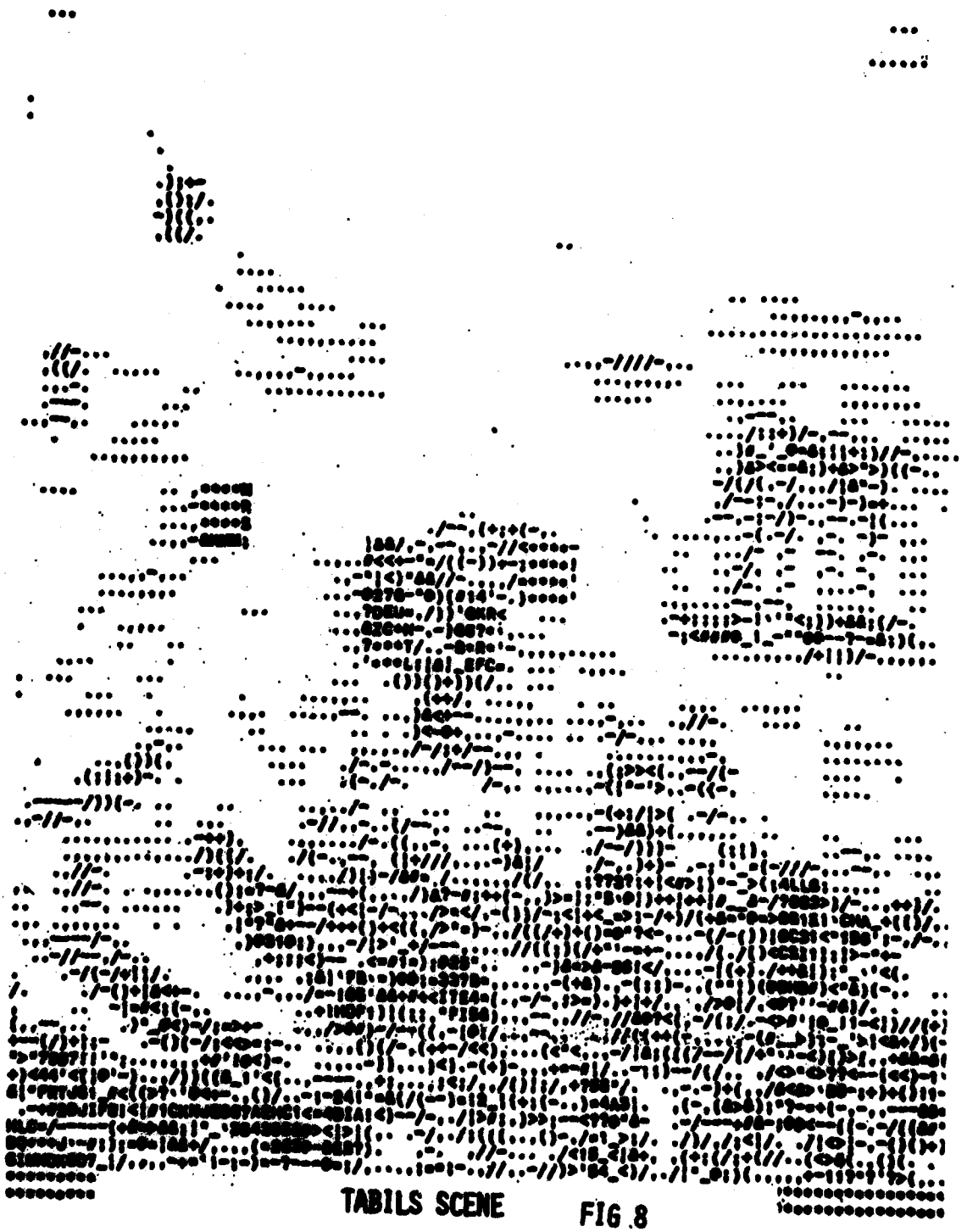


FIG 7 TABILS SCENE RADIANCE DATA



TABILS SCENE FIG 8



FIG 9 TABILS SCENE-GATED M2 + SPATIAL

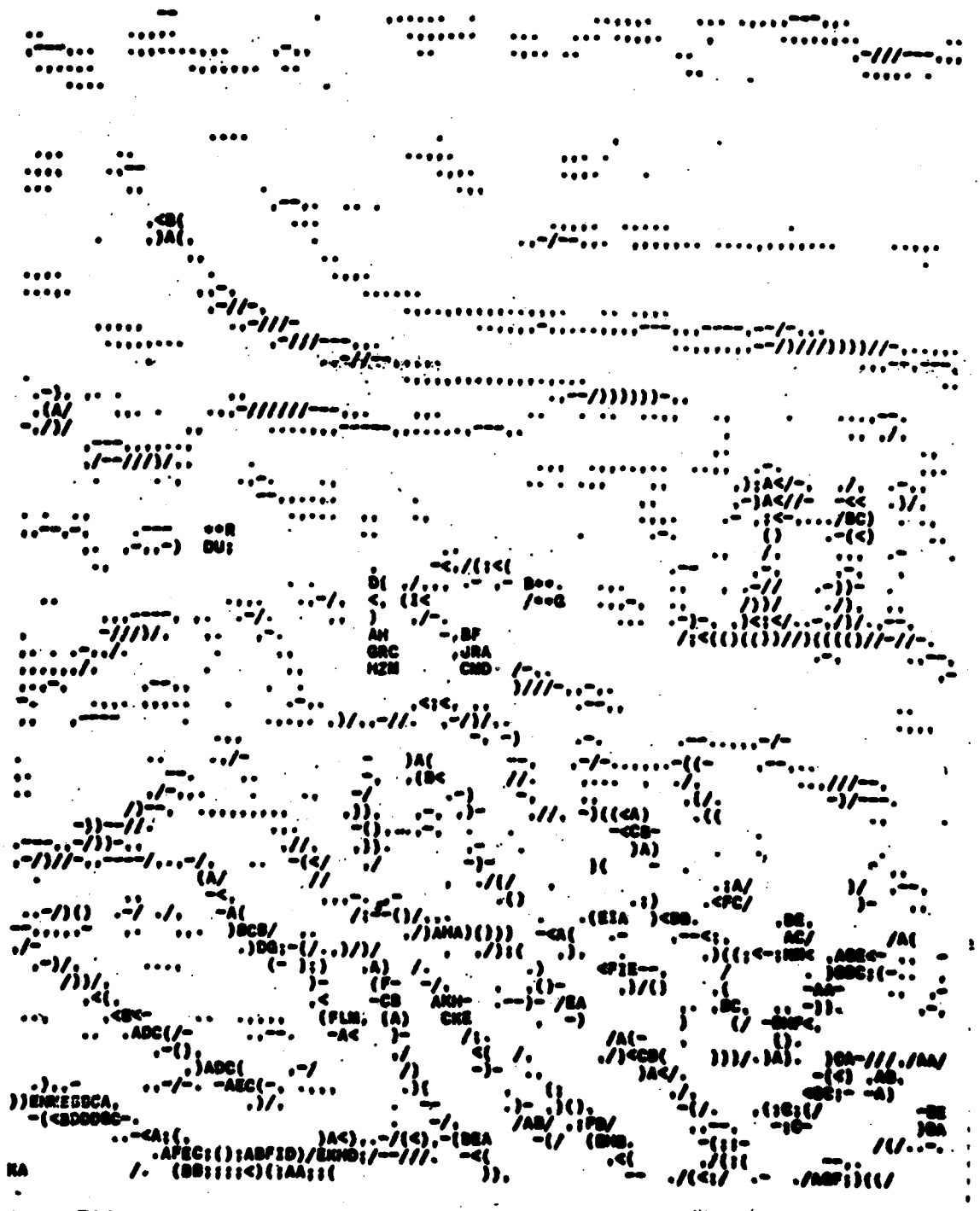


FIG 10 TABILS SCENE - MATRIX MULTIPLICATION

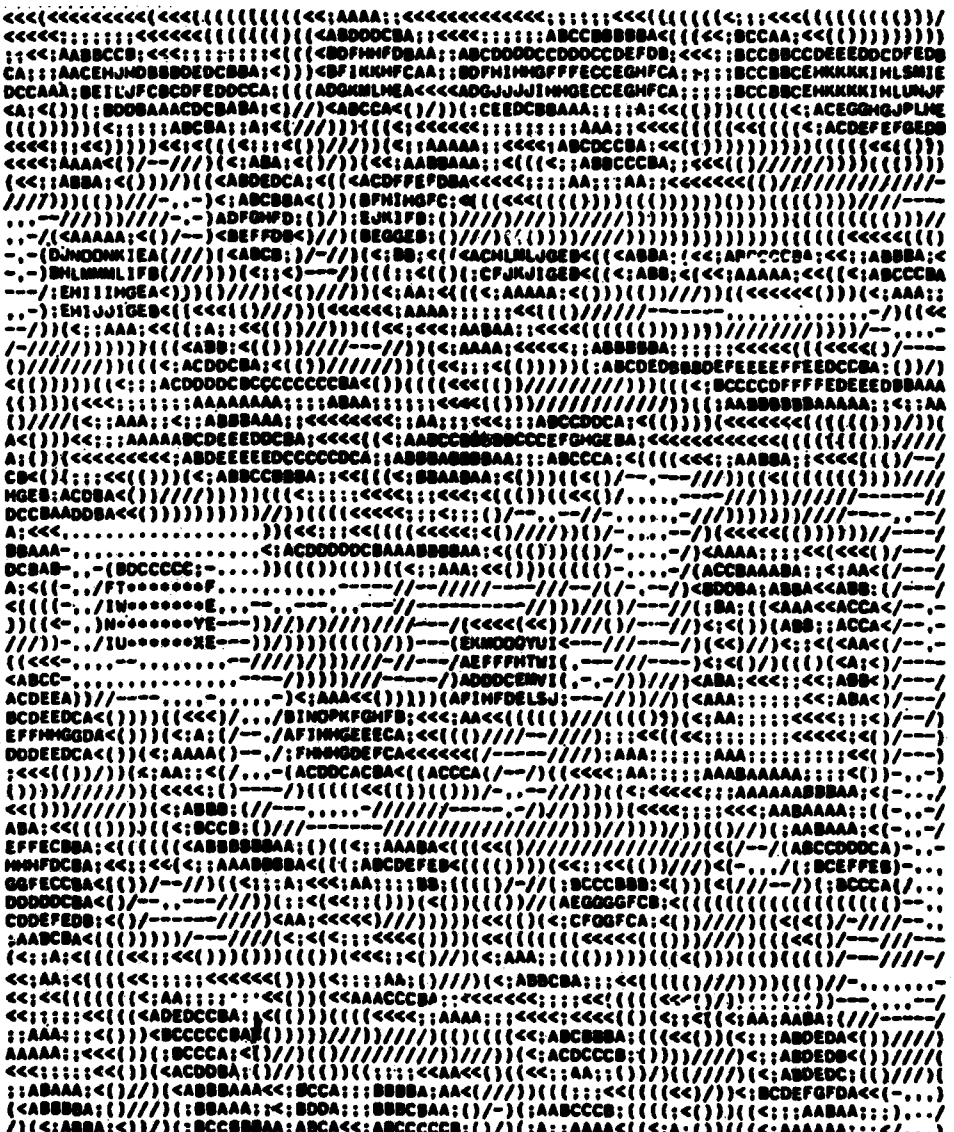


FIG 11 TABILS SCENE - FIRST MOMENT

LOAL ALGORITHM PERFORMANCE ON TABILS IMAGERY

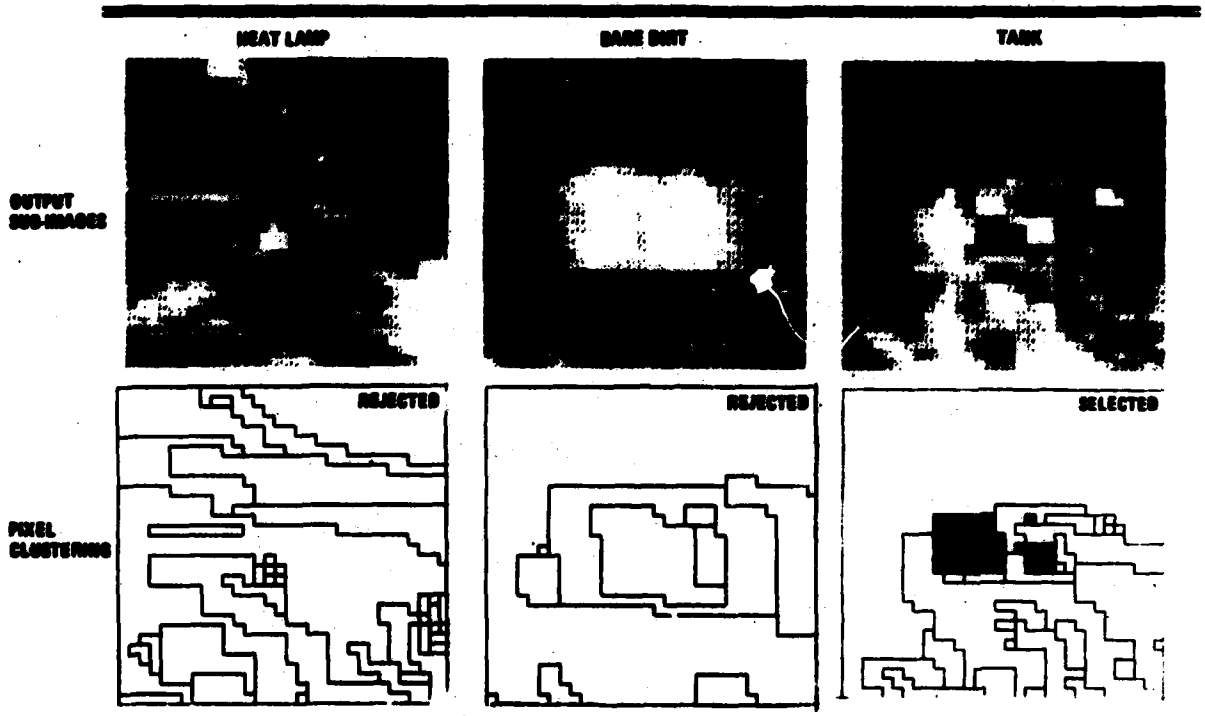


FIG 12

background. Ground tests of a seeker using these algorithms confirmed the existance of the problem areas predicted by the computer analysis and thus captive flight tests were not conducted.

6. CONCLUSIONS: Various algorithms have been evaluated against several hundred IR scenes. Some of these algorithms have been found to be more successful at autonomous target detection than others. Generally, those which use the second moment operator find more real targets. Those using the first moment are more successful when they use a second algorithm to segment the image. LCWSL, using these analyses, has obtained valuable insight into the capabilities and limitations of not only the algorithms used by the AIFS contractors but also of other contractors who have proposed algorithms to the Government. The ability to captive flight test only seekers with the most promising algorithms provides the Army the opportunity to conserve valuable resources. Thus the usefulness of predicting probable success or failure of target detection algorithms from a computer analysis using various IR data sources has been demonstrated. Additional insights into algorithm performance allows the Government to better specify future autonomous seekers as well as to design test programs which will accurately predict the system performance of the fielded weapon system.

7. REFERENCES:

1-Parenti, R., Otazo, J. and Tung, E. "The Design of a Digital Filter for Resolved Stochastic Targets" Presented at 1980 National IRIS Meeting.

DEVELOPMENT OF A NEW DESIGN PROCEDURE FOR OVERLAND FLOW SYSTEM

C. JAMES MARTEL, ENVIRONMENTAL ENGINEER
U.S. ARMY COLD REGIONS RESEARCH AND ENGINEERING LABORATORY
HANOVER, NEW HAMPSHIRE 03755

INTRODUCTION

In 1971 the U.S. Army Corps of Engineers conducted five comprehensive studies on the feasibility of regional wastewater management systems for large urban areas. Each study identified land application as a viable means for wastewater treatment and disposal. However, the Corps soon determined that the technology for designing and operating land application systems was inadequate. As a result, the Corps initiated a land treatment research and development program in 1972, and the Cold Regions Research and Engineering Laboratory (CRREL) was designated as the lead laboratory. This program was successfully completed in 1980 and much of the research provided information for updating the Corps/EPA/USDA/DOI Process Design Manual for Land Treatment of Municipal Wastewater (16). Although this research program has been completed, the Corps continues to support a center of land treatment expertise at CRREL for continued technology transfer and assistance to Divisions and Districts.

The three methods of land application studied under the research and development program were slow rate, rapid infiltration and overland flow. The slow rate method is quite similar to agricultural irrigation in that wastewater is applied to farms, fields and forests at a rate which produces no runoff. The rapid infiltration method is similar to natural sand filtering and can only be used when sandy or gravelly soils are available.

The three methods of land application studied under the research and development program were slow rate, rapid infiltration and overland flow. The slow rate method is quite similar to agricultural irrigation in that wastewater is applied to farms, fields and forests at a rate which produces no runoff. The rapid infiltration method is similar to natural sand filtering and can only be used when sandy or gravelly soils are available. The overland flow method is best suited in areas with less permeable soils such as clays and clay loams. When the research program

was initiated in 1972, overland flow was the least developed method of land application.

An overland flow system consists of a series of grassy terraces which are carefully graded so that wastewater flows downslope in a thin sheet. Gated pipe, troughs or sprinklers are used to uniformly distribute the wastewater at the top of each terrace. Renovation occurs as the wastewater travels over the soil surface, and the volume of wastewater is reduced because of evapotranspiration and percolation. The remaining runoff is collected in a shallow ditch at the base of the terrace and discharged to a receiving stream. When properly designed and managed, the runoff water quality from an overland flow system can easily meet secondary effluent standards.

Overland flow was initially developed in this country back in the early fifties by the Campbell Soup Co., which used this process for treating high-strength cannery wastes. Their success encouraged the Environmental Protection Agency (EPA) to evaluate overland flow as a process for treating municipal wastewater. Studies conducted at the EPA Laboratory in Ada, Oklahoma, by Thomas et al. (15) confirmed that the process was effective in renovating municipal wastewater. However, little was known about the operational limits of overland flow and the only procedure available for design was based on general guidelines and rules-of-thumb. Thus, a primary objective of the Corps research effort was to test the limits of overland flow performance and develop a more rational procedure for design. This research effort was conducted at both CRREL and the U.S. Army Engineer Waterways Experiment Station (WES).

The new design procedure developed as a result of this research is based on reactor kinetics, a concept familiar to most environmental engineers. In the case of overland flow, the reactor is the soil surface where various physical, biological and chemical reactions take place. As in conventional process design, the controlling parameter is detention time. For overland flow, detention time is the average time a unit volume of water takes to travel from the top to the bottom of the terrace. The desired level of treatment can be achieved by controlling the length of time that wastewater remains in contact with the soil surface. With this approach, overland flow systems can be constructed for a wide range of site conditions as long as detention time requirements are met.

DESCRIPTION AND OPERATION OF CRREL OVERLAND FLOW TEST SITE

The data used to develop the hydraulic and kinetic design relationships were obtained from the CRREL overland flow test site. This site, which has been in operation since June 1977, is 30.5 m long x 8.8 m wide (0.03 ha) and graded to a 5% slope. It is subdivided into three equal

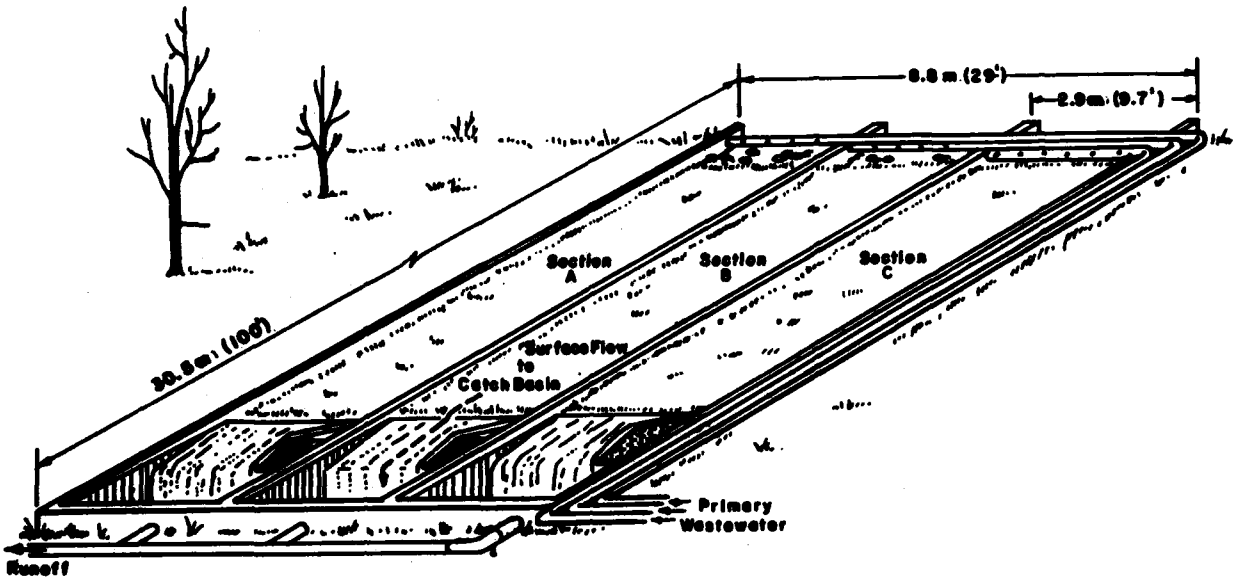


Figure 1. Schematic of CRREL overland flow test site.

sections designated A, B and C so that parallel studies can be conducted. Underlying the soil at a depth of 15 cm is a 30.0-mil-thick rubber membrane, which was installed to prevent downward percolation. The grass cover on the site is a mixture of many species including K-31 tall fescue, orchardgrass, Kentucky bluegrass and quackgrass (10). The grass was harvested on the average of once every six weeks during the growing season. A schematic of the site is shown in Figure 1.

Undisinfected primary effluent was applied to the overland flow test site during the entire study. Perforated plastic pipe was used to distribute wastewater along the top of each section, and a bed of crushed stone placed beneath the pipe helped to uniformly disperse the flow. The quality of the primary effluent is shown in Table 1.

The application rate was monitored and controlled by means of a constant head weirbox. Five application rates ranging from 0.35 to 1.20 m^3 hr^{-1} were tested. The application cycle was 7 hr on, 17 hr off for 5 days per week. At these application rates and this cycle, the equivalent hydraulic loading rates were 13.8 to 46.7 cm wk^{-1}. Each application rate was evaluated for a period of approximately 6 weeks. All sections were operated simultaneously at the same application rate. Because of leaks in the membrane along the outside boundary, wastewater applications to section A were discontinued during 1979.

Table 1. Quality of applied primary effluent.

Parameter	Mean	Standard deviation	No. of observations
BOD (mg L^{-1})	72	23	58
Total suspended solids (mg L^{-1})	59	30	98
Ammonia (mg L^{-1} as N)	24	6	99
Total phosphorus (mg L^{-1} as P)	6.6	2.2	33

Runoff was collected at the base of each section in individual galvanized steel catch basins. A small submersible sump pump located in each basin discharged the runoff into a drainage ditch. The volume of runoff was recorded by flowmeters attached to the discharge lines. During this study, the average runoff rate was 75, 87 and 89% of the application rate for sections A, B and C, respectively.

All measurements of detention time and water quality sampling were conducted during periods of hydraulic steady-state operation. The hydraulic steady-state period began when the runoff rate stabilized, and it terminated when application was stopped. The amount of time needed to reach hydraulic steady state varied depending on antecedent moisture conditions.

Hydraulic detention time was determined by measuring the travel time of a chloride tracer. Chloride was selected because it is conservative and easily analyzed. A tracer solution was made by dissolving 94.6 g of sodium chloride in 3 L of distilled water. The sodium chloride solution was added as a "slug addition" to the distribution chamber in the constant head weirbox. Composite samples were taken of the runoff at various intervals and analyzed for chloride. Chloride concentrations were then plotted vs time, and the peak of the response curve was chosen to represent detention time. An example of a chloride response curve is shown in Figure 2. The detention time in this case was 40 min for an application rate of 0.6 m^3 hr^{-1}. Altogether, 50 detention times were measured at the CRREL site during this study.

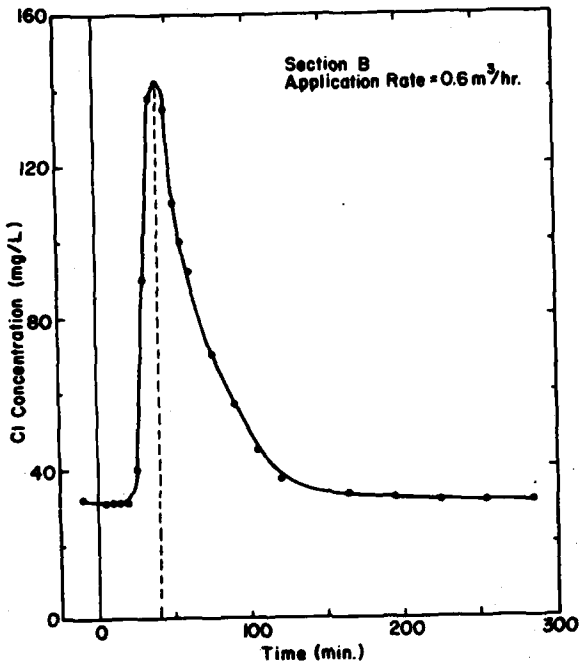


Figure 2. Typical chloride response curve for measuring detention time.

DEVELOPMENT OF THE HYDRAULIC DETENTION TIME RELATIONSHIP

At a well designed and operated overland flow site, water flows downslope as a thin sheet until it freefalls into a runoff collection ditch. Under these conditions, the overland flow system can be considered to operate in the laminar flow regime (4). For the simplest case of overland flow over a smooth surface, the average velocity v_s can be described by the following equation (8):

$$v_s = \frac{g\ S\ d^2}{3\ \nu} \quad (m\ s^{-1}) \tag{1}$$

where g = gravitational constant, 9.81 m s^{-2}
S = slope, m m^{-1}
d = average depth of flow, m
ν = kinematic viscosity, $m^2\ s^{-1}$.

For an actual overland flow system, resistance to flow will be greater because of the grass and vegetative litter. Therefore, the aver-

age overland flow velocity V will be lower than the smooth surface velocity v_s and can be expressed as

$$V = \alpha\, v_s \ (m\ s^{-1}) \tag{2}$$

where α, the resistance coefficient, is less than 1.0. Substituting eq 2 into eq. 1, the velocity of flow over an overland flow terrace can be described by

$$V = \alpha \left[\frac{g\ S\ d^2}{3\ \nu}\right] \ (m\ s^{-1}) \tag{3}$$

If one assumes that most of the water flows in a straight path downslope, the velocity V can be expressed as

$$V = \frac{L}{\bar{t}} \ (m\ s^{-1}) \tag{4}$$

where L is the length of terrace in meters, and $\bar{t}$ the hydraulic detention time in seconds. Also, from the continuity equation, the average depth of flow d can be determined by

$$d = \frac{Q\ \bar{t}}{L\ W} \ (m) \tag{5}$$

where Q is the average overland flow rate ($m^3\ s^{-1}$) and W the width of the terrace in meters. Substituting eq 4 and 5 into eq 3 and rearranging terms, detention time can be calculated from the relationship

$$\bar{t} = \left[\frac{3\ \nu\ W^2}{\alpha\ g\ S\ Q^2}\right]^{1/3} L. \tag{6}$$

In more convenient terms, where the average detention time is described in minutes ($\bar{T}$) and the average overland flow rate (q) in $m^3\ hr^{-1}\ m^{-1}$ of width, eq 6 becomes

$$\bar{T} = 5.65 \left[\frac{\nu}{\alpha\ g}\right]^{1/3} \frac{L}{S^{1/3}\ q^{2/3}} \ . \tag{7}$$

Assuming a kinematic viscosity of $0.112 \times 10^{-5}\ m^2\ s^{-1}$ (at 15.6°C) and substituting the value of the gravitational constant g, eq 7 becomes

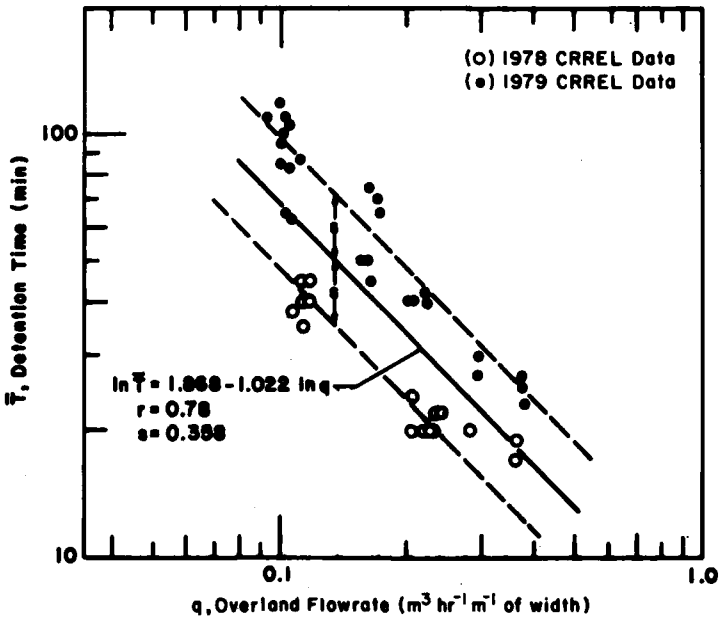


Figure 3. Overland flow rate vs detention time for CRREL overland flow test site.

$$\bar{T} = 0.0274 \frac{L}{\alpha^{1/3} S^{1/3} q^{2/3}} \quad . \tag{8}$$

To determine the resistance coefficient α, eq 8 was evaluated using data obtained from the CRREL overland flow test site. For each CRREL test section, the values of L and S are 30.5 m and 0.05 m m^{-1} respectively. Substituting these values, eq 8 can be simplified to

$$\bar{T} = \frac{2.27}{\alpha^{1/3} q^{2/3}} \quad . \tag{9}$$

By plotting detention time vs the average overland flow rate on log-log paper, α can be determined from the line of best fit. This was done for the CRREL data as shown in Figure 3. A regression analysis indicates good correlation ($r = 0.78$) between application rate and detention time. However, the standard deviation is large, indicating that detention time varied considerably for a given overland flow rate. Most of this deviation appears to be caused by a difference in results obtained between the 1978 and 1979 growing seasons. The equation for the line of best fit shown in Figure 3 is

$$\ln \bar{T} = 1.868 - 1.022 \ln q \tag{10}$$

or

$$\bar{T} = \frac{6.48}{q^{1.022}} \ . \tag{11}$$

Substituting eq 11 for $\bar{T}$ in eq 9, an expression for the resistance coefficient is

$$\alpha = 0.043\ q^{1.066} \approx 0.043\ q \ . \tag{12}$$

This expression indicates that the resistance coefficient α increases in direct proportion to the average overland flow rate. This relationship can be explained by the fact that, as the flow rate increases, the depth of flow also increases. On the irregular surface of most overland flow terraces, increasing the depth causes more surface area to be wetted, which increases the resistance to flow. This hypothesis is consistent with visual observations at the CRREL site and several other overland flow sites.

Substituting eq 12 back into eq 8, the final form of an empirical relationship for predicting detention time is

$$\bar{T} = \frac{0.078\ L}{S^{1/3}\ q} \ . \tag{13}$$

This equation indicates that T is directly proportional to L and inversely proportional to q. Slope, being to the one-third power, is less significant although it cannot be considered negligible. For example, assuming L = 50 m and q = 0.2 $m^3\ hr^{-1}\ m^{-1}$, an increase in slope from 2 to 12% would decrease detention time from 72 to 40 minutes, a decrease of 44%.

To determine the validity of eq 13, detention times were measured at two other overland flow sites. The first site, located near Utica, Mississippi, was a research facility operated by the U.S. Army Engineer Waterways Experiment Station (WES). This site (no longer in operation) had 24 terraces, each 45 m long by 4.5 m wide and slopes of 2, 4, and 8% (11). The second site is located indoors at the University of California at Davis. Each laboratory scale terrace is 6 m long x 1.5 m wide and set at a 4% slope (14). A combined total of 40 detention time measurements were taken at both sites.

Statistical analysis of these data indicated that the average difference between predicted and measured detention times was only 8 minutes.

In most cases the measured detention time was longer than predicted, which allows an extra margin of safety in the design. In a Student's t distribution, the difference between measured and predicted detention time was not significant at the 95% level. Therefore, eq 13 appears to adequately describe the average hydraulic characteristics of overland flow. However, dat-to-day differences varied considerably. This is understandable, considering the variability of the surface microtopography from one terrace to another. Construction techniques, patterns of vegetative growth and harvesting operations are also factors which can change the hydraulic detention time.

DEVELOPMENT OF THE KINETIC RELATIONSHIPS

Kinetic relationships describing removal of biochemical oxygen demand (BOD), total suspended solids (TSS), ammonia (NH_3-N) and total phosphorus (total P) were developed by taking several detention time measurements during each application period. The average detention time ($\overline{T}$) was then calculated along with the average percent removal on a mass basis for each constituent.

BOD removal

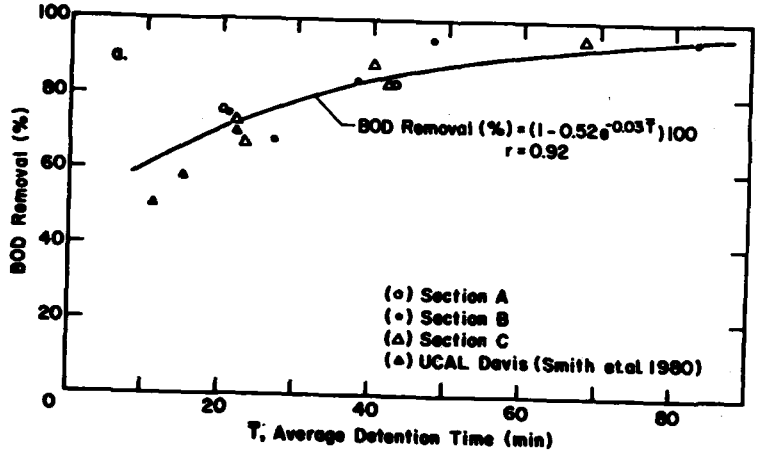
BOD is removed by sedimentation, filtration and biological oxidation (16). The first two mechanisms are responsible for removing particulate BOD. The soluble BOD is oxidized by microorganisms which are probably similar to the attached biomass found in trickling filters. However, some soluble organic compounds are released from the plant-soil system and, as a result, runoff BOD concentrations below 3 to 5 mg L^{-1} should be expected (9).

Temperature also has an effect on runoff BOD concentrations. Martel et al. (5) found that BOD concentrations in the runoff exceeded 30 mg L^{-1} at soil temperatures at or below 4°C. However, temperature effects should not be a significant problem at full-scale facilities if wastewater is stored during the winter. In this study, temperature effects were nullified by selecting performance data obtained during the growing season only (April through October).

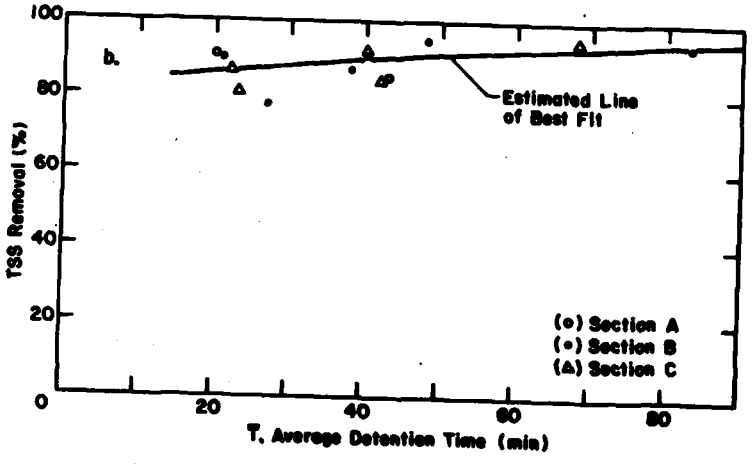
Data obtained at both CRREL and the U. of California, Davis (Fig. 4a) indicate that BOD removal can be expressed as a first-order equation in the form

$$\text{Percent removal} = \left(1 - A\, e^{-k\overline{T}}\right) 100 \ . \qquad (14)$$

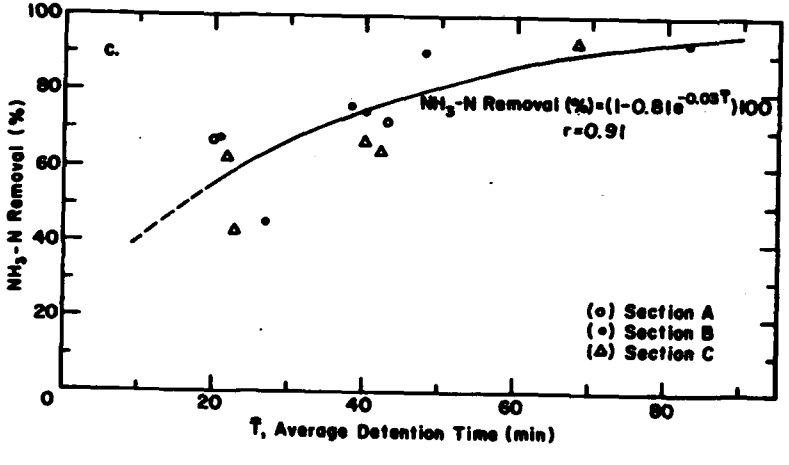
The coefficients A and k, obtained by a least-squares fit to the data, were 0.52 and 0.03 min^{-1}, respectively. The coefficient k is the average



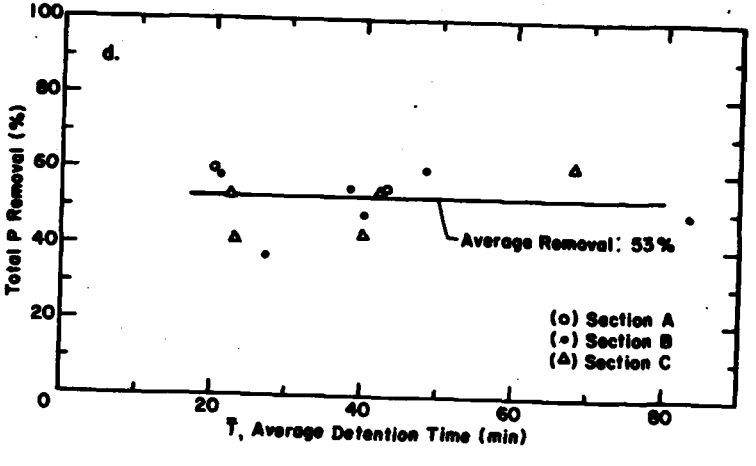
a. BOD.



b. TSS.



c. NH_3-N.



d. Total P.

Figure 4. Kinetic relationships for BOD, TSS, NH_3-N and Total P removal.

kinetic rate constant. The coefficient A can be interpreted as the non-settleable fraction of BOD in the applied wastewater while the remaining settleable fraction (0.48) is removed during the first few meters or minutes after wastewater is applied.

TSS Removal

Total suspended solid (TSS) removal vs average detention time from the CRREL site is shown in Figure 4b. The flat slope of the estimated line of best fit indicates that TSS removal changed little over the range of detention times tested. For example, at a detention time of 20 minutes, TSS removal was 86%. A three-fold increase in detention time (60 min) only increased removal by 6%.

The high solids removal efficiency of the overland flow process is due to the shallow depth of water and the long travel distance to the end of the terrace. Even minute particles with slow settling velocities are able to settle out before reaching the collection ditch. Also, grass and vegetative litter help to entrap and filter out particles.

The solids removal relationship developed in this study (Fig. 4b) applies to fecal types of solids only. Removal of algal solids found in lagoon effluent is more difficult to predict. Data from the Easley, South Carolina, site indicated that algae removal by overland flow is marginal (12). However, Peters et al. (11) report good removal of algae at low application rates.

Nitrogen Removal

A number of mechanisms are involved in nitrogen removal, including volatilization, nitrification-denitrification, adsorption, plant uptake and soil storage. The ammonia form of nitrogen can be removed by any of the above mechanisms. Most of the organic nitrogen is initially removed by sedimentation and then incorporated into the soil or converted to ammonia by saprophytic bacteria. Nitrate is the most difficult form of nitrogen to remove (13, 18). Nitrate ions have little affinity for soil particles and thus are not retained on the overland flow terrace.

This study focused on the kinetics of ammonia removal because it is the nitrogen form of most concern in discharge limitations. The correlation between ammonia removal and detention time obtained from CRREL data is shown in Figure 4c. The first-order equation which closely fits these data ($r = 0.91$) is also shown in Figure 4c. For ammonia removal the coefficients A and k were 0.81 and 0.03 min^{-1}, respectively.

It is interesting to note that both BOD and ammonia removal equations (see Fig. 4a and 4c) contain the same kinetic rate constant ($k = 0.03$ min^{-1}), suggesting that both BOD and ammonia removal are controlled by the

same rate-limiting step. It is unlikely that both substrates would have the same removal rate constant; a more likely explanation is that removal rate is mass transport limited. In other words, the rate of mass transport from the bulk liquid to the active biomass and adsorption sites is the mechanism governing removal rate. This reasoning is reinforced by the fact that overland flow operates in a laminar flow regime, which reduces the opportunities for substrate contact with reactive sites.

Phosphorus removal

Phosphorus is removed primarily by sorption to soil particles. On overland flow terraces only surface exchange sites are available because most of the wastewater passes over the soil surface rather than through it. As a result, the exchange sites are used up rather quickly, and the removal of phosphorus by overland flow systems is limited. Plant uptake is another mechanism capable of removing phosphorus. Palazzo et al. (10) reported that forage grasses removed 54% of the applied phosphorus at the CRREL site.

As shown in Figure 4d, phosphorus removal did not change significantly over the range of detention times tested. Percentage removals ranged between 37 and 61% and averaged 53%. Analyses of runoff samples indicated that most of the total phosphorus was in the "ortho" form, which indicates that the phosphorus removed was tied up with particulate matter. As discussed earlier (see TSS removal), particulate matter was easily removed by overland flow.

Validation

The kinetic relationships for removal of BOD, TSS and NH_3-N were validated by comparing the predicted removal to the actual removal reported at seven full-scale systems. Statistical analysis of these data revealed that the average differences between predicted and actual BOD, TSS and NH_3-N removal were only 1.9, -2.0 and 2.8%, respectively, for systems receiving primary or raw wastewater (see Table 2). These results confirm that the kinetic relationships for BOD and TSS removal are valid for overland flow systems receiving primary or raw wastewater. The lowest applied concentration of BOD and TSS where these relationships hold is estimated to be 45 mg L^{-1}. Different kinetic relationships need to be developed for overland flow systems receiving secondary or pond effluent.

CONCLUSION

The hydraulic and kinetic relationships developed during this study can be used as the basis of a rational procedure for design of overland flow systems. The three basic steps in this procedure are

Table 2. Predicted vs actual removal efficiencies on overland flow systems receiving primary or raw wastewater.

System	Calculated detention time (min)	Runoff frac.	Applied conc. (mg L^{-1})			Predicted removal (%)			Actual removal (%)			Predicted-actual removal (%)		
			BOD	TSS	NH_3-N	BOD	TSS	NH_3-N	BOD	TSS	NH_3-N	BOD	TSS	NH_3-N
Ada, Okla. (15)	222	0.50	150	160	17.0	99+	95+	99+	96	97	98	3	-2	1
	195	0.50	150	160	17.0	99+	95+	99+	97	98	96	2	-3	3
	171	0.50	150	160	17.0	99+	95+	99+	97	98	97	2	-3	2
Pauls Valley, Okla. (2)	294	0.50	117	105	17.0	99+	95+	98	96	97	91	3	-2	7
Werribee Farm, Aust. (13)	626	0.80	507	233	31.0	99+	95+	99+	98	93	20*	3	2	--
Easley, S.C. (12)	59	0.70	200	186	19.4	91	92	86	91	97	85	0	-5	1
Paris, Tex. (1)	138	0.60	480	181	--	99	95+	--	99	96	--	0	-1	--
Mean												1.9	-2.0	2.8
Std. deviation												1.3	2.2	2.5

* Wastewater was applied during the winter when crops were not actively growing.

1. Determine the detention times required to remove pollutants specified in the discharge permit (Figs. 4a, b, c, d).

2. Calculate the application rate needed to satisfy the longest or most critical detention time (eq 13).

3. Calculate the land area required from the application rate and system design flow.

These three steps are discussed in further detail by Martel et al. (7). Examples of how to use the procedure can be found in the new Process Design Manual for Land Treatment of Municipal Wastewater (16) and Martel et al. (6, 7).

LITERATURE CITED

1. C.W. Thornthwaite Associates (1969) An evaluation of cannery waste disposal by overland flow spray irrigation. Climatology, vol. 22, no. 2, Elmer, New Jersey.

2. Hall, D.H., J.E. Shelton, C.H. Lawrence, E.D. King and R.A. Mill (1979) Municipal wastewater treatment by the overland flow method of land application. EPA Research Reporting Series EPA-600/2-79-178.

3. Jenkins, T.F., C.J. Martel, D.A. Gaskin, D.J. Fisk and H.L. McKim (1978) Performance of overland flow land treatment in cold climates. In State of Knowledge in Land Treatment of Wastewater, Proceedings of an International Symposium, Hanover, New Hampshire, vol. II, p. 61-70.

4. Kirkby, M.J. (1978) Hillslope hydrology. New York: John Wiley and Sons.

5. Martel, C.J., T.F. Jenkins and A.J. Palazzo (1980a) Wastewater treatment in cold regions by overland flow. CRREL Report 80-7. ADA084489.

6. Martel, C.J., D.D. Adrian, T.J. Jenkins and R.E. Roberts (1980b) Rational design of overland flow systems. Proceedings of 1980 National Conference on Environmental Engineering, ASCE, New York.

7. Martel, C.J., T.J. Jenkins, C.J. Diener and P.L. Butler (1982) Development of a rational design procedure for overland flow. USA Cold Regions Research and Engineering Laboratory, CRREL Report 82-2.

8. Nakano, Y., R.A. Khalid and W.H. Patrick, Jr. (1978) Water movement in a land treatment system of wastewater by overland flow. Proceedings, International Conference on Development in Land Methods of Wastewater Treatment and Utilization, Melbourne, Australia, p. 14/1-14/21.

9. Overcash, M.A., D.M. Covil, J.W. Gilliam, C.W. Westerman and F.J. Humenik (1976) Overland flow pretreatment of wastewater. Water Resources Research Institute, Univ. of North Carolina, Report 117.

10. Palazzo, A.J., C.J. Martel and T.F. Jenkins (1980) Forage grass growth on overland flow systems. Proceedings of 1980 National Conference on Environmental Engineering, ASCE, New York.

11. Peters, R.E., C.R. Lee and D.J. Bates (1981) Field investigations of overland flow treatment of municipal lagoon effluent. U.S. Army Engineer Waterways Experiment Station, Vicksburg, Mississippi, Technical Report EL-91-9.

12. Pollock, T.E. (1979) Design and operation of overland flow systems: The Easley overland flow facility. Proceedings of Workshop on Overland Flow for Treatment of Municipal Wastewater. Continuing Engineering Education, Clemson University, Clemson, South Carolina.

13. Scott, T.M. and D.M. Fulton (1979) Removal of pollutants in the overland flow (grass filtration) system. Progress in Water Technology, vol. 11, nos. 4/5, p. 301-313.

14. Smith, R.G., E.D. Schroeder and G. Tchobanoglous (1980) Demonstration of the overland flow process for the treatment of municipal wastewater: Phase I pilot studies. Department of Civil Engineering, University of California, Davis.

15. Thomas, R.E., B. Bledsoe and K. Jackson (1976) Overland flow treatment of raw wastewater with enhanced phosphorus removal. EPA Technology Transfer Series, EPA-600/2-76-131.

16. U.S. Environmental Protection Agency, U.S. Army Corps of Engineers, U.S. Department of Agriculture, U.S. Department of Interior (1981) Process design manual for land treatment of municipal wastewater. EPA Technology Transfer Series, EPA 625/1-81-013 (COE EM 1110-1-501).

17. Walters, J.N. (1974) Phosphate and nitrate removal by a grass filtration system for final treatment of municipal wastewater. Masters Thesis, Pennsylvania State University, NTIS PB-269 171.

BIOLOGIC DOSIMETRY FOR NUCLEAR ENVIRONMENTS BY ELECTRON PARAMAGNETIC RESONANCE (EPR) METHODS (U)

*MICHAEL J. McCREERY, Ph.D., CPT
CHARLES E. SWENBERG, Ph.D.
MICHAEL J. BASSO
JAMES J. CONKLIN, M.D., Lt Col
JEN-SHU HSIEH, Ph.D.
ARMED FORCES RADIOBIOLOGY RESEARCH INSTITUTE
BETHESDA, MARYLAND 20814
US ARMY COMBAT SURVEILLANCE AND TARGET ACQUISITION LABORATORY
FT. MONMOUTH, NEW JERSEY 07703

INTRODUCTION

To date there is no standard method by which the total dose of ionizing radiation to which a person has been exposed can be reliably ascertained by the measurement of some biological parameter. We are currently investigating the possible use of electron paramagnetic resonance (EPR) spectroscopy of teeth and bones for biologic dosimetry. We propose that the radiation-absorbed dose for a person can be obtained as a function of the quantity of long-lived free radicals present in teeth and bones as measured by EPR. In this paper, we present data in support of this proposal. The data show that exposure levels having a dynamic range from 13 rads up to 10 megarads can be measured with as little as a 10-25 mg chip of tooth enamel. A further proposal is made suggesting the use of a finger for EPR dosimetry that is completely noninvasive and nondestructive.

The advent of nuclear weapons compelled the US Army to develop the capability to monitor radiation resulting from the detonation of such devices. Improvements in modern missile and artillery guidance systems, the reduction of weapon yield, and the development of enhanced neutron weapons have all combined to motivate the Army to further develop methods for the measurement of radiation-absorbed dose in battlefield or other coincident personnel. These factors have changed the relative importance of blast, thermal, and ionizing radiation effects produced by detonation of a nuclear device in producing combat casualties. For example, a nuclear weapon delivery system having a circular-error-probability (CEP) of 1 kilometer has been determined to require a 12-kiloton warhead to ensure a 50% probability of target destruction. Reduction of the CEP

to 500 meters requires only a 1.4-kiloton device to achieve the same probability of target destruction. This change also significantly alters the radiation yield at the CEP boundary where radiation doses for a 12- and 1.4-kiloton standard-fission weapon are 2,000 and 10,000 rads, respectively. Therefore, both intitial and prompt nuclear radiations become a significant factor in casualty production relative to blast and thermal effects, and are of greater concern to the battlefield commander.

Battlefield dosimetry has two recognized functions: (a) to inform the commander of troop exposure levels for timely decisions with respect to combat effectiveness, and (b) to assess casualties for triage and medical treatment. The Army qualitative Research Requirement (QRR) states in Section 14.2, a 1:

> "The current method of determining the future combat effectiveness of troops following exposure to nuclear radiation has many shortcomings. A unit is classified in one of three broad radiation status categories based on averaged dose measurements from a few tactical dosimeters. The reliability of this method is limited by the inaccuracies associated with tactical dosimetry such as the orientation and location of personnel wearing the dosimeters, and the inaccuracy of the dosimeter itself. Additionally, the variation in human response is considerable, further limiting the usefulness of this method. In short, unreliable casualty estimates and an unaffordably inefficient use of remaining personnel assets."

The need for biologic dosimetry is clear since the only sure thing that a soldier carries is his body. Development of a reliable biologic dosimeter may preclude the need for more expensive personnel dosimeters. In addition, biologic dosimetry may eliminate the uncertainties of estimating the exposure of the soldier by the dose received by the dosimeter where problems of placement, shielding, and dose distribution will always be significant. The Army QRR Section 14.2b outlines the following specific military requirements for biologic dosimetry: "(1) Identification of biochemical parameters that will provide an early, accurate indication of the degree of radiation injury sustained. (2) A clinical test giving absorbed radiation doses based on those biochemical parameters and suitable for use in field medical facilities. (3) A method for use by medical personnel in the field which quickly identifies supralethally irradiated personnel to the triage surgeon."

The nature of the radiation-induced EPR signal in calcified tissues makes it amenable for use in biologic dosimetry. This signal is associated with lattice defects or electron traps within bones and teeth by impinging radiation (1). The asymmetric singlet at 2.0023 is extremely stable in vitro. A recent study by Hennig et al. (2) showed that it decays exponentially with a time constant describable by the Arrhenius relationship:

$$\tau = \tau_o \exp(E_a/kT) \quad (1)$$

where $\tau_o = 3.8 \times 10^{-14}$ sec and the trap depth $E_a = 1.65 \pm 0.1$ eV. For storage at 15°C, Eq (1) gives the time constant, τ, as 9×10^7 years; at 37°C, $\tau = 1.7 \times 10^5$ years. Thus, in the absence of degradative chemical processes, at least for in vitro calcified tissue, the EPR signal does not decay in the time domain in which we are interested.

Bone tissue is made up of a collagenous matrix having the shape of the intact bone intercalated with mineral desposits of mainly calcium phosphate, carbonates, fluorides, hydroxides, and citrates. Bone contains mainly salts of magnesium, sodium, and smaller amounts of potassium. These minerals occur in a crystalline array, and most of them belong to the group of hydroxyapatites of the approximate composition $Ca_{10}(PO_4)_6(OH)_2$. Other divalent cations can replace Ca^{2+} in the hydroxyapatite crystal structure. Although teeth are composed of salts qualitatively the same as bone, teeth have a much higher percentage of hydroxyapatite. Tooth enamel is composed of almost exclusively crystalline hydroxyapatite with an average crystal size much larger than that found in bone (3).

In neither the collagen matrix nor the crystalline fraction of bone and teeth is the microscopic nature of the radiation-induced defects known with certainty. Degassed powdered bone exposed to ionizing radiation is known to generate two predominant paramagnetic species at room temperature: one is associated with the collagen matrix and the other is derived from the crystalline fraction of mineralized tissue. The EPR spectrum consists of a symmetric doublet with weak outer components that may belong to a quartet. The splitting between the two central lines is about 26G, characterized by g = 2.0032. These radicals are stable for months in degassed samples but disappear in 3-6 days in powdered samples and in the presence of atmospheric oxygen. The radicals generated in the collagen phase are thought to arise from direct loss of hydrogen atoms from amino acid linkages to collagen, presumably glycine and alanine (4).

In contrast to the relatively short-lived defects in the organic matrix, defects in the crystal lattice of hydroxyapatite, as already noted, are very stable and account for the majority of the EPR signals produced by ionizing radiation. The molecular origin of these signals is not altogether clear, but there is evidence to suggest the presence of hydrogen atoms, oxygen, phosphate, and carbonate radicals. The most informative study of these defects is on synthetic single crystals. For example, X-ray-irradiated single crystals of calcium hydroxyapatite at 6°K exhibit three prominent paramagnetic centers: (a) trapped hydrogen atom, (b) a trapped hole shared by OH^- and a neighboring PO_4^{3-} ion (the I center), and (c) O^- ion (the A center), with the last resulting from the extraction of hydrogen from an OH^- ion (5,6). Neither the I center nor the trapped hydrogen is stable at room temperature. Unfortunately, the presence of native defects in enamel that are not present in the synthetic hydroxyapatite can obscure the A center at g = 2.0023 in irradiated powdered human tooth at low dose.

METHODS

All spectra were gathered on a Varian E-109 spectrometer run at X-band in the absorption mode. Experiments with human and canine nails reported here were single sweeps of 16 minutes. All other spectra illustrated are the sums of multiple sweeps collected and stored on either a Varian E-900 or Nicolet 1180 computer system interfaced with the EPR spectrometer. Except for the low-temperature studies, spectra were performed at room temperature, approximately 25°C, which varied as much as ±3°C over the period of months in which the data were collected. However, within a given experiment, room temperature variation was less than ±1°C. Additional details concerning spectrometer settings and experimental conditions are given in the figure legends.

Sample Preparation

Tooth: Human teeth free of decay were obtained from a local dentist. In the experiments involving enamel and dentine, separation was accomplished by chipping off the enamel layer with bone rongeurs. Splitting the tooth in half revealed the dentine layer underneath, which was also chipped. Chips were no smaller than the size required to fit into a 3-mm EPR tube. All other tooth samples were prepared by powdering whole tooth in a brass pulverizer with a hammer blow. Experiments to compare the efficacy of the different radiation sources in producing EPR defects used samples from the same source: two human teeth that had been powdered and mixed thoroughly.

Bone: Bovine bone (femur) was obtained from a local butcher and thoroughly cleaned. Pieces were then cut from the shaft of the bone (diaphysis) and mechanically powdered. Rat bones (femurs) for the *in vivo* studies were dissected from laboratory rats (Sprague Dawley, male, approx. 250 g) that either had been irradiated with ^{60}Co gamma or were nonexposed. These bones were cleaned, cut into transverse sections, chipped, and placed randomly in the EPR tubes.

Nails: Human fingernails or canine (beagle) toenails were obtained within the laboratory. These were cut into pieces small enough to fit into a 3-mm EPR tube with random orientation.

Irradiations

All irradiations of samples were performed at AFRRI except for the blind samples, which were irradiated at Oak Ridge National Laboratory. The sources used at AFRRI were (a) 135,000-Ci ^{60}Co source, (b) Varian model V-7725 linear accelerator for electrons, (c) Phillips Industrial X-ray generator, and (d) TRIGA reactor. Most samples were placed in screw-cap test tubes, which were held in a linear array by a [illegible] holder for exposures. Dosimetry was performed in the same experimental [illegible] just before sample exposure, using TLD's or ionization chamber techniques. Doses reported are cumulative.

Low-temperature studies

The top of a human incisor was cut off (that part above the gum line), powdered, and placed in a 3-mm quartz EPR tube. The end was sealed in air with flame. The tube was inverted, and the sealed end containing the powder was placed in a styrofoam block, which had been hollowed out in the shape of a cylinder of approximately 40 mm diameter to hold liquid N_2. This setup was placed in the path of the LINAC beam 350 cm from the port. The opposite end of the quartz tube used later for measurement was shielded by lead and polystyrene bricks to stop electrons and bremsstrahlung radiation. Liquid N_2 was continuously delivered to the styrofoam container during the experiment. After exposure, the EPR tube was inverted and immediately placed in a dewar containing liquid N_2 where it was maintained until measurement. All spectra were recored at 77°K using a liquid N_2 dewar insert in the microwave cavity.

Oak Ridge Unknown Samples

A single human cuspid was powdered and divided equally into each of three test tubes. These samples were irradiated at Oak Ridge National Laboratory (ORNL) Dosimetry Applications Research Facility (DOSAR) using the Health Physics Research Reactor (HPRR) in the pulse mode. Each sample was irradiated separately at room temperature in air. Each test tube was positioned 1.5 meters above the floor and 3 meters from the HPRR by using a horizontal bar clamped to a ring stand for support. The irradiations were made for three different shielding configurations. These configurations are listed in Table I along with their known neutron/gamma flux and their corresponding reference gamma doses (7).

RESULTS AND DISCUSSION

A typical room-temperature EPR signal produced by exposure of tooth to ionizing radiation is illustrated in Fig. 1. Although there are contributions to this signal from the collagenous matrix of these tissues as noted above, the most prominent feature is the asymmetric singlet at g = 2.0023. Although the spectrum shown in Fig. 1 is that of a powdered human tooth, a similar spectrum is observed for powdered bone. The top spectrum displays the EPR signal 1 hour after irradiation, and shows that additional lines at low field relative to the main signal are easily discernible. The CO_3^- radical has been suggested for the weak absorption at low field (1), although this assignment is still uncertain. These lower field lines decay significently with time, as is evident in the bottom spectrum of Fig. 1, which was taken 26 hours postirradiation. Concomitant with this decay is an increase in the amplitude of the main signal. Allowing this same sample to sit at room temperature for 1 week produced a spectrum in which only this asymmetric signal was present. Longer times (measured up to 1 month) postirradiation produced no further changes in the stable signal.

In order to elucidate the decay of the different free radical species produced in calcified tissues by ionizing radiation, we irradiated and examined

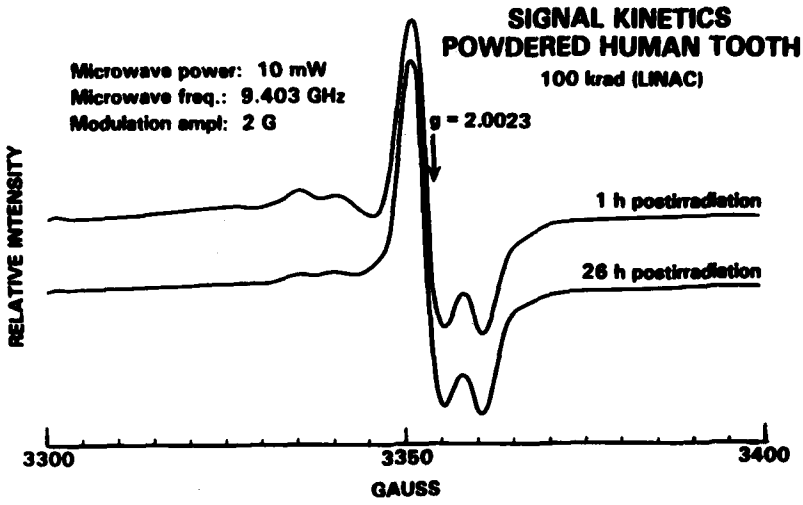


FIGURE 1. Room-temperature annealing of low-field lines with time after electron irradiation. Radiation: 14.5-MeV electrons at 15.9 rads/pulse. Experimental parameters: scan time, 10 min; scans, 10; time constant, 0.064 sec.

samples at 77°K. The experimental protocol followed is outlined above. The bold trace in Fig. 2 shows that the spectrum is considerably more complex than the one obtained for room-temperature exposure. The main signal observed at room temperature at g = 2.0023 is also present at 77°K. In addition, a doublet overlies this signal centered at g = 2.011. It has a hyperfine splitting of 26 gauss and is also overlapped by a singlet thought to derive from glycine and alanine radicals in the collagen matrix. Not shown in this spectrum is another doublet centered about the main signal and having a hyperfine splitting of 504 gauss. These lines are very weak and, as noted in the introduction, are due to atomic hydrogen (1).

These spectra were recorded at 0.2 mW to prevent saturation. The spectrum, indicated by the dotted line in Fig. 2, was recorded at 77°K after the sample was allowed to sit at room temperature for 5 hours. It is evident that the doublet at g = 2.011 has annealed considerably, with the singlet from the collagen decaying at a slower rate. Although exact quantitation has not yet been performed, it appears likely that at least a fraction of the free radical species represented by these signals have undergone disproportionation to the signal at g = 2.0023. The results of allowing the sample to anneal at room temperature for 5 days postirradiation are indicated by the lighter solid line in Fig. 2. This spectrum demonstrates further decay of the less stable species, and is very much like the signal of those samples both irradiated and measured at room temperature. The doublet representing atomic hydrogen was no longer observable under these conditions.

Although the in vitro stability of the radiation-induced signal at g = 2.0023 is well documented, its rate of decay in a functioning animal is not clear. In vivo

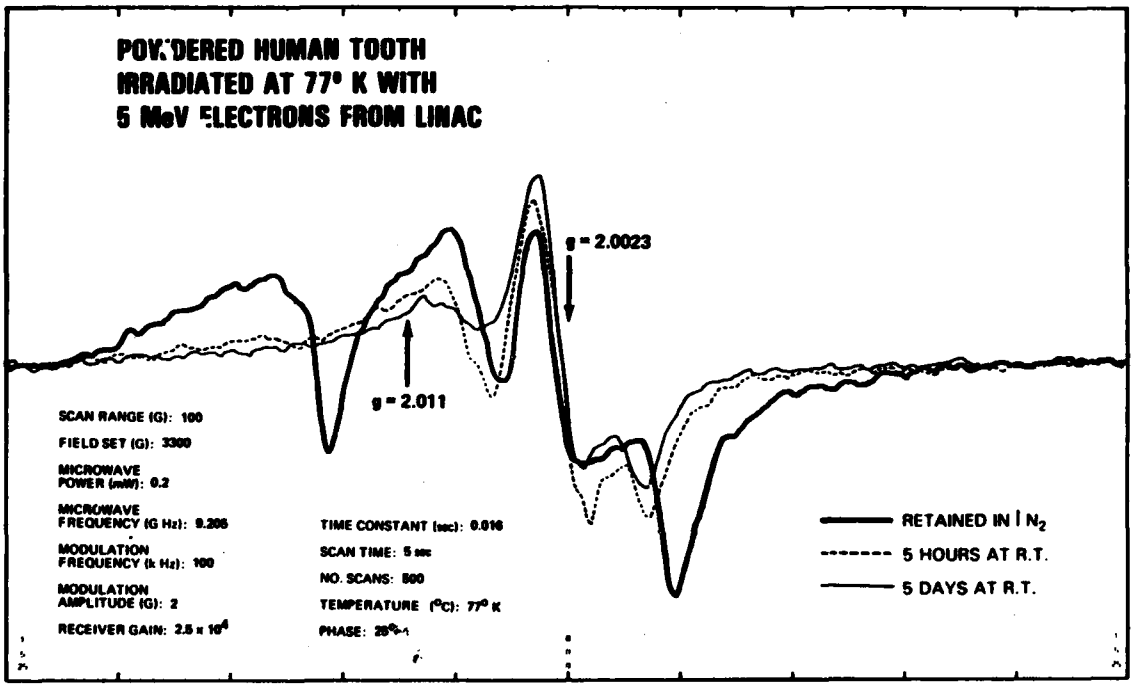


FIGURE 2. Electron paramagnetic spectrum of powdered human tooth irradiated and recorded at liquid nitrogen temperature. Experimental parameters are as indicated.

mineral deposits in bone are in a dynamic state, exchanging rapidly with ions within the extracellular fluids. This dynamic property of bone introduces the possibility of in vivo annealing of the radiation-induced paramagnatic centers. In fact, Slager, Zucker and Reilly (8) reported a disappearance of the EPR signal with the ingrowth of new tissue from the host in bone-graft studies in dog within 35 days. In contrast to these results, Koberle et al. (9) reported no loss of signal amplitude in the vertebrae of rat tails 21 days postirradiation. Our preliminary experiments on gamma-irradiated bone implanted into the peritoneal cavity of littermate rats support the notion of decreased stablity of paramagnetic radicals in functioning animals compared to the in vitro stability. However, loss of signal after 1 month appears to be less than 20%. This annealing is probably due to the fast exchange caused by the exposure of the vast surface area of bone to extracellular fluid. However, in spite of the dynamic state of bones and teeth, their composition normally remains constant provided the dietary intake of salts and vitamins is sufficient.

The signal intensity is extremely dependent on the precentage of hydroxyapatite per unit weight composing the sample as well as crystal sizes of the hydroxyapatite. This is demonstrated in Fig. 3, where the signal in tooth enamel is compared with that of its underlying dentine. The signal is approximately an order of magnitude greater, which reflects the higher hydroxyapatite content and greater crystallinity.

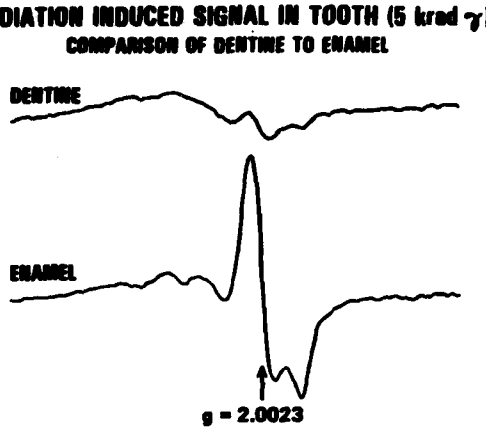


FIGURE 3. Comparison of EPR signal intensity of chipped dentine and enamel from human bicuspid. Spectra not corrected for sample weight. Experimental parameters: range, 100 gauss; power, 10 mW; modulation amplitude, 2 gauss at 100 KHz; scan time, 5 sec; number of scans, 150; microwave frequency, 9.520 GHz; field set, 3390 gauss.

Different parts of the same bone also vary in their dose-response sensitivities (unpublished results). As might be expected, Ostrowski (1) showed that the degree of crystallinity in humerous was 1.6 times higher in the diaphysis than the epiphysis. Hydroxyapatite crystals are about three times larger in enamel than in compact bone (10).

The stable radiation-induced signal has a wide dynamic range. Numerous workers have shown a linear relationship between dose and signal intensity from several hundred rads up to the megarads region. Ikeya and Miki (11) have suggested signal amplitude saturation at doses between 10^6 and 10^7 rads. Our studies have yielded similar results. Moreover, our data show that doses into the tens of rads region are discernible.

Fig. 4 illustrates a powdered whole-tooth sample irradiated with increasing cumulative doses of accelerated electrons. Doses of 10 krads and above yield a signal that is easily quantitated. Examination of the spectrum at lower doses gives evidence of a much broader paramagnetism, which overlaps the sharper radiation-induced signal. This broad singlet is native to these tissues since it is present in nonirradiated samples as shown in the trace labeled "control" in this figure. It is not the result of mechanically created defects since it also exists in undisturbed whole bone. The native signal represents a major limitation in detecting doses lower than 10 rads. Panepucci (12) suggested that the native signal could be reduced in amplitude by heating, but this method has not been successful in our laboratory. Curve analysis and simulation of the experimental spectra by summation of lorentzian lines may prove useful in the analysis of lower doses. When plotted, the data in Fig. 4 exhibit linear characteristics down to doses of about 100 rads. Doses below this appear to yield a more hyperbolic relationship, as shown in Fig. 5. This is possibly due to a radical species that saturates at low exposures.

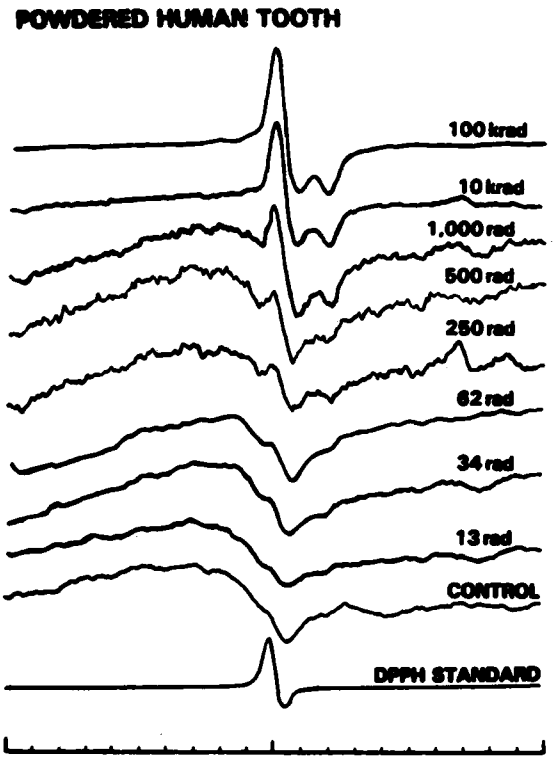


FIGURE 4. Room-temperature EPR signal as function of total dose for powdered human cuspid. Radiation: linear accelerator; 14.5-MeV electrons from water scatter at 15.9 rads/pulse. Experimental parameters: scan range, 100 gauss; modulation, 2 gauss; time constant, 0.064 sec; modulation frequency, 100 kHz; power, 10 mW; scan time, 30 sec. Receive gain and number of scans, variable.

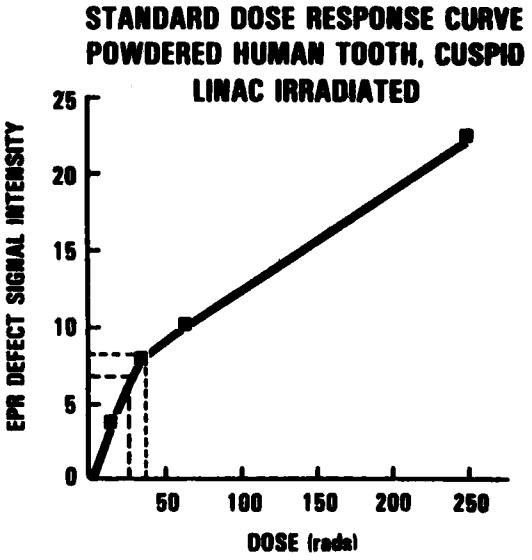


FIGURE 5. Low dose-response dependence of radiation-induced EPR signal intensity for electron-irradiated powdered human tooth. Signal intensity corrected for sample weight and background substracted. See Fig. 4 legend for experimental parameters. Dotted lines discussed in text.

If biologic dosimetry requires removal of a physiological sample from a person, then consideration should be given to the accessibility of the sample. To this end, human fingernails and dog toenails were investigated for radiation-induced changes in their EPR spectra. Fig. 6 shows the EPR results of dog toenails exposed to ^{60}Co gamma. This tissue, composed of cornified epithelial cells, does not exhibit the asymmetric singlet as in bone. Rather, there is a multiplet, the relative peak heights of which change with radiation dose. The ratios of the high-field upward peaks of control, 25 krads and 50 krads, are 3.2, 5.0, and 6.1, respectively. Although this function is dose dependent, its sensitivity permits analysis only in the krads range. A similar conclusion was reported by Brady (13).

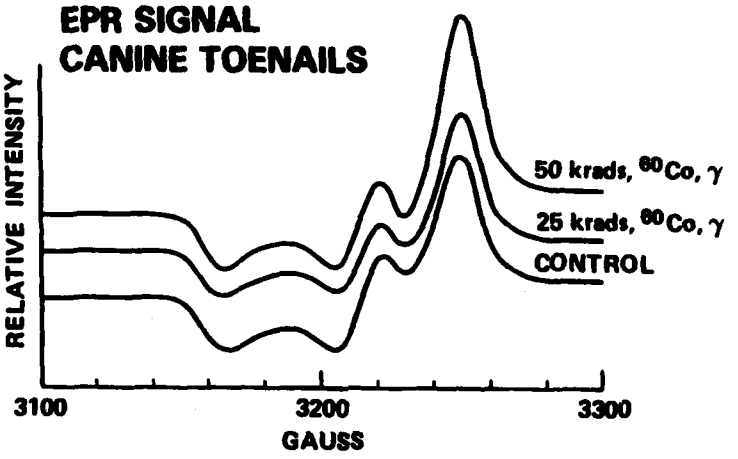


FIGURE 6. Room-temperature EPR signal for ^{60}Co-gamma-irradiated canine toenails. Experimental parameters: microwave frequency, 9.115 GHz; field set, 3000 gauss; modulation amplitude, 10 gauss; modulation frequency, 100 kHz; power, 100 mW; time constant, 0.5 sec; scan, 16 min; gain variable.

Ideally, a dosimetric method not only would yield dose information but also would discriminate among various types of ionizing radiation. We investigated this possibility using the EPR method by dividing two powdered teeth into equal samples, which were then exposed to different sources of radiation. The EPR signature appeared to be identical for the gamma, accelerated electrons, and X rays, except for the signal intensity response with dose. These data are displayed in Fig. 7. The slopes of the lines for gamma, electrons, and X rays are 0.0094, 0.024, and 0.040, respectively.

In this set of experiments, the efficacy of neutrons in producing trapped defects in tooth was also investigated. The above samples were exposed to a TRIGA spectrum of radiation shielded by a 6"-lead barrier to produce a neutron/gamma ratio of 21.6 at total doses the same as those in Fig. 7. The asymmetric singlet characteristic of the other forms of ionizing radiation was not clearly apparent until the total dose was greater than 2000 rads. This peak is most likely due to the gamma dose component. In this region of the spectrum, neutrons

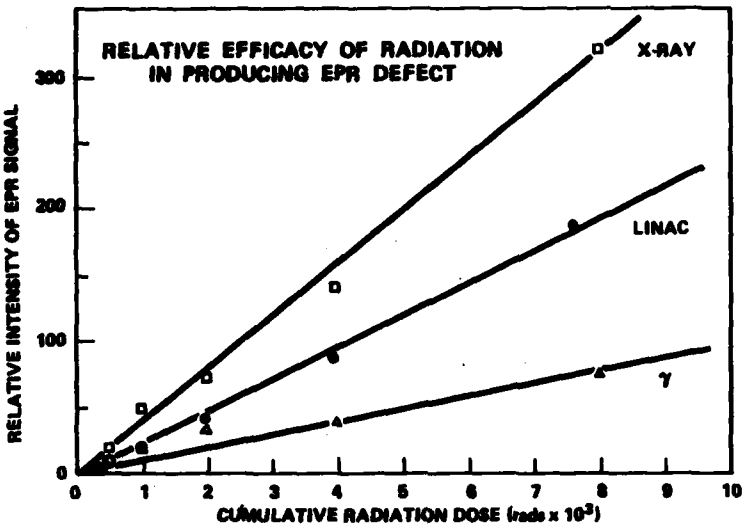


FIGURE 7. Room-temperature EPR signal height, corrected for sample weight, as function of dose for whole powdered human tooth (incisor). Radiation: X rays, 54 KeV at 1832 rads/min; LINAC, 14.5-MeV electrons from water scatter at 11 rads/pulse; gamma from ^{60}Co source. Experimental parameters: field sweep, 100 gauss; modulation amplitude, 2 gauss at 100 kHz; time constant, 0.016 sec; power, 10 mW; 500 scans at 5 sec/scan; gain normalized. Solid lines are least squares fits of data to $y = mx$.

seemed to only slightly increase the amplitude and linewidth of the native signal, but a dose dependency was not observed. However, a very broad signal (~200 gauss) centered around 2550 gauss was unique to the neutron-irradiated tooth spectrum. Further investigation is needed to evaluate this method for neutron dosimetry. Interestingly, these tissues are radiation activated when they are exposed to neutrons. This is not surprising, considering the traces of heavy metals present in bones and teeth. This observation also requires additional study to explore its use in biologic neutron dosimetry.

In order to access the possibility of using the radiation-induced EPR signal intensity as a dosimeter, a preliminary experiment was performed by irradiating samples of powdered human tooth and inferring the dose received from the measured EPR signal intensity and standard dose-response curves.

The unknown gamma dose delivered to all samples was determined in the same fashion as the following procedure for sample #2. Illustrated in Fig. 8 is the EPR signal measured for sample #2; this corresponds to the 12-cm lucite shielding configuration. Because of the difference in sample mass, the measured defect signal at $g = 2.0023$ was corrected by reference to a sample of known mass. This was done by scaling the radiation-induced signal to the native signal, the latter signal taken to be measured at approximately 10 gauss to the low-field side of the radiation-induced signal. This procedure gives the normalized signal intensity

$$hc = h\,S_0/S'$$

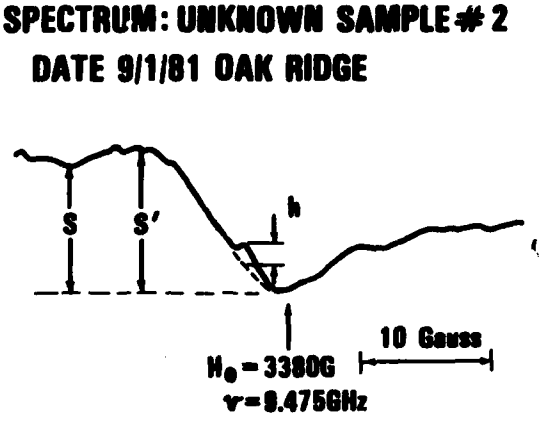


FIGURE 8. Room-temperature EPR spectrum of whole powdered human tooth. Radiation source: HPRR Oak Ridge National Laboraotry, neutron/gamma = 1.2 ± 0.1. Experimental parameters: time constant, 0.064 sec; power, 10 mW; modulation amplitude, 2 gauss at 100 kHz; 150 scans at 30 sec/scan. Dotted line denotes background of nonirradiated sample; h, EPR signal intensity; S and S', two choices for background EPR intensity. S_0 is native EPR intensity of 34-rad standard (see Fig. 5).

where h is the actual measured signal height, S_0 is the background height of the endogenous signal measured for the 34-rad sample of the standard dose-response curve, and S' is the background height for the unknown sample, as shown in Fig. 8.

From the hc value and the standard dose-response curve given in Fig. 5, the dose the sample received was inferred. The two dotted lines in Fig. 5 correspond to two choices for S', the background height of the native signal for sample #2.

The calculated doses determined by this procedure for the three samples are given in Table I.

TABLE I. Comparison of reference gamma dose by Oak Ridge Laboratory with dose determined by EPR method

Radiation Configurations	N/γ Ratio	Reference γ-Dose (Rad)	EPR Dose (Rad)
Unshielded	6.2 ± 0.9	45	23
Shielded by 12 cm lucite	1.2 ± 0.1	38	35
Shielded by 20 cm concrete	2.2 ± 0.8	18	30

The results illustrate favorable agreement with the known gamma doses delivered by the HPRR (7). The neutron dose was not estimated in these samples since, as discussed above, their effects for small neutron doses do not greatly perturb the gamma EPR signal.

CONCLUSIONS

A number of stable and unstable free radicals are produced by exposure of calcified tissues to ionizing radiation. Identification of the exact molecular nature of these radicals and their kinetic annealing properties awaits further investigation. However, the high stability and the dose-response characteristics of the signal at g = 2.0023 make it very promising for use in biologic dosimetry.

A 10-25 mg sample of enamel is enough to determine absorbed dose from 10-10^7 rads of gamma, X-ray, and accelerated electron exposures. This sample, which is no more than a chip from the ridge of a human tooth, does not expose the living portion of the tooth and so is relatively noninvasive. Investigations to establish a standard method for this procedure and to evaluate neutron dose is in progress.

Although the procedures outlined here are relatively noninvasive, this technique affords other approaches that are even less invasive. Unlike optical methods, the sample for EPR analysis does not have to be transparent. It is not even necessary to suspend the sample as a fine powder. As long as the bone or tooth sample is within the dimensional limits of the tuned EPR cavity, it can be analyzed intact. This fact makes in vivo analysis feasible. We are currently making attempts to enlarge the EPR cavity with modification of the corresponding electronic components so that a finger can be analyzed. If this approach is successful, the expedient evaluation of large numbers of casualties might be possible by a method that is noninvasive and nondestructive.

REFERENCES

1. K. Ostrowski, A. D. Gochawskz, and W. Stachowicz. In Free Radicals in Biology, Vol. IV. W. A. Pryor, ed. Academic Press, New York, 1980, p. 321.

2. G. J. Hennig, W. Herr, E. Weber, and N. I. Xirotiris. Nature 292: 533, 1981.

3. A. White, P. Handler, and E. L. Smith. Principles of Biochemistry. McGraw-Hill Book Co., New York, 1968.

4. R. A. Patten and W. Gordy. Radiat. Res. 22: 29, 1964.

5. D. M. Close, M. Mengeot, and O. R. Gilliam. J. Chem. Phys. 74: 5497, 1981.

6. M. Mengeot, R. H. Bartram, and O. R. Gilliam. Phys. Rev. B11: 4110, 1975.

7. C. S. Sims and H. W. Dickson. Health Physics 37: 687, 1979.

8. U. T. Slager, M. J. Zucker, and E. B. Reilly. Radiat. Res. 22: 556, 1964.

9. C. Koberle, C. Terrile, H. C. Panepucci, and S. Mascarenhas. An. Acad. Brasil. Cienc 45: 158, 1973.

10. K. Ostrowski and A. Dziedzic-Goclawska. In The Biochemistry and Physiology of Bone, Chapter 7. G. M. Bourne, ed. Academic Press, New York, 1976, p. 303.

11. M. Ikeya and T. Miki. Science 207: 977, 1980.

12. H. C. Panepucci, quoted in: W. Stachowicz, J. Michalik, A. Dziedzic-Goclawsk, and K. Ostrowski. Nuckeonika 19: 767, 1974.

13. J. M. Brady, N. O. Aarestod, and H. M. Swartz. Health Physics 15: 43, 1968.

THE EFFECTS OF NERVE AGENTS ON BEHAVIORAL PERFORMANCE AND THEIR MODIFICATION WITH ANTIDOTES AND ANTIDOTE COMBINATIONS

JOHN H. MCDONOUGH, JR., PH.D., MAJ, MSC
U.S. ARMY MEDICAL RESEARCH INSTITUTE OF CHEMICAL DEFENSE
ABERDEEN PROVING GROUND, MARYLAND 21010

The nerve agents are highly toxic organophosphate (OP) anticholinesterase compounds. It is generally agreed that they exert their primary toxic actions by inhibition of the cholinesterase family of enzymes which results in build-up of the neurotransmitter acetylcholine in central nervous system and at peripheral effector muscles and organs (18). This results in cholinergic overstimulation which in turn results in their toxic symptoms: bronchospasm and associated hypersecretions of the respiratory tract; vasodilation and consequent fall in blood pressure contributing to circulatory failure; weakness of the diaphragm and intercostal muscles; violent convulsions of central origin which interfere with ventilatory efficiency and causing an increased oxygen demand on an already overtaxed respiratory system; a failure of the central respiratory centers (3). Drug therapy for poisoning by nerve agents mainly consists of liberal administration of cholinergic antagonists, most notably atropine, to block the symptoms of poisoning and prompt administration of oximes, such as 2 PAM-Cl, to reactivate inhibited cholinesterase enzyme before the enzyme-inhibitor complex "ages" and becomes refractory to reactivation (10). The nerve agent that poses the greatest problem for medical treatment is soman (pinacolyl methylphosphonofluoridate) or GD. The enzyme-inhibitor complex formed by this agent rapidly (<5 min) "ages" and becomes refractive to oxime reactivation (4).

For many years there have been widescale research efforts to develop more efficatious drugs for the treatment of nerve agent poisoning. These efforts have relied almost entirely on toxicologial techniques to evaluate new treatment drugs or drug combinations. This approach relied on a single biological variable, usually survival, as a measure of effectiveness. In these procedures a test drug was appraised of its ability to decrease the per cent lethality in animals exposed to an empirically chosen standard agent dose or to increase the dose required to produce a fixed percentage of fatalities in an experimental group. These criteria are only applicable for acute toxicity. Sublethal doses of nerve agents produce effects on judgement and cognition at low doses and higher levels produce

long-term physical incapacitation even if the casualty is administered standard drug treatment (17). Either of these effects could seriously compromise military performance. As a result, it is now recognized that behavioral testing procedures can add a significant new dimension to the evaluation of new nerve agent treatment compounds. Early in the development of improved nerve agent treatment compounds test drugs should be evaluated in animals for their ability to reverse nerve agent induced changes in behavior and physical performance as well as ability to counteract physiological toxicity. Some initial success in this direction by this laboratory has been reported elsewhere (7,8). The present studies describe the selection and validation of one such test which responds to the detrimental effects of the nerve agent soman on animal performance and then demonstrates the sensitivity of the test to detect and analyze the therapeutic capabilities of a drug mixture.

A review of the experimental literature reveals that both carbamates or OP compounds less toxic than the nerve agents produce depressions in avoidance performance of rats trained to perform active avoidance tasks (2,12,13,14,15,16). The two-way shuttle avoidance task in particular has been used to study the behavioral toxicology of several OP compounds (i.e. paraoxon, diisopropyl-fluorophosphate (DFP), armin). The changes in normal avoidance performance with low doses have been considered behavioral indicators of central cholinergic dysfunction (2). When higher doses of these compounds were administered observable toxic symptoms resulted and performance decrements were attributed to physical incapacitation rather than cognitive disruption (2,16). In addition to examining the effects of OP compounds *per se* on learned shuttle avoidance performance, Rosic and colleagues (14,16) have studied how these performance decrements can be modified by oxime cholinesterase reactivators. In general, they reported that the oximes 2 PAM-Cl, Toxogonin, or TMB-4 were not capable of reversing the depression in avoidance behavior produced by a number of OP compounds, but do lead to a faster partial recovery of performance. There have been no studies reported on the ability of centrally acting anticholinergic drugs such as atropine, either alone or in combination with oximes, to antagonize this depression of avoidance performance. Neither have there been any reports as to the capability of anticholinergics and/or oximes to antagonize the physical incapacitation produced by higher doses of OP compounds.

GENERAL METHODS

Subjects: Adult male Sprague-Dawley X Wistar rats, at least 120 days old and weighing 300-500 g served as subjects(*). Animals were housed two to a cage with food and water available *ad libitum* in a well ventilated room with temperature maintained at 68° and a controlled 0600-1800 light 1800-0600 dark cycle.

Apparatus: Two BRS-LVE model 146-04 rat toggle floor shuttle cages, with no center dividers, served as testing apparatus. These rectangular cages measure 10-3/4" H x 8-1/2" W x 19" L, with a grid floor to present shock,

and have a 28V cue light affixed to each end wall and a speaker centered in the lid. They were housed in sound attenuating chambers equipped with exhaust fans for continuous air exchange. Shock source was two Grason Stadler E1064GS shock generators. Programming and data recording was accomplished using BRS-LVE solid state logic modules and counters.

Procedure: The rats were randomly assigned to one of the two test apparatus. Training consisted of a single test session per day, five days a week. A session consisted of 50 shock avoidance trials, presented on a variable interval 20 sec schedule. A trial began with the illumination of the stimulus lamp in the side of the cage the animal was in at the start of the trial, and the sounding of an $\simeq$ 70db warning tone through the speaker (compound light + tone (conditioning stimulus). These stimuli stayed on for 4 sec. If the animal did not shuttle to the opposite end of the cage within 4 sec (an avoidance response) the CS stimuli were continued and an 0.6 ma scrambled shock applied to the grid floor. At this time the subject could escape the shock (escape response) by shuttling to the opposite end. If the subject failed to respond the shock was only applied for a maximum duration of 4 sec. A response anytime during a trial turned off all stimuli (i.e. light, tone, and shock if on). Thus, on a given trial there were three possible outcomes: an avoidance of the shock, an escape once the shock was turned on, or a failure to escape the shock. In addition, a record was kept of the total number of crossings. Criterion performance was defined as at least two consecutive sessions with 90% or greater avoidance responses (i.e. avoids shock on 45 of 50 trials). Once criterion performance was reached drug tests began.

Drugs: Soman was diluted in cold saline to a concentration of 50 µg/ml soman. Saline was the control vehicle for soman. All soman injections were administered i.v. in the tail vein while the rat was briefly restrained in a surgical towel 20 min prior to behavioral testing. Therapy drugs were the anticholinergics atropine SO_4 and benactyzine HCl and the oxime TMB-4. These compounds comprise the nerve agent antidote TAB. These compounds, either alone or as drug mixtures were administered i.m. in the thigh of the animal's hind leg. When only therapy drugs were given (Experiments II A+B) it was 20 min before testing. Therapy drugs were given 30 sec after soman challenge (Experiments III + IV). Therapy injection volumes were 1cc/kg; the vehicle was distilled water for injection, USP, with 0.5% methyl paraben and 0.05% propyl paraben (w/v) added for stabilization, and adjusted to pH $\simeq$ 3.0 with HCl. Therapy drugs were prepared in lots and stored under refrigeration.

Data Analysis: All three dependent measures, avoidances, escapes, and escape failures were expressed as percentages. All percentages were then transformed using the following recommended (19) formula: $X = 2 \arcsin \sqrt{\%}$. These transformed scores were then subjected to analysis of variance (ANOVA) procedures with reliable results analyzed by Newman-Kuels tests.

Experiment I

The dose-effect and duration of effect for producing depressions in

avoidance performance varies considerably from one OP compound to another (2). For example, 0.5 Xs the lethal dose 50 (LD50 - the dose of a compound that produces 50% fatalities) of paraoxon depresses performance by about 50% at the time of maximal intoxication, and the effect lasts less than 12 hr. A less toxic dose of DFP (0.25 X LD50) however produces a greater depression in avoidance performance which can persist for up to 24 hr. The purpose of the first experiment was to characterize the dose-effect relationship for decrements in two-way shuttle avoidance produced by the nerve agent soman. To test the effect of various doses on soman on learned avoidance performance the subjects were randomly assigned to one of six treatment groups: saline, 10, 20, 30, 40 or 50 μg/kg soman. Since the higher doses of soman were lethal within 48 hr in some of the subjects additional animals were assigned to these groups until all group sizes N=10 subjects which survived 48 hr after soman administration. The doses of soman were chosen to range from no effect to lethal effects just beyond the 48 hr LD50.

The performance variables were analyzed by a 6 (dose group: saline vs 10 vs 20 vs 30 vs 40 vs 50 μg/kg soman) X 5 (test sessions: criterion session 1 vs criterion session 2 vs treatment session vs post-treatment session 1 vs posttreatment session 2) ANOVA with repeated measures on test sessions. This design allows statistical comparisons within a group across test sessions (i.e. 1 Test session/day: 2 "criterion" sessions, and sessions 20 min, 24 hr and 48 hr after soman challenge), and between groups (soman dose) on a given test session.

RESULTS

The most consistently elicited signs in the order of increasing severity were: licking and chewing, hypersalivation, tremor, and convulsions. Symptoms were dose dependent and latency to symptoms was 1 min. In all cases where subjects exhibited toxic symptoms there was some degree of obvious performance decrement upon testing. In some of the animals that showed no toxic signs there were also performance deficits, while in others there were no changes from criterion performance. The LD50 for soman under the conditions of this experiment was calculated to be 41.37 μg/kg iv (48.15 - 35.87 μg/kg = 95% fiducial limits) by probit analysis. These data were based on a 48 hr observation period and included all animals used in the study.

Performance Data: The performance data is shown in Figure 1. In analyzing the data it became apparent that the arcsin transformation made no allowances for maximum (100%) or minimum (0%) performance on any of the three measures. Such "ceiling" and "floor" effects on a number of performance measures became quite apparent in the data (i.e. on criterion sessions most subjects had 0% fail escape scores, and many had 100% avoidance scores. For purposes of data analysis the maximum and minimum scores allowed by the transformation (for 99.99 and 0.01% respectively) were assigned these scores: 100% = 3.08, 0% = 0.06. It was felt these scores most accurately reflected the performances observed.

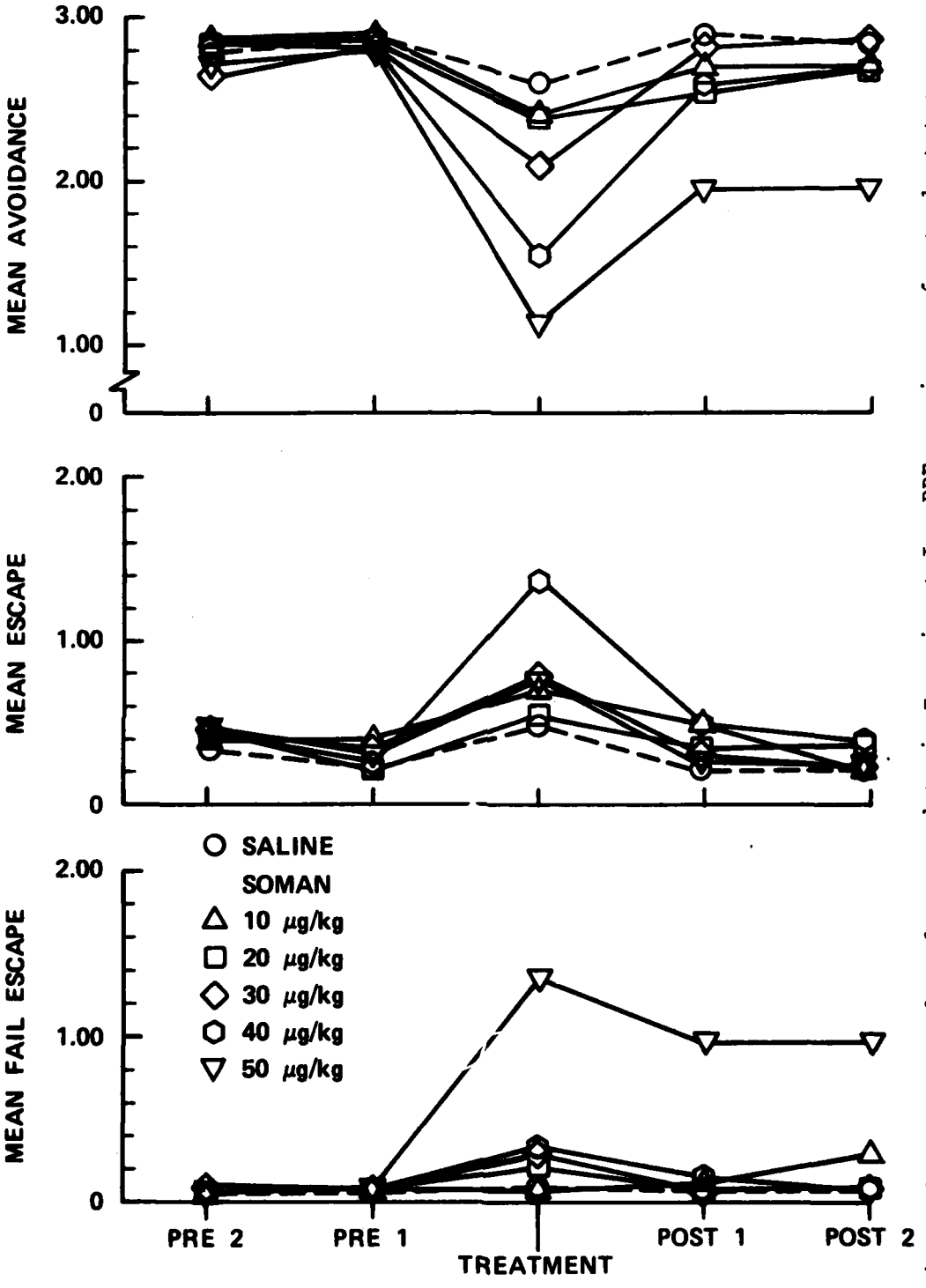


Figure 1. Summary of performance data in Experiment I. PRE sessions refer to last two criterion sessions, POST sessions were 24-and 48 hrs after treatment session. Note the pattern of normal performance during criterion sessions: high avoidance scores, low escape scores, and virtually no fail escapes.

Avoidances: Analysis of the transformed avoidance scores revealed highly reliable effects for treatment ($F(5,54) = 4.65$, $p < .01$), sessions ($F(4, 216) = 23.12$, $p < .001$), and the treatment X sessions interaction ($F(20,216) = 2.65$, $p < .01$). Only the interaction effect will be described in detail. Analysis of the treatment X sessions interaction with a Newman-Keuls test showed the avoidance performances of the 30, 40 and 50 µg/kg soman groups were significantly depressed (p's $< .05, .01, .01$ respectively) on the treatment session when compared to their own criterion performance, or the performance of any other group on any other session. Additionally, the avoidance performance of the 50 µg/kg soman group remained depressed on the two sessions following soman administration, when compared to their pretreatment performance, or the performance of any other group on those sessions (p's $< .01$).

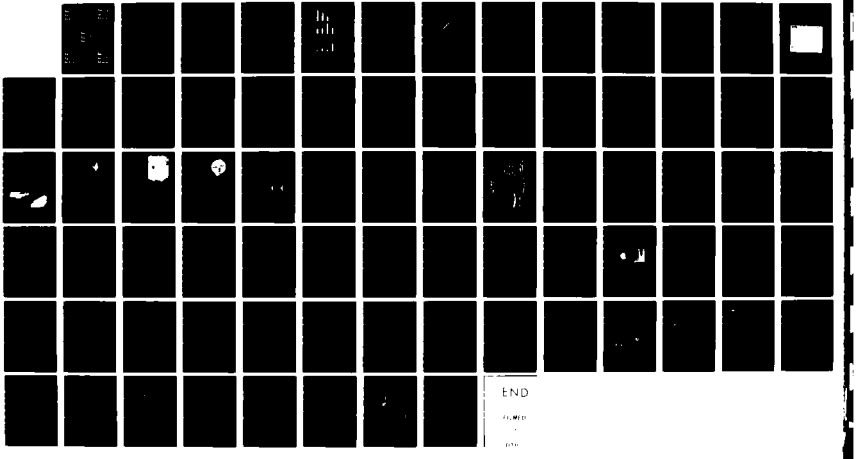
Escape: The results of the analysis of transformed escape scores showed significant effects for treatment ($F(5,54) = 2.70$, $p < .05$), sessions ($F(4, 216) = 21.84$, $p < .001$), and the treatment X sessions interaction ($F(20,216 = 1.77$, $p < .05$). Only the interaction was considered further. This analysis revealed that on the soman treatment session escape responses increased (p's $< .05$) for the 30, 40 and 50 µg/kg groups when compared to other groups or their own pretreatment criterion performance. Also, the 40 µg/kg soman treatment group had a greater increase in escape responding than any other group (p's $< .01$) on this session.

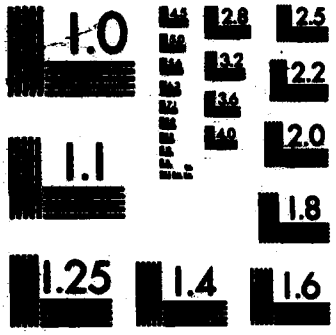
Escape Failure: The same pattern of significant results appeared from this analysis as from the other two measures, viz. reliable effects for treatments ($F(5,54) = 2.89$, $p < .01$), sessions ($F(4,216) = 8.44$, $p < .01$), and the treatment X sessions interaction ($F(20,216) = 3.22$, $p < .01$). Analysis of the interaction effect showed a large increase in the failure to escape measure by the 50 µg/kg soman group on the day of treatment, and these scores remained elevated on the two days following treatment (p's $< .01$).

DISCUSSION

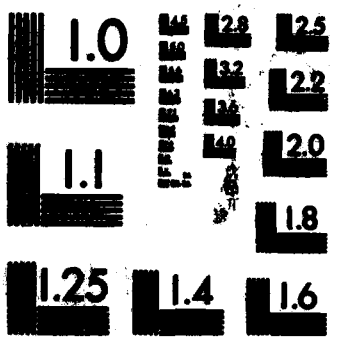
The results of this study confirm that sublethal doses of the nerve agent soman disrupt performance of a well learned shuttle avoidance task by rats. Performance decrements were significant at a dose (30 µg/kg) that was lethal to only 9% (1 subject) of the animals tested. Four subjects in this group showed no toxic signs yet displayed performance decrements. The decrement in avoidance performance showed a positive relationship to the dose of soman administered, with progressively greater decrements becoming apparent as the challenge dose was raised. The reduction in avoidance performance in the 30 and 40 µg/kg soman groups was accompanied by a corresponding increase in escape responding along with a modest, but statistically insignificant increase in failures to escape. These data, and casual observation of the animals, leads to the conclusion that while many may have been exhibiting moderate toxic symptoms they were physically capable of performing a shuttle response and the shock still retained aversive properties to motivate such a response. The

AD-A120 812 PROCEEDINGS OF THE 1982 ARMY SCIENCE CONFERENCE HELD AT 5/5
THE UNITED STATES.. (U) DEPUTY CHIEF OF STAFF FOR
RESEARCH DEVELOPMENT AND ACQUISITIO.. 18 JUN 82
UNCLASSIFIED F/G 5/2 NL

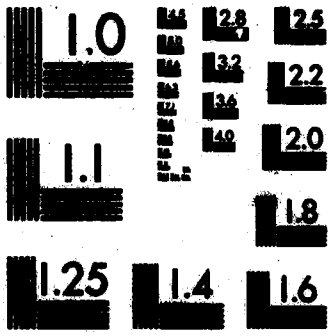




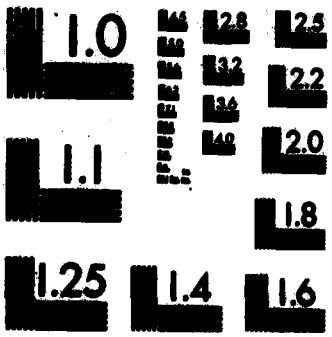
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A



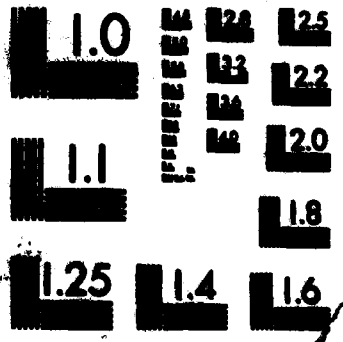
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

changes in avoidance performance at these doses is attributed to a pharmacologically induced disruption of the neural processes which mediate this behavior. At higher dose levels of soman such an explanation for the decrement in avoidance performance is not warranted. The 50 µg/kg soman group experienced the severest disruption in avoidance performance, accompanied by a moderate increase in escape responding and a dramatic increase in the failure to escape measure. This failure to escape the shock is primarily due to physical incapacitation since these animals were usually prostrate in the apparatus and displaying strong symptoms of OP intoxication. Although some symptoms of intoxication (convulsions, salivation) had subsided by the day after treatment some animals remained prostrate and incapable of escaping shock which is reflected in the elevated fail escape scores on the two post-treatment days. However at the lower dose levels avoidance performance had returned to normal 24 hrs after treatment.

The dose at which significant decrements in avoidance performance were first observed, 30 µg/kg soman, corresponds to 0.73 LD50 Soman. This indicates a very narrow range between the psychopharmacological effects as measured by changes in performance of this task and the lethal effects of this compound.

Experiment II

Before studying the capability of anticholinergic and/or oxime compounds to reverse the behavioral depression produced by soman, control experiments were required to define the effects of these compounds per se on avoidance performance. Review of the experimental literature indicated that centrally acting anticholinergics such as atropine, benactyzine, and scopolamine enhance acquisition of two-way shuttle performance, but do not disrupt well-trained performance (1). Rosic (14) reports that large doses of oximes can produce a non-specific depression of two-way shuttle avoidance performance. Since mixtures of anticholinergics and oxime were to be tested for antagonism of soman induced performance decrements, Experiment IIA was conducted to investigate the effect of such mixtures on avoidance performance. When it became apparent that the highest doses of the mixture of produced decrements, Experiment IIB studied the effects of the individual drug components. The drug mixture termed TAB consisted of the anticholinergic drugs atropine SO_4 and benactyzine HCl, both of which have been shown to antagonize soman lethality in animals studies (11), and the oxime TMB-4, which has been studied for antagonism of avoidance depression produced by paraoxon and armin (14,16). In experiment IIA there were four groups: vehicle, TAB-1, TAB-2, TAB-3 (N=10/group). The TAB doses were spaced 0.5 log units apart. The dosages of drug found in the three TAB mixtures is shown in Table 1.

TABLE 1

Drug dosages of the therapeutic TAB mixtures used in experiment IIA. The doses of the individual components in TAB-3 were used in experiment IIB. Drugs were prepared in the vehicle to yield concentrations of 1 cc/kg.

TABLE 1 con't

Component		TAB-1 (mg/kg)	TAB-2 (mg/kg)	TAB-3(mg/kg)
T	TMB-4	2.820	9.00	28.20
A	atropine SO_4	.282	0.90	2.82
B	benactyzine HCl	.175	0.55	1.75

In experiment IIB there were three dosage groups (N=10/group) each of which received one of the components of the highest dose level of TAB. The percentage scores for avoidance, escape, and fail escape performance were transformed as in Experiment I and subjected to single-factor repeated measures ANOVAs using scores from the two criterion sessions, treatment session, and the two post-treatment sessions.

RESULTS

Experiment IIA. There were no reliable effects on any measure of performance with the vehicle, TAB-1, or TAB-2. TAB-3 however produced a reliable decrement in avoidance behavior ($F(4,48) = 11.03$, $p < .01$) accompanied by a corresponding increase in escape behavior ($F(4,48) = 9.92$, $p < .01$) on day of treatment. There was no reliable effect on fail escape performance.

Experiment IIB. The high doses of atropine or benactyzine did not effect any measure of performance. The 28.20 mg/kg dose of TMB-4, however, reproduced the effects seen with the TAB-3 dose: a significant reduction in avoidances ($F(4,36) = 21.64$, $p < .01$) and a significant increase in escapes ($F(4,36) = 21.28$, $p < .01$), with no effect on the fail escape measure.

DISCUSSION

The results of Experiment II are in agreement with the reports that anticholinergic drugs, and in this case mixtures of drugs, do not affect well trained shuttle avoidance performance. The finding that the TAB-3 mixture produced avoidance decrements can be attributed to the high dose of TMB-4 found in this mixture since this drug alone produced equivalent avoidance decrements. Similar depressant effects on avoidance performance have been reported with large (> 25 mg/kg) doses of toxogonin, an oxime similar to TMB-4 (14). For antagonism studies of the soman-produced decrements in performance the TAB-2 dosage levels of the three compounds were chosen since this mixture failed to affect any measure of avoidance performance. All further references to TAB or the individual components refer to the TAB-2 doses shown in Table 1.

Experiment III

The central question of these studies was whether compounds that are successful in treating the lethal effects of OP intoxication are also capable of reversing the decrements in avoidance performance and/or physical incapacitation. As already demonstrated in Experiment I soman produces physical incapacitation at higher doses than those doses which produce decrements in avoidance performance. Experiment II demonstrated that the therapeutic mixture TAB itself produces no effects on avoidance performance. Experiment III was designed to study the abiity of the TAB therapy

mixture to modify the decrements in performance produced by a range of soman doses.

METHOD

Subjects: Eighty adult male rats were used as subjects.

Procedure: The same apparatus and training procedures as described in Experiment I were used. When a subject met criterion performance it was randomly assigned to one of eight groups (N=10/group). Four experimental groups respectively received 30, 40, 50 or 60 µg/kg soman plus TAB therapy. The four control groups received the same doses of soman plus saline therapy. The transformed scores for avoidance, escape, and fail escape were analyzed separately using a 4 (dose: 30 vs 40 vs 50 vs 60 µg/kg soman) X 2 (therapy: TAB vs saline) X 2 (sessions: last criterion session vs treatment session) analysis of variance with the sessions factor treated as a repeated measure. In cases where an animal died as a result of soman treatment prior to behavioral testing, the subject was assigned the following performance scores for data analysis: avoidance = 0%, escape = 0%, fail escape = 100%. It was reasoned these scores would most accurately reflect the subject's performance had it survived to be tested.

RESULTS

Probit analysis of the data of the soman + saline groups gave an LD50 = 44.9 µg/kg (53.2 - 38.2 µg/kg = 95% fiducial limits). This LD50 falls well within the 95% fiducial limits obtained in Experiment I. A similar LD50 could not be established for the soman + TAB groups since mortality was only observed in the 60 µg/kg soman + TAB group. In most cases the TAB therapy was successful in suppressing the convulsant actions of soman within approximately 5 min after being administered.

Performance Data: Figure 2 displays the performance data of all groups on the day of treatment.

Avoidance: Results of the analysis of transformed avoidance scores showed reliable effects for dose ($F(3,72) = 32.45$, $p < .01$), therapy ($F(1, 72) = 12.12$, $p < .01$), sessions ($F(1,72) = 499.87$, $p < .01$), the dose X session interaction ($F(3,72) = 22.73$, $p < .01$), and the therapy X session interaction ($F(1,72) = 6.86$, $p < .01$). The dose X therapy interaction ($F(3, 72) = 0.76$) and the dose X therapy X session interaction ($F(3,72) = 1.35$) were not significant. The results show that soman depressed avoidance performance ($p < .01$) in a dose related fashion. Although there were no reliable differences in performance at a given dose between the groups given either saline or TAB [illegible] all doses were considered together TAB treated subjects had signific[illegible]etter performance than those treated with saline ($p < .01$).

Escape: All factors of the analysis of escape performance showed reliable effects (dose: $F(3,72) = 3.56$; $p < .05$; therapy: $F(1,72) = 171.6$, $p < .01$; sessions: $F(1,72) = 59.52$, $p < .01$; dose X therapy: $F(3,72) = 5.63$; $p < .01$; dose X session: $F(3,72) = 3.73$, $p < .05$; therapy X session: $F(1,72) = 17.56$, $p < .01$; dose X therapy X session: $F(3,72) = 4.24$, p .01). Only the dose X therapy X session interaction will be con-

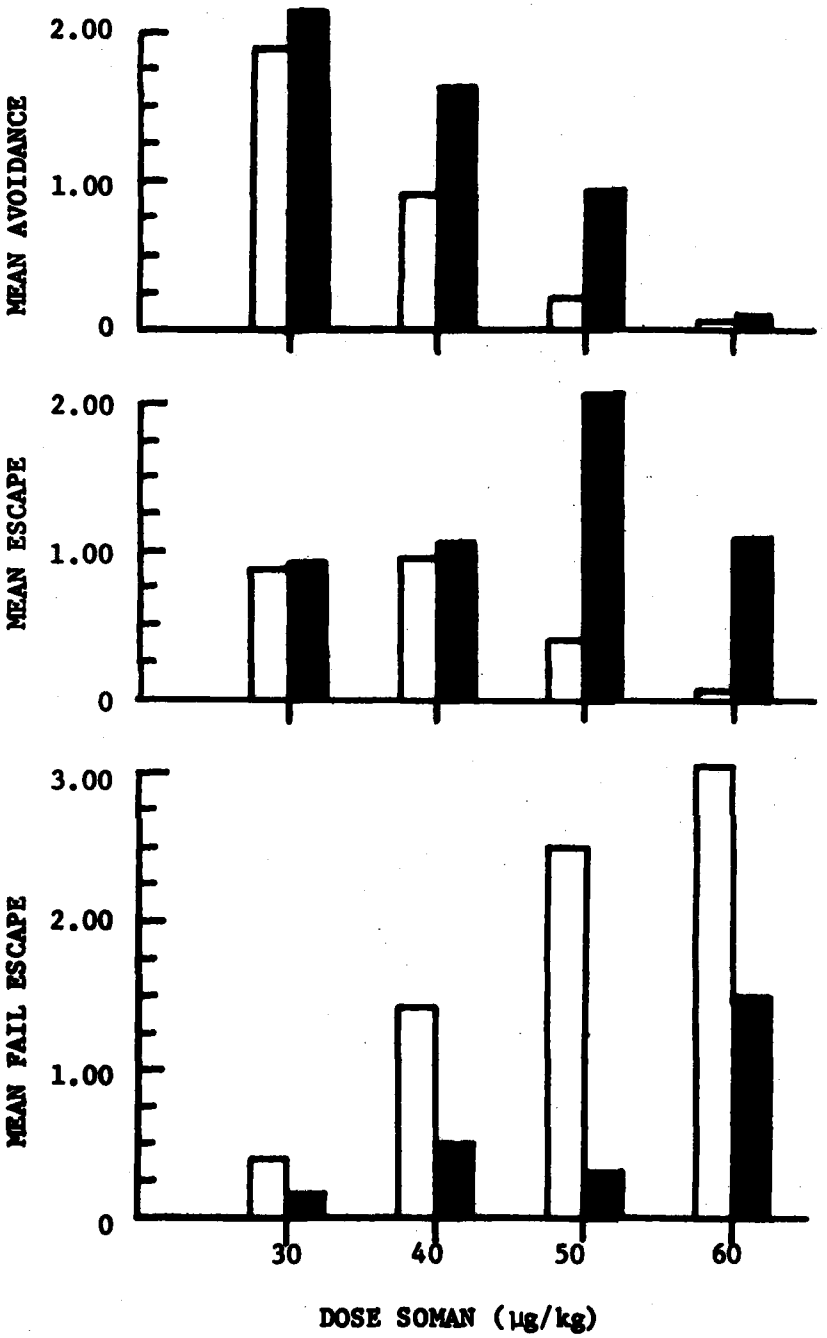


Figure 2. Summary of performance data in Experiment III. Open bars refer to subjects receiving saline therapy, shaded bars represent subjects that received TAB therapy.

sidered further. At the 30 μg/kg challenge dose neither saline nor TAB treated subjects showed any reliable increase in failure to escape compared to their control performance. As the challenge dose of soman was raised fail escape scores were progressively elevated over control values (p's < .01), with respect to one another (30 μg/kg < 40 μg/kg < 50 μg/kg < 60 μg/kg soman; p's < .01), and with respect to performance of TAB treated subjects at the same dose (40, 50 and 60 μg/kg soman; p's < .01). The fail escape scores of TAB therapy subjects in the 30, 40 and 50 μg/kg soman groups did not differ from their respective control values. The group treated with 60 μg/kg soman + TAB therapy had a higher fail escape score than those of other TAB therapy subjects (p's < .01), yet this was still reliably less than that of the 60 μg/kg soman + saline therapy group (p < .01).

DISCUSSION

The analysis of avoidance scores indicates that the TAB therapy mixture was unable to reverse the decrements in avoidance performance at any given level of soman challenge. The reliable therapy X session interaction indicates that TAB provided some marginal, yet significant, antagonism of the depression of avoidance behavior produced by soman. The TAB mixture however was quite successful in antagonizing the physical incapacitation as indexed by the fail escape measure. This protection against incapacitation parallels the ability of TAB to protect against soman lethality. Since all groups had 0% fail escape scores on the control sessions it is possible to calculate the dose of soman which yielded 50% incapacitation (ID50: i.e. 50% fail escape): soman + saline ID50 = 41.0 μg/kg; soman + TAB ID50 = 57.8 μg/kg. The relationship of the ID50 to the LD50 in untreated subjects (ID50/ LD50) shows the ID50 = .9% LD50. The ID50 of TAB treated subjects is 1.42 X ID50 of saline treated subjects.

Experiment IV

Because TAB therapy is a mixture of three drugs, it was of some interest to determine which of these three components contributes to the antagonism of both soman lethality and incapacitation. Ability to discriminate the drug or drugs primarily responsible for this effect would lend further weight to the value of using this behavioral test to screen therapy compounds. Since the TAB mixture could antagonize fully both the lethal and incapacitating effects of 50 μg/kg soman this dose was chosen as a standard challenge with the individual drugs, binary combinations of the TAB components, and TAB itself serving as the therapeutic treatments. Animals were assigned at random to one of the following therapy groups after reaching criterion performance: saline, T(TMB-4, 9.0 mg/kg), A(atropine, 0.9 mg/kg), B(benactyzine, 0.55 mg/kg), A+B, A+T, B+T, T+A+B (N=10/group). These drug(s) were administered 30 sec after 50 μg/kg soman and animals were tested 20 min later. Since the fail escape scores accurately described the degree of physical incapacitation only this analysis will be described in detail.

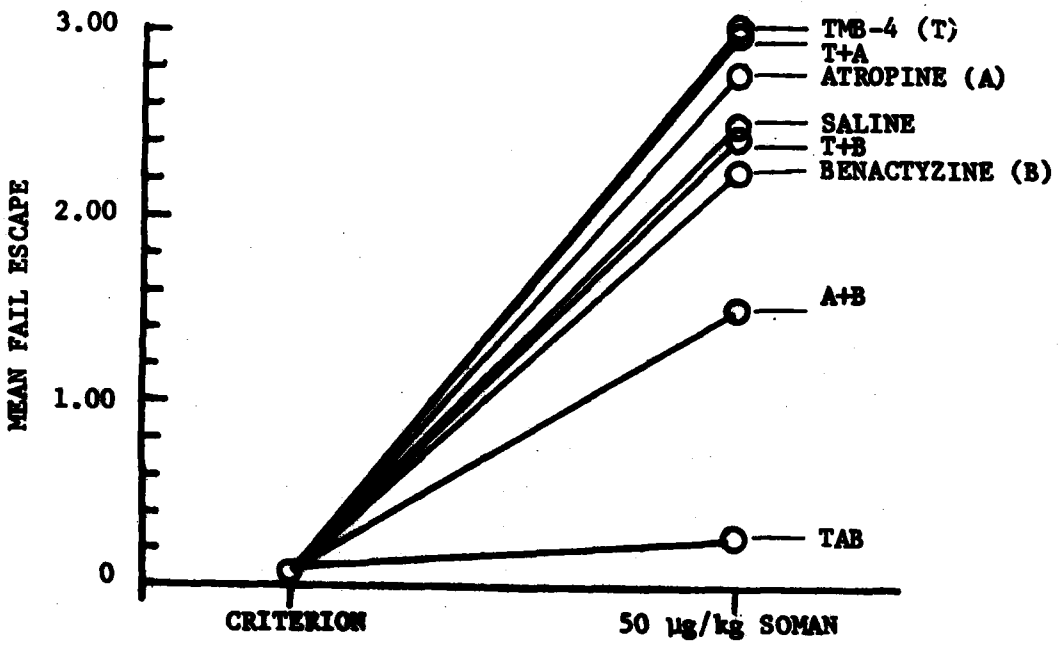


Figure 3. The differential effects of TAB and its various components in reversing the physical incapacitation produced by 50 µg/kg of soman on avoidance performance. Note that higher scores represent greater incapacitation.

RESULTS

From the probit analysis of Experiments I and III 50 µg/kg of soman should result in a mortality rate of ≃ 60-70%. The 24 hr mortality rates observed were: saline = 60%, TMB-4 = 90%, atropine = 70%, benactyzine = 30%, A+B = 40%, A+T = 90%, B+T = 40%, T+A+B = 0%.

Performance Data: Analysis of the experimental day fail escape scores showed there were highly reliable differences between therapy treatments ($F(7,72) = 13.69$, $p < .001$) (Figure 3). Both the TAB and A+B therapy groups experienced significantly less incapacitation than all other groups which did not differ from one another. In addition the TAB therapy group experienced significantly less incapacitation than the group receiving A+B therapy.

DISCUSSION

The results of this study further support and extend the results of Experiment III showing that TAB therapy is successful in fully reversing both the lethal and incapacitating effects of 50 µg/kg soman and identifies the combination of the two anticholinergic drugs as being primarily responsible for the effect. Neither atropine nor benactyzine by itself or in binary combinations with the oxime TMB-4 provided this amount of protection against lethality or incapacitation. The results also highlight the uniqueness of the triple drug combination in that the oxime

showed significant therapeutic effects when combined with both anticholinergic drugs, but not when given alone or in the binary combinations with them. The results also support the observation that potent centrally active anticholinergic drugs such as benactyzine, when combined with atropine, afford a significant enhancement in the ability to treat the toxic actions of nerve agents over either compound alone (11). It may be speculated that when this enhanced anticholinergic protection is provided, thus dampening symptoms, conditions become optimal for observing the therapeutic benefit of the oxime. This has been demonstrated with more oxime-sensitive nerve agents such as sarin and VX (5,6,11). In these cases oxime administration in and of itself does not provide protection but does enhance that provided by atropine alone, and protection is further increased by the addition of another anticholinergic.

The results of these experiments demonstrate the general utility of applying behavioral testing procedures to evaluate experimental drugs developed in treatment of nerve agent poisoning. The techniques are also relevant to studying the effects of therapy drugs alone since behavioral side-effects which decrement peformance take on critical importance in the military setting. Experiment II demonstrated that a therapy alone can decrement task performance. The drug dose levels in most self-aid antidotes represent a deliberate trade-off between therapeutic efficacy in the event of poisoning and the potential to decrement military performance if administered mistakenly in the absence of poisoning. This is especially critical since many treatment compounds are in themselves potent psychoactive drugs (9). The ideal drug treatment against nerve agent poisoning would offer significant protection against lethality, reverse both physical and behavioral agent incapacitation, yet produce few or no side effects when taken in the absence of poisoning. Recent research in this Institute using behavioral testing techniques suggests that behavioral methods will contribute to identifying such a treatment regimen (7,8).

(*)The experiments reported here were conducted according to the Guide for Care and Use of Laboratory Animals (1978) as prepared by the Committee on Care and Use of Laboratory Animals, National Research Council, DHEW Pub. No. (NIH) 80-23.

REFERENCES

1. Bignami, G., Amorico, L., Frontali, M., & Rosic, N. Physiol. Beh., 1970, 7, 461.
2. Bignami, G., Rosic, N., Michalek, H., Milosevic, M., & Gatti, G. In: Behavioral Toxicology, B. Weiss & V. Laites, Eds., Plenum Press: N.Y., 1975, 155-210.
3. DeCandole, C., Douglas, W., Evans, C., Holmes, R., Spencer, K., Torrence, R., & Wilson, K. Br. J. Pharmacol., 1953, 8, 446.
4. Fleisher, J. & Harris, L. Biochem. Pharmacol., 1965, 14, 641.
5. Gordon, J., Leadbeater, L., & Maidment, M. Toxicol. Appl. Pharmacol., 1978, 43, 207.
6. Harris, L., Flisher, J., Innerebner, T., Cliff, W., & Sim, V.

Toxicol. Appl. Pharmacol., 1969, 15, 216.
7. Harris, L., McDonough, J., Stitcher, D., & Monroe, R. Pharmacologist, 1980, 22, 239.
8. Harris, L., Lennox, W., Stitcher, D., McDonough, J., Talbot, B., & Barton, J. Pharmacologist, 1981, 23, 224.
9. Headley, D. Mil. Med., 1982, 147, 122.
10. Hobbinger, F. Br. J. Pharmacol., 1955, 10, 356.
11. Jovic, R. & Milosevic, M. Eur. J. Pharmacol., 1970, 12, 85.
12. Kurtz, P. Tox. Appl. Pharmacol., 1977, 42, 589.
13. Kurtz, P. Pharmacol. Biochem. Behav., 1977, 6, 303.
14. Rosic, N. Arch. Int. Pharmacodyn. Ther., 1970, 183, 139.
15. Rosic, N. & Milosevic, M. Acta Nerv. Sup., 1971, 13, 241.
16. Rosic, N., Bignami, G., & Gatti, G. In: The Correlation of Adverse Effects in Man with Observations in Animals, Proceedings of the European Society for the Study of Drug Toxicity, Vol. 12, S. Baker Ed., Excerpta Medica: Amsterdam, 1971, 242-246.
17. Sidell, F. Clin. Tox., 1974, 7, 1.
18. Wills, J. In: Anticholinesterase Agents, A. Karczmar Ed., Pergamon Press, N.Y., 1970, 367-395.
19. Winer, B.J. Statistical Principles in Experimental Design. 2nd Ed., 1971.

PLASMIDS OF *LEGIONELLA* SPECIES

*PERRY MIKESELL, CPT
GREGORY B. KNUDSON, PhD
U.S. ARMY MEDICAL RESEARCH INSTITUTE OF INFECTIOUS DISEASES
FORT DETRICK, FREDERICK, MARYLAND 21701

Plasmids are extrachromosomal genetic elements of bacteria that are not essential for survival of the host under most circumstances but often carry genes that allow the host to compete more successfully in adverse environments. Plasmids have been shown to carry genes which code for toxins, such as botulinum, diphtheria, streptococcus erythrogenic, and the enterotoxin of *Escherichia coli* (1). In addition to carrying genes responsible for metabolism, resistance to metals and fertility factors, plasmids also encode for drug resistance factors. These latter genetic elements are an important consideration in the treatment of infectious diseases, particularly in treating pathogens responsible for nosocomial infections. The sudden outbreak of a highly infectious disease caused by a multiply drug-resistant organism could severely impact on the planning and progress of a given military operation. Areas of high troop concentrations, such as basic training units, are particularly susceptible to the rapid spread of an infectious disease.

There are approximately 800,000 cases of pneumonia in the United States each year for which no known viral or bacterial agent can be identified (2). An outbreak of pneumonia occurred in 1976 at the Bellevue-Stratford Hotel in Philadelphia during an American Legion convention (3). The etiological agent was identified as a fastidious, aerobic, gram-negative bacterium and given the name *Legionella pneumophila*. The number of reported cases of legionellosis is rising rapidly and it is probable that a significant number of here-to-fore undiagnosed cases of pneumonia can be attributed to this novel pathogen.

The possible involvement of plasmids in legionellosis was suggested by reports of cytotoxin (4) and β-lactamase production (5) by *L. pneumophila* and by reports of virulent to avirulent conversion of the organism through serial passage on artificial media. Another strong indicator that plasmids may be involved in the pathogenesis of legionellosis was suggested by the very narrow spectrum of antibiotics effective in the treatment of the disease. We were interested in examining strains of *Legionella* for

the possible presence of plasmid DNA and determining what if any role these elements play in the pathogenesis of the disease.

MATERIALS AND METHODS

In the initial phase of our study, 16 strains from the six serogroups of L. pneumophila were examined for the presence of extrachromosomal DNA by a modified cleared lysate technique. Bacteria were harvested by scraping 10 charcoal yeast extract agar plates, which exhibited confluent growth, using a bent glass rod and 3 ml per plate of saline containing 0.1% tryptose, pH 7.0. Each strain was also harvested from 400 ml of yeast extract broth culture at ca. 2×10^9 cells/ml by centrifugation at 10,000 x g for 10 min at 4°C. The pellet was resuspended in 25 ml of 25% sucrose solution made up in 0.05 M tris (hydroxymethyl)aminomethane (Tris), pH 8.0, at room temperature. Spheroplasts were generated by the addition of 3.0 ml of freshly prepared lysozyme solution (10 mg/ml in 0.25 M Tris, pH 8.0). The suspension was mixed and placed on ice for 5 min. A 3.0-ml portion of 0.25 M ethylenediaminetetraacetate (EDTA), pH 8.0, was then added, and the mixture was chilled in an ice bath for an additional 5 min with occasional swirling. Lysis was brought about by adding 30 ml of 2.0% Triton X-100 made up in 0.025 M EDTA and 0.05 M Tris, pH 8.0. The final mixture was incubated at 37°C for 15 min and gently swirled until clearing occurred. The viscous lysate was centrifuged at 4°C for 30 min at 48,000 x g to sediment most of the chromosomal DNA and cellular debris. The plasmid DNA remained in the supernatant; it is referred to as cleared lysate.

The supercoiled plasmid DNA in the cleared lysate was further purified by CsCl-ethidium bromide (EtBr)-density gradient centrifugation (6). One gram of solid CsCl and 0.1 ml of EtBr solution (5 mg of EtBr per ml in TE buffer; 10 mM Tris, 1 mM EDTA, pH 8.0) was added per milliliter of lysate. The EtBr-DNA complex was protected from light to prevent light-activated dye nicking of the covalently closed circular (CCC) DNA. The bands were located by using long-wave ultraviolet illumination and were removed through the side of the tube with an 18-gauge needle and syringe. The dense satellite bands containing the CCC plasmid DNA were then pooled, and the EtBr was extracted three times with an equal volume of isopropanol (presaturated with water and CsCl). The samples were then dialyzed against three 1-liter changes of 15 mM NaCl and 1.5 mM sodium citrate and stored at 4°C.

Samples of 40 to 100 µl of plasmid preparation were mixed with 20 µl of tracking dye solution (0.7% bromophenol blue, 7% sodium dodecyl sulfate, and 16.5% glycerol in water) and subjected to electrophoresis, similar to the method described by Meyers et al. (7). DNA was electrophoresed in both 0.7 and 0.8% agarose with Tris-borate buffer (89 mM boric acid), pH 8.3, on a vertical slab gel. Electrophoresis was carried out at 120 V and 20°C for 2.5 h, or until the dye reached gel bottom. For comparative purposes, samples were also electrophoresed at 30 V for 16 h. The gels were stained for 30 min in an aqueous EtBr solution (0.5 µg/ml) and then washed in

water for 20 min. Plasmid DNA bands were visualized using ultraviolet light.

The second phase of our study was to examine another group of bacteria which were classified as Legionella-like organisms (Table 1). These organisms possess various degrees of phenotypic or genotypic relatedness to L. pneumophila based in part on DNA homology studies. They have since been classified as representing five species of the genus Legionella. Plasmid elements could not be detected in the Legionella-like strains by the protocol used for plasmid isolation from the Atlanta-1 and -2 strains of L. pneumophila. We therefore developed an isolation procedure which incorporated the central features of several well-established methods.

Table 1. Legionella-like Strains

Legionella pneumophila OLDA	Legionella bozemanii WIGA, MI-15
Legionella micdadei TATLOCK, HEBA, PPA	Legionella dumoffii TEX-KL, NY-23
Legionella gormanii LS-13	

Legionella-like bacteria were cultured on chemically defined medium according to established parameters of growth for L. pneumophila (11). Cells from 100 ml of exponential-phase cultures were harvested by centrifugation and washed once in 10 ml of 10 mM sodium phosphate buffer (pH 7.0). Washed cells were suspended in 3.0 ml of 25% sucrose in 50 mM Tris (pH 8.0), lysozyme (3.0 mg/ml) was added, and the suspension was incubated at 37°C in a shaker-incubator. After 15 to 20 min of incubation, 3.0 ml of 250 mM EDTA (pH 8.0) was added, and the cells were chilled on ice for 5 min. Cell lysis was achieved by the addition of 1.5 ml of 20% sodium dodecyl sulfate followed immediately by incubation in a 55°C water bath for 5 min with gentle agitation. Freshly prepared 3 N NaOH was added drop-wise until the pH was 12.1 to 12.4. The pH was immediately reduced to 8.5 to 9.0 with 2 M Tris (pH 7.0). Denatured chromosomal DNA and cellular debris were precipitated by the addition of 1.5 ml of 20% sodium dodecyl sulfate and 3.0 ml of 5 M NaCl followed by overnight storage at 4°C. The following day, the lysate was centifuged for 30 min at 17,000 x g at 4°C. The precipitate was discarded, and ribonuclease (2 mg/ml in distilled water, heated to 100°C for 5 min) was added to the supernatant to a final concentration of 100 µg/ml and incubated for 30 min at 37°C. Plasmid DNA was precipitated by the addition of 0.05 volume of 3 M sodium acetate and two volumes of cold 95% ethanol and stored at -20°C for at least 4 h. Plasmid DNA was concentrated for 30 min at 17,000 x g, and the resultant pellet was suspended in 100 to 200 µl of Tris-borate buffer. Samples were subjected to electrophoresis in 0.8 and 1% agarose, using Tris-borate running buffer and tracking dye, and stained as previously described by Meyers et al. (7). Samples were electrophoresed at 2 mA for 60 min followed by 50 mA

for 90 to 210 min depending on the degree of band separation desired.

RESULTS

The migration patterns of purified plasmid DNA from the OLDA, WIGA, and TEX-KL isolates are shown in Fig. 1. Molecular weight estimates were determined from the relative migration rates of plasmid bands in agarose gels (Fig. 2). The OLDA strain of L. pneumophila contained a single covalently closed circular plasmid species, pLP3 (Fig. 1D), with an estimated molecular weight of 59.8×10^6. This was the largest of the five plasmid isolates. The WIGA bacterium contained two plasmid species, pLB1

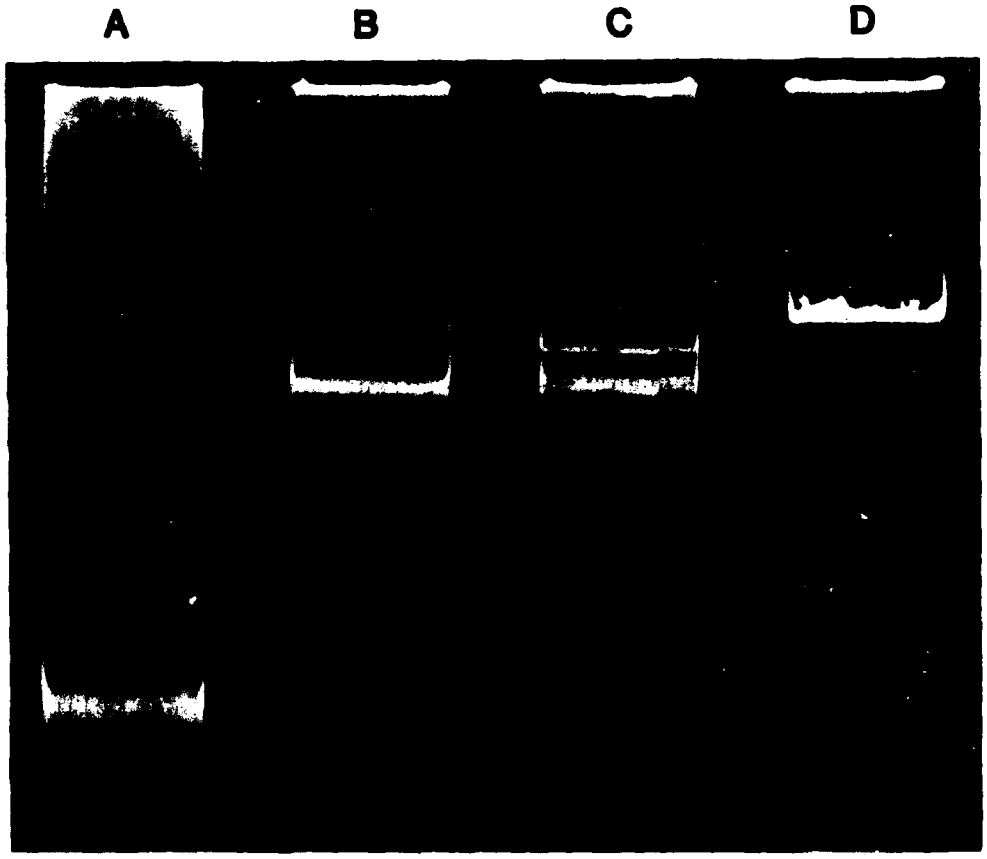


Figure 1. Gel electrophoresis of purified plasmids from Legionella-like bacteria. Purified DNA (30 µl) was mixed with 40 µl of tracking dye. The DNA-dye mixture (30 µl) was applied to agarose well. DNA samples were subjected to electrophoresis in 1% agarose at 2 mA for 60 min followed by 50 mA for 210 min. (A) Pseudomonas aeruginosa, PU21, control; (B) TEX-KL strain; (C) WIGA isolate; (D) OLDA strain of L. pneumophila.

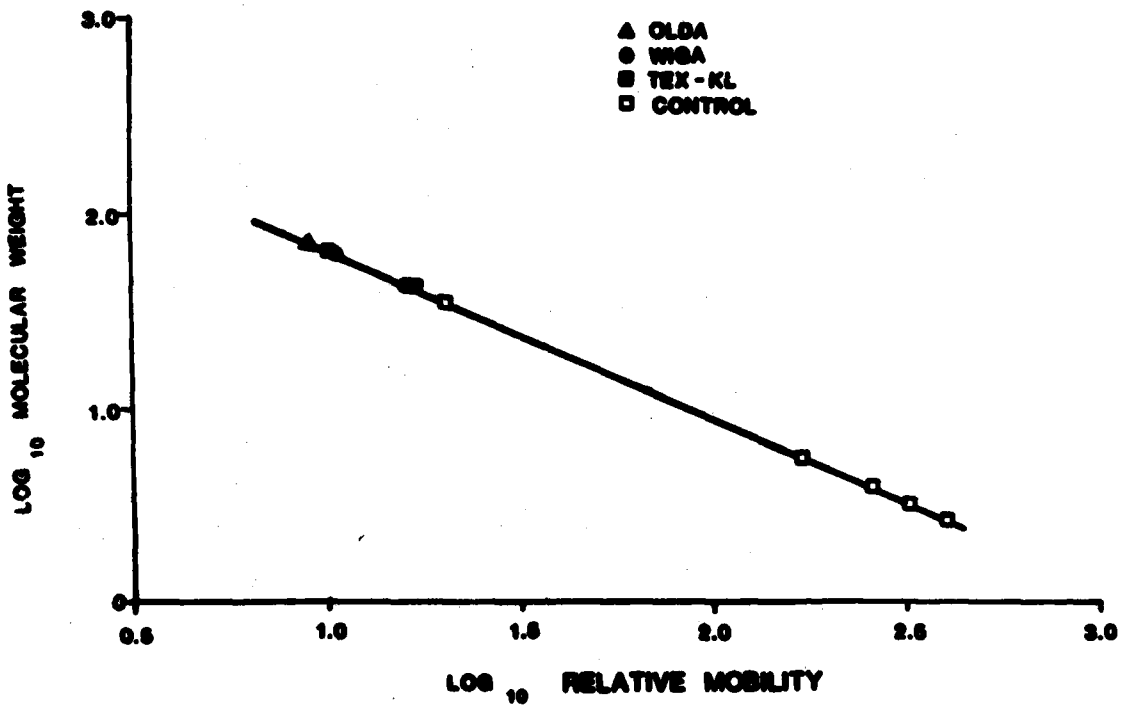


Figure 2. Least-squares regression analysis of plasmid DNA. Plasmid molecular weights were calculated as described by Hansen and Olsen (8). Plasmid DNA from E. coli V517 purified by CsCl-EtBr buoyant density centrifugation was used as a marker strain. This strain contains eight plasmid species. Plasmid molecular weights of marker strain are 35.8×10^6, 4.8×10^6, 3.7×10^6, 3.4×10^6, and 2.6×10^6. The three smaller plasmid species of the marker strain (2×10^6, 1.8×10^6, and 1.4×10^6) were not retained on the gel under the stated electrophoretic conditions.

(molecular weight, 54.3×10^6) and pLB2 (molecular weight, 47.6×10^6) (Fig. 1C), the smaller of the two having a double-band appearance. The third intermediate band is believed to be a catenated form of the smaller of the two plasmids (pLB2) in the WIGA isolate and not an open circular form of pLB2 or a third distinct plasmid species. The TEX-KL organism also had two plasmid species, pLK1 and pLK2, with molecular weights of 58.6×10^6 and 46.6×10^6, respectively (Fig. 1B). The alteration of selected electrophoretic parameters shows that all plasmid isolates are unique entities. Plasmid DNA has also been purified from two additional strains of Legionella-like bacteria (data not shown). LS-13 contains a single plasmid species with a molecular weight of $50\text{-}60 \times 10^6$. The NY-23 strain contains three distinct plasmid elements with molecular weights ranging from 35 to 60×10^6. Plasmid DNA was not detected in the HEBA,

TATLOCK, MI-15, or five Pittsburgh pneumonia agent isolates. The failure to isolate extrachromosomal DNA from these organisms may have been due to shortcomings in our technique, and alternate methodologies may eventually establish the presence of plasmid DNA in these organisms. The small 20 x 10^6-molecular-weight cryptic plasmid of the control P. aeruginosa strain was not observed (Fig. 1A). It is possible that this culture was cured of this smaller plasmid, since extrachromosomal DNA from E. coli in the molecular weight range of 10 x 10^6 to 20 x 10^6 was successfully resolved by our protocol (data not shown).

The recovery of plasmid DNA from cells grown in complex medium was very low compared with recovery of plasmid material from cells grown in chemically defined medium. Although these plasmid species do not have high molecular weights, they appear to be highly susceptible to shear forces, which may explain our initial failures in attempting to isolate extrachromosomal DNA by conventional procedures.

DISCUSSION

The isolation of extrachromosomal DNA from members of the Legionella genus is not surprising, considering the ubiquitous nature of plasmid elements. Our results indicate that members of this genus, like other human pathogenic microorganisms, are able to maintain stable plasmid DNA as part of their total genetic complement. Since we have found naturally occurring plasmids in seven strains of Legionella, it seems reasonable that these organisms could support the replication of other plasmids. In recent years, there have been epidemics of typhoid fever in Vietnam and Mexico due to Salmonella typhi having plasmid-mediated resistance to chloramphenicol (9) and the emergence of Haemophilus influenzae resistant to ampicillin (10). Indigenous cryptic plasmids, such as those found in Legionnella may acquire transposons which code for drug resistance. The presence of an R-plasmid coding for resistance to erythromycin would severely restrict the present treatment of legionellosis. Although rapid progress has been made in the diagnosis and treatment of this disease, there is still much to be understood concerning the genetics of Legionella: the reason for the difference between the pneumonic and nonpneumonic forms, its precise etiological niche, its nutritional supply, and its phylogenetic relationship to other gram-negative bacteria.

There have been two documented outbreaks of legionellosis in the military population (12, 13). These two epidemics were called "Fort Bragg fever." The etiologic agent was described as rickettsia-like; however, it was never shown conclusively that this organism was responsible for the 1942-43 Fort Bragg epidemics. The organism originally designated TATLOCK, after the individual who isolated the bacterium, has been subsequently classified as a strain of L. micdadei (14). It has been estimated that there are approximately 26,000 undiagnosied cases of legionellosis in the United States each year (2). In view of these morbidity estimates and in consideration of the amount of time which has passed since the organism

was implicated in a military epidemic, there is a significantly high probability that many cases of respiratory disease in military communities have been or are caused by members of the genus Legionella.

The significance of our research with Legionella has not had an immediate impact on the diagnosis and treatment of legionellosis in the military. Our research program was designed to isolate and characterize naturally occurring and genetically engineered plasmids in pathogens of military importance. We successfully established a technological base for this program, in high containment laboratories, using an organism about which virtually nothing was known when we initiated the studies.

The ultimate goal of the program is to contribute to the diagnosis, treatment and prophylaxis of diseases considered to be of military importance. As an extension of our current research program, we have recently demonstrated an association between plasmid DNA and toxin production in Bacillus anthracis. We anticipate that these findings and the application of recombinant DNA technology will contribute to our attempts to develop a more efficacious, long-lasting human anthrax vaccine.

REFERENCES

1. Maas, W. K. 1977. Genetics of toxin production by bacteria. In: Perspectives in Toxinology (A. W. Bernheimer, ed.). John Wiley and Sons, New York, pp. 1-13.
2. Fraser, D. W. and J. E. McDade. 1979. Legionellosis. Sci. Am. 241 (4):82-99.
3. Center for Disease Control. 1976. Follow-up on respiratory illness--Pennsylvania. Morbid. Mortal. Weekly Rep. 25:308.
4. Friedman, R. L., B. H. Iglewski, and R. D. Miller. 1980. Identification of a cytotoxin produced by Legionella pneumophila. Infect. Immun. 29:271-274.
5. Thornsberry, C. and L. A. Kirven. 1978. β-Lactamase of the Legionnaires' bacterium. Curr. Microbiol. 1:51-54.
6. Radloff, R., W. Bauer, and J. Vinograd. 1967. A dye-buoyant-density method for the detection and isolation of closed circular duplex DNA: the closed circular DNA in HeLa cells. Proc. Natl. Acad. Sci. U.S.A. 57:1514-1521.
7. Meyers, J. A., D. Sanchez, L. P. Elwell, and S. Falkow. 1976. Simple agarose gel electrophoretic method for the identification and characterization of plasmid deoxyribonucleic acid. J. Bacteriol. 127:1529-1537.
8. Hansen, J. B., and R. H. Olsen. 1978. Isolation of large bacterial plasmids and characterization of the P2 incompatibility group plasmids mMG1 and pMG5. J. Bacteriol. 135:227-238.
9. Butler, T., N. N. Linh, K. Arnold, and M. Pollack. 1973. Chloramphenicol-resistant typhoid fever in Vietnam associated with R factor. Lancet ii:983-985.

10. Jacobson, J. A., J. B. McCormick, P. Hayes, C. Thornsberry, and L. Kirven. 1976. Epidemiologic characteristics of infections caused by ampicillin-resistant Hemophilus influenzae. Pediatrics 58:388-391.
11. Ristroph, J. D., K. W. Hedlund, and S. Gowda. 1981. Chemically defined medium for Legionella pneumophila growth. J. Clin. Microbiol. 13:115-119.
12. Daniels, W. B. and H. A. Grennan. 1943. Pretibial fever--an obscure disease. JAMA 122:361-365.
13. Tatlock, H. 1944. A rickettsia-like organism recovered from guinea pigs. Proc. Soc. Exp. Biol. Med. 57:95-99.
14. Hébert, G. A., C. W. Moss, L. K. McDougal, F. M. Bozeman, R. M. McKinney, and D. J. Brenner. 1980. The rickettsia-like organisms TATLOCK (1943) and HEBA (1959): bacteria phenotypically similar but genetically distinct from Legionella pneumophila and the WIGA bacterium. Ann. Intern. Med. 92:45-52.

FLIGHT INSTABILITIES OF SPINNING PROJECTILES HAVING NON-RIGID PAYLOADS

Miles C. Miller
Chemical Systems Laboratory
US Army Armament Research and Development Command
Aberdeen Proving Ground, Maryland 21010

PREFACE

These studies were supported by funding under the ARRADCOM Chemical Systems Laboratory In-House Independent Research (ILIR) and Munitions Division programs.

INTRODUCTION

The use of an artillery delivered smoke screen has received new impetus as a means of reducing the vulnerability of friendly forces to enemy weapons, particularly guided missiles. The XM761, 155mm artillery projectile was intended to provide an improved smoke screening capability. The round consisted of an M483A1-type projectile body containing a pre-loaded payload canister assembly illustrated in Figure 1.

The thin-walled, aluminum payload canister shown in Figure 2 had an inside length of 51.76 cm and an inside diameter of 12.06 cm. An aluminum cruciform baffle ran the length of the cylinder effectively dividing the interior into four quadrants. A total of 48 cotton patio torch wicks was placed in the canister in three tiers with 16 wicks per tier. A single wick (also shown in Figure 2) measured about 16.5 cm in length and weighed about 17-grams. The canister was then filled with liquid WP, resulting in the wicks being saturated with WP. The fully loaded canister had a mass of 12.25 kg including 5.8 kg of WP, giving a total projectile weight of about 46.27 kg.

The wicks retained the WP, whether in a liquid (hot) or solid (cold) state. Upon expulsion from the projectile over the target, the WP saturated wicks would be dispersed onto a relatively large area of the ground. Each wick would spontaneously ignite providing a series of point sources of smoke resulting in a rapidly formed, dense smoke screen of relatively long duration.

The XM761 WP/wick projectile produced the desired smoke screening performance. However, during the subsequent development program, flight tests revealed a serious flight stability problem (References 1, 2, and 3).

When fired at a low temperature with the WP in a solid state, the projectiles had stable flights under all launch conditions. Figure 3 contains yaw sonde data for a typical stable flight. The peak-to-peak amplitude of the SIGMA N trace denotes the total angle of yaw possessed by the projectile. A reduction of yaw with time represents a stable projectile flight whereas an increasing yaw with time would indicate an instable condition.

At elevated temperatures where the WP was in a liquid state, the projectile experienced a severe flight instability. The unique feature of this instability was that both a large increase in yaw angle and severe loss in spin rate were suffered by the projectile, causing the round to fall short of its intended range. This problem was particularly critical at the Zone 4 (transonic muzzle velocity) firing condition where the projectile possesses minimum aeroballistic stability and experiences relatively large initial yaw angles. Figure 4 presents yaw sonde data from a typical unstable flight and clearly depicts the increase in yaw angle and simultaneous loss in spin rate. The test data indicated that the nutation rate and angle are the critical motions with regard to this unstable behavior.

A similar type of flight instability which also produces coupling between yaw and spin had been recently identified and documented. Aeroballistic analysis of that problem indicated that rotational motion of the projectile in flight caused rigid components internal to the projectile to assume specific motions relative to the projectile body producing an inertial moment which affected the flight stability of the projectile. Murphy (Reference 4) subsequently produced a detailed theory describing the specific cause of this inertial effect which was based on rigid, well-defined components located within the projectile.

The flight instabilities of the XM761 projectile appeared to be similar, but were more severe. However, because of the non-rigid composition of the payload components of the XM761 projectile (i.e., partial solid/partial liquid), no theory existed for this particular situation.

Accordingly, a special ground test fixture was built which caused an actual, full-scale projectile payload assembly to undergo the basic spin and nutation motion of the projectile in flight (Reference 5). The internal payload motion could then be reproduced, the resulting despin moment measured, and the data used to evaluate the potential flight stability of the projectile carrying the payload configuration tested.

This paper describes the use of this laboratory test fixture to determine the critical performance characteristics of the despin resulting from the relative motion of the XM761-type non-rigid payload; the subsequent use of the fixture to evolve a stable WP smoke round payload configuration; and the results of studies with simple homogeneous, viscous liquid fills which demonstrated their similarity to the more complex partial solid/partial liquid non-rigid payloads in producing this type of flight instability.

SYMBOLS

I_C	Canister and payload axial moment of inertia
I_P	Projectile total axial moment of inertia
M	Despin moment due to non-rigid payload
M_{AERO}	Despin moment due to aerodynamic effects
M_F	Despin moment due to canister bearing friction
M_A	Despin moment due to air turbine
P	Spin rate of projectile
SIGMA N	Angle of projectile centerline relative to sun direction
t	Time
α	Nutation yaw angle
γ	Specific mass density
θ	Test fixture canister coning angle
ν	Kinematic viscosity
Ω	Test fixture canister coning rate
ω	Test fixture canister spin rate
ω_N	Nutation frequency

$(\dot{})$	Denotes time derivative
WP	White phosphorus

LABORATORY TEST FIXTURE RATIONALE

The rationale for the test fixture is based on the assumption that both the yaw moment and spin moment produced by this instability are components of a single payload induced moment. The spin fixture measures only the despin component of the total destabilizing moment induced by the non-rigid payload. Assessment of the destabilizing potential of a particular payload is based on the premise that a payload induced destabilizing yawing moment is present that is proportional to the magnitude of the payload induced despin moment measured on the test fixture.

A better understanding of this situation can be gained by comparison with other types of flight instabilities. In the case of a Magnus instability which is due to an external aerodynamic effect, the resulting instability moment acts about an axis which is perpendicular to the spin (longitudinal) axis of the projectile. Thus, only the yaw of the projectile is affected. A Stewartson-type instability (Reference 6) is caused by an internal pressure wave effect in a pure liquid payload. However, the resulting instability moment axis is also perpendicular to the projectile spin axis and only affects the projectile yawing motion. The moment produced by an inertial instability due to an internal moving part acts about an axis which is not perpendicular to the projectile longitudinal axis. Thus, the component of this moment perpendicular to the longitudinal axis affects the yaw of the projectile, while the component parallel to the longitudinal axis acts to despin the projectile. Because of the similarities in the unstable projectile motion of the XM761 and that produced by an inertial instability, it was assumed that the single moment effect was a common factor in both types of instabilities even though the exact mechanism producing the moment might not be the same.

Figure 5 compares the motion and moments acting on the projectile in flight with those experienced by the payload canister in the laboratory test fixture. As shown in Figure 5(a), the projectile has a spin rate about its longitudinal axis and is assumed to be undergoing simple coning motion with a nutation frequency and at a nutation angle. The rotational motion assumed by a spinning projectile in flight includes both precessional (slow frequency) and nutational (fast frequency) epicyclic motion. However, it had been determined from flight test data that the nutational motion appeared to be the major factor in the instability. Thus, simple nutational coning motion was regarded as valid and sufficient representation of the projectile flight motion with regard to its influence on the payload motion and its resulting despin moment. In flight, the spin rate of the projectile and inclosed payload canister is a function of two

moments: the external aerodynamic spin damping moment and the payload induced despin moment, as shown.

The laboratory test fixture arrangement is shown in Figure 5(b). Since the instability being studied is due solely to an internal payload effect the external projectile shaping is not required. It is only necessary to match the payload elements and their internal constraints. These characteristics are present because an actual canister and payload assembly are used. The canister is attached to the fixture frame by bearings, allowing it to spin freely about its longitudinal axis, while the fixture frame is forced to spin about a vertical axis. This results in the canister assuming the desired simultaneous spinning and coning motion similar to that of the projectile in flight. In the case of the test fixture, the spin rate of the canister is the net result of three moments: an air turbine moment which causes the canister to attain its spin, a friction moment due to the canister bearings, and the payload induced despin moment. The friction moment and turbine moment are effects which are only present in the test fixture, and their values are determined by calibration. Therefore, the spin deceleration of the coning canister can be used to determine the payload induced despin moment.

DESCRIPTION AND OPERATION OF THE LABORATORY TEST FIXTURE

Figure 6 contains a photograph of the laboratory test fixture showing the system details. An actual, full-scale canister and inclosed payload assembly from a 155mm projectile is mounted between two clevis assemblies, each containing a bearing. These clevises are mounted to the test fixture frame so that the canister longitudinal axis can be oriented at a predetermined angle to the vertical. This angle represents the nutational coning angle. The canister can be set at angles from 0 to 20-degrees in 5-degree increments. The upper clevis contains a nozzle/air turbine arrangement providing spin torque to the canister. The table, frame, and attached canister are rotated about a vertical axis by means of an electric motor located on the lower section of the test fixture. The canister spin rate is measured by means of a magnetic tachometer located on the lower clevis, whereas the table spin (coning) rate is indicated by a photo tachometer located beneath the table. Weights located on opposite corners of the frame eliminate dynamic imbalance effects due to the canister angle.

A typical test run sequence is as follows: with the canister mounted at a particular angle, the air turbine spins the canister up to the spin rate of the projectile for the specific conditions being tested; approximately 60 seconds being required to achieve 6,000 rpm. When the desired canister spin rate has been achieved, the electric motor spins the table with the attached frame and spinning canister up to the desired nutational coning frequency. When the canister has reached a constant coning rate,

the air turbine is cut-off allowing the canister to spin down due to the combined effects of bearing friction and the payload induced despin moment. The canister spin rate is recorded as a function of time. The total canister despin moment at a particular canister spin rate and coning rate can be determined by multiplying the spin deceleration measured at that condition by the total canister axial moment of inertia. The despin moment due to friction is known from previous calibration as a function of the canister spin rate, coning rate, and coning angle. Therefore, any additional despin moment is due to a payload induced effect. Similarly, tests could be conducted over a range of constant coning rates at each fixed coning angle, thus encompassing nutation rates corresponding to various firing zones and initial yaw angles.

EVALUATION OF THE XM761 PAYLOAD

A series of tests were conducted on the test fixture using an actual XM761 canister assembly containing 48 patio torch wicks, the X-longitudinal baffle and calcium nitrate as a simulant for WP. Calcium nitrate is the commonly used simulant for WP because it has the same mass density and melting temperature. The canister was evaluated with the WP simulant in both a solid (cool) and a liquid (hot) state.

For a given canister angle, the spin moment was plotted as a function of the coning rate as shown in Figure 7. Data from several runs for the same canister configuration are shown. The numbers beside each symbol indicate the spin rate of the canister for the particular data point. The despin moment was not a function of canister spin rate, provided a sufficient canister spin rate is present (> 1,000 rpm) and can be represented by a single curve as shown. The data indicate that the despin moment is a non-linear function of the coning rate. It should also be noted that tests with a solid payload produced no payload induced despin moment. This testing procedure was repeated at canister angles of 15, 10, and 5 degrees, resulting in the data shown in Figure 8. These data are presented in a different form in Figure 9 which contains the despin moment as a function of canister coning angle for various coning rates and indicates that the despin moment in a non-linear function of the coning angle.

The test fixture results revealed that the payload induced despin moment is not dependent on set-back (longitudinal acceleration) at launch. This latter face was important in that payload asymmetries created by the projectile yaw acceleration at launch had initially been considered to be a possible cause of the instability. Also, the payload induced despin could be reproduced even after a liquid/solid/liquid cycle and thus, is not a one-time, irreversable effect.

The despin data obtained from the fixture showed good correlation with projectile spin loss measured during flight tests, providing confidence that

the fixture could be used to access the potential flight performance of candidate payload configurations.

EVOLUTION OF XM825 CONFIGURATION

With the basic capabilities demonstrated, the test fixture was used to evaluate several candidate payload canister configurations in support of the ARRADCOM, Chemical Systems Laboratory (CSL) efforts to evolve an improved 155mm WP smoke round.

An initial investigation was performed to evolve a simulant for WP that would be liquid at room temperature. The result was a blend of two inert and non-toxic E-series Freons providing a close match of the basic physical properties of liquid WP, including the specific density and specific viscosity (Table 1), and did not require the elevated temperatures for liquid-state tests as does the calcium nitrate.

SUBSTANCE	MELTING TEMPERATURE (°C)	SPECIFIC DENSITY (GM/CM^3)	VISCOSITY (CENTI-POISE)	SPECIFIC VISCOSITY (CENTI-STOKES)
White Phosphorous (60°C)	44.1	1.73	2.60	1.50
Calcium Nitrate (60°C)	42.7	1.71	58.1	34.0
Blended Freon (25°C)	N/A	1.73	2.42	1.40
Water (25°C)	0	1.00	1.15	1.15

TABLE 1. Physical Properties of Liquid White Phosphorous Simulants

Fixture tests conducted with the nominal XM761 payload configuration, but using the blended Freon simulant, produced identical despin moments to those obtained with the same payload configuration using calcium nitrate as the simulant thereby validating the blended Freon as a simulant for liquid WP. The use of the blended Freon and a reusable canister greatly simplified the testing procedure and reduced the time involved in the subsequent test program.

Various payload configurations were evaluated on the fixture. Despin moments due to the non-rigid payloads were measured on the fixture even with payload configurations which had demonstrated stable flight characteristics. It would appear that the projectile possesses sufficient inherent aero-ballistic stability to overcome a certain degree of this payload induced stability. However, if the payload induced effect is large enough, it will dominate and cause the projectile to go unstable.

A configuration was evolved by the ARRADCOM CSL composed of 92 1-inch thick F7-type felt wedges which were stacked between the X-longitudinal baffle as illustrated in Figure 10. The felt wedges served the same function as the cotton wicks, but the denser material and tighter packing of the felt wedges restricted their movement inside the canister. Test fixture results for this payload configuration are shown in Figure 11 and indicate a despin moment considerably lower than that measured for the nominal XM761. Data are shown for a coning rate of 500 rpm which is the projectile nutational frequency for the critical Zone 4 (transonic launch) condition. Subsequent flight tests of the felt wedge configuration using actual liquid WP resulted in stable flights under firing conditions which cause the nominal XM761 WP/wick configuration to be unstable (Reference 7).

The test fixture was then used to evaluate alternate felt wedge sizes, materials, and packing densities. A payload configuration composed of 120 3/4-inch thick, higher density F3-type felt wedges resulted in a significantly lower despin moment (Figure 11) indicating improved stability characteristics. This latter performance was also subsequently demonstrated in instrumented flight tests (Reference 8). Based on the improved flight stability this latter configuration was entered into development as the XM825 projectile.

EVALUATION OF HOMOGENEOUS, VISCOUS LIQUID FILLS

Although the original goal of the laboratory spin fixture was achieved in the evolution of the improved smoke screening projectile, the exact source of the instability associated with this partial solid/partial liquid-type of non-rigid payload was still not understood. The fixture was next used to generate a data base to provide further insight into the physical phenomena responsible for the instability.

Vaughn (Reference 9) proposed a hypothesis that the combined solid/liquid XM761 payload acts as a homogeneous, highly viscous liquid. An extensive series of controlled experiments were conducted with the laboratory test fixture using liquid fills of various viscosities and densities both with and without the baffle (References 10 and 11). The liquids were selected to represent various viscosity ranges. Different viscosity values for a given liquid were achieved by preconditioning the liquid to a specific temperature. These tests indicate that a cylindrical canister, completely

filled (i.e., no void) with a homogeneous viscous liquid produces a measurable payload induced despin moment under the combined spinning and coning motion of the test fixture. Figure 12 contained representative data for the canister (without baffle) filled with corn syrup having a specific mass density of 1.4 and a kinematic viscosity of 200,000 CS. The despin moment was found to be independent of canister spin rate for the range of spin rate values considered. Also, the payload induced moment is a non-linear function of both the coning rate and coning angle. The similarities of these results with the data for the XM761 payload are apparent.

Figure 13 shows the despin moment as a function of the liquid fill kinematic viscosity for a 20-degree canister coning angle. It had been determined from these tests that the despin moment, is directly proportional to the liquid density. Accordingly, the despin moment has been normalized to the mass density of water. These data show that for the 100% filled cylindrical canister (without baffle) the despin increases with the liquid fill viscosity, achieving a maximum value in the 10^5 CS range, thereupon diminishing to zero at very large values of viscosity. The presence of the X-longitudinal baffle with a given liquid results in a despin moment equivalent to that of a pure liquid of higher viscosity and remains at a constant level at the lower viscosity range.

Figure 14 presents selected test fixture data for several homogeneous, viscous liquids as well as for the partial liquid/partial solid payloads of the XM761 and XM825. These data illustrate additional characteristics which are not apparent in the generalized, single coning angle data of Figure 13. Note that for the homogeneous liquids having viscosities less than or greater than the 10^5 CS range of the peak design value, the payload induced despin moment is only present at the larger coning angles.

The magnitude of the despin moment measured for the canister filled with corn syrup (ν = 200,000 CS) is nearly identical to that of the XM761 payload configuration. Thus, based on these results, a projectile filled with corn syrup should experience a similar flight instability. Subsequent instrumented flight tests of full-scale 155mm projectiles having identical viscous liquid payloads to those evaluated on the test fixture were conducted by Ballistics Research Laboratory (BRL) (References 12 and 13). These tests showed good qualitative correlation to the fixture results. The similarity between the unstable flight motion of a corn syrup filled projectile and that of the XM761 is apparent from Figure 15.

This similarity could mean that the instabilities are due to a common source. If true, a theoretical analysis based on a well defined homogeneous viscous liquid fill might be more tractable than for the more difficult to quantify partial liquid/partial solid type non-rigid payloads.

CONCLUSIONS

1. A laboratory test fixture was designed and fabricated which allowed full-scale 155mm non-rigid payload assemblies to undergo the combined spinning and simple coning motion of the projectile in flight. The resulting despin moment produced by the non-rigid payload was used to evaluate the potential flight stability of a projectile carrying the non-rigid payload.

2. Test fixture evaluation of the partial solid/partial liquid payload indicates that the payload induced despin moment increases non-linearly with both coning frequency and coning angle. The despin moment is independent of canister spin rate providing a minimum spin rate of about 1,000 rpm is present. Below this spin rate, the despin moment is markedly decreased. Modifying the shaping and composition of the solid elements of the payload to restrict their movement reduced the despin moment and the associated flight instability potential.

3. Fixture tests with homogeneous, viscous liquid fills produced similar despin characteristics to those obtained with the partial solid/partial liquid pyalodas. These data indicate increasing instability with increasing liquid viscosity with a maximum effect at a kinematic viscosity of about 10^5 CS, whereupon the instability decreases with increasing viscosity. The despin moment was found to be directly proportional to the liquid mass density.

4. The similarity in the instability characteristics due to both the simple homogeneous, viscous liquids and the more complex partial solid/partial liquid non-rigid payloads could aid in the development of a general theory describing the source of the instability and establishing design criteria for future projectile applications.

REFERENCES

1. D'Amico, W. P., "Early Flight Experiments With the XM761," BRL-MR-2791, Sep 77.

2. D'Amico, W. P., "Field Tests of the XM761: First Diagnostic Test," BRL-MR-2792, Sep 77.

3. D'Amico, W. P., "Field Tests of XM761: Second Diagnostic Test," MR-02806, Jan 78.

4. Murphy, C. H., "Influence of Moving Internal Parts on Angular Motion of Spinning Projectiles," Journal of Spacecraft and Rockets, VOL 1, Mar-Apr 78, PP 117-122.

5. Miller, M. C., "Flight Instability Test Fixture for Non-Rigid Payloads," US Army Research and Development Command, ARRADCOM-SP-79005, Jan 79.

6. Stewartson, K., "On the Stability of a Spinning Top Containing Liquid," Journal of Fluid Mechanics, VOL 5, Part 4, Jun 59.

7. D'Amico, W. P., "Aeroballistic Testing of the XM785 Projectile: Phase I;" BRL-MR-02911, Mar 79.

8. D'Amico, W. P., and Oskay, V., "Aeroballistic Testing of the XM825 Projectile: Phase II," BRL-MR-679, Mar 80.

9. Vaughn, H. R., "Flight Dynamic Instabilities of Fluid Filled Projectiles," Sandia Laboratories, Albuquerque, NM, SAND 78-0999, Jun 78.

10. "In-House Laboratory Independent Research (ILIR) Program - Annual Review FY78," US Army Chemical Systems Laboratory Special Report 79003, PP3, I-9, Oct 78.

11. Miller, M. C., "Flight Instabilities of Spinning Projectiles Having Non-Rigid Projectiles," Journal of Guidance, Control and Dynamics, VOL 5, Mar-Apr 82, PP 151-157.

12. D'Amico, W. P., and Miller, M. C., "Flight Instability Produced by a Rapidly Spinning, Highly Viscous Liquid," Journal of Spacecraft and Rockets, VOL 16, Jan-Feb 79, PP 62-64.

13. D'Amico, W. P., and Clay, W. H., "High Viscosity Liquid Payload Yawsonde Data for Small Launch Yaws," BRL-MR-03029, Jun 80.

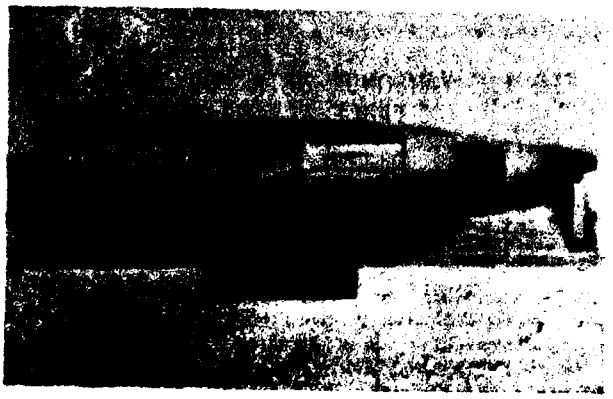


FIG. 1. XM761 PROJECTILE AND PAYLOAD ARRANGEMENT



FIG 2. XM761 PAYLOAD CANISTER AND WICK

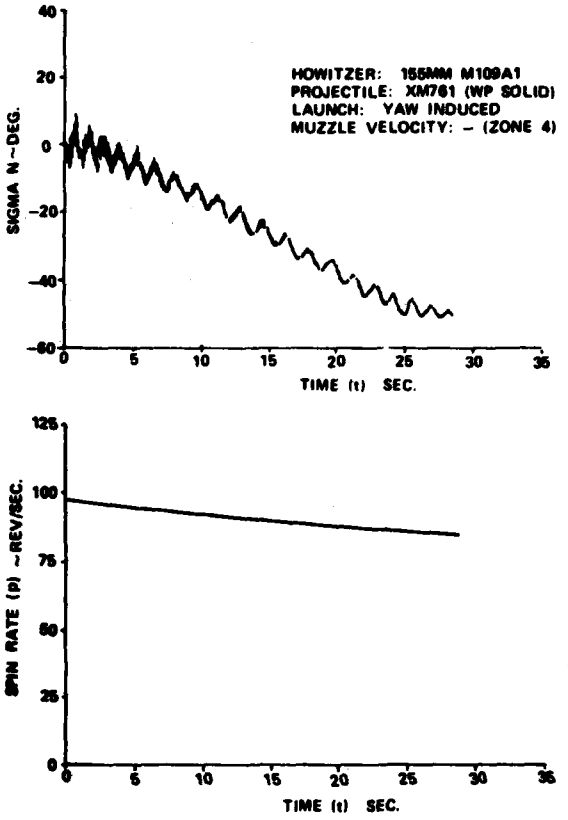


FIG 3. YAWSONDE DATA FOR XM761 STABLE FLIGHT (WP SOLID)

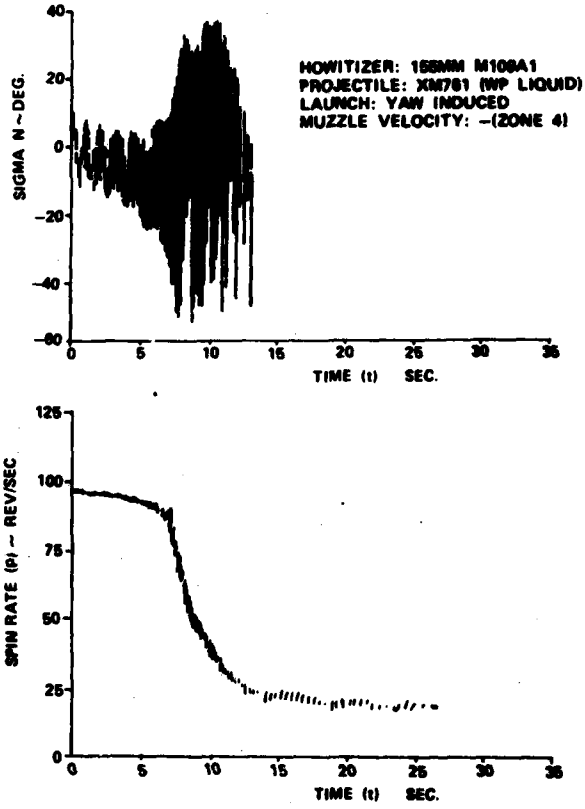


FIG 4. YAWSONDE DATA FOR XM761 UNSTABLE FLIGHT (WP LIQUID)

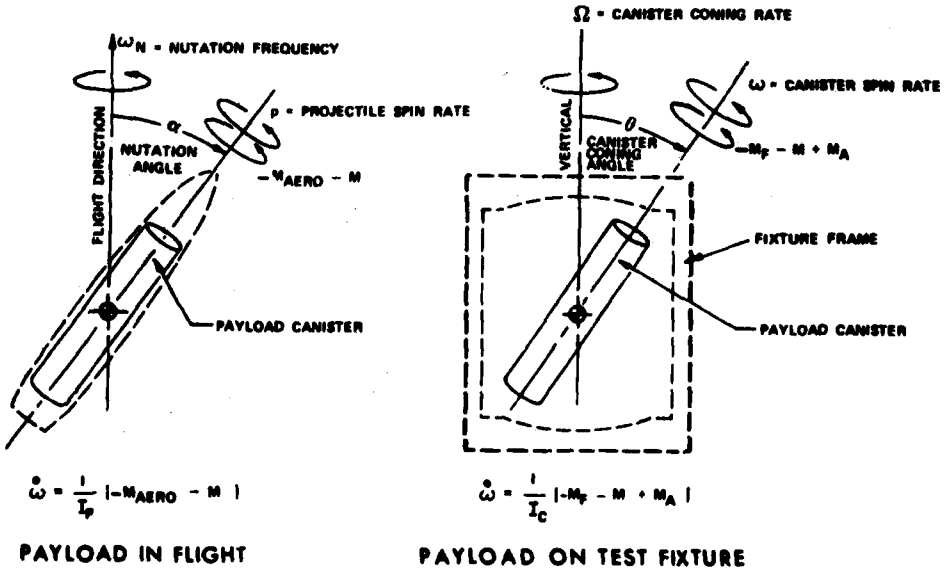


FIG 5. TEST FIXTURE RATIONALE

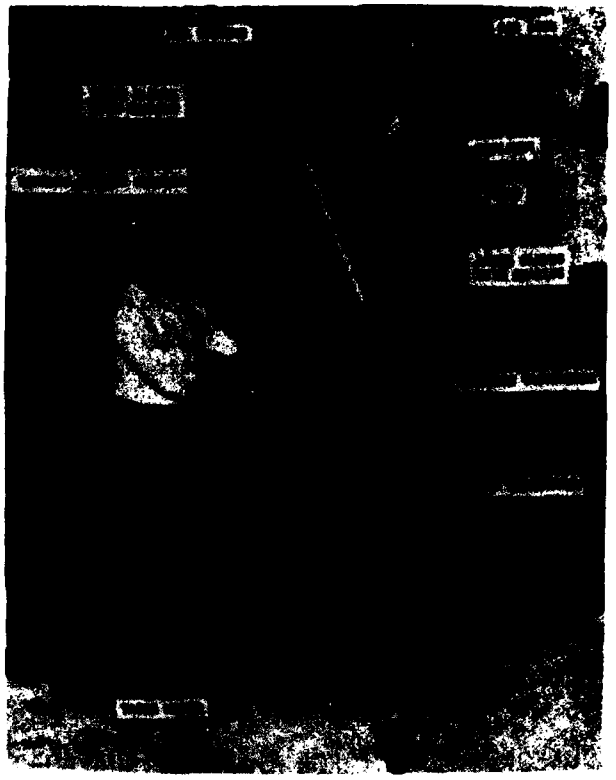


FIG 6. LABORATORY TEST FIXTURE

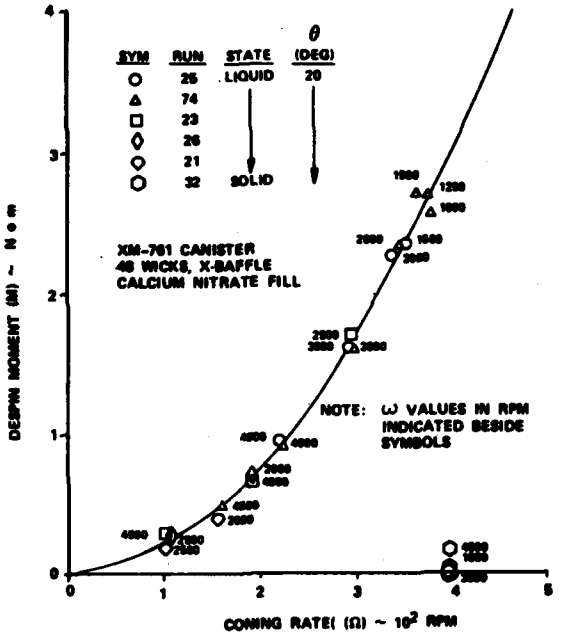


FIG 7. TEST FIXTURE DATA FOR THE XM761 PAYLOAD

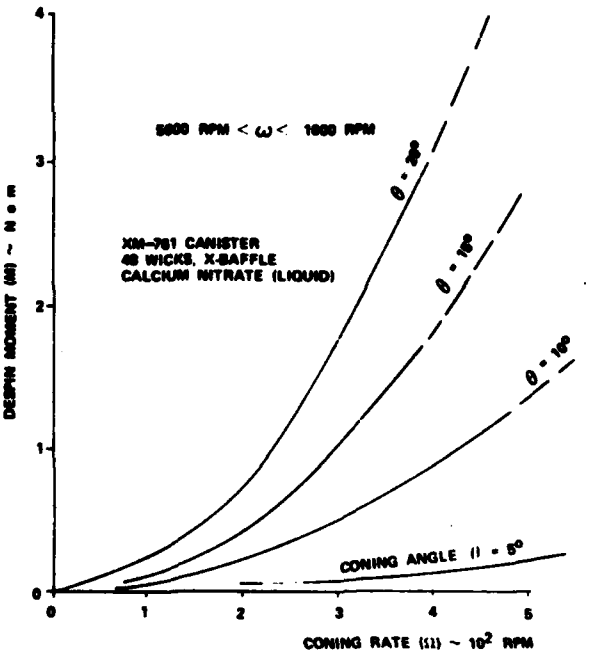


FIG 8. DESPIN MOMENT FOR THE XM761 PAYLOAD VS CONING RATE AND CONING ANGLE

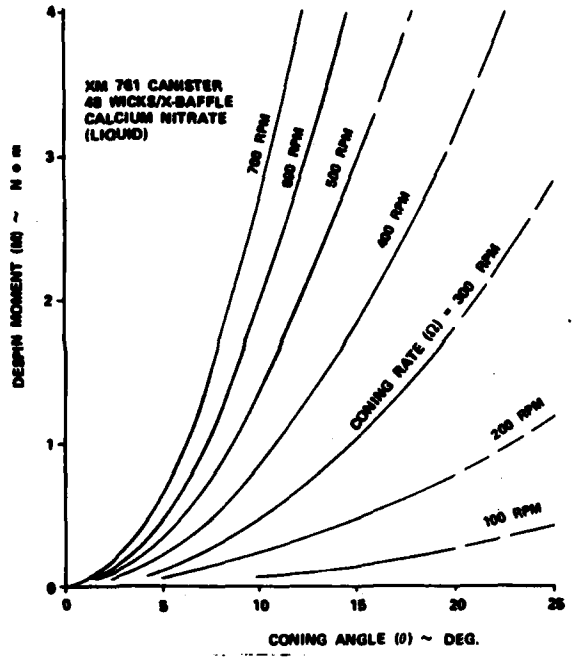


FIG 9. DESPIN MOMENT FOR THE XM761 PAYLOAD VS CONING ANGLE AND CONING RATE

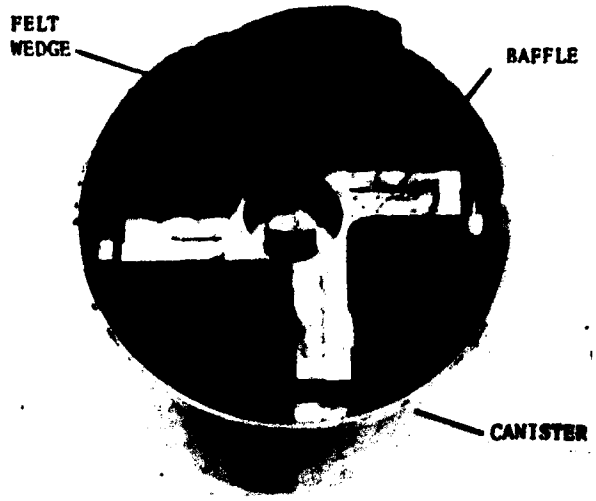


FIG 10. XM825 FELT WEDGE PAYLOAD CONFIGURATION

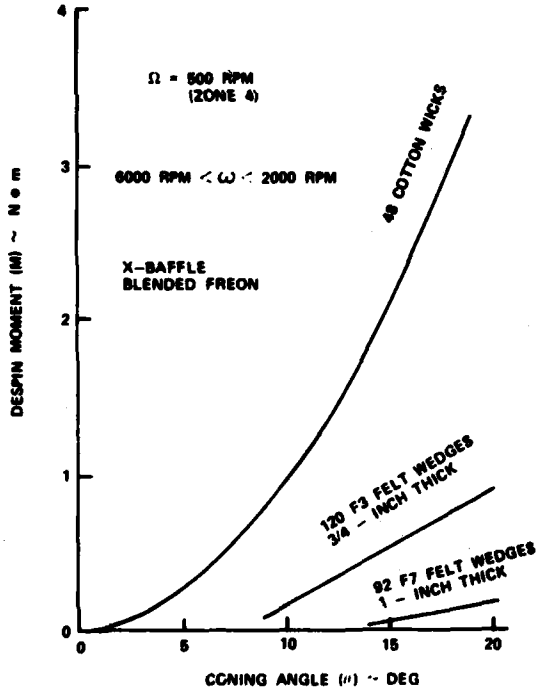


FIG 11. DESPIN MOMENT FOR COTTON WICK(XM761) AND FELT WEDGE (XM825) PAYLOADS

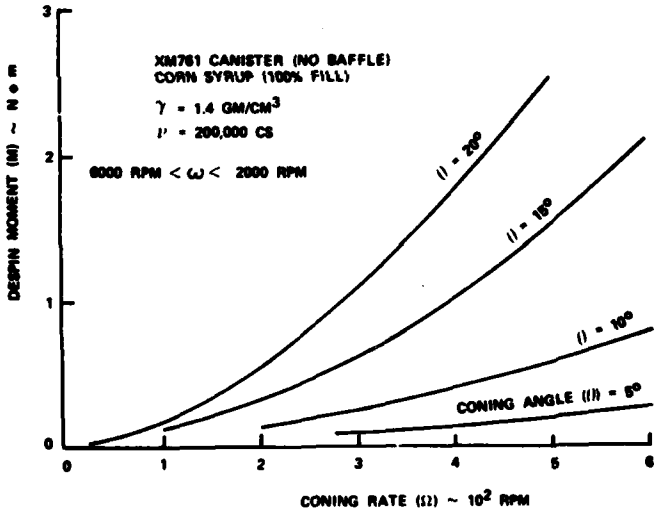


FIG 12. DESPIN MOMENT FOR HOMOGENEOUS, HIGH VISCOSITY LIQUID VS CONING RATE AND CONING ANGLE

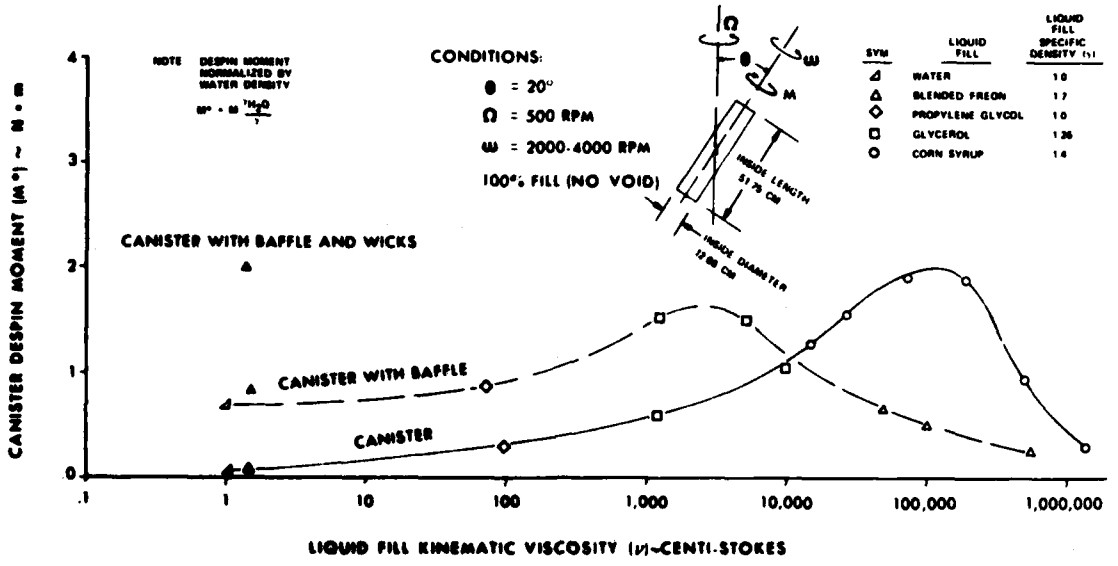


FIG 13. DESPIN MOMENT AS A FUNCTION OF LIQUID FILL VISCOSITY FOR A 20° CONING ANGLE

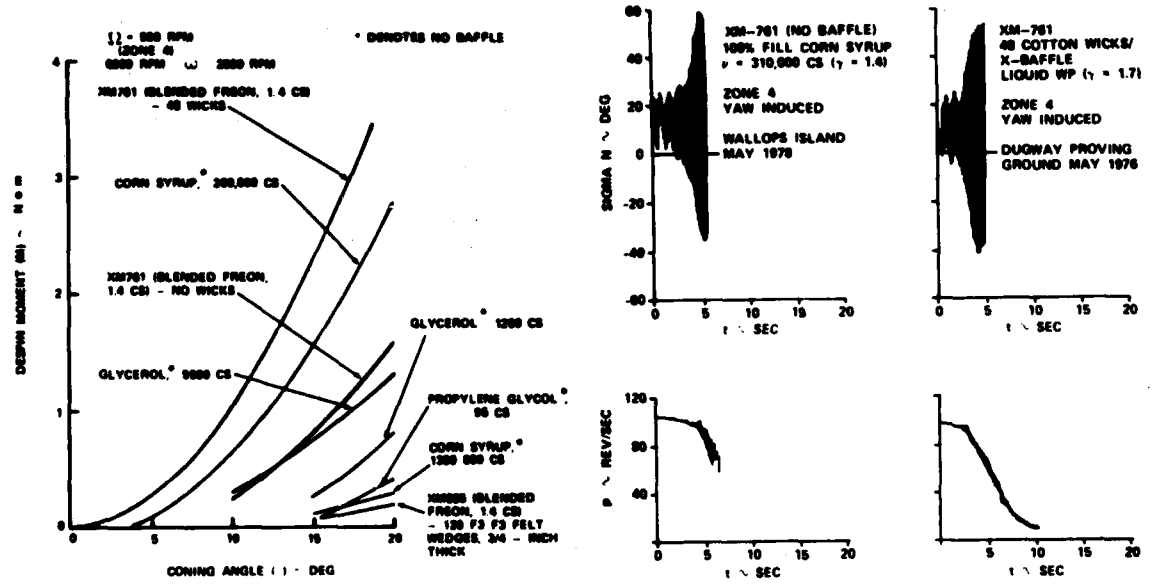


FIG 14. DESPIN MOMENT AS A FUNCTION OF CONING ANGLE FOR NON-RIGID PAYLOAD CONFIGURATIONS

FIG 15. UNSTABLE FLIGHT MOTION DUE TO A HOMOGENEOUS, VISCOUS LIQUID FILL AND PARTIAL LIQUID/PARTIAL SOLID PAYLOAD

ARMORED COMBAT VEHICLE TECHNOLOGY (ACVT) PROGRAM MOBILITY/AGILITY FINDINGS (U)

NEWELL R. MURPHY, JR., MR.
MOBILITY SYSTEMS DIVISION, U. S. ARMY ENGINEER WATERWAYS EXPERIMENT STATION
VICKSBURG, MISSISSIPPI 39180

PART I: INTRODUCTION

Background

The ACVT program was a joint Army and Marine Corps project to develop technology for designing and building armored vehicles in the post-1985 time frame. The program examined potentials for improving weapon systems, armor, and mobility/agility performance particularly as advances in these areas could be combined to produce combat vehicles of greater battlefield lethality and survivability. This paper is concerned only with the mobility and agility part of the program, which consisted of three closely related activities:

a. Careful testing of two special test chassis, plus the General Motors XM1 automotive test rig (ATR), the M113A1 Armored Personnel Carrier (APC), and the M60A1 Main Battle Tank (MBT), to develop quantitative data relating various measures of performance to a wide range of vehicle design parameters and terrain conditions and to driver behavior.

b. Development or refinement of analytical models for predicting vehicle performance, and validation of these models using the data base derived from a above.

c. Use of the validated analytical models to conduct broad parametric studies, to support war games which integrated mobility/agility, weapon systems, and armor considerations, and to evaluate concept designs for lightweight combat vehicles based on present and near-future component technology.

The principal special test chassis was the HIMAG vehicle whose gross weight, center of gravity, suspension spring and damping rates, and wheel travels could be widely varied. The HIMAG was used for ride and shock tests, dash tests, and traverse tests with various drivers. The second special test chassis was the twin-engine M113 developed for research

purposes by the U. S. Army Engineer Waterways Experiment Station (WES) and referred to as the M113 HOTROD. The M113 HOTROD (86 gross horsepower per ton (hp/ton) and the ATR (36 gross hp/ton) were used in special tests to examine the effects of high speeds on the motion resistance offered by soils.

Objectives and Issues

The long-range objectives of the ACVT mobility/agility program were to develop an extensive, reliable data base on combat vehicle performance and an array of validated analytical models to make these data readily available to vehicle designers and evaluators.

The immediate objectives were to address the following issues:

a. What are the effects of various vehicle chassis design parameters upon the attainment of high mobility/agility?
b. Are there any risk areas associated with high-speed travel in the area of vehicle-soil physics?
c. What is the fraction of available mobility used by a crew?
d. What is the M1 MBT level of mobility?
e. What is the mobility/agility performance of the HIMAG test bed versus that of conventional armored vehicles, the M1 MBT, and lightweight concept vehicles?
f. Can lightweight combat vehicles be designed with mobility/agility equal to or greater than that of the M1 MBT?
g. Does the attainment of high mobility/agility provide a payoff in survivability?

Mobility/Agility Tests

The test work was designed to develop quantitative data relating specific measures of vehicle performance to the engineering characteristics of a vehicle configuration and of the terrain and to driver behavior. Emphasis was placed upon obtaining a wide range of variations in vehicle and terrain so that trends could be seen clearly and analytical models could be checked as broadly as possible.

More than 1900 mobility/agility tests were conducted with 21 high-performance and 2 contemporary vehicles. Eighteen distinct configurations of the HIMAG variable high mobility/agility test bed were tested to explore mass and suspension effects on performance. Tests were also conducted with the ATR with and without the 13.5-ton turret, the M113 HOTROD and two contemporary vehicles, the M60A1 MBT, and the M113A1 APC. The test vehicles provided a range in gross vehicle weight from 9 to 52 tons, in gross hp/ton ratios from a low of 14 for the M60A1 MBT to a high of 86 for the M113 HOTROD, and in sprocket hp/ton ratios from a low of 8.4 for the M60A1 MBT to 28.9 for the M113 HOTROD. The vehicles were appropriately instrumented to measure and record the data of interest in each test.

Seven types of tests were conducted in quantitatively defined test areas to produce the desired data base and vehicle-terrain-driver relations (1). Five principal types of engineering tests were run--acceleration-deceleration (dash), ride dynamics, obstacle-impact response (shock), turning, and controlled-slalom (maneuver). Two types of tests were conducted to test tactical performance--a 20-km traverse test through many quantitatively defined terrain types for vehicle speed and driver response evaluation, and hit-avoidance tests to determine the survivability attributed to vehicle mobility/agility. The majority of tests were conducted at Fort Knox, Kentucky, with some special soft-soil tests conducted in a floodplain near Vicksburg, Mississippi. In these latter tests, trafficability, mobility, and agility data were obtained from cross-country and soft-soil tests at speeds more than twice those ever before achieved.

Mobility/Agility Models and Simulations

Concurrently with the field tests, turning, maneuver, and traverse models were developed to describe the mobility/agility performance along any specified path through any terrain (2). Field test results validated these new models (3) and revalidated the basic Army Mobility Model (AMM) (4), and its dynamics module VEHDYN (5) as well. The several validated models provided the analytical tools needed to predict mobility/agility performance and to conduct meaningful parametric studies.

These models were used to compare the performance in quantitatively defined German and Middle East terrains of more than 30 concept combat vehicles designed by the engineers at the U. S. Army Tank-Automotive Command to meet specific Army and Marine Corps requirements, plus the M1 MBT, the M3 Cavalry Fighting Vehicle (CFV), the M60A1 MBT, and the M113A1 APC.

PART II: RESULTS OF HIMAG CHASSIS TESTS

Mobility/agility performance depends on design balance, terrain, weather, and a specified mission profile. It cannot be assessed on the basis of a single vehicle parameter.

Figure 1 illustrates the principal factors that affect mobility/agility performance. The results of the ride and shock tests showed that the effects of suspension jounce travel (i.e., the vertical travel of a roadwheel from its static equilibrium position to the bump stop) depended on the degree of suspension damping, suspension spring rate, vehicle weight, and surface roughness. Reduced jounce travel combined with soft springs and low damping caused a progressive increase in suspension "bottoming" (roadwheels striking the bump stops) as the surface roughness or obstacle

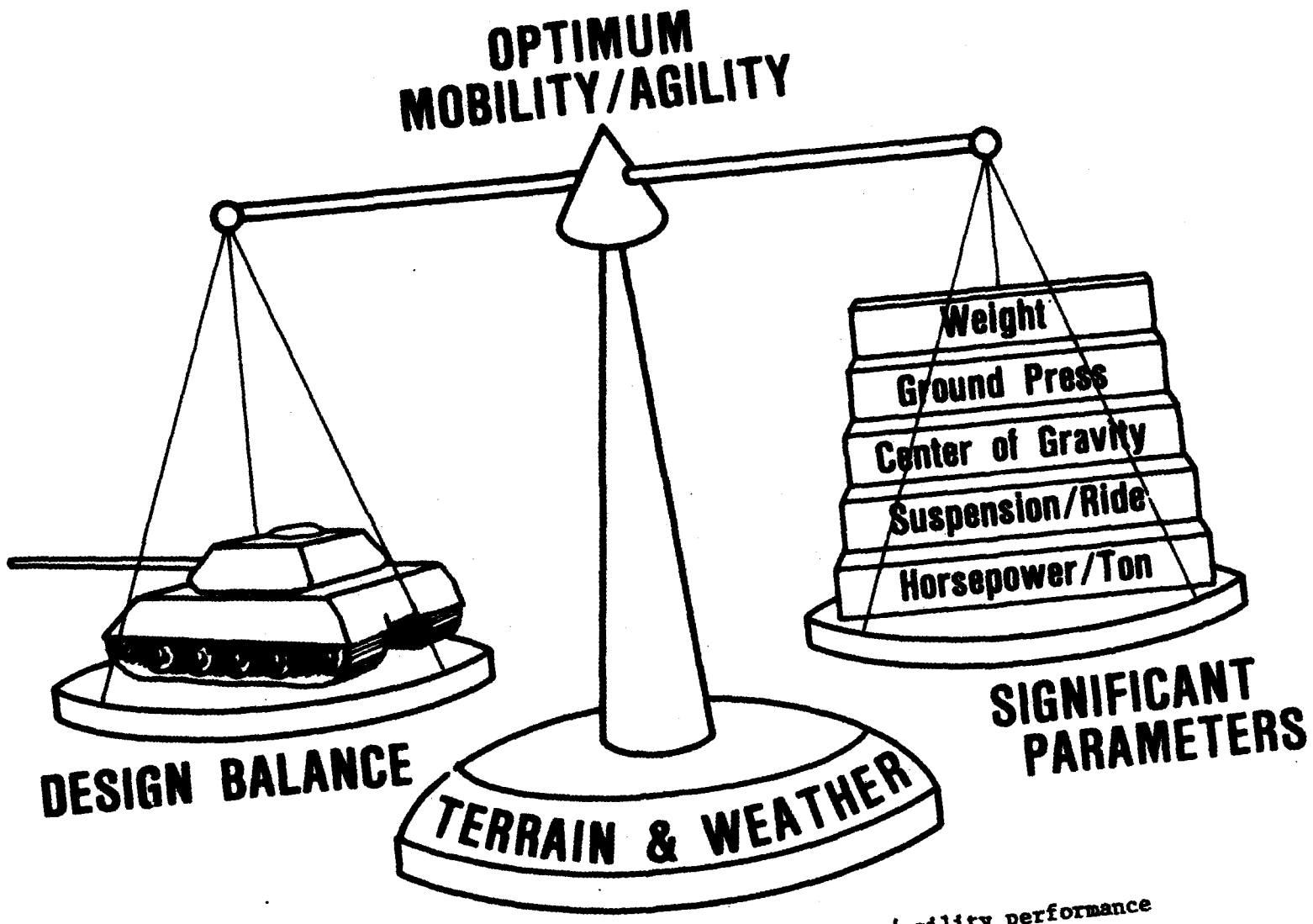


Figure 1. Principal factors that influence mobility/agility performance

height increased. This condition became worse for this type of suspension if the vehicle weight was increased. However, the shock effects caused by suspension bottoming could be effectively reduced with increased damping. The test results showed that light damping provides the best ride on smoother terrains and progressively heavier damping is required as the surface roughness increases. The performance patterns demonstrated the potential value of adaptive suspensions that could sense changing conditions and automatically alter the damping levels to optimize the ride over all terrain.

Ride performance is a function of surface roughness, and shock performance is a function of obstacle height. Consequently, the distribution of surface roughness and of obstacles in the area of operation is a significant factor on overall ride and shock performance. It is wasted time and money to design a vehicle to perform well in very rough terrains and over high obstacles if such conditions are only rarely found in the intended operating areas.

Sprocket horsepower per ton is definitely a prominent factor in mobility/agility performance. Yet it is obvious that a vehicle with high sprocket horsepower per ton and poor suspension will be able to use that power only on smooth terrain surfaces where ride and shock are not limiting factors. Likewise, the mobility/agility advantages of high horsepower per ton are quickly diminished in deformable soils if the vehicle's ground pressure does not provide sufficient flotation to prevent excessive sinkage and soil motion resistance; or in curves and sharp turns during evasive maneuvers if the vehicle's center of gravity is too high for stability; or if the vehicle's dimensions prevent effective maneuvering in the dense forests, such as those found in Germany and certain tropical areas of military interest.

There are no risk areas with vehicle-soil physics.

A principal concern was to determine if the soil motion resistance increased significantly at high speeds in a manner similar to the exponential increase in resistance offered by water to high-speed boats. If the increase was significant, there would be practical limits on power trains beyond which large increases in motion resistance would largely offset power increases, resulting in only small gains in speed. Until this program, power trains in cross-country vehicles had not permitted speeds where such soil resistance rate effects, if they existed, were encountered.

Figure 2 illustrates the effects of soil motion resistance on speeds. The plots depict the motion resistance coefficients R/W (motion resistance to gross vehicle weight ratio) as a function of speed for two distinct soil conditions. The total resistance, the resistance on a firm, level surface, and the resistance due solely to the soil are shown for both the measured and predicted relations. The most important observation from these data

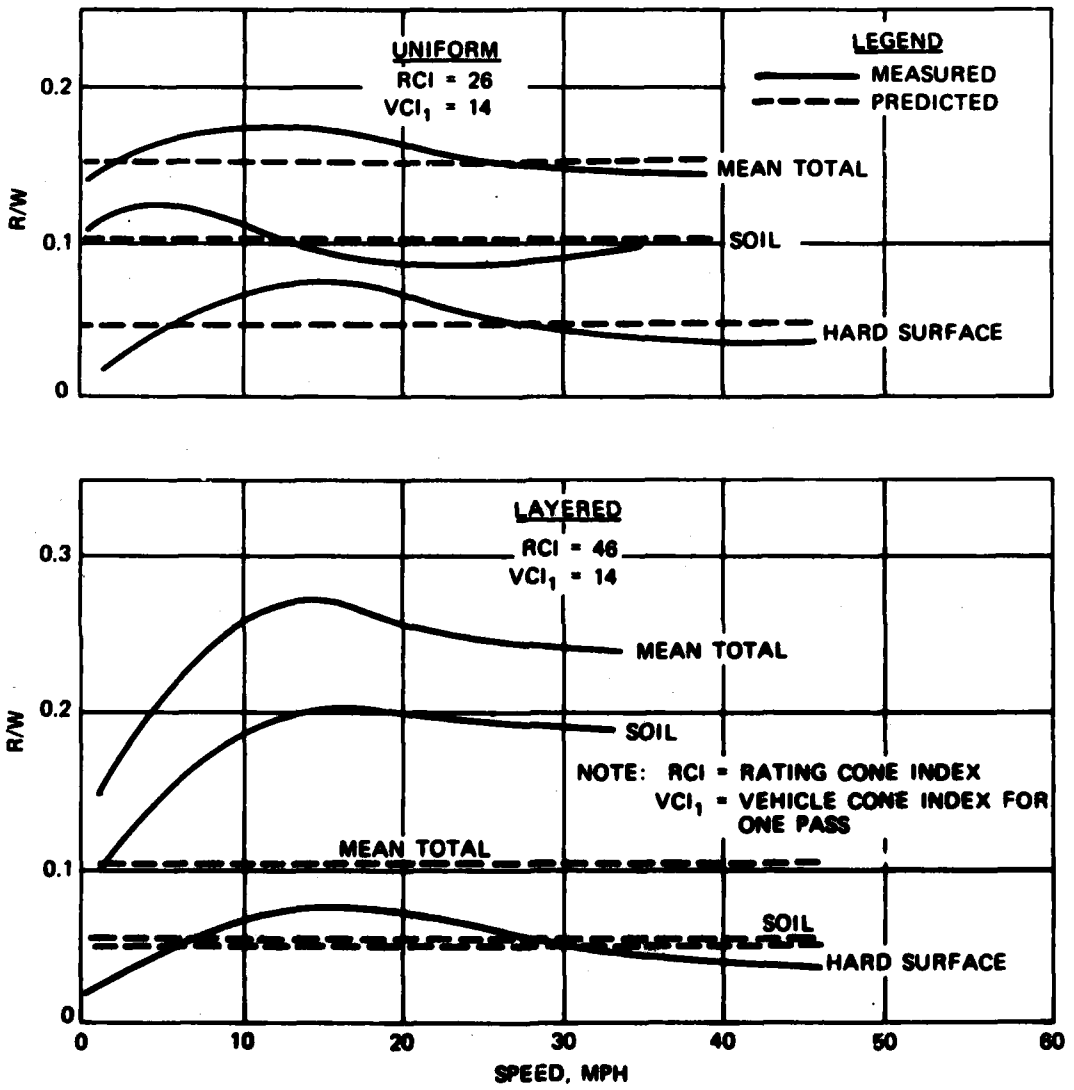


Figure 2. Comparison of measured and predicted motion resistance coefficients versus speed for two soil types (M113 HOTROD)

is that over the speed range from 10 to 30 or 40 mph there appears to be no significant increase in motion resistance; i.e., up to at least 40 mph there is no evidence that increased power will not provide proportionally increased speeds in normal weak soil conditions. The upper plot in Figure 2 shows the average results of four tests run in a soft, sticky soil (rating cone index (RCI) = 26) and compares with the prediction made using the present vehicle cone index (VCI) methods and relations (6). The correlation is good. It is not good for the medium-strength layered soil condition (RCI = 46) shown in the lower plot. It is predicted that R/W = 0.10 . The test logs note that in addition to a firm layer about 5 in. below the surface, the soil in these tests was extremely sticky and tended to clog the tracks and that the test area (continuously flooded for several months just before testing) was "spongy," suggesting that some viscoelastic response of the soil was absorbing substantial energy. These factors would increase the actual measured motion resistance.

Trained military tank drivers will use the increased mobility available in high-performance tracked vehicles.

A comparison of the performance between the professional WES drivers and military drivers was used to determine the degree that trained military drivers would exploit the increased mobility capability of the HIMAG chassis. It would be wasted effort and money to design and build vehicles with 50 or 60 percent increase in mobility capability if military drivers only use 10 or 15 percent of the increased capability. Because the WES professional drivers had been driving the HIMAG vehicle in the previous engineering tests for more than five months just before these tests, they were considered able to exploit the maximum performance capability of the vehicle. Their performance was used as the reference for comparing performance of the military drivers. The evaluations were made from tests with two of the best HIMAG configurations, along with the M60A1 MBT and M113A1 APC for reference, over a rugged 20-km test course, which was composed of 189 distinctly different segments of terrain and 5 general terrain types. There were two groups of military drivers--a group that was familiar with the test course and a group that had never seen the course. All were equally well trained in driving the HIMAG chassis.

The following tabulation shows a comparison of the percent of the WES drivers' speeds achieved by the drivers familiar with the course and the drivers unfamiliar with the course in each terrain type and over the entire 20-km course. For the two HIMAG configurations the familiar drivers reached 90 to 95 percent of the speeds achieved by the WES drivers over the entire course. The unfamiliar drivers achieved 87 percent for the HIMAG 5, but the somewhat unstable behavior of the lighter, tail-heavy HIMAG 2 had a significant influence on those drivers not familiar with the course and they achieved only 79 percent of the WES drivers' speeds.

Percent of WES Drivers' Speeds Achieved by Familiar and Unfamiliar Drivers

Vehicle	Secondary Road		Dirt Trails		Hog Hollow		Pipe-line		Tank Trails		Entire Course	
	F	U	F	U	F	U	F	U	F	U	F	U
HIMAG 5	95	85	92	78	94	79	98	81	98	80	95	79
HIMAG 2	96	99	99	83	88	83	92	85	82	80	90	87
M113A1 APC	97	95	90	90	68	68	89	83	89	84	89	86
M60A1 MBT	95	92	88	81	92	82	86	76	79	75	88	81

The relative performance results show that the military drivers actually exploited more of the available mobility capability from the two HIMAG configurations than from the two contemporary vehicles. These results clearly illustrate that trained military drivers (familiar with the area) will use 90 to 95 percent of the available HIMAG level mobility in tracked vehicles.

Level of mobility is a multiparameter definition.

Level of mobility is a multiparameter definition. The single parameter definition, horsepower-per-ton, often misused to describe the M60A1 MBT, M1 MBT, and HIMAG levels of mobility, is not adequate. Eleven principal factors that include both vehicle and terrain characteristics have been identified that limit mobility. These factors listed below are also used by the AMM to predict speed (4).

a. Insufficient soil strength.
b. Insufficient traction.
c. Obstacle interference.
d. Combination of terrain factors.

} NOGO factors

e. Ride (surface roughness).
f. Soil/slope resistance.
g. Visibility.
h. Maneuverability (through forests or around obstacles).
i. Vegetation (override resistance).
j. Shock (obstacle negotiation).
k. Linear features (streams, ditches, embankments, etc.).

An example of specifying a given level of mobility based on seven performance criteria that involve the eleven factors above is shown in Table 1. Comparisons of performance between the M1 MBT and the HIMAG 5 are shown for those criteria where data were available. A specified level of mobility, such as the M1 MBT mobility and the HIMAG mobility, can be rather accurately defined in terms of the combined minimum acceptable levels for each of the seven performance criteria, but not by using any single criterion.

PART III: RESULTS OF MODEL SIMULATIONS

Lightweight combat vehicles can be designed with mobility/agility equal to or greater than that of the M1 MBT.

Tables 2 and 3 compare the relative performance of eight selected vehicles for four distinct types of mobility in both dry and wet conditions in West Germany and Middle East terrains, respectively. The vehicles are ranked in each case according to performance. The four mobility types--dash, traverse, maneuver, and cross-country--are representative of those often encountered in tactical situations. The relative performance of the vehicles varies according to the type of mobility, the area of operations, and the terrain conditions. The variations show no particular pattern with respect to gross vehicle weight or sprocket horsepower per ton. Observe that:

a. Generally the HIMAG was a top performer except in the German cross-country terrain where its size severely restricted maneuverability through the denser German forests. The M1 MBT encountered the same problem.
b. In most cases all the lighter concept vehicles outperformed the M1 MBT.
c. The M1 MBT demonstrated excellent maneuver performance (in open, level terrain) except in wet German terrain where its performance fell below that of the CFV.
d. The M113A1 APC and the M60A1 MBT were consistently the worst performers.

These results demonstrate that the M1 MBT always outperforms the two contemporary vehicles but generally falls below the performance levels of the HIMAG and the still lighter concept vehicles. However, subsequent war gaming indicated these differences between the M1 MBT and the concept vehicles were not tactically significant. The results also reflect that vehicle performance depends upon the combined effects of the vehicle, the mission, and the terrain and does not vary directly with weight or horsepower. Consequently, with proper attention to design, a lightweight armored vehicle in the 16- to 20-ton range can achieve or surpass the mobility/agility performance of the M1 MBT.

High mobility/agility provides an increased hit-avoidance capability, but the reduced effectiveness to fire-on-the-move while maneuvering violently may result in only a marginal payoff in survivability.

The results of the hit-avoidance tests revealed that a vehicle capable of performing fast, quick maneuvers can gain an additional measure of hit avoidance (7). The major payoff in high mobility/agility vehicles is in

performing fast, dash-to-cover tactics. The principal components in reducing hit probability are the time available to engage the target and the aiming errors. Figure 3 shows a comparison of the dash performance between the M113 HOTROD and the contemporary M113A1 APC in a medium-strength soil (RCI = 46). The maximum speed of the M113 HOTROD was 49 mph compared with 23 mph for the contemporary M113A1 APC. More important is the significant difference in the rate of acceleration. For example, the time required to accelerate from a standing start to 20 mph is only 2.9 sec for the M113 HOTROD compared with 33.0 sec for the M113A1 APC. This quick acceleration permits abrupt speed changes, rapid stops to return fire, and quick starts, and may be more important than the maximum achievable speed per se.

In the simplest sense, against opposing guns, a maneuvering vehicle moves out of the way of a projectile already in flight causing what is referred to as target-induced error. Likewise, a fast, agile target affects the ability of a gunner to accurately track the target in his sight. The gun turret drive and fire-control computer system are also affected. This type of error, which occurs before the round is fired, is referred to as system-induced error. Finally, the fast, agile target reduces exposure time to opposing gunners. These three factors--increased target-induced error, increased system-induced error, and decreased exposure time created by a fast, agile maneuvering vehicle decrease the probability of being hit. Further, a maneuver that minimizes exposure time while maximizing accelerations seen by the firer could be considered optimal (8). However, reduced capability to fire-on-the-move effectively while maneuvering violently may significantly counter the gains in hit-avoidance and result in little net payoff for the latter tactic.

PART IV: CONCLUSIONS

Based on the information presented in this study, it is concluded that:

a. With careful attention to design balance, lightweight combat vehicles can be developed with mobility/agility performance equal to or greater than the M1 MBT.
b. Increases in performance beyond M1 MBT levels, possible with power train components available in the near future, are relatively small and not tactically significant.
c. Increased installed horsepower, up to at least 29 hp/ton at the sprocket, will pay off in increased mobility/agility performance, even in relatively weak soils, provided other design features are kept in balance.
d. Such increases can be achieved in properly designed combat vehicles ranging from 16- to 58-ton gross weight.
e. Properly trained military drivers will achieve more than 90 percent of the mobility available in high-performance tracked vehicles.

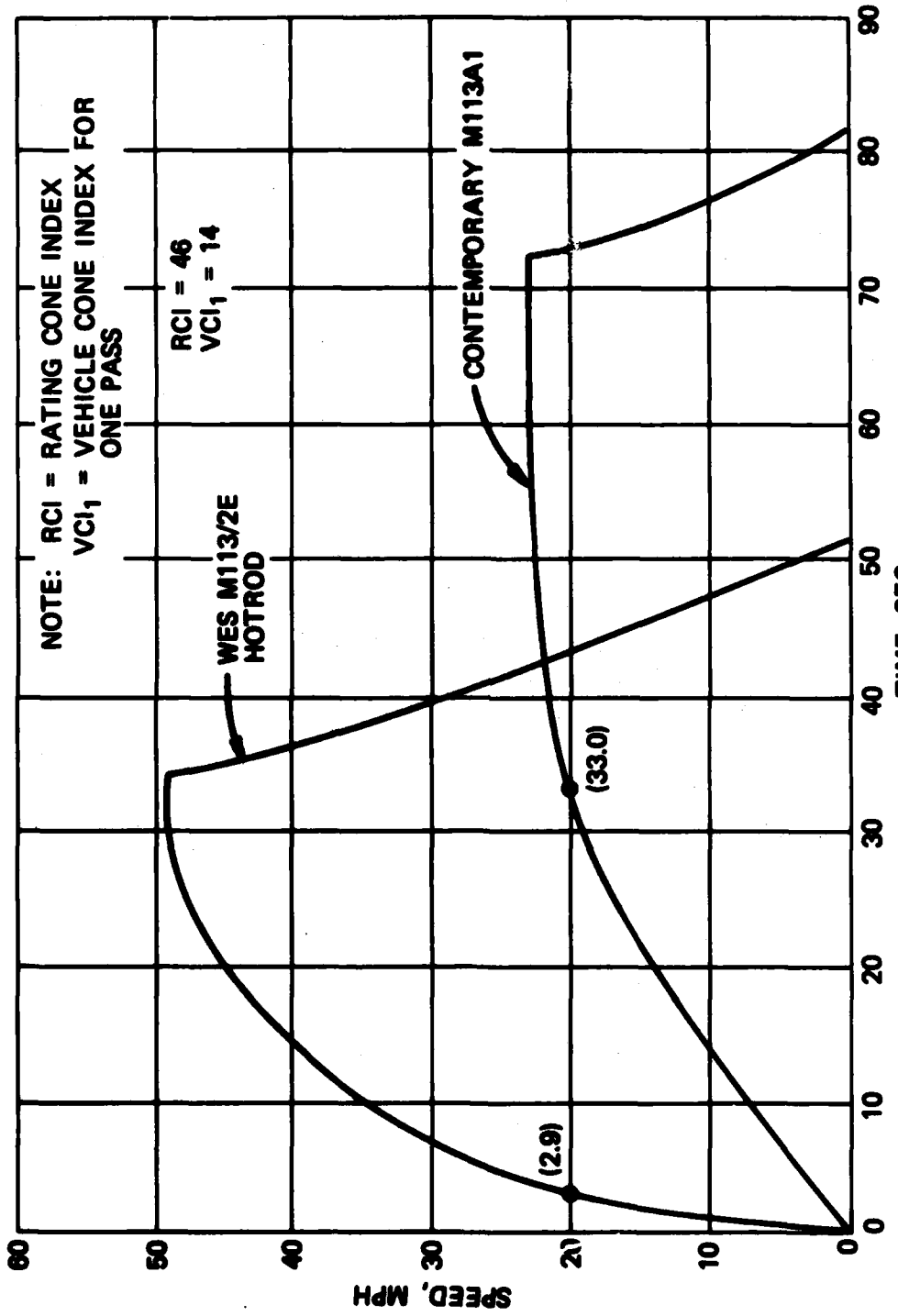


Figure 3. Comparison of acceleration-deceleration performance of WES M113/2E HOTROD and standard M113A1 on the same soil strength

f. A fast, agile maneuvering vehicle provides an increased hit-avoidance capability, but the reduced effectiveness of fire-on-the-move while maneuvering violently may result in only a marginal payoff in survivability.

REFERENCES

1. Schreiner, Barton G., and Green, Charles E. 1980. "Test Areas for High Mobility/Agility (HIMAG) Chassis Test Program," Technical Report GL-80-7, U. S. Army Engineer Waterways Experiment Station, CE, Vicksburg, Miss.
2. Baladi, George Y., and Rohani, Behzad. 1979. "A Terrain-Vehicle Interaction Model for Analysis of Steering Performance of Track-Laying Vehicles," Technical Report GL-79-6, U. S. Army Engineer Waterways Experiment Station, CE, Vicksburg, Miss.
3. Murphy, Newell R., Jr. 1981. "Armored Combat Vehicle Technology (ACVT) Program: Volume I, Results of HIMAG Chassis Tests," (C) Technical Report GL-81-13, U. S. Army Engineer Waterways Experiment Station, CE, Vicksburg, Miss.
4. Nuttall, Clifford J., Jr., and Randolph, Donald D. 1976. "Mobility Analysis of Standard- and High-Mobility Tactical Support Vehicles (HIMO Study)," Technical Report M-76-3, U. S. Army Engineer Waterways Experiment Station, CE, Vicksburg, Miss.
5. Murphy, Newell, R., Jr., and Ahlvin, Richard B. 1976. "AMC-74 Vehicle Dynamics Module," Technical Report M-76-1, U. S. Army Engineer Waterways Experiment Station, CE, Vicksburg, Miss.
6. Rula, Adam A., and Nuttall, Clifford J., Jr. 1971. "An Analysis of Ground Mobility Models (ANAMOB)," Technical Report M-71-4, U. S. Army Engineer Waterways Experiment Station, CE, Vicksburg, Miss.
7. BDM Corporation. 1981. "Hit-Avoidance Test," Monterey, Calif.
8. BDM Corporation. 1981. "Armored Combat Vehicle Technology (ACVT) Program - Test Results on Mobility and Agility and Their Effects on Survivability: Executive Summary," (C) AD No. TRADOC TRMS No. 1-[illegible] TRADOC ACN 36615, Fort Knox, Ky.

Table 1

Minimum Performance Criteria for Specifying a Given Level of Mobility

Principal Vehicle Factors	Performance Criteria	M1	HIMAG 5
Suspension, power train	Ride: Speed, mph, over surface roughness		
	0.5 rms, in.	45.8	55.0
	1.0 rms, in.	45.8	53.0
	1.5 rms, in.	31.5	43.0
Suspension, power train	Shock: Speed, mph, over obstacle height		
	8-in.	45.8	55.0
	10-in.	45.8	55.0
Power train (hp/ton at sprockets)	Dash: Time, sec, for 500-m dash on hard surface	32.4	31.6
GVW, power train, track-ground contact area	Soft-soil: VCI_1	24	18
GVW, power train	Slope: Negotiate a 60% dry slope	Yes	Yes
Vehicle width (1.5 × width = NOGO)	Maneuver (forests): Speed, mph, through forest with 15-ft average spacing	*	*
Center of gravity, tread, track length on ground, power train	Maneuver (agility): Speed-made-good, mph, for maneuver (5 m by 100 m) on hard surface	35	*

Note: GVW = gross vehicle weight; rms = root-mean-square elevations; and VCI_1 = vehicle cone index (minimum soil strength for one pass).

* No experimental data available at this time.

Table 2

Comparison of the Relative Performance of Selected Vehicles at Several Levels of Mobility Based on Speed Predictions in West Germany Terrains

Dry Condition				Wet Condition			
Vehicles	GVW tons	SHP/Ton	Avg Speed mph	Vehicles	GVW tons	SHP/Ton	Avg Speed mph
Dash (500 m)							
HIMAG 5	42	21	38.0	HIMAG 5	42	21	32.2
CON 22 (PIP)*	21	26	33.6	CON 22 (PIP)*	21	26	30.7
CON 3	16	15	32.7	CON 3	16	15	27.4
CON 22	21	15	32.1	CON 22	21	15	26.8
M1	58	18	31.3	M1	58	18	25.3
CFV	23	16	27.9	CFV	23	16	24.5
M113A1	11	12	26.3	M113A1	11	12	18.9
M60A1	52	8	22.1	M60A1	52	8	16.8
Traverse (25 km)							
CON 22 (PIP)*	21	26	21.8	CON 22 (PIP)*	21	26	16.4
CON 3	16	15	19.0	HIMAG 5	42	21	14.7
HIMAG 5	42	21	18.9	CON 22	21	15	14.0
CON 22	21	15	18.1	CON 3	16	15	13.4
M1	58	18	17.0	M1	58	18	13.0
CFV	23	16	16.6	CFV	23	16	12.7
M113A1	11	12	14.4	M113A1	11	12	11.4
M60A1	52	8	12.2	M60A1	52	8	9.2
Maneuver (5 m by 100 m)**							
HIMAG 5	42	21	46.6	CON 22 (PIP)*	21	26	34.3
M1	58	18	43.9	CON 22	21	15	33.0
CON 22 (PIP)*	21	26	41.8	CON 3	16	15	32.5
CON 3	16	15	41.6	HIMAG 5	42	21	32.5
CON 22	21	15	41.2	CFV	23	16	31.6
CFV	23	16	34.0	M1	58	18	29.2
M113A1	11	12	20.5	M113A1	11	12	27.8
M60A1	52	8	21.9	M60A1	52	8	17.8
Cross-Country (AMM, V_{90})†							
CON 22 (PIP)*	21	26	18.7	CON 22 (PIP)*	21	26	14.8
CON 3	16	15	17.3	CON 22	21	15	13.2
CON 22	21	15	17.1	CON 3	16	15	12.2
HIMAG 5	42	21	16.0	HIMAG 5	42	21	11.9
M1	58	18	13.8	CFV	23	16	8.7
CFV	23	16	13.7	M1	58	18	7.3
M113A1	11	12	10.6	M60A1	52	8	4.1
M60A1	52	8	9.0	M113A1	11	12	2.0

Note: GVW = gross vehicle weight, and SHP/Ton = sprocket horsepower per ton.

* Denotes up-powered version of CON 22.

** Denotes maneuvers of 5-m amplitude and 100-m wavelength on only level terrain with mild to medium surface roughness.

† V_{90} represents the average speed in the area after eliminating the worst 10 percent of the terrain.

Table 3

Comparison of the Relative Performance of Selected Vehicles at Several Levels of Mobility Based on Speed Predictions in Middle East Terrains

Dry Condition				Wet Condition			
Vehicles	GVW tons	SHP/Ton	Avg Speed mph	Vehicles	GVW tons	SHP/Ton	Avg Speed mph
			Dash (500 m)				
HIMAG 5	42	21	41.4	HIMAG 5	42	21	38.6
CON 22 (PIP)*	21	26	34.0	CON 22 (PIP)*	21	26	33.2
CON 3	16	15	33.6	CON 3	16	15	32.6
CON 22	21	15	33.3	CON 22	21	15	32.3
M1	58	18	31.5	M1	58	18	30.6
CFV	23	16	28.0	CFV	23	16	27.4
M113A1	11	12	26.4	M113A1	11	12	25.7
M60A1	52	8	22.9	M60A1	52	8	21.9
			Traverse (25 km)				
HIMAG 5	42	21	24.5	HIMAG 5	42	21	24.6
CON 22 (PIP)*	21	26	21.9	CON 22 (PIP)*	21	26	22.5
M1	58	18	21.3	M1	58	18	20.9
CON 3	16	15	19.9	CON 3	16	15	19.4
CON 22	21	15	19.5	CON 22	21	15	19.4
CFV	23	16	17.0	CFV	23	16	16.8
M113A1	11	12	14.3	M60A1	52	8	13.6
M60A1	52	8	13.9	M113A1	11	12	13.4
			Maneuver (5 m by 100 m)**				
CON 22 (PIP)*	21	26	40.3	HIMAG 5	42	21	40.8
HIMAG 5	42	21	40.2	M1	58	18	39.3
M1	58	18	40.2	CON 22 (PIP)*	21	26	37.9
CON 22	21	15	39.7	CON 22	21	15	37.1
CON 3	16	15	39.3	CON 3	16	15	37.1
CFV	23	16	34.0	CFV	23	16	33.6
M113A1	11	12	28.0	M113A1	11	12	28.0
M60A1	52	8	24.5	M60A1	52	8	23.9
			Cross-Country (AMM, V_{90})†				
HIMAG 5	42	21	23.3	HIMAG 5	42	21	23.9
CON 22 (PIP)*	21	26	21.4	CON 22 (PIP)*	21	26	21.8
CON 3	16	15	21.2	CON 22	16	15	21.6
CON 22	21	15	21.0	CON 3	21	15	20.6
M1	58	18	19.8	M1	58	18	19.3
CFV	23	16	16.4	CFV	23	16	16.3
M113A1	11	12	14.1	M113A1	11	12	13.6
M60A1	52	8	13.2	M60A1	52	8	12.6

Note: GVW = gross vehicle weight, and SHP/Ton = sprocket horsepower per ton.

* Denotes up-powered version of CON 22.

** Denotes maneuvers of 5-m amplitude and 100-m wavelength on only level terrain with mild to medium surface roughness.

† V_{90} represents the average speed in the area after eliminating the worst 10 percent of the terrain.

COMPUTATIONS OF PROJECTILE MAGNUS EFFECT AT TRANSONIC VELOCITIES

CHARLES J. NIETUBICZ*, Mr., WALTER B. STUREK, Dr.,
and KAREN R. HEAVEY, Mrs.
U.S. Army Ballistic Research Laboratory/ARRADCOM
Aberdeen Proving Ground, Maryland 21005

I. INTRODUCTION

The accurate prediction or experimental determination of projectile aerodynamics is of significant importance to the shell designer and ballistician. The shell designer requires accurate aerodynamic data for the overall development of new shell. The ballistician is concerned with the development of aiming data and therefore relies heavily on accurate aerodynamic data. Experimental costs have sky rocketed in recent years and have contributed significantly to overall system development costs. Computational techniques are beginning to show promise as a means to alleviate or at least temper these rising development costs by providing relatively low cost computer analysis of new designs. As computer technology increases and machines become faster with larger memory, the use of computational methods in design becomes more of a reality.

The means to compute projectile aerodynamics for all Mach number regimes covered by a given projectile in its flight history has been an area of research actively pursued by the Aerodynamics Research Branch of the Ballistic Research Laboratory. Early work had focused on the supersonic flight regime and, in particular on the accurate prediction of the Magnus force. The Magnus force, which is very small in magnitude (on the order of 1/10 the normal force), is a critical parameter in determining the dynamic stability of shell. The Magnus force is generated by a spin induced distortion of the boundary layer; therefore, correct modeling of the viscous/inviscid interation is critical for accurate computations. The work of Sturek, et. al.[1] has shown that accurate results in the supersonic regime can be obtained for ogive-cylinder projectile shapes. This technique involved separate computations of the turbulent, viscous boundary layer and the outer inviscid flow field. As the projectile shapes were generalized to include boattails, more sophisticated computational techniques had to be employed. These new methods, which solved the thin-layer Navier Stokes equations, were successfully applied to ogive-cylinder-boattail shapes by Sturek and Schiff[2,3]. The solution of the Navier Stokes

equations allows for the simultaneous computation of the viscous/inviscid flow field and thus provided the basis for good Magnus force prediction.

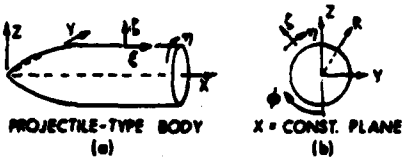
A region of critical aerodynamic behavior occurs in the high transonic regime, $0.90 < M < 1.15$, where aerodynamic coefficients have been found to increase by as much as 100%. This flight velocity regime is both experimentally and computationally difficult. Thus, only a small amount of experimental data are available for design studies and only limited computational studies have been made. An initial attempt to develop a computational capability suitable for Magnus prediction at transonic velocity was made by Nietubicz, et. al[4] whereby the thin-layer Navier Stokes computational technique was applied to standard and hollow projectile shapes at zero degrees angle of attack. Computational results were also obtained for non-spinning projectiles at angle of attack[5]; however limitations of computer resources (CDC 7600) became apparent. In this paper, recent results are presented for Magnus force computations using the Cray 1S computer. Comparisons are made with the earlier CDC computer results and are further compared to some limited experimental data. A discussion of the numerical technique is included. These results represent the first computations of the Magnus effect on projectile shapes in the transonic flight regime.

II. GOVERNING EQUATIONS AND COMPUTATIONAL TECHNIQUE

The set of equations which govern fluid motion are the Navier-Stokes equations. This set of highly non-linear partial differential equations have proven extremely difficult to solve in their entirety. In most instances approximations had to be made before a solution could be attempted. For example, the well known boundary layer equations are derived by applying approximations to the Navier Stokes equations. The present solution technique also makes use of an approximation. The Navier-Stokes equations solved here make use of the thin-layer approximation. That is, the viscous terms are neglected in both the longitudinal and circumferential directions. The viscous terms are retained, however, in a direction nearly normal to the surface where large flow field gradients exist. This formulation retains the momentum equations in all three coordinate directions. The retention of the three momentum equations allows for the computation of separated flow and thus differs significantly from boundary layer assumptions.

The equations solved here are written in a generalized coordinate system. This allows a wide variety of body shapes to be computed using the same basic numerical technique. The notation for the physical and transformed coordinate systems are shown in Figure 1. The three dimensional, transformed, thin-layer Navier-Stokes equations, written in non-dimensional, strong conservation law form are[6]

$$\partial_\tau \hat{q} + \partial_\xi \hat{E} + \partial_\eta \hat{F} + \partial_\zeta \hat{G} = Re^{-1} \partial_\zeta \hat{S} \qquad (1)$$



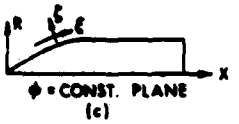


Figure 1. Physical and Transformed Coordinate System

The general coordinate transformations are defined as

$\xi = \xi(x,y,z,t)$ - longitudinal coordinate

$\eta = \eta(x,y,z,t)$ - circumferential coordinate

$\zeta = \zeta(x,y,z,t)$ - near Normal coordinate

$\tau = t$ - time

and

$$\hat{q} = J^{-1} \begin{bmatrix} \rho \\ \rho u \\ \rho v \\ \rho w \\ e \end{bmatrix} \qquad \hat{E} = J^{-1} \begin{bmatrix} \rho U \\ \rho uU + \xi_x p \\ \rho vU + \xi_y p \\ \rho wU + \xi_z p \\ (e+p)U - \xi_t p \end{bmatrix}$$

$$\hat{F} = J^{-1} \begin{bmatrix} \rho V \\ \rho uV + \eta_z p \\ \rho vV + \eta_z p \\ \rho wV + \eta_z p \\ (e+p)V - \eta_t p \end{bmatrix} \qquad \hat{G} = J^{-1} \begin{bmatrix} \rho W \\ \rho uW + \zeta_z p \\ \rho vW + \zeta_y p \\ \rho wW + \zeta_z p \\ (e+p)W - \zeta_t p \end{bmatrix}$$

The viscous matrix, $\hat{S}$, is written as

$$\hat{S} = J^{-1}\begin{bmatrix} 0 \\ \mu(\zeta_x^2+\zeta_y^2+\zeta_z^2)u_\zeta+(\mu/3)(\zeta_x u_\zeta+\zeta_y v_\zeta+\zeta_z w_\zeta)\zeta_x \\ \mu(\zeta_x^2+\zeta_y^2+\zeta_z^2)v_\zeta+(\mu/3)(\zeta_x u_\zeta+\zeta_y v_\zeta+\zeta_z w_\zeta)\zeta_y \\ \mu(\zeta_x^2+\zeta_y^2+\zeta_z^2)w_\zeta+(\mu/3)(\zeta_x u_\zeta+\zeta_y v_\zeta+\zeta_z w_\zeta)\zeta_z \\ \{(\zeta_x^2+\zeta_y^2+\zeta_z^2)[\mu/2(u^2+v^2+w^2)_\zeta+\kappa Pr^{-1}(\gamma-1)^{-1}(a^2)_\zeta] \\ +(\mu/3)(\zeta_x u+\zeta_y v+\zeta_z w)(\zeta_x u_\zeta+\zeta_y v_\zeta+\zeta_z w_\zeta)\} \end{bmatrix}$$

The velocities

$$\begin{aligned} U &= \xi_t + \xi_x u + \xi_y v + \xi_z w \\ V &= \eta_t + \eta_x u + \eta_y v + \eta_z w \\ W &= \zeta_t + \zeta_x u + \zeta_y v + \zeta_z w \end{aligned} \tag{2}$$

represent the contravarient velocities. The non-dimensional velocities U, V, and W are those components in the direction of the transformed coordinates ξ, η, and ζ, respectively. The Cartesian velocity components u, v, w together with the density ρ and total energy per unit volume e are retained as the dependent variables. The local pressure, p, is determined using the relation

$$p = (\gamma - 1)(e - .5\rho(u^2 + v^2 + w^2)).$$

The velocities are non-dimensionalized by the free stream speed of sound a_∞, the density by ρ_∞, and the total energy by $\rho_\infty a_\infty^2$. The additional parameters appearing in equation 1 are: (a) coefficient of thermal conductivity, κ; (b) dynamic viscosity, μ; (c) Reynolds number based on body diameter, Re; (d) Prandtl number, Pr, (e) λ which, based on Stokes hypothesis, is $-2/3\ \mu$.

As mentioned earlier, these equations are written in transformed coordinates; therefore, the various body shapes are introduced through determination of the metric terms ξ_x, η_x, ζ_z, etc.. These terms are formed by a combination of the derivative terms x_ζ, y_ζ, z_ζ, etc., and, together with the transformation Jacobian, allow for variable body geometries. Thus, one of the first steps in performing a computation is the generation of a computational grid which provides the x, y, z points for the metric determination. These points are determined prior to the computations and are not changed with time. Examples of the computational grid used in this study are shown in Figures 2 and 3. A two dimensional slice of the overall

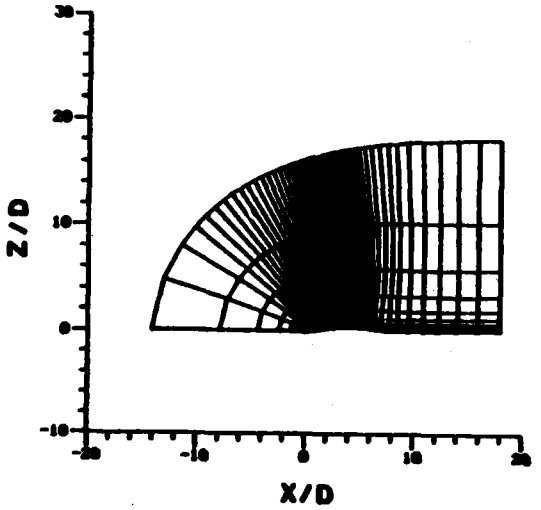


Figure 2. Physical Grid - Total Flow Field

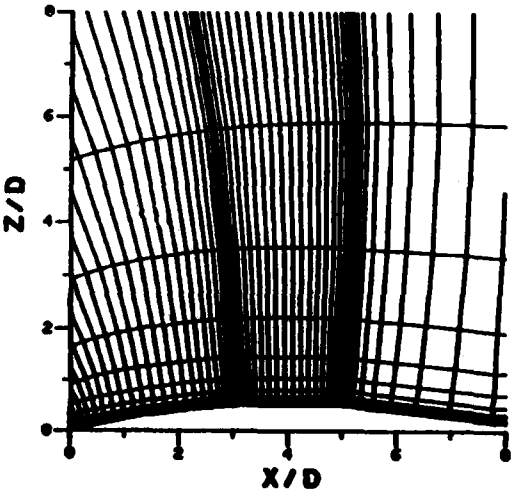


Figure 3. Physical Grid - Expanded View Near Projectile Surface

grid is shown in Figure 2. The upstream, downstream and outer flow field computational boundaries extended approximately 18 body diameters from the body surface. At this distance the flow field should be uniform and the imposed boundary conditions are considered valid. Figure 3 shows an expanded view near the body. The clustering of grid points near the body surface is required in order to resolve the viscous components of the flow field near tne body surface. Due to the lack of sufficient computer storage, judicious use must be made of the limited grid points available. In regions where the viscous effects are not predominant and the flow field changes slowly, the grid points are sparse. Additional grid clustering is used in the longitudinal direction where flow field gradients are expected. The two dimensional grid shown in Figure 2 was rotated about the body axis in 10 degree increments in order to obtain the three dimensional grid required for computations at angle of attack.

As mentioned in the introduction, the Magnus effect is produced by a spin induced distortion of the boundary layer. The computation must therefore be fully three dimensional since no plane of symmetry exists. The boundary conditions used for the computations are:

(1) inner boundary, body surface

$U = W = 0$
$V = \omega$ = non-dimensional angular velocity
ρ - first order extrapolation
p - calculated using the three transformed momentum equations

(2) outer boundary

constant free stream values are used for all variables

(3) downstream boundary

$M < 1$ pressure is fixed at p_∞ and all other variables are extrapolated
$M > 1$ first order extrapolation on all variables

The numerical scheme used for the solution of equation 1 is a fully implicit, approximately factored, finite difference algorithm in delta form as analysed by Beam and Warming[7]. This scheme can be first or second order accurate in time and second or fourth order accurate in the three spatial directions. The solution of the three dimensional equations is implemented by an approximate factorization which allows the system of equations to be solved in three coupled one dimensional steps. This procedure has been utilized in previous applications[2-5] with a high degree of success. Additional details of the numerical method, computational algorithm and boundary conditions can be found in Reference 6.

The turbulence model employed is an algebraic eddy viscosity model as developed by Baldwin and Lomax[8]. This same model has yielded excellant results for Magnus effects at supersonic velocities[3].

As indicated in equation 1, this solution technique involves solving the time dependent Navier-Stokes equations. The procedure is started by assuming a uniform, free stream solution for all grid points in the computational domain. The calculation then marches in time until a steady state solution is obtained. The implicit technique used here allows for large time steps to be taken which helps to reduce the computation time.

III. MODEL GEOMETRY AND EXPERIMENT

One means of establishing the computational accuracy of a numerical scheme is through comparisons with available experimental data. The model used for the experiment and computational study presented here is an idealization of a realistic artillery projectile geometry. The experimental model shown in Figure 4 consists of a 3 caliber (1 caliber = maximum body diameter) sharp secant-ogive nose, a two caliber cylindrical mid-section, and a 1 caliber 7° conical afterbody or boattail. A similar model was used for the computational studies with the only difference being a 5% rounding of the nose tip. The nose tip rounding was done for computational efficiency and is considered to have little impact on the final integrated forces.

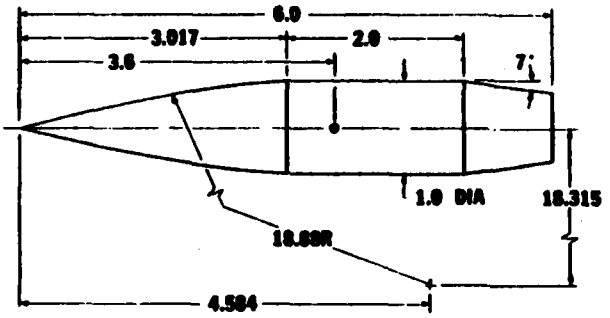


Figure 4. Model Details

The surface pressure experimental data[9] used for comparison in this paper was obtained in the NASA Langley 8 Ft. Pressure Tunnel. The test conditions of 1 atmosphere supply pressure and 320°K supply temperature resulted in a Reynolds number of 4.5×10^6 based on model length. The model was instrumented with pressure ports at 15 longitudinal stations. Pressure data and aerodynamic force data were obtained at Mach numbers of 0.91 to 1.2 and angles of attack from 0 to 10 degrees.

Additional tests were conducted at the Naval Surface Weapons Center for similar tunnel conditions. This test utilized the same model which now included the capability for spin. The model was spun to 500 rev/sec; and Magnus force and moment measurements were obtained while the model coasted down to zero spin. Aerodynamic force measurements were obtained at Mach = 0.91 for α = 0.0° to 10°.

IV. COMPUTER RESOURCES

The initial computations were carried out on the BRL CDC 7600 computer. This machine has a maximum large core storage capability of approximately 380,000 useable words. This limited the maximum computational grid to be 21,600 points (60 longitudinal, 20 normal and 18 circumferential). As previously mentioned the solution is marched in time until a steady state condition is obtained. A typical converged solution required approximately

1000 time steps. The CDC computation ran at a speed of 12.96×10^{-4} sec/time step/point. This resulted in 7.78 hour computation time for one converged solution at one set of conditions. As will be shown later this computational grid did not provide adequate grid resolution.

In order to obtain improved grid resolution and faster computational speed, the computer code was placed on the Cray 1S computer at Kirtland Air Force Base. Operationally, the connection to the Cray was made via the ARPANET from Aberdeen Proving Ground. The ARPANET is a Department of Defense digital switching network which allows terminals and geographically separated computers to communicate. The Cray 1S computer is a vector processor and has a demonstrated speed advantage over the CDC 7600 of 10 in the vector mode. The present computations have not yet taken full advantage of the vector capability, however. The same computation described earlier ran at a speed of 5.2×10^{-4} sec/time step/point on the Cray which resulted in a 2.5x increase in speed for the same grid. An additional advantage in using the Cray is the increased storage capability. The present configuration allows for 1 million words of storage with an expansion capability to 4 million words. The computational grid was therefore increased to 51,840 grid points (60 longitudinal, 24 normal and 36 circumferential). It was felt that increased grid resolution in the circumferential direction was required for determination of the Magnus force components. With the increased grid and same convergence criterial used earlier, the final computational time was 7.49 hours on the Cray. With full vectorization, it is anticipated that this time can be reduced to 1.5 hours. With the advent of even faster machines (Cray 2, with 32 million words of storage and speeds 6-12 times faster than the Cray 1) the eventual computational time can potentially be reduced to 15 minutes.

V. RESULTS

The results to be presented are in the form of pressure coefficients, velocity profiles and aerodynamic coefficients. The computed results are compared with the experimental data for $M = 0.91$, $\alpha = 2.0°$, and $Re = 4.5 \times 10^6$. The Magnus measurements were compared at $PD/U_\infty = 0.39$ where P = angular velocity, D = maximum body diameter and U_∞ = free stream velocity. Surface pressure comparisons are made using the Langley data, while the aerodynamic coefficient comparisons are made with the NSWC data[10].

a. Surface Pressure Coefficient

The surface pressure coefficient, $C_p = (p - p_\infty)/.5\rho_\infty U_\infty^2$, on the leeward ray is shown as a function of longitudinal position in Figure 5. The experimental data, and computational results from both the CDC 7600 and Cray 1S computers are shown. The CDC results (dashed line) show marginal

agreement with the experimental data (circles) over the entire projectile surface. The expansions and recompressions which occur near the ogive-cylinder and cylinder-boattail junction have not been adequately captured by the computation. The solid line in Figure 5 is the results from the computation on the Cray 1S computer where the grid has been expanded to 36 points circumferentially, 24 points normal to the surface and the 60 longitudinal points have been redistributed. The agreement with the experimental data has improved significantly. Some discrepancy is still apparent however, on the cylinder and on the boattail. As can be seen from Figure 3 the grid points in the vicinity of the boattail have been severely stretched and is the apparent cause of the discrepancy. All remaining results to be presented were obtained from the Cray 1S. A comparison of the windward and leeward pressure distribution is shown in Figure 6. The ogive experiences high pressure along the windward side whereas the high pressure for the boattail is on the leeward side. On the ogive, the high pressure on the windward ray causes an upward force. On the boattail, the high pressure on the leeward ray causes a negative or downward force. This condition forms a couple about the center of gravity and contributes to the critical aerodynamic behaviour which occurs at transonic velocities.

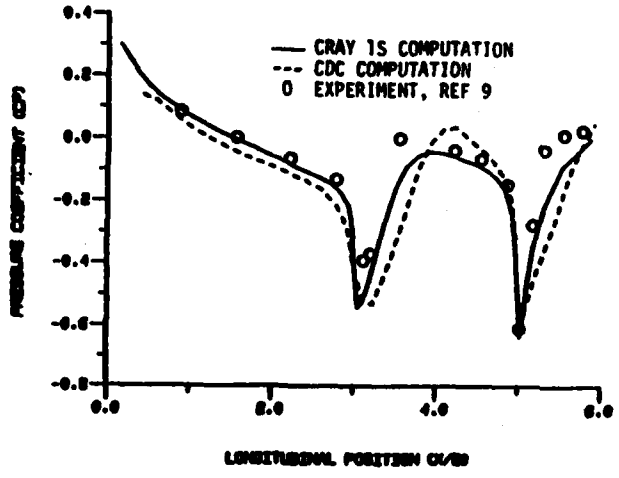


Figure 5. Surface Pressure Coefficient, Experiment and Computations, M = 0.91, α = 2.0°

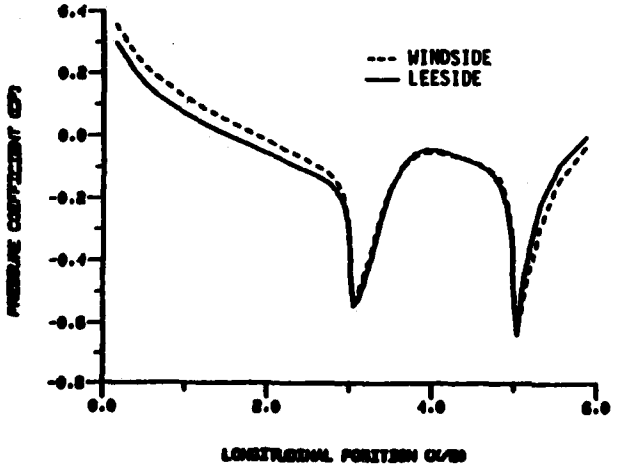


Figure 6. Windward and Leeward Surface Pressure Coefficient, M = 0.91, α = 2.0°

Comparisons between computation and experiment for the circumferential surface pressure distribution are shown in Figures 7 and 8 for X/D stations 1.56 and 5.19 respectively. As seen in these figures, the computation predicts the correct trend of the data while the actual magnitude shows a deviation. Upon close evaluation of the circumferential pressure distribution, a slight asymmetry is detected about the $\phi = 180$ plane. This condition contributes to the expected side force. The grid used for these calculations is by no means optimum and additional computational experimentation is required. The results do, however, indicate the potential for obtaining satisfactory aerodynamic coefficients. This optimism is based on results which have been obtained[5] previously for an axisymmetric computation. Figure 9 shows a comparison between experimental data and a computation using a grid consisting of 80 longitudinal points and 40 points in the near normal direction. The agreement is excellent and indicates the quality of results which should be possible for the three dimensional cases given adequate grid resolution.

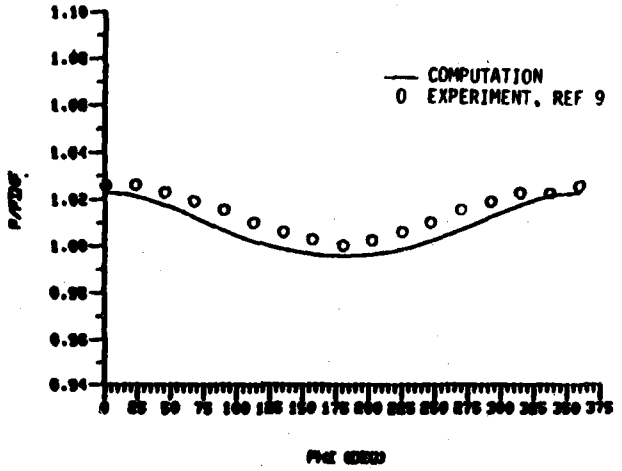


Figure 7. Circumferential Pressure Distribution, $M = 0.91$, $\alpha = 2.0°$, $X/D = 1.56$

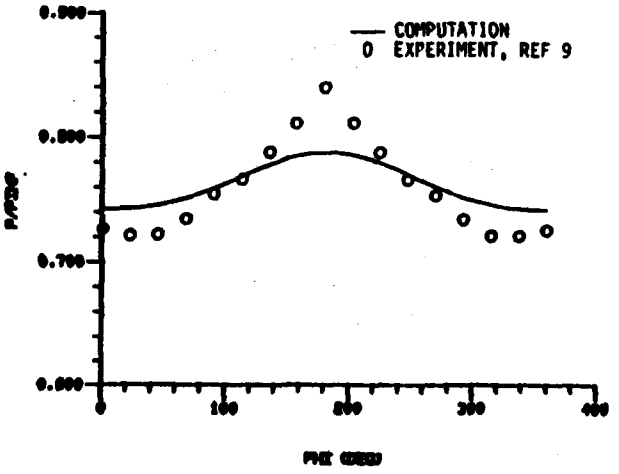


Figure 8. Circumferential Pressure Distribution, $M = 0.91$, $\alpha = 2.0°$, $X/D = 5.19$

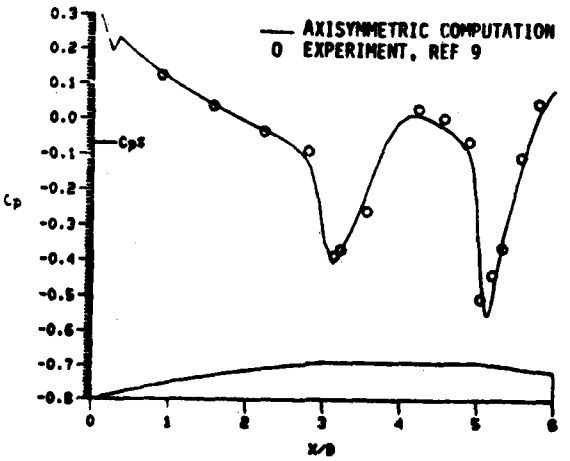


Figure 9. Surface Pressure Coefficient, Experiment and Axisymmetric Computation, $M = 0.96$, $\alpha = 0.0°$

A primary purpose of this research effort is the development of a capability for the prediction of aerodynamic coefficients and, in particular, the capability to compute the Magnus effect. As noted, the Magnus effect is a viscous phenonena associated with the spinning projectile. Therefore, in order for a computational technique to predict this effect, it must adequately compute the longitudinal and circumferential wall shear stress for the spinning projectile at angle of attack. The experimental determination of the u, v, and w velocity distribution is especially difficult at transonic velocities. Although no experimental data are available for comparison, the computed circumferential velocity distributions are shown in Figures 10 and 11 for X/D = 4.22 and X/D = 5.50, respectively. A significant asymmetry can be seen in the velocity distributions at ϕ = 90° and

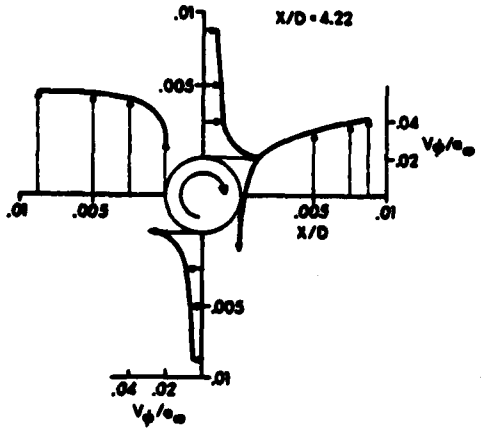


Figure 10. Circumferential Velocity Profiles for ϕ = 0°, 90°, 180°, 270° at M = 0.91, α = 2.0°, X/D = 4.22

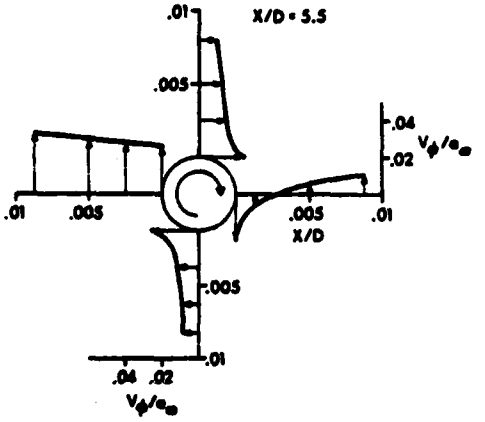


Figure 11. Circumferential Velocity Profiles for ϕ = 0°, 90°, 180°, 270° at M = 0.91, α = 2.0°, X/D = 5.5

270°. At ϕ = 90° cross flow velocity caused by the angle of attack is in the same direction as the wall velocity. At ϕ = 270°, the outer cross flow velocity opposes the wall velocity. The circumferential velocity of profiles at ϕ = 0° and ϕ = 180° are equally affected by the surface spin. Figure 11 shows the velocity distribution at a station midway on the boattail. The circumferential velocity distribution at ϕ = 90° and ϕ = 270° has changed significantly from that shown in Figure 10. On the boattail, the decreasing body diameter results in the surface velocity decreasing in magnitude. However, the boundary layer thickness in this region increases and the effect of surface spin is seen to persist further out.

VI. PITCH PLANE AND MAGNUS COEFFICIENTS

An integration of the pressure and viscous forces has been carried out in order to determine the aerodynamic coefficients. The sign convention used for the coefficients is shown in Figure 12. The results in Figures 13-17 are plotted as a function of longitudinal position and thus show how the force develops over the length of the projectile. Figure 13 is a plot of the normal force coefficient and shows the rapid increase in normal force which occurs on the ogive portion of the projectile. The cylinder portion should produce no significant additional normal force; however, the computation indicates a slight increase in normal force here. The reversal in the direction of the force on the boattail can be clearly seen as the accumulated normal force decreases over the length of the boattail. The experimental normal force coefficient, indicated by the circle, shows very good agreement with the computation.

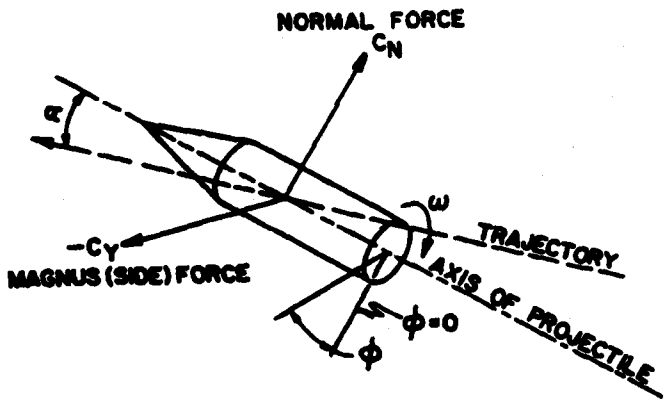


Figure 12. Aerodynamic Coefficient Sign Convention

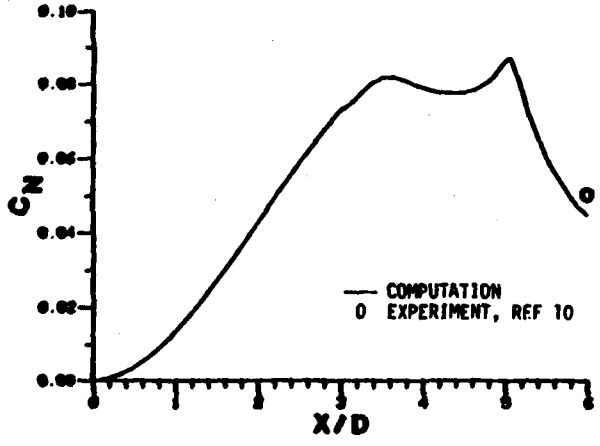


Figure 13. Normal Force Coefficient Along the Projectile, Computation and Experiment, M = 0.91, α = 2.0°

The spin rate of typical artillery shell is of the order of 300-500 rev/sec. As mentioned previously, the Magnus effect results from a spin induced distortion of the viscous boundary layer which occurs for artillery shell at angle of attack. Previous studies[1,3] have shown that the Magnus effect consists of the sum of the boundary layer displacement effect (asymmetric surface pressure distribution) plus the viscous wall shear stress contributions

$\tau_x = \mu(\frac{du}{dy})_{y=0}$ and $\tau_\phi = \mu(\frac{du}{dy})_{y=0}$. The development of all three components of the Magnus force are shown in Figures 14, 15, and 16 respectively, as a function of longitudinal position. Both the longitudinal and circumferential components (Figures 14 and 15) are seen to be of the order of 10^{-5} and 10^{-4}, respectively. The pressure component (Figure 16) is of the order 10^{-2}. The dominant component of the transonic Magnus effect is, therefore

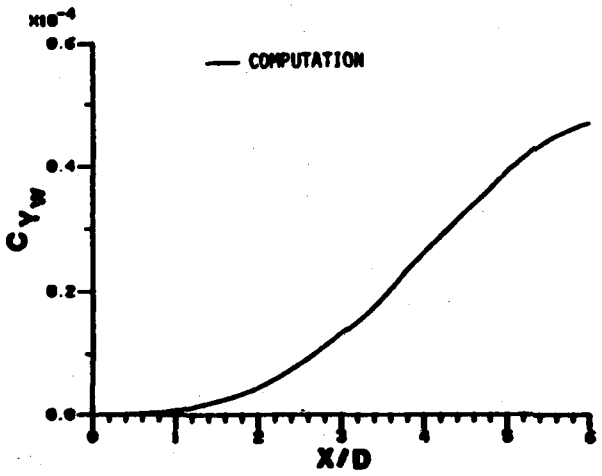


Figure 14. Circumferential Wall Shear Contribution to Magnus Force, M = 0.91, α = 2.0°, ω = 333 rev/sec

seen to be the boundary layer displacement effect, Cy_p.

Additionally, the largest portion of the total Magnus effect is seen to develop on the boattail where the viscous boundary layer reaches its maximum thickness. This is the same qualitative behavior reported by Sturek, et al[1,3] for supersonic flow. The total Magnus force ($Cy = Cy_u + Cy_w + Cy_p$) is shown in Figure 17 compared to the experimental measurement. Considering the small magnitude of the Magnus force and the agreement achieved for the normal force, the quantitative agreement between the computation and experiment is regarded as very good. The experimental Magnus force measurements were obtained in a wind tunnel not specifically designed for transonic flow and are considered to be of good qualitative value but of questionable quantitative value. Additional computations at various transonic Mach numbers and reliable experimental data are required before a full assessment of the computational technique can be made. This first result, however, for predicting the Magnus effect at transonic velocity is considered very encouraging.

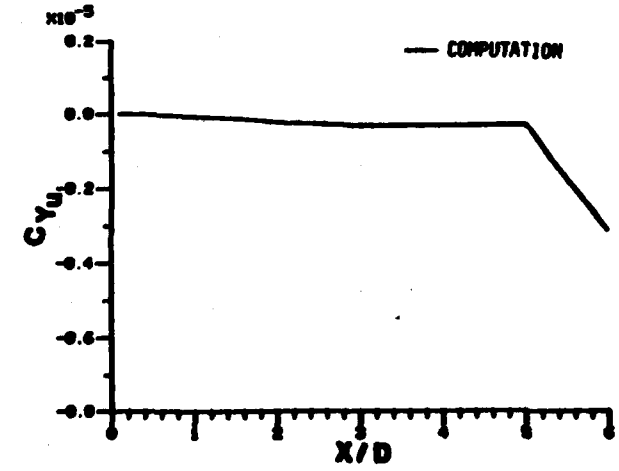


Figure 15. Longitudinal Wall Shear Contribution to the Magnus Force, M = 0.91, α = 2.0°, ω = 333 rev/sec

VII. SUMMARY

The research effort presented in this paper is part of an overall program to develop a sophisticated predictive capability for projectile aerodynamics. The pacing requirement for this capability is the determination of the Magnus force in the transonic flight regime.

An implicit finite difference code, which solves the unsteady thin-layer Navier-Stokes equations, has been applied to a projectile shape at α = 2.0°,

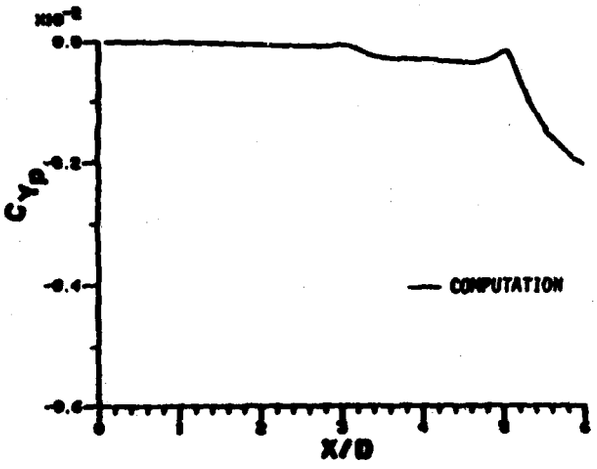


Figure 16. Circumferential Pressure Contribution to the Magnus Force, M = 0.91, α = 2.0°, ω = 333 rev/sec

M = 0.91. The solution was marched in time until a steady state result was obtained. Computations were first performed on a CDC 7600 using a finite difference grid of 21,600 points and required 7.78 hours of computer time. Increased grid resolution with faster computational speed per grid point was obtained by performing the computation on a Cray 1S vector computer.

The computations have been compared to experimental surface pressures and aerodynamic force coefficients. The circumferential velocity distribution, presented for two axial locations, showed the significant interaction between the cross flow velocity resulting from angle of attack and the body surface velocity. Experimental velocity profile data, which are very difficult to obtain for a spinning model at transonic speeds, are required to fully assess the computational results. The normal and Magnus force coefficients have been shown to be in good quantitative agreement with experimental data. The individual components of the Magnus force have additionally been presented and indicate qualitatively good results. The need for additional grid resolution or adaptive grid techniques[11] have been identified as a further requirement to achieve more accurate predictions. Good quality experimental, transonic Magnus data is also required for future code validation.

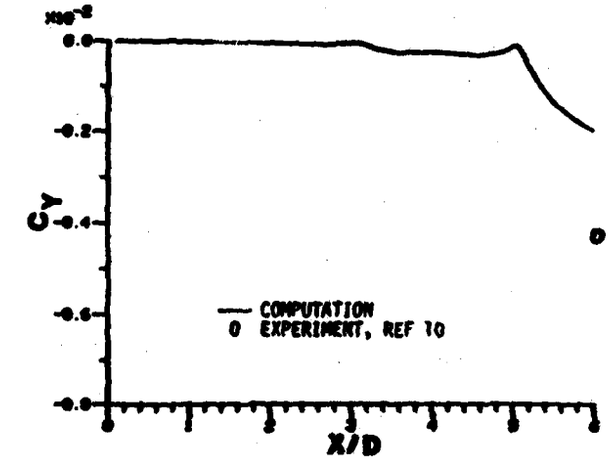


Figure 17. Total Magnus Force Along Projectile, Computation and Experiment, M = 0.91, α = 2.0°, ω = 333 rev/sec

The present results indicate that the thin-layer Navier-Stokes computational technique, in conjunction with enhanced computer technology, has the potential of providing the capability to accurately predict the aerodynamic behavior of spinning shell at transonic velocities, including the Magnus effect.

REFERENCES

1. Sturek, W.B., et al., "Computations of Magnus Effects for a Yawed, Spinning Body of Revolution", AIAA Journal, Vol. 16, No. 7, July 1978, pp. 687-692.

2. Schiff, L.B., and Sturek, W.B., "Numerical Simulation of Steady Supersonic Flow Over Cone Ogive-Cylinder-Boattail Body", AIAA Paper No. 80-0066, January 14-16, 1980.

3. Sturek, W.B., and Schiff, L.B., "Computations of the Magnus Effect for Slender Bodies in Supersonic Flow", AIAA Paper No. 80-1586-CP, August 11-13, 1980.

4. Nietubicz, C.J., Pulliam, T.H., and Steger, J.L., "Numerical Solution of the Azimuthal-Invariant Thin-Layer Navier-Stokes Equations", AIAA Journal, Vol. 18, No. 12, December 1980, pp. 1411-1412.

5. Nietubicz, C.J., "Navier-Stokes Computations for Conventional and Hollow Projectile Shapes at Transonic Velocities", AIAA Paper No. 81-1262, AIAA 14th Fluid and Plasma Dynamics Conference, Palto Alto, CA, 1981.

6. Pulliam, T.H., and Steger, J.L., "On Implicit Finite-Difference Simulations of Three-Dimensional Flow", AIAA Journal, Vol. 18, No. 2, February 1980, pp. 159-167.

7. Beam, R., and Warming, R.F., "An Implicit Factored Scheme for the Compressible Navier-Stokes Equations", AIAA Paper No. 77-645, June 1977.

8. Baldwin, B.S., and Lomax, H., "Thin Layer Approximation and Algebraic Model for Separated Turbulent Flows", AIAA Paper No. 78-257, January 1978.

9. Kayser, L.K., and Whiton, F., "Surface Pressure Measurements on a Boattailed Projectile Shape at Transonic Speeds", to be published as a BRL Report.

10. Kayser, L.D., Ballistic Research Laboratory/ARRADCOM, Aberdeen Proving Ground, Maryland 21005, private communications.

11. Dwyer, H.A., Kee, R.J., and Sanders, B.R., "Adaptive Grid Method for Problems in Fluid Mechanics and Heat Transfer", AIAA Journal, Vol. 18, October 1980. pp. 1205-1212.

EXPLOSIVES DETECTION SYSTEMS EMPLOYING BEHAVIORALLY MODIFIED RATS AS SENSORY ELEMENTS

RAYMOND V. NOLAN, D.Sc.
US ARMY MOBILITY EQUIPMENT RESEARCH AND DEVELOPMENT COMMAND
FORT BELVOIR, VIRGINIA

I. INTRODUCTION

Synopsis of the Experiment. The research program was concerned with improving the operational reliability of animals employed as the sensory element in explosives detection systems. Four basic theses were investigated and the research results in each of these areas represented an original contribution to the developing field of biosensor detection systems. The theses are: (A) It is postulated that rats can detect, via their olfactory sensory function, some component of the military explosive 2, 4, 6-trinitrotoluene (TNT) in minute quantities, (B) It is postulated that if A above is true, then proper behavioral conditioning can cause the animals to operantly signal their awareness of the presence of some effluent of TNT in the ambient air, (C) It is postulated that if B is true, then it should be possible to train several subjects simultaneously using a semi-automatic test station employing operant and classical conditioning paradigms, (D) It is postulated that if B is true, then it should be possible to classically condition the test subjects to anticipate the reward that was heretofore forthcoming only as a consequence of an operant. The state of cerebral arousal resulting from the anticipation should then result in an anticipatory evoked spectral change (AESC) in the cortical electroencephalogram (EEG) of subjects so conditioned.

The research protocol spanned several disciplines including physiology, behavioral science, electronics engineering, biochemistry and explosives technology. Unfortunately, space limitation precludes significant discussion of the relative merits of Physico-Chemical, Electromagnetic (the existing explosives detecting techniques), and Biosensor systems. This limit also narrows the discussion of the overall experiment to a rather short discussion of the results. The matters of subject selection, behavioral training and instrumentation are amply discussed in MERADCOM Report 2343 dated December 1981, entitled "Investigation of Behaviorally Modified

Rats for Use in Explosives Detection Systems" and in the author's doctoral dissertation (same title) dated July 1981 GW University, Washington, DC and these matters will not be treated in any depth here.

It must suffice to state that the test subjects were chosen to be small, inexpensive, non-controversial, clonelike mammals which manifested macrosmatic characteristics and which evidenced no emotional dependence upon the human race. The subject most closely meeting these specifications was a strain of albino rats (Sprague-Dawley) bred for laboratory use over many generations. For greatest unit-to-unit similarity, only male rats were used and then only then whose body weight was 500$\pm$ 100 grams.

Experiments with dogs at MERADCOM and elsewhere revealed that the detection performance of animals which are simply behaviorally modified and then sustintatively rewarded varies hourly and diurnally and thus the user is never really certain that the ultimate performance is being extracted from the animal in any given search. Further there is absolutely no reason to believe that these animals are always willing or even able to perform at maximum sensitivity and specificity.

It was desirable therefore to devise a scheme whereby an animal biosensor could be induced to sense target substances with unprecedented singleness of purpose, while at the same time maximally exercising its natural ability to discriminate the desired target substance in the presence of extremely heavy clutter.

Thus, the four postulates and the experimental program were devised to demonstrate that a reliable biosensor detector could be reduced to practice in a variety of Army and civilian explosives detection scenarios.

II. EXPERIMENTAL PROTOCOL

Anticipation is the basis of conditioned behavior regardless of whether the subject anticipates a pleasant or aversive result from a known pattern of events. The effectiveness of any conditioning paradigm is directly related to the intensity of anticipation and thus one must seek that rewarding stimulus which most intensely effects a particular class of experimental subjects for the longest period of time. The anticipation effect exists, at least in all mammals, regardless of whether the behavioral conditioning regimen is based upon operant (instrumental) or classical (Pavlovian) principles.

It is not possible in this short paper to discuss the details of operant and classical conditioning but these matters are well known and exhaustively treated in the vast literature of behavioral science. It is sufficient here to report that, in this research, an operant regimen was established to determine the "conditionability" of the subjects and to

establish for the subject the relationship between the presence of TNT vapor (the olfactory stimulus) and the reward stimulus. Following this initial training phase, the subject was classically conditioned in order to maximize the anticipation of reward and to eliminate the need for any operant response. This approach rendered the subject to the status of a biosensor detector element in a system where only intense anticipation was required of the subject. Since this is essentially a sub-conscious response to the initiating stimulus (TNT Vapor) there is no opportunity for the subject to willfully modify this response to the odorant and, in essence, a subject operating in this mode "cannot tell a lie" to gain a reward and it cannot refuse to annunciate detection of the target substance (as dogs are known to do on occasion).

The ultimate rewarding stimulus for rats was defined by Dr. James Olds in the 1950's as electrical stimulation of the Medial Forebrain Bundle (MFB) in either cerebral hemisphere. Consequently, no research time was expended in investigating this well established fact and all subjects surgically received a chronic electrode implant in this brain area. The reward stimulus was then termed electrical brain stimulus or EBS. Because of the evident pleasure manifested by all subjects so stimulated, Olds called the MFB the "pleasure center" and this terminology is still in use.

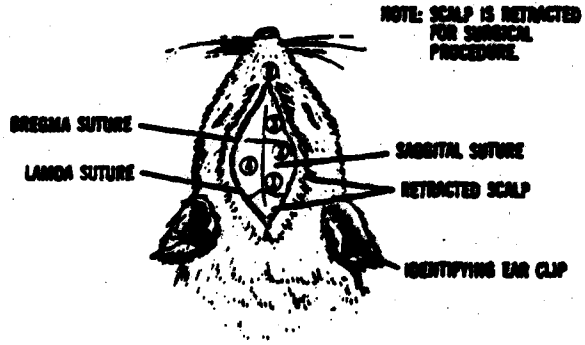
It should be noted that EBS results in no change in homeostasis as is the case with sustentative rewards such as food or water and thus the sensorium is much more constant than that of nutritionally deprived subjects. Even more important is the fact that EBS does not satiate as do sustentative rewards and thus anticipation remains constant even in long (>2hours) search sessions. This factor alone is sufficient to demand EBS as a reward stimulus!

Thus, the training paradigm prepared the subject to intensely anticipate an extraordinary pleasure whenever he detected the odor of TNT. However, the final test procedure allowed for no EBS during TNT stimulus runs since 100% reward tends to result in habituation which ultimately dulls anticipation and thereby actually lessens the probability of detection in a biosensor system of this nature. In an operational detection scenario using a rat as the sensor element, EBS would be forthcoming on a random basis in 70% of those events where an AESC was evident since, as described later, this percentage of reward resulted in maximal anticipation which, I believe, in turn resulted in maximum AESC which then gave optimal probability of detection.

Electrode Sites. The EBS electrode site was defined a priori but the definition of the electroencephalograph (EEG) electrode sites required some effort. First, the electrode had to be large (> 1 mm^2) since no specific locus for the AESC would be surmised. An epidural location seemed most appropriate since this tissue is quite rugged whereas the gross electrodes

would surely cause avulsion and necrosis in direct contact with cortical tissue. Thus the matter of penetration was resolved and the EEG electrodes (0-80 stainless steel screws) were threaded through the skull and juxtaposed to the dura mater. Since the limbic system must at least mediate if not originate the pleasure reaction, a locus proximal to the cingulate gyrus was chosen as the principal electrode site. A parietal electrode was implanted to acquire signals from the associative areas. Figs 1 and 2 show these loci.

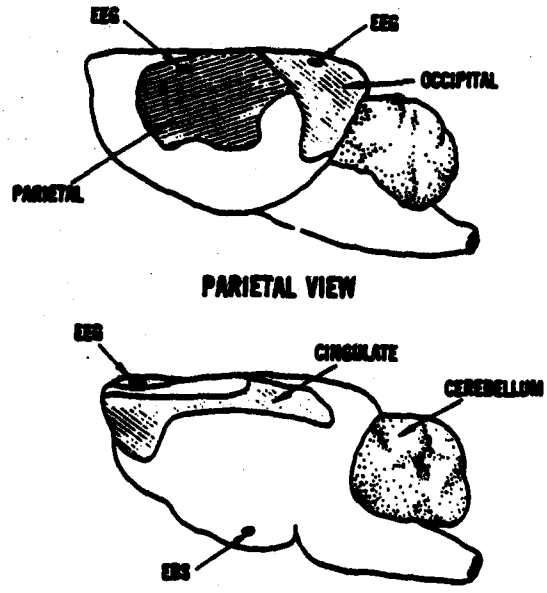
Stimulus Delivery. Fig 3 depicts the system used throughout the program. The subject was placed in severe restraint (to minimize motor signal artifacts in the EEG) in a copper-shielded enclosure -- a Faraday Box -- which also contained the EEG preamplifiers (PAR Model 113). Odorants were delivered in two classes: TNT and Neutral. Neutral stimuli were random selections of asphalt, cedar and pine. The odorant delivery system used "zero" air to maintain a constant concentration of odorant from day to day but the ambient air in the Faraday Box was simply laboratory air. The Faraday Box was operated at a slight negative pressure so that TNT build-up did not occur in the proximal to the subject. The effluents were exhausted external to the laboratory and the subjects gave no evidence - ever - of awareness of TNT contamination of the laboratory. Constant cleanliness was essential and heroic efforts were used to assure continuing freedom from TNT build-up in any of the laboratory or domiciliary areas. The negative pressure source also assured a rapid lowering of ambient odorant concentration during non-stimulus periods between odorant presentations.



1 = OCCIPITAL ZONE (EEG)
2 = PARIETAL LOBE (EEG)
3 = CINGULATE GYRUS (EEG)
4 = MEDIAL FOREBRAIN BUNDLE (EBS)
R = NASION (INDIFFERENT ELECTRODE)

EEG & EBS ELECTRODE SITES (TREPHINE LOCATIONS)

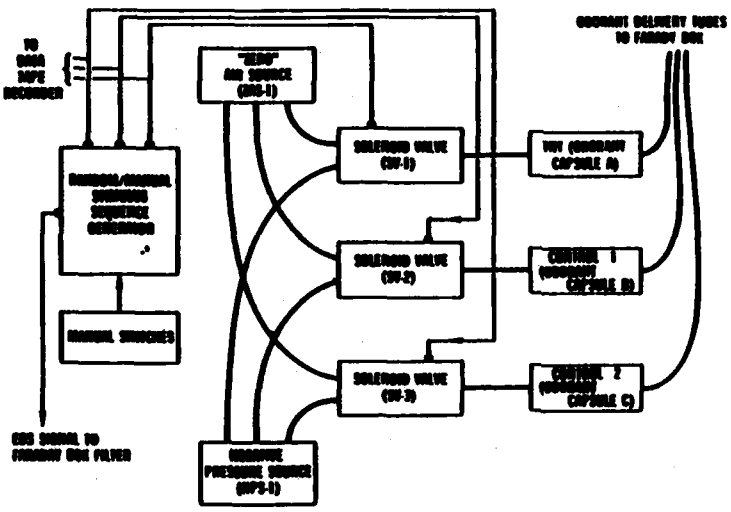
FIGURE 1



PARIETAL & SAGGITAL VIEWS OF EEG & EBS ELECTRODE SITES

FIGURE 2

All stimuli - including EBS - could be either randomly or manually sequenced. During most test runs the sequencing was manual so that even in short runs approximately the same number of TNT and Neutral stimuli were delivered, whereas long conditioning test runs were generally accomplished by use of the automatic random sequencer.

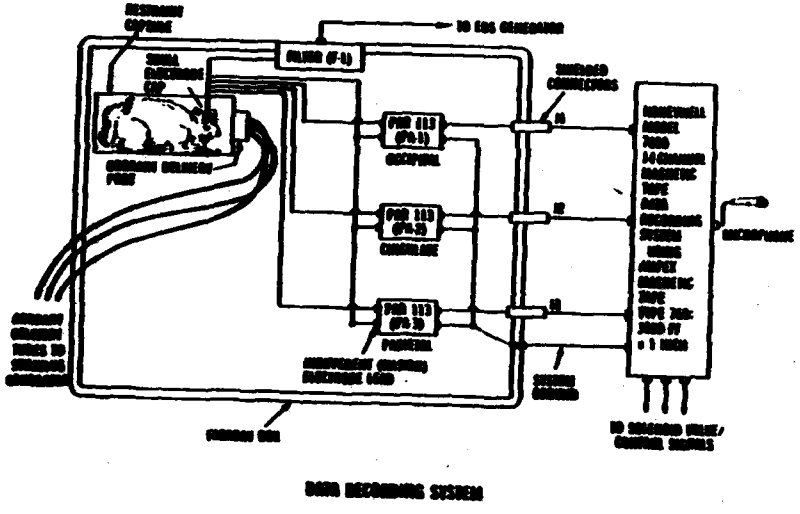


EBS & ODORANT DELIVERY SYSTEMS
FIGURE 3

No attempt was made to quantify the TNT molecular concentration due to the extreme difficulty of such measurements even in analytical chemistry laboratories. In a simple physiology laboratory the task was determined to be virtually impossible. Also, knowledge of exact concentration of odorants was not necessary to the research program. Ultimately, sensitivity must be quantified relative to dogs but a dual dog/rat program was out of the question at all times during the research period.

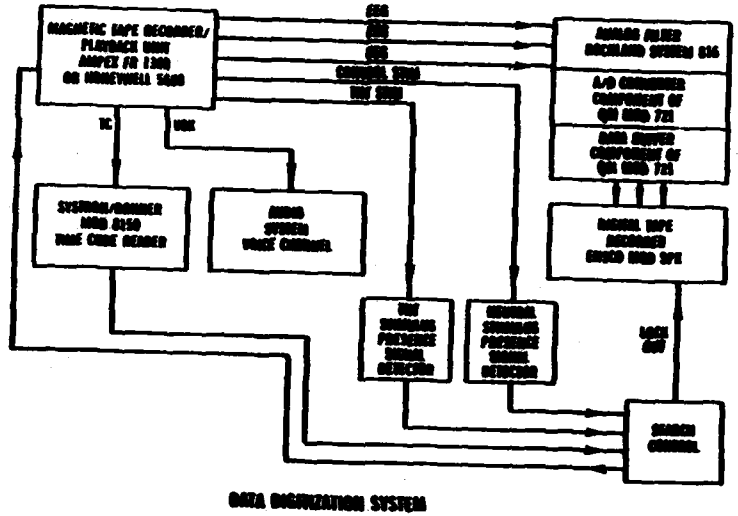
Data Systems. The data recording system is depicted in Fig 4 and the data playback system is shown in Fig 5. The PAR EEG preamplifiers were set to 6 DB bandwidth points of 0.3H_z and 300 H_z with gain set at 10,000. Thus, the EEG signals (which were found to range between 20 and 150 microvolts at the dura mater) were delivered at levels of 200 to 1000 millivolts to the analog magnetic tape recorder (Honeywell Mod 7600). In addition to EEG signals, the recorder also acquired an IRIG time code, voice commentary and the control signals from the odorant delivery solenoid valves in the stimulus delivery system. The solenoid control signals were of extreme importance in the final data analysis programs, since timing of odorant bolus arrival in the Faraday Box was thereby effected. Data recording was usually continuous in 2 to 3 hour segments so that a great mass of "resting" background EEG data were available if needed.

The digitization of the analog EEG data was carried out with an analog system bandwidth of DC to 312 Hz (IRIG Low-Band), digital system bandwidth of DC to 200 Hz with 512 data points per second and a density of 1600 bits per inch.



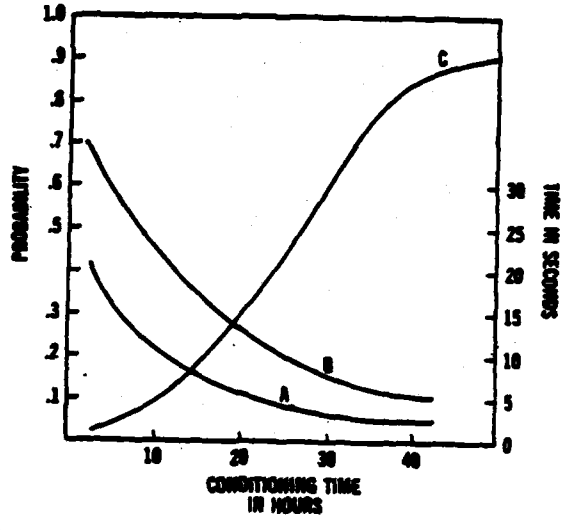
DATA RECORDING SYSTEM

FIGURE 4



DATA DIGITIZATION SYSTEM

FIGURE 5



A = TREADLE PRESS TIME FOLLOWING TERMINATION OF ODORANT FLOW
B = PROBABILITY OF FALSE RESPONSE
C = PROBABILITY OF TRUE RESPONSE

RAT PERFORMANCE VS CONDITIONING TIME

FIGURE 6

III. RESULTS

Verification of Theses A, B & C. Prior to this research effort, there were no known data which could unequivocally demonstrate that rats were able to detect TNT, even though assumptions were voiced to the effect that this capability should exist. It was essential, then, that Thesis A be proven valid at the earliest possible moment in the history of the research and it became apparent in the formulation of the experiment that Thesis A was most convincingly proved by the concurrent proof of Thesis B.

A total of 818 3-second data epochs were extracted from a total population of 1300 similar epochs for the purpose of final data analysis (the unused 482 epochs were contaminated by excessive noise or motor artifacts in early data runs). Spectral analysis of 100 epochs revealed that the experimental 312 Hz recording bandwidth was quite adequate. Indeed, the data evidenced little activity in any epoch above 100 Hz and the average data epoch revealed an EEG with maximum spectral density occurring in human EEG's.

Fig 6 summarizes the results of the behavioral analyses of the data mass. It was soon apparent that the conditioning regimen was quite effective as curve "C" demonstrates. Forty hours past the state of total naivity, the average subject would annunciate by treadle press (in the first or operant phase of conditioning) his awareness of the presence of the target substance TNT. At the same training landmark, the subjects evidenced a greatly diminished tendency to falsely signal TNT presence.

Concurrent with the diminishment of false annunciation was the decrease in the tendency to attempt to prolong the pleasure of EBS by treadle pressing after the disappearance of TNT-laden air from the stimulus delivery tubes (Curve A). This factor is important since the low probability of continued treadle pressing at 40 hours of training implies that the subject has clearly recognized that EBS is absolutely related to TNT presence and that EBS is not available in any other case. The early high rates of signalling in the absence of TNT was most probably due to the intense desire for EBS of "no reward without TNT odor" since diminishment of this activity occurred as subject "awareness" developed.

As noted earlier, some extinguishment of behavioral conditioning invariably accompanies any operational protocol in which a conditioned subject is rewarded each time a correct operant or classical response is effected in response to stimulus. Such behavioral response (habituation) can seriously alter the course of an experiment, since these effects will probably be of indeterminate magnitude a posteriori, thus leaving an unknown degree of bias in the overall experimental results. To preclude this untoward eventuality, fifteen rats were studied for the effects of unity EBS/odorant presence ratio. Habituation was generally evident in those

subjects which had passed the 40-hour point in their conditioning regimen (during which time a unity ratio was the rule) by about 20 hours. Although the diminishment of the detection performance -- as evidenced at this stage of training by treadle press -- was highly variable between subjects, the average diminishment observed in all fifteen subjects was found to be a drop to about 0.7 probability of detection. Since all the selected subjects had evidenced 0.9 or better probability of detection near the end of their initial 40-hour conditioning regimen this diminishment of performance was quite significant from the standpoint of practical application in explosives detection service.

As the EBS/odorant presence ratio was varied from unity to 0.3 by withholding EBS in a pseudo-random manner, the average subject began initially to increase his probability of detection (due to enhanced anticipation) and then extinguishment again appeared. The curve of Fig 7 represents the average performance of the test subjects.

It is important to stress that the exact nature of the olfactory cue was never a matter for investigation in this research. Certainly military grade TNT contains several substances in addition to 2,4,6-trinitrotoluene, but these substances are always present in approximately the same proportions in all samples of military grade TNT, and thus it is of no consequence to this research if the animal identifies one component or all components of this explosive, since detection of the aggregate substance is the desired experimental goal. Further, the matter of distractants was not pertinent, and hence this avenue of investigation was not explored in this research.

Similarly, no extensive effort was made to find the "ultimate" neutral stimuli since, in effect, <u>any</u> olfactory stimulus other than TNT was, by definition, neutral, and degrees of neutrality were of no interest. In the long term, the matter of distractants (or non-neutral, non-desired) stimuli will be of interest to the overall program of the sponsor, but such investigations were totally beyond the scope and intent of the instant program.

At the end of a forty-hour test sequence, the average subject clearly verified Theses A & B by his behavior. To summarize: In the absence of TNT vapor, the subject generally ignored the treadle after about thirty hours of conditioning, except for an occasional random press (perhaps this represented an attempt to hasten EBS, or just "to see if anything good will happen"). When TNT vapor was present at the 40 hour point, the animal began furiously to press the treadle in an attempt to extract the maximal pleasure from what it may have come to recognize as an ephemeral event.

Fig 6 graphically demonstrates the validity of Theses A & B and these early results demanded that the program continue in an attempt to prove the very important Thesis D.

The learning patterns, while exhibiting variations due to the ever-present physiological and psychological differences in subjects, were, overall, similar to the general trends shown in Fig 6, and at the end of the 40-hour training periods, all subjects were performing as the curves depict within limits of about ±30%.

Thesis C was proved in achieving the proof of Theses A & B by the use of a group of five behaviorally shaped subjects which were totally naive to the EBS/TNT relationship at the start of testing. Each test subject was given identical TNT/EBS pairings, thus simulating sequentially what would have occurred simultaneously had sufficient laboratory facilities been available.

It is pertinent to note here that the continuously observed indifference evidenced by all test subjects toward man en masse or as an individual argues strongly in favor of the concept of automatic training, since there would be no resulting impact upon the depth of conditioning even if the test operators were changed several times during training.

This final factor was of great significance not only to the immediate test results, but in consideration of future systems employing rats as biosensor detector elements, since the optimal biosensor animal must be one which requires no emotional ties to any human element employed in overall system operation.

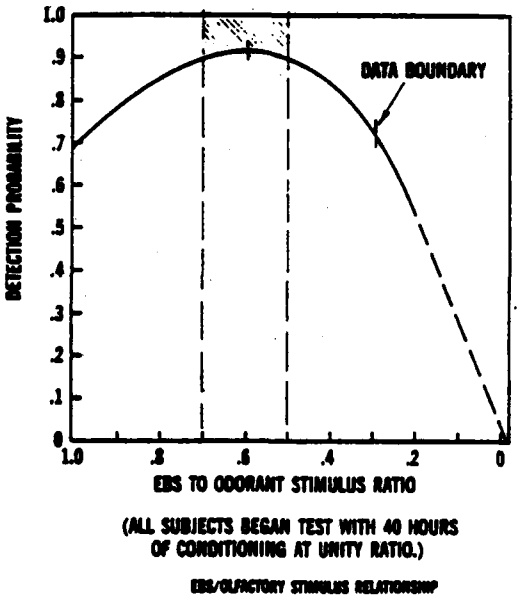


FIGURE 7

Verification of Thesis D

One of the two successful signal analysis - or feature extraction - techniques used for the proof of Thesis D was a metric in the general form of the Covariance process. The covariance process will not be described here since it is well defined in the literature. It is necessary to note that the covariance of the raw data was first examined with a scanning window width of 11 data points and then these covariances were in turn examined with a window width of 31 data points. Thus, the final covariance data used in the proof of Thesis D was the "covariance of the covariance".

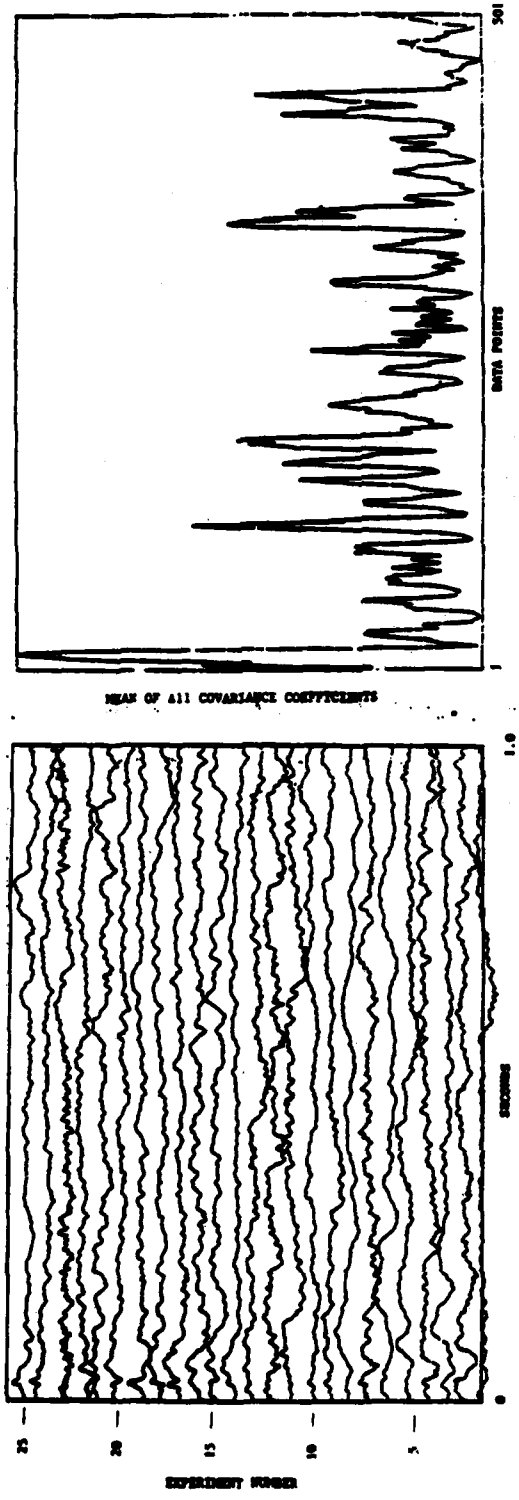


FIGURE 8

25 BIPOLAR RAW DATA EPOCHS FOR INTERVAL T6 FOR RAT C.

FIGURE 9

MEAN OF A11 COVARIANCE COEFFICIENTS FOR INTERVAL T6 FOR 25 DATA EPOCHS RAT C.

FIGURE 10

A31 COVARIANCE MEAN OF A11 COVARIANCE MEANS OF INDIVIDUAL RATS DURING INTERVAL T6.

FIGURE 11

A31 COVARIANCE MEAN OF A11 COVARIANCE MEANS OF INDIVIDUAL RATS DURING INTERVAL NA.

The final form of the data epochs subjected to analysis consisted of three one-second intervals defined as follows: [FN]

a. Interval T - 1: The period of data occurring in the one-second interval immediately prior to the trigger.

b. Interval Tø: The period of data occurring in the one-second interval immediately subsequent to the trigger.

c. Interval T + 1: The period of data occurring in the one-second interval immediately subsequent to the Tø interval.

d. Similarily N-1, Nø, and N+1 represent the same time intervals for the case of neutral stimuli.

With these terms in mind, it will now be of interest to examine Figs 8, 9, & 10. Fig 8 depicts a small segment of the raw analog monopolar signals obtained from the cingulate electrode of Rat C. Here are presented twenty-five epochs of the interval Tø with the relative amplitudes of the signals maintained in true relationship to the original signals.

Fig 9 is a plot of the Δ11 covariance coefficients versus time (in data points/second) for the period covering the first second subsequent to a TNT odorant bolus release. The first large peak (about data point 22) occurred at about 40 milliseconds subsequent to the activation of the TNT odorant solenoid valve relay. During initial data analysis this event was assumed to be perhaps fortuitous, but after examination of similar data epochs across the rat population, it became apparent that this peak was related to the solenoid valve relay activation in a relatively invariant manner. Such a repetitive event was seen to exist also in all Nø plots (where the neutral stimulus odorant-delivery solenoid activation pulses were used to define the time period of analysis. The most reasonable explanation for this ubiquitous signal feature is that it represents the effect of the relay "click" on the auditory processing circuits in the rat brain!

Fig 9 and similar plots for other epoch sets shows some evidence of olfactory stimulus recognition, (if one knows exactly where to look) but, standing alone, this set of plots would not suffice for verification of Thesis D.

FN The trigger, as noted earlier, is defined as the leading edge of the DC pulse applied as the activation signal to the odorant-delivery solenoid valve relays. Δ11 is the symbol used to designate those covariance processes in which the window width was 11 data points and Δ31 is the symbol used to designate those covariance processes in which the window width was 31 data points.

Since the transit time of the odorant bolus was known to be 500 ± 70 milliseconds, a careful examination of the mid-abcissa region of the Tø plot in Fig 9 may reveal some evidence of unusual covariance at about data point 220, but this feature is not as immediately apparent as one might wish and, in order to make the feature at data point 220 more evident, a second application of the covariance process was effected on all Δ11 data sets using a window width of 31 data points.

Fig 10 is a plot of the results of this process. This plot is of great significance since it represents graphical evidence of the validity of Thesis D!

To recapitulate the specific events displayed in Fig 10:

a. 87 separate experiments derived from the data for Rats A, B, C, D, 98, 100 and 103 were analyzed over the Tø interval.

b. Δ11 covariance means were obtained for the ensemble of 87 experiments.

c. Δ31 covariance means were derived for the summed results of the Δ11 calculations.

d. The unique features seen in Fig 10 were obtained only for TNT stimuli.

Fig 11 is the neutral olfactory stimulus equivalent to Fig 10 in the Nø interval. One may observe the clearly evident auditory signal near data point 24 in the Nø interval. However, there is no significant pairing such as that seen in Fig 10.

In Fig 11, the earliest Nø peak which could possibly be related to an olfactory stimulus occurs at about data point 290 which represents a time beyond the upper measured limit of odorant arrival time. Thus, while the single large peak near the center of Nø ensemble may or not represent some olfactory related event, there is little doubt that the twin peaks seen in Fig 10 are unique evidence of an event related only to the presence of TNT vapor.

Based upon the foregoing analysis, it is virtually certain that a unique feature appears in the output of the Covariance Process when -- and only when -- the input EEG is derived from a fully conditioned test subject during a period of olfactory stimulation by TNT vapor.

The Segmentation Process. This metric is fully described in the author's report (MERADCOM 2343) and in his doctoral dissertation. Basically this metric places a time scanning iterative process into the

category of an adaptive feature extraction technique.

Fig 12 depicts the result of application of the Segmentation Process to a single TNT stimulus experiment from the data set for Rat B. In this and similar plots the scanning window width was set - as it was throughout the final iteration of this process - at 128 data points, and the value of β (a reference threshold) was set by trial to increase the fine structure of the plot while maintaining relative freedom from random event peaks. Note that the abcissa is the time axis, and, in terms of the Covariance Process terminology, intervals T - 1, TØ and T + 1 are contiguous in presentation in this figure. Since there is a 500 ± 70 millisecond delay between the solenoid release signal -- which occurs in this figure at 1.0 on the Time axis -- and the arrival of the odorant bolus, those events which occur at or near 1.5 on the Time axis are worthy of close scrutiny.

Note the double peak near time 1.5 in Fig 12. Here β has been lowered to the experimentally determined limit of 0.175.

The similarity to the plot of Fig 10 is interesting since the same general result was obtained from two greatly different metrics applied to two different data sets.

In Fig 13, the Segmentation Process output is subjected to spectral analysis which elucidates two dominant peaks and one apparently spurious peak. The significant peak at 8 Hz may possibly represent those physiological processes associated with the so-called Alpha rhythms in mammals (8 to 13 Hz) or it may have a more obscure meaning here. The peak at 20 Hz occurred in one subject only and may be an artifact which would be insignificant in a larger data sample. The most significant grouping of peaks in this figure is that existing between about 10 Hz and 17 Hz.

Fig 14 also manifests a dominant peak at about 7 to 8 Hz, and it probably has the same significance as it does with TNT data plots. A secondary broad peak centered at about 10 Hz is seen as well as very small peaks out to 125 Hz.

Clearly, the AESC appears in the plots of Fig 13 as the frequency/density peak seen between the 8 Hz cluster and the single 20 Hz peak. The composite structure is so different from that generated by the neutral stimulus data presented in Fig 14 that the peak cluster in Fig 13 must be called unique.

This, then, is the final argument in support of the postulate of Thesis D. Two totally different processes -- Covariance and Segmentation -- present strong evidence of the presence of unique cortical spectral features occurring in close time coincidence with the application of TNT

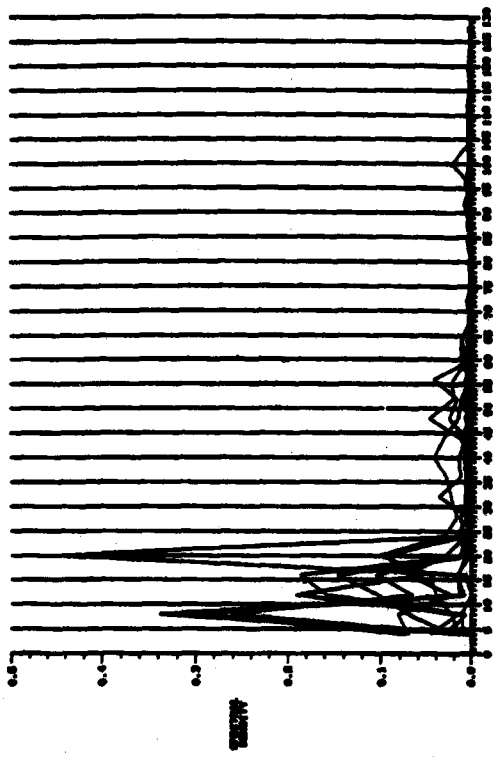


FIGURE 13

COMPOSITE SPECTRAL DENSITY FOR ALL RATE, TINT STIMULUS.

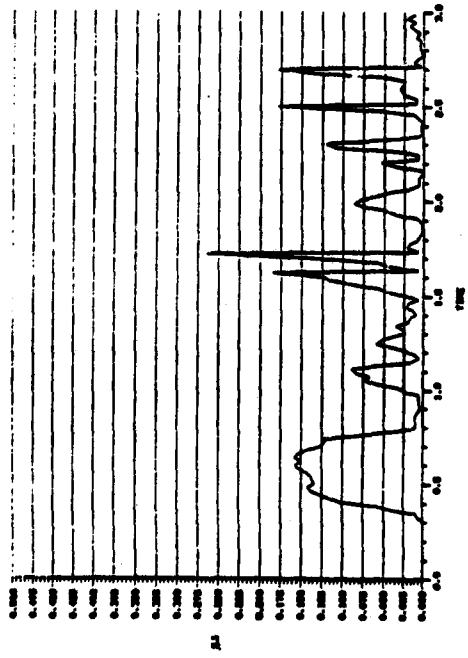


FIGURE 12

ERROR FUNCTION VERSUS TIME IN DATA EPOCH FOR $\beta = 0.178$, TINT STIMULUS.

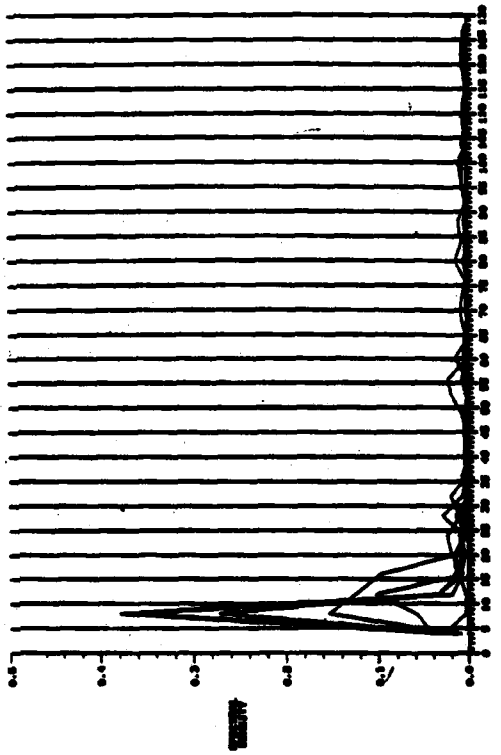


FIGURE 14

COMPOSITE SPECTRAL DENSITY FOR ALL RATE, NEUTRAL STIMULUS.

olfactory stimuli in the specially trained test population. Taken individually, each process presents a strong case for the occurrence of the AESC; taken collectively, there is little doubt that an identifiable feature does exist as a direct consequence of the TNT stimulus, and there is little doubt that this feature can be identified by existing microprocessor technology, especially if three or more experiments can be summed before recognition is demanded by the detection system operating protocol. It appears therefore, that Thesis D is proved to be valid.