AD-A259 319

DTIC
ELECTE
JAN 5 1993
S D
C

Army Science Conference Proceedings

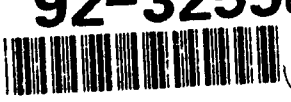
22-25 June 1992

Volume II

Principal Authors H through M

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

92-32558



92 12 22 136

Assistant Secretary of the Army
(Research, Development and Acquisition)
Department of the Army

DEPARTMENT OF THE ARMY
OFFICE OF THE ASSISTANT SECRETARY
WASHINGTON, DC 20310-0103

SARD-TR

3 0 JUL 1992

MEMORANDUM FOR SEE DISTRIBUTION

SUBJECT: Proceedings of the 1992 Army Science Conference

The 18th Army Science Conference was held at the Hyatt Hotel, Orlando, Florida, 22-25 June 1992. The conference presented a cross section of the many significant scientific and engineering programs carried out by the Department of the Army (DA). Additionally, it provided an opportunity for DA civilian and military scientists and engineers to present the results of their research and development efforts before a distinguished and critical audience.

These Proceedings of the 1992 Army Science Conference are a compilation of all papers presented at the conference and the supplemental papers that were submitted.

Our purpose for soliciting these papers was to:

a. Stimulate the involvement of scientific and engineering talent within the Department of the Army;

b. Demonstrate Army competence in research and development;

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

d. Promote the interchange of ideas among members of the Army scientific and engineering community as well as management and technical personnel from industry and other government agencies.

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.

George T. Singley III
Deputy Assistant Secretary
For Research and Technology

DISTRIBUTION:
Office of the Under Secretary of Defense for Research and Advanced Technology, Wash, DC 20310-3080
Office of the Assistant Secretary of the Army (RD&A), ATTN: SARD-TR, Wash, DC 20310-0103
HQDA, Assistant Chief of Staff for Intelligence, ATTN: DAMI-ZA, Wash, DC 20310
HQDA, Deputy Chief of Staff for Logistics, ATTN: DALO-ZX, Wash, DC 20310-0500
HQDA, Deputy Chief of Staff for Operations & Plans, ATTN: DAMO-FD, Wash, DC 20310-0400
HQDA, Director of Information Systems for C4, ATTN: SAIS-ZA, Wash, DC 20310-0400
HQDA, Deputy Chief of Staff for Personnel, ATTN: DAPE-ZA, Wash, DC 20310-0300
Office of the Chief of Engineers, ATTN: CERD-M, 20 Massachusetts Avenue, N.W., Wash, DC 20314-1000
Office of the Surgeon General, ATTN: DASG-ZA, 5109 Leesburg Pike, Falls Church, VA 22041-3258
COMMANDERS/DIRECTORS:
USA Concepts Analysis Agency, 8120 Woodmont Avenue, Bethesda, MD 20814-2797
USA Operational Test & Evaluation Agency, 5600 Columbia Pike, Falls Church, VA 22041-5115
USA Information Systems Engineering Command, Ft. Huachuca, AZ 85613-5000
USA Strategic Defense Command, P. O. Box 15280, Arlington, VA 22215-0150
USA Materiel Command, ATTN: AMCLD, 5001 Eisenhower Avenue, Alexandria, VA 22333-0001
USA Materiel Systems Analysis Activity, ATTN: AMXSY-D, APG, MD 21005-5071
USA Armament, Munitions & Chemical Command, ATTN: AMSMC-CG, Rock Island, IL 61299-6000
USA Armament RD&E Center, ATTN: SMCAR-TD, Picatinny Arsenal, NJ 07806-5000
Fire Support Armament Cen, ATTN: SMCAR-FS, PA, NJ 07806-5000
Armament Engineering Dir, ATTN: SMCAR-AE, PA, NJ 07806-5000
Close Combat Armament Cen, ATTN: SMCAR-CC, PA, NJ 07806-5000
Chemical RD&E Center, ATTN: SMCCR-TD, APG, MD 21010-5423
USA Aviation Systems CMD, ATTN: AMSAV-GTD, 4300 Goodfellow Blvd, St. Louis, MO 63120-1798
USARTA, Ames Research Center, ATTN: SAVRT-D, Moffett Field, CA 94035-1099
USARTA, Aeroflightdynamics Dir, ATTN: SAVRT-AF, Moffett Field, CA 94035
USARTA, Aviation Applied Technology Dir, ATTN: SAVRT-TY, Ft Eustis, VA 23604
USARTA, Propulsion Dir, ATTN: SAVRT-PN, Lewis Research Center, 21000 Brook Park Rd, Cleveland, OH 44135
USARTA, Aerostructures Dir, ATTN: SAVRT-SD, Hampton, VA 22665
Avionics R&D Activity, ATTN: SAVAA, Ft Monmouth, NJ 07703-5000

USA Communications-Electronics Cmd, ATTN: AMSEL-TDD, Ft Monmouth, NJ 07703-5001
Center for Command, Control & Communications System, ATTN: AMSEL-RD-C3, Ft Monmouth, NJ 07703-5202
Center for Night Vision & Electro-Optics, ATTN: AMSEL-RD-NV-D, Ft Belvoir, VA 22060-5677
Electronics Warfare/RSTA Center, ATTN: AMSEL-RD-EW-D, Ft Monmouth, NJ 07703-5303
Signals Warfare Lab, ATTN: AMSEL-RD-SW-D, Vint Hill Farms Station, Warrenton, VA 22186-5100
USA Laboratory Command, ATTN: AMSLC-TD, 2800 Powder Mill Road, Adelphi, MD 29783-1145
Atmospheric Sciences Lab, ATTN: SLCAS-DD, WSMR, NM 88002-5501
Ballistic Research Lab, ATTN: SLCBR-OD, APG, MD 21005
Electronics Technology & Devices Lab, ATTN: SLCET-D, Ft Monmouth, NJ 07703-5000
Electronic Warfare Vulnerability Assessment Lab, ATTN: SLCEW-M-D, WSMR, NM 8802-5513
Harry Diamond Labs, ATTN: SLCHD-D, Adelphi, MD 20783-1197
Human Engineering Lab, ATTN: SLCHE-H, APG, MD 21005-5001
Materials Technology Lab, ATTN: SLCMT-D, Watertown, MA 02172-0001
Army Research Office, ATTN: SLCRO-D, RTP, NC 27709-2211
Army Research, Development & Standardization Group (Europe) Box 65, FPO New York 09510
USA Missile Command, ATTN: AMSMI-R, Redstone Arsenal, AL 35898-5240
USA Tank-Automotive Cmd, ATTN: AMSTA-CG, Warren, MI 48397-5000
  Aberdeen Proving Ground, ATTN: STEAP-OC, APG, MD 21005-5001
  Dugway Proving Ground, ATTN: STEDP-CO, Dugway, UT 84022-5000
  Electronic Proving Ground, ATTN: STEEP-DO, Ft Huachuca, AZ 85613-7110
  Aviation Development Test Activity, ATTN: STEBG-CO, Ft Rucker, AL 36362-5276
  Combat Systems Test Act, ATTN: STECS-CO, APG, MD 21005-5059
  White Sands Missile Range, ATTN: STEWS-CG, WSMR, NM 88002-5000
USA Aviation Systems Cmd, ATTN: AMSAV-G, 4300 Goodfellow Blvd, St. Louis, MO 63120-1798
Belvoir RD&E Center, ATTN: STRBE-Z, Ft. Belvoir, VA 21005-5606
Natick RD&E Center, ATTN: STRNC-Z, Natick, MA 01760-5000
US Army Corps of Engineers
  Cold Regions Rsch & Eng Lab, ATTN: CRREL-TD, P. O. Box 282, Hanover, NH 03755-1290
  Construction Eng Rsch Lab, ATTN: CERL-Z, P. O. Box 4005, Champaign, IL 61820-135
  Engineer Topographic Labs, ATTN: ETL-TD, Ft Belvoir, VA 22060-5546
  Waterways Experiment Station, ATTN: CEWES-IM-MI, P. O. Box 631, Vicksburg, MS 39180-0631
USA Medical R&D Command, ATTN: SGRD-ZA, Ft Detrick, Frederick, MD 21701-5012
Aeromedical Rsch Lab, ATTN: SGRD-UAC, Ft Rucker, AL 36362-5292
Inst of Dental Rsch, WRAMC, ATTN:SGRD-UDZ, Wash, DC 20307-5300
Inst of Surgical Rsch, ATTN: SGRD-USZ, Ft Sam Houston, TX 78234-6200

Letterman Army Inst of Rsch, ATTN: SGRD-ULZ, Presidio of San Francisco, CA 94129-6800
Biomedical R&D Lab, ATTN: SGRD-UBZ, Frederick, MD 20701-5010
Medical Rsch Inst of Chemical Defense, ATTN: SGRD-UV-ZA, APG, MD 21010-5425
Medical Rsch Inst of Environmental Medicine, ATTN: SGRD-UEZ, Natick, MA 01760-5007
Medical Rsch Inst of Infectious Diseases, ATTN: SGRD-UIZ-A, Ft Detrick, Frederick, MD 21701-5011
Walter Reed Army Inst of Rsch, ATTN: SGRD-UWZ, Washington, DC 20307-5100
USA Health Services Command, Ft. Sam Houston, TX 78234-61000
USA Environmental Hygiene Agency, ATTN: HSHB-Z, APG, MD 21010-5422
USA Research Institute for the Behavioral and Social Sciences, ATTN: PERI-ZT, 5001 Eisenhower Ave, Alexandria, VA 22333-5600
ARI Field Unit, ATTN: PERI-IJ, POB 2086, Ft Benning, GA 31905
ARI Field Unit, ATTN: PERI-SB, POB 6057, Ft Bliss, TX 79906
ARI Field Unit, ATTN: PERI-SH, Ft Hood, TX 76544-5065
ARI Field Unit, ATTN: PERI-IK, Steele Hall, Ft Knox, KY 40121
ARI Field Unit, ATTN: PERI-SL, POB 290, Ft Leavenworth, KS 66027
ARI Field Unit, ATTN: PERI-IR, Ft Rucker, AL 36362-5000
ARI Field Unit, ATTN: PERI-IO, POB 5787, Presidio of Monterey, CA 93940
USA Training and Doctrine Command, ATTN: ATDO-ZT, Ft Monroe, VA 23651-5000
Combined Army Test Act, ATTN: ATCT-CG, Ft Hood, TX 76544-5065
COMMANDANTS:
  USA Air Defense Cen, ATTN: ATZC-CG, Ft Bliss, TX 79916-5000
  USA Armor Cen, ATTN: ATZK-CG, Ft Knox, KY 40121-5000
  USA Aviation Cen, ATTN: ATZQ-CG, Ft Rucker, AL 36362-5000
  USA Chemical Cen, ATTN: ATZN-CG, Ft McClellan, AL 36205-5000
  USA Combined Arms Cen, ATTN: ATZL-SC, Ft Leavenworth, KS 66027-5000
  USA Engineer Cen, ATTN: ATZA-CG, Ft Belvoir, VA 22060-5000
  USA Infantry Cen, ATTN: ATZB-CG, Ft Benning, GA 31905-5000
  USA Intelligence Cen, ATTN: ATSI-SA, Ft Hauchuca, AZ 85613-7000
  USA Ordnance Cen, ATTN: ATSL-CMT, APG, MD 21005-5201
  USA Signal Cen, ATTN: ATZH-CH, Ft Gordon, GA 30905-5000
  USA Field Artillery Center, ATTN: ATZR-C, Ft Sill, OK 73503
SUPERINTENDENT:
  US Military Academy, ATTN: Technical Library, West Point, NY 10996
COPIES FURNISHED:
  Defense Advanced Research Projects Agency, 1400 Wilson Blvd, Arlington, VA 22209-2308
  Defense Logistics Agency, Cameron Station, Alexandria, VA 22304-6183
  Defense Technical Information Center, Cameron Station, Alexandria, VA 22304-6145
  Lawrence Livermore National Lab, ATTN: L-191, POB 808, Livermore, CA 94550
  Los Alamos National Lab, ATTN: Dir for Energy, Rsch & Tech, Los Alamos, NM 87545

NASA HQS, Suite 4237, 400 Maryland Ave, SW, Wash, DC 20546
National Science Foundation, 1800 G Street, NW, Wash, DC 20550
United Nations Library, ATTN: Acquisition Section, Room L-138A, New York, NY 10017

US NAVY

Naval Air Systems Command, Code AIR-03-D, 1411 Jefferson Davis Hwy, Arlington, VA 22202-3000
Naval Rsch Lab, ATTN: Dir of Rsch, Washington, DC 20375
Office of Naval Rsch, Code 10, 800 North Quincy St, Arlington, VA 22217-5000
HQ US Marine Corps, Code RD-1, Washington, DC 29380-0001

US AIR FORCE

Air Force Systems Command, ATTN: Technical Director, Andrews AFB, Washington, DC 20334-5000
Air Force Office of Scientific Research/XOT, ATTN: Distribution, Bolling AFB, Washington, DC 20332-6448

Accession For	
NTIS GRA&I	☑
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

PROCEEDINGS

OF THE

1992 ARMY SCIENCE CONFERENCE

ORLANDO, FLORIDA

22 - 25 JUNE 1992

VOLUME II

PRINCIPAL AUTHORS H THROUGH M

TABLE OF CONTENTS

PROCEEDINGS OF THE 1992 ARMY SCIENCE CONFERENCE

AUTHOR	TITLE	VOL	PAGE
Alabran, David M.	See Willingham, Reginald	III	369
Alster, Jack	See Nicolich, Steven	III	1
Andersen, Steven L.	See Martin, Rodger K.	II	301
Atha, Larry C.	Advanced Composite Interceptor Kill Vehicle Structures	I	5
Azrin, Morris	See Beatty, John H.	I	51
Bahder, Thomas B.	Analytic Dispersion Relations Near the Gamma Point in Strained Zincblende Semiconductors	I	19
Baker, Paul K.	Lightweight Exo-Atmospheric Projectile (LEAP)	I	25
Balekdjian, Armand	See Harvey, James F.	II	45
Barrows, Austin W.	Development of Laser-Based Ignition Systems for Large Caliber Guns	I	37
Beatty, John H.	Correlation of Ballistic Performance to Shear Instability Studies in High Strength Steels	I	51
Beaudry, William T.	See Wagner, George W.	III	247
Behrens, Richard, Jr.	Thermal Decomposition of RDX and HMX in the Condensed Phase: Isotope Scrambling and Deuterium Isotope Effect Studies	I	65
Bennett, Kelly W.	See Brody, Philip S.	I	175
Berger, Henry	A New Understanding of Power Transmission and Reflection in Remote and Near-In Sensing	I	83
Beyer, Richard A.	Small Scale Experiments and Modeling of Laser Ignition of Gun Propellants	I	97

AUTHOR	TITLE	VOL	PAGE
Beyer, Richard A.	See Barrows, Austin W.	I	37
Bhagat, Hitesh R.	See Reid, Robert H.	III	35
Bhat, Ramachandra K.	See Guadagno, Janine	I	471
Bhattacharya, P.K.	See Tober, Richard L.	III	221
Billingsley, Daniel	See Watkins, Wendell R.	III	275
Binder, Michael	Method for Increasing Breakdown Voltages in Polymer Film Capacitors	I	111
Birk, Avi	See McQuaid, Michael J.	II	353
Bishop, Steven S.	Development of the Mine Resistant Vehicle Using Advanced Computational Methods	I	119
Bissett, Frank H.	See Willingham, Reginald	III	369
Bliesener, Gary	See McQuaid, Michael J.	II	353
Blum, Maxi	See Juhasz, Arpad	II	159
Board on Army Sc. & Technology (BAST)	Strategic Technologies for the Army of the Twenty-First Century	I	1
Boedeker, Edgar C.	See Reid, Robert H.	III	35
Boesch, H. Edwin,Jr.	Trapping Characteristics of Radiation-Generated Charge in Silicon-on-Insulator Buried Oxides	I	135
Bowden, Charles M.	See Crenshaw, Michael E.	I	319
Bowden, Charles M.	See Harvey, James F.	II	45
Bowman, Phillip D.	See Schuschereba, Steven	III	119
Bradshaw, John L.	See Tobin, Mary S.	III	231
Braun, Christopher	Advanced Electronic Detonator Modules for Live-Fire Missile Testing	I	151

AUTHOR	TITLE	VOL	PAGE
Brimfield, A. A.	Catalytic Antibodies as Biological Scavengers for Organophosphorus Poisons	I	167
Brody, Philip S.	Photovoltaic Effect in Thin Ferroelectric Films for Nonvolatile Memory	I	175
Broomfield, C. A.	See Brimfield, A. A.	I	167
Brown, Steven D.	See Harper, Alice M.	II	31
Brown, C. David	Scene Segmentation and Definition for Autonomous Robotic Navigation Using Structured Light Processing	I	189
Bryant, Garnett W.	See Tober, Richard L.	III	221
Bryant, Garnett W.	See Tobin, Mary S.	III	231
Bryzik, Walter	Advanced Ground Propulsion System Phenomena	I	205
Bulusu, Suryanarayana	See Behrens, Richard Jr.	I	65
Burchanowski, John	See Guadagno, Janine	I	471
Carrillo, Alejandro	See Horner, David A.	II	109
Cartland, Harry E.	Time-Resolved Infrared Diode Laser Spectroscopy: A State Specific Probe of Atom-Molecule Reaction Dynamics	I	221
Casas, L.	See Lee, Howard S.	II	253
Chambers, James P.	See Valdes, James J.	III	239
Chang, Wayne H.	See Cooke, P. W.	I	291
Chang, Wayne H.	See Choi, K. K.	I	253
Chang, Wayne H.	See Lux, Robert A.	II	295
Chang, Wayne H.	See Shen, Hongen	III	133

AUTHOR	TITLE	VOL	PAGE
Chang, Wayne H.	See Dutta, Mitra	I	399
Checkai, Ronald	See Wentsel, Randall	III	345
Chin, Ernest S. C.	Dynamic Impact Response of Titanium Aluminide Composites	I	237
Choi, K. K.	High Sensitivity Infrared Hot-Electron Transistors	I	253
Christensen, Charles	Optical Mapping of HgCdTe Detector Electronic Properties	I	267
Chu, Shih C.	A More Rational and Rigorous Analysis of In-Bore Projectile	I	277
Clarke, Frederick W.	See Christensen, Charles	I	267
Clayton, Michael	See Sjogren, Maria H.	III	161
Cohen, Arthur	See Beyer, Richard A.	I	97
Cohen, Arthur	See Barrows, Austin W.	I	37
Cole, Melanie W.	See McLane, G.F.	II	339
Cole, Melanie W.	See Lee, Howard S.	II	253
Cole, Melanie W.	See Harvey, James F.	II	45
Colston, James T.	See Valdes, James J.	III	239
Condie, Richard M.	See Sjogren, Maria H.	III	161
Cooke, P. W.	Strained Quantum Well Field Effect Lasers for Optoelectronic Integration	I	291
Cornell, John H.	See Willingham, Reginald	III	369
Cotariu, Steven S.	Transient Energy Transfer in Wide Gap Semiconductors	I	305
Crenshaw, Michael E.	New Optical Switch Stemming from Dipole-dipole Interactions in Dense Media	I	319
Crow, Samuel B.	See Watkins, Wendell R.	III	275

AUTHOR	TITLE	VOL	PAGE
Cullen, John W.	See Willingham, Reginald	III	369
Danielson, Eugene	See Bryzik, Walter	I	205
Daschka, Paul	Ceramic Matrix Fiber Reinforced Composites	I	333
Davis, David	See Reid, Robert H.	III	35
DeAnni, Anthony	See Lux, Robert A.	II	295
DeAnni, Anthony	See Lee, Howard S.	II	253
DeCamp, Brian C.	MAXIMIZE Automatic Target Recognition Algorithm Development Environment	I	349
DeFrank, Joseph J.	See Harvey, Steven P.	II	59
DiDomenico, Leo	See Mizan, Muhammad	II	397
Dillon, Tom M.	A Framework for the Environmental Risk Assessment of Contaminated Sediments at DoD Installations	I	357
Don, R. C.	Ballistic Performance--Material Property Relationships of Glass Armor Versus Shaped Charge Jets	I	373
Doughty, David T.	See Morrison, Robert W.	II	409
Dowding, Robert J.	Current Status of Tungsten Alloy Research	I	383
Dubey, Madan	See Harvey, James F.	II	45
Dudley, M.	See Fazi, C.	I	421
Dutta, Mitra	See Harvey, James F.	II	45
Dutta, Mitra	A Novel High Contrast Optical Modulator in a GaAs-AlAs Structure	I	399
Dutta, Mitra	See Shen, Hongen	III	133

AUTHOR	TITLE	VOL	PAGE
Eckart, D.W.	See McLane, G.F.	II	339
Ehlen, Judy	See Hansen, John V.E.	II	1
Eicke, John S.	See Wiles, George C.	III	357
Eidelman, Shmuel	See Bishop, Steven S.	I	119
Eldefrawi, Moyhee D.	See Thompson, Roy G.	III	205
Ennis, Douglas H.	See Atha, Larry C.	I	5
Esch, Frederick H.	Imagery Exploitation System/Balanced Technology Initiative	I	407
Evaldsscn, P. A.	See Cooke, P. W.	I	291
Evans, Timothy D.	See Hansen, John V.E.	II	1
Ezzell, John W.	See Friedlander, Arthur	I	445
Fazi, C.	Observation of rf-Induced Failures in Silicon Bipolar Junctions Using Synchrotron X-Ray Topographic Imaging	I	421
Ferrando, Ronald E.	See Schuschereba, Steven	III	119
Fisher, Jamie M.	Characterizing the Response of Solid Rocket Motors to Fragment Impact	I	429
Forch, Brad E.	See Barrows, Austin W.	I	37
Franz, David R.	See Sjogren, Maria H.	III	161
Friedlander, Arthur	Post-exposure Prophylaxis Against Experimental Inhalation Anthrax	I	445
Gatto, Lawrence	See Melloh, Rae	II	381
Gentry, Mary K.	See Lebeda, Frank J.	II	237
Gillespie, James B.	Battlefield Environmental Effects on Ultraviolet Atmospheric Sensing Propagation	I	459

AUTHOR	TITLE	VOL	PAGE
Gordon-Smith, D.	See Fazi, C.	I	421
Gorman, Francis J.	See Soicher, Haim	III	193
Grisham, John A.	See Christensen, Charles	I	267
Grogl, Max	See Martin, Rodger K.	II	301
Guadagno, Janine	Landmine Detection by Backscatter Radiation Radiography	I	471
Gupta, N.	Optical Neural Network Device Architecture	I	485
Hack, Dallas C.	See Lebeda, Frank J.	II	237
Han, W.Y.	See Lee, Howard S.	II	253
Hansen, John V.E.	Mine Detection in Dry Soils Using Radar	II	1
Haraburda, Scott S.	Developmental Research for Designing a Microwave Electrothermal Thruster	II	15
Hardaway, Mike	See Welsh, James P.	III	335
Harnden, Phillip G.	See Shires, Dale R.	III	147
Harper, Alice M.	Artificial Intelligence for Chemical and Biological Detectors	II	31
Harvey, James F.	See Lux, Robert A.	II	295
Harvey, James F.	Physics and Device Applications of Silicon Microclusters	II	45
Harvey, Steven P.	Biodegradation of Chemical Warfare Agents: Demilitarization Applications	II	59
Hayes, Pamela G.	Vulnerability of Protective Structures to Artillery Fire	II	71
Hepfinger, Lisa B.	Laser-Induced Luminescence: Soldier Vulnerability	II	83

AUTHOR	TITLE	VOL	PAGE
Hevenor, Richard A.	See Hansen, John V.E.	II	1
Higgins, Thomas P.	See Lux, Robert A.	II	295
Hock, Vincent F.	See Ruzga, Richard J.	III	63
Holly, Frank F.	See Shires, Dale R.	III	147
Hoock, Donald W.	An Approach to Mitigating Atmospheric Effects on Image-Based Pattern Recognition by Neural Networks	II	95
Horner, David A.	Optimization of Particle Model for Large-Scale Computing	II	109
Howe, Gerald B.	See Friedlander, Arthur	I	445
Huff, William L.	See Hayes, Pamela G.	II	71
Hunt, Robert E.	Host Respiratory Protection for Lethal Staphylococcal Enterotoxin B	II	119
Hursh, Steven R.	Modeling Human Performance to Predict Unit Effectiveness	II	129
Ivins, Bruce E.	See Friedlander, Arthur	I	445
Jasper, L.	See Kim, A.	II	197
Jasper, L.	See Kim, A.	II	185
Jett, Marti	See Hunt, Robert E.	II	119
Johnson, Anthony J.	See Hunt, Robert E.	II	119
Johnson, Daniel L.	See Patterson, James H.	III	25
Johnson, John C.	High Bandwidth, Wide Field of View Lasercom Demonstration	II	145
Johnson, M.A.	See Meisel, L.V.	II	365
Jones, K.A.	See Lee, Howard S.	II	253
Juhasz, Arpad	Solid Propellant Electrothermal Gun Propulsion	II	159

AUTHOR	TITLE	VOL	PAGE
Kaplan, Zvi	See Juhasz, Arpad	II	159
Katechis, James C.	Dormancy for the Ground Based Interceptor	II	173
Kim, Hie-Joon	See Ross, Edward	III	49
Kim, A.	Monolithic, Photoconductive Impulse-Generating Device	II	197
Kim, A.	High-Power Impulse Generation Using a Photoconductively Switched Radial Transmission Line	II	185
King, Wendell C.	See Witten, Alan J.	III	383
Klohn, Ken	See Mizan, Muhammad	II	397
Klopcic, J Terrence	See Walbert, James N.	III	261
Kodak, James A.	See Reid, Robert H.	III	35
Komisar, Jack L.	See Hunt, Robert E.	II	119
Koscica, Thomas E.	See Lux, Robert A.	II	295
Kumar, Ashok	See Daschka, Paul	I	333
Kumar, Pramod	See Valdes, James J.	III	239
Kustin, Kenneth	See Ross, Edward	III	49
Kyle, Dennis E.	See Martin, Rodger K.	II	301
Ladas, Andrew P.	See Wiles, George C.	III	357
LaiHing, Kenneth	See McDonald, Joseph K.	II	327
Lane, Gerald R.	Establishing Integratged Two-Man Crew Station (ITCS) Requirements	II	209
Lanzerotti, Yvonne	Power Spectral Characterization of Fracture Surfaces of TNT, Composition B, and Octol	II	223
Lareau, Richard T.	See McLane, G.F.	II	339

AUTHOR	TITLE	VOL	PAGE
Lareau, Richard T.	See Harvey, James F.	II	45
Lareau, Richard T.	See Lee, Howard S.	II	253
Lareau, Richard T.	See Schauer, Stephen N.	III	105
LaSala, John E.	See Cotariu, Steven S.	I	305
Lavene, Bernard	See Binder, Michael	I	111
Lawrence, Wade B.	See Friedlander, Arthur	I	445
Leavitt, Richard P.	See Tobin, Mary S.	III	231
Leavitt, Richard P.	See Gupta, N.	I	485
Lebeda, Frank J.	Thermodynamic Analyses of Transmembrane Channel Formation by Botulinum Toxin	II	237
Lee, Howard S.	See Schauer, Stephen N.	III	105
Lee, Howard S.	See McLane, G.F.	II	339
Lee, Howard S.	Ohmic Contacts to Heavily Carbon-Doped p^+-GaAs Using Ti/Si/Pd	II	253
Lemire, George W.	See Sausa, Rosario C.	III	91
Lenz, D. E.	See Brimfield, A. A.	I	167
Lepore, A.	See McLane, G.F.	II	339
Li, W. Q.	See Tober, Richard L.	III	221
Lieb, Robert J.	A Small Angle Neutron and X-Ray Scattering Study of the Onset and Nature of Fracture of Uniaxially Compressed Gun Propellants	II	267
Liston, John E.	Exoatmospheric Intercept: GBI Program and the ERIS FTV Flight Tests	II	281
Lottati, Isaac	See Bishop, Steven S.	I	119

AUTHOR	TITLE	VOL	PAGE
Lowe, John R.	See Friedlander, Arthur	I	445
Lu, Y.	See Lee, Howard S.	II	253
Lukasavage, William	See Nicolich, Steven	III	1
Lukaszek, Ted	See Mizan, Muhammad	II	397
Lund, David J.	See Schuschereba, Steven	III	119
Lux, Robert A.	Optical Control of Resonant Tunnel Diode Devices	II	295
Lux, Robert A.	See Harvey, James F.	II	45
Mammone, Robert J.	See Binder, Michael	I	111
Martin, Rodger K.	*Plasmodium falciparum:* Characterization of Multiple Drug Resistance Phenotypes by Reversal Modulators	II	301
Matthew, Candace B.	An Animal Model of Drug-Induced Thermoregulatory and Endurance Decrements	II	315
Maxwell, D. M.	See Brimfield, A. A.	I	167
Mays, Brian T.	See Wiles, George C.	III	357
McDonald, Joseph K.	Nonlinear Optical Properties of Layered Composites	II	327
McKee, Kelly T.	See Sjogren, Maria H.	III	161
McLane, G.F.	Magnetron Ion Etching for GaAs Device Processing	II	339
McNally, Richard E.	See Hursh, Steven R.	II	129
McQuaid, Michael J.	Imaging and Analysis of Combusting Liquid Propellant Sprays	II	353
McQueen, Charles E.	See Reid, Robert H.	III	35
Meisel, L.V.	Multifractal Analysis of Chaotic Point Sets	II	365

AUTHOR	TITLE	VOL	PAGE
Melloh, Rae	River and Lake Ice Conditions as Interpreted from Microwave Imagery	II	381
Melnik, D.	See Juhasz, Arpad	II	159
Menking, Darrel E.	See Thompson, Roy G.	III	205
Meyyappan, M.	See McLane, G.F.	II	339
Mikesell, Perry	See Friedlander, Arthur	I	445
Milhous, Wilbur K.	See Martin, Rodger K.	II	301
Mizan, Muhammad	Microwave Solid-State Transmitter for Miniature Multiband Beacon Transponder	II	397
Miziolek, Andrzej W.	See Barrows, Austin W.	I	37
Miziolek, Andrzej W.	See Sausa, Rosario C.	III	91
Morrison, Robert W.	Chromium-free Carbon, a New Adsorbent for Chemical Warfare Agent Filters	II	409
Morton, David C.	See Harvey, James F.	II	45
Mozo, Ben T.	See Patterson, James H.	III	25
Nakashima, Masato	See Willingham, Reginald	III	369
Namaroff, M.	See McLane, G.F.	II	339
Nelson, William R.	See Patterson, James H.	III	25
Newberry, Joyce E.	See Barrows, Austin W.	I	37
Newman, Peter G.	See Dutta, Mitra	I	399
Newman, Peter G.	See Lux, Robert A.	II	295
Nickolaisen, Scott	See Cartland, Harry E.	I	221
Nicolich, Steven	Quantitative Synthesis of HMX via the GARDEC HMX Process	III	1

AUTHOR	TITLE	VOL	PAGE
Norvelle, F. Raye	Using Iterative Orthophoto Refinements to Correct Digital Elevation Models	III	13
Nuzum, Edwin O.	See Martin, Rodger K.	II	301
Overholt, James L.	See Wehage, Roger A.	III	307
Palacios, Fernando	See Watkins, Wendell R.	III	275
Pamulapati,Jagadeesh	See Dutta, Mitra	I	399
Pamulapati,Jagadeesh	See Harvey, James F.	II	45
Paolella, Arthur C.	See Lux, Robert A.	II	295
Parmelee, Robert	See Wentsel, Randall	III	345
Patterson, James H.	New Human Exposure Limits for Freefield Artillery Blast Overpressure	III	25
Pearson, Earl F.	See McDonald, Joseph K.	II	327
Pennise, Christine	See Boesch, H. Edwin, Jr.	I	135
Peters, Clarence J.	See Sjogren, Maria H.	III	161
Peters, John F.	See Horner, David A.	II	109
Petreanu, John	See Daschka, Paul	I	333
Pham, John	See Gupta, N.	I	485
Pham, John	See Tobin, Mary S.	III	231
Phillips, Yancy Y.	See Patterson, James H.	III	25
Piekarz, Richard C.	See Harvey, James F.	II	45
Pinto, James J.	See Lanzerotti, Yvonne	II	223
Pitt, Louise	See Hunt, Robert E.	II	119
Pitt, Margaret L.M.	See Friedlander, Arthur	I	445
Quong, Julie A.	See Schuschereba, Steven	III	119

AUTHOR	TITLE	VOL	PAGE
Rael, Eppie D.	See Valdes, James J.	III	239
Rapacki, E. J., Jr.	See Don, R. C.	I	373
Reid, Michael	See Bryzik, Walter	I	205
Reid, Robert H.	Oral, Microencapsulated, CFA/II Vaccine Against E. coli Diarrheal Disease: Preclinical Evaluation	III	35
Remy, David E.	See Willingham, Reginald	III	369
Richmond, Donald R.	See Patterson, James H.	III	25
Ripple, Gary	See Patterson, James H.	III	25
Roach, Joseph F.	See Willingham, Reginald	III	369
Roberson, William E.	See Liston, John E.	II	281
Rod, B. J.	See Brody, Philip S.	I	175
Rogers, Kim R.	See Thompson, Roy G.	III	205
Rondeau, Ernest	See Binder, Michael	I	111
Rose, Kenneth J.	See Friedlander, Arthur	I	445
Rosen, David L.	See Gillespie, James B.	I	459
Ross, Edward	Kinetic Model for Predicting Bacterial Destruction from Intrinsic Chemical Marker Formation in Thermally Processed Foods	III	49
Ruble, David	See Hunt, Robert E.	II	119
Ruzga, Richard J.	See Daschka, Paul	I	333
Ruzga, Richard J.	The Infiltration of Cotton Fibers with Sol-Gel Solutions, Forming Oxide and Non-Oxide Ceramics	III	63

AUTHOR	TITLE	VOL	PAGE
Sadoff, Gerald C.	Cloning, Expression, Production and X-ray Crystallographic Structure of Acetylcholinesterase	III	79
Saphier, D.	See Juhasz, Arpad	II	159
Sasserath, J.	See McLane, G.F.	II	339
Sau, Keya	See Reid, Robert H.	III	35
Sausa, Rosario C.	Laser-Based Sensitive Detection of Trace Atmospheric Vapors of Military Interest	III	91
Scalora, Michael	See Crenshaw, Michael E.	I	319
Schauer, Stephen N.	Accurate Measurement of Metal Diffusion in Electronic Materials Using Secondary Ion Mass Spectrometry and Scanning Auger Microscopy	III	105
Schiff, Lewis B.	See Weinacht, Paul	III	321
Schuschereba, Steven	Basic Fibroblast Growth Factor Accelerates Repair and Prevents Degeneration in Retinal Laser Lesions	III	119
Schwarz, Ernest	See Bryzik, Walter	I	205
Sergi, Sergio A.	See Bishop, Steven S.	I	119
Shafferman, Avigdor	See Sadoff, Jerald C.	III	79
Shattuck, Edgar	See Ross, Edward	III	49
Shen, Hongen	Non Contact, Non Destructive Optical Method for Measuring Electric Fields in GaAs, AlGaAs and InGaAs Heterostructures	III	133
Shen, Hongen	See Harvey, James F.	II	45
Shen, Hongen	See Dutta, Mitra	I	399

AUTHOR	TITLE	VOL	PAGE
Shires, Dale R.	High-Ratio Bandwidth Reduction of Video Imagery for Teleoperation	III	147
Silman, Israel	See Sadoff, Jerald C.	III	79
Simeonsson, Josef B.	See Sausa, Rosario C.	III	91
Simonis, G. J.	See Gupta, N.	I	485
Sjogren, Maria H.	Safety and Kinetics of a Despeciated Equine $F(ab)_2$ Heptavalent Botulinum Antitoxin in Volunteers	III	161
Smith, Doran D.	Scaling Properties of the Magnetoexciton Problem	III	169
Smith, Leonard A.	Cloning, Characterization, and Expression of Animal Toxin Genes for Vaccine Development	III	179
Smith, Stan	See Juhasz, Arpad	II	159
Soicher, Haim	Regression/Correlation Analysis of Ionospheric Parameters at Midlatitudes	III	193
Soreq, Hermona	See Sadoff, Jerald C.	III	79
Starks, Michael W.	See Walbert, James N.	III	261
Stead, M. R.	See Gupta, N.	I	485
Sturek, Walter B.	See Weinacht, Paul	III	321
Sturzebecher, Dana	See Lux, Robert A.	II	295
Sussman, Joel	See Sadoff, Jerald C.	III	79
Tanton, George A.	See Christensen, Charles	I	267
Taub, Irwin	See Ross, Edward	III	49
Taylor, Thomas L.	See Boesch, H. Edwin, Jr.	I	135
Taylor, G. W.	See Cooke, P. W.	I	291

AUTHOR	TITLE	VOL	PAGE
Taysing-Lara, M.	See Choi, K. K.	I	253
Thompson, R.J.	See Lee, Howard S.	II	253
Thompson, Roy G.	Chem-Biodetection: Potentiometric and Fiber Optic Biosensor Test Beds	III	205
Thomson, David J.	See Lanzerotti, Yvonne	II	223
Thornton, Clarence	See Harvey, James F.	II	45
Tober, Richard L.	Electric Field Effects on the Optical Properties of Layered Semiconductor Structures	III	221
Tobin, Mary S.	Simultaneous Electron and Hole Tunneling in Coupled Quantum Wells	III	231
Tressler, L.	See Don, R. C.	I	373
Trevino, Samuel F.	See Lieb, Robert J.	II	267
Tseng, Lee-Ying	See Reid, Robert H.	III	35
Tseng, Jeenan	See Hunt, Robert E.	II	119
Tsu, Raphael	See Harvey, James F.	II	45
Ursic, James R.	See Witten, Alan J.	III	383
Valdes, James J.	Detection of Picogram Quantities of Botulinum Toxin-B (BoTX) Using the Light Addressable Potentiometric Sensor	III	239
Valdes, James J.	See Thompson, Roy G.	III	205
Vanderbeek, Richard	See Harper, Alice M.	II	31
Vargas, Joseph A.	See Schuschereba, Steven	III	119
Veney, David W.	See Cartland, Harry E.	I	221
Wade, William L.	See Binder, Michael	I	111

UTHOR	TITLE	VOL	PAGE
agner, George W.	Solid-State ^{31}P MAS NMR Study of G-Agent Simulants Adsorbed on Synthetic Resin Catalysts	III	247
albert, James N.	A Taxonomy for the Vulnerability/Lethality Analysis Process	III	261
alker, John J.	See Katechis, James C.	II	173
ang, S.	See Fazi, C.	I	421
ard, J. Richard	See Wagner, George W.	III	247
atkins, Wendell R.	Camouflage Flicker: A Highly Observable Low Observable	III	275
atson, Jerry L.	Propellant Response to Shaped Charge Jet Impacts	III	291
ehage, Roger A.	New Directions in Computational Dynamics at TACOM	III	307
einacht, Paul	Navier-Stokes Predictions of Pitch Damping for Axisymmetric Shell Using Steady Coning Motion	III	321
einer, M.	See Kim, A.	II	197
einer, M.	See Kim, A.	II	185
elkos, Susan L.	See Friedlander, Arthur	I	445
elsh, James P.	Review of Environmental Research Specific to Smart Weapons Operability Enhancement for the Battlefield Environment	III	335
entsel, Randall	Novel Techniques to Determine Ecological Effects at U.S. Army Sites	III	345
est, Wade	See Welsh, James P.	III	335
iles, George C.	Projectile Tracking Device Using GPS	III	357
ilhelmsen,Catherine	See Hunt, Robert E.	II	119

AUTHOR	TITLE	VOL	PAGE
Wilhelmsen,Catherine	See Reid, Robert H.	III	35
Willingham, Reginald	Synthesis of Metallo-Tetrabenzporphyrins Possessing High Third-Order Optical Nonlinearity for Military Laser Eye Protection	III	369
Witten, Alan J.	High Resolution Image Processing of Geophysical Data with Diffraction Tomography	III	383
Wittig, Curt	See Cartland, Harry E.	I	221
Wolfe, Allan	See Lanzerotti, Yvonne	II	223
Woolsey, Patrick	See Chin, Ernest S. C.	I	237
Worsham, Patricia L.	See Friedlander, Arthur	I	445
Yalamanchili, Rao	Aerodynamic Heating of Unsteady Hypervelocity Projectile	III	399
Yang, Janet M.	See Schuschereba, Steven	III	119
Yang, L.W.	See Lee, Howard S.	II	253
Youmans, R.	See Kim, A.	II	197
Youmans, R.	See Kim, A.	II	185
Zeto, R.	See Kim, A.	II	197
Zeto, R.	See Kim, A.	II	185
Zhou, Weimin	See Harvey, James F.	II	45
Zukas, Walter X.	Chemistry of Adhesives at Surfaces	III	417

Mine Detection in Dry Soils Using Radar

Mr. John V.E. Hansen*, Dr. Judy Ehlen, Mr. Timothy D. Evans
and Mr. Richard A. Hevenor
U.S. Army Topographic Engineering Center
Research Institute
Fort Belvoir, Virginia 22060-5546

Introduction

This project resulted from two independent developments. First, minefields could be anticipated as a major part of the defense system in any military action in the Middle East. Second, previous work showed that subsurface features and former waterways were detectable by spaceborne radars in arid soils. [1,2,3] Other investigators[4,5,6,7] provided similar supporting evidence of ground-penetrating radar capabilities, particularly in arid regions. Although a number of previous studies have been conducted to remotely detect mines, none of these attempted to exploit the penetrating capabilities of long-wavelength radar for this application in arid soils. Based on these developments, an effort was launched to demonstrate the feasibility of detecting minefields in dry soils using L-band imaging synthetic aperture radar (SAR). Buried mines were not detected under the conditions of the experiment, but a surprising result was a strong signal of disturbed soil obtained under some conditions. These findings provide a number of opportunities warranting further research.

Background

The project plan developed by the Army's Topographic Engineering Center (TEC) involved: (1) locating a site that contained dry, sandy soils, (2) constructing a minefield representative of a typical threat, (3) imaging the site with an airborne radar system, and (4) assessing the usefulness of the information to Army image analysts and others.

The U.S. Marine Corps Air Ground Combat Center (MCAGCC) at Twentynine Palms, California was selected for the test site because it offered several advantages. These included: (1) dry soils, (2) a secluded test site that could be protected, (3) few problems with respect to air traffic control, since MCAGCC is off-limits to commercial aircraft, and (4) Marine Corps assistance in terms of personnel and equipment to perform the test.

The site selected is near Gypsum Ridge (coordinates 78/01; 34°20'30"N,

116°9'W). It is indistinguishable from many sites in the Middle East. Of the many variables affecting the results of radar ground-penetration studies, moisture is probably the most important. Water is an efficient absorber of microwave energy, and its presence reduces penetration. Dry soils are thus most desirable for such experiments, and soil moisture content of less than 2 percent is an acceptable upper limit (G.G. Schaber, U.S. Geological Survey, personal communication, 1991). Surface soil moisture contents between 0.33 percent and 0.50 percent for Twentynine Palms are thus acceptable.

Soil samples from the Middle East were located from various sources. It was not possible to determine soil moisture content on the samples from the Middle East, because they were not collected under controlled conditions. However, citations in the literature suggest soil moisture is significantly below 1 percent.[7] These samples were compared to soils from Twentynine Palms with respect to particle size and were determined to be adequately similar, although the soils from the Middle East are slightly coarser grained and contain more gravel than those from Twentynine Palms.

A test plan reflecting the type of mines and minefields anticipated in Middle East scenarios was developed. Personnel at Belvoir Research, Development and Engineering Center (BRDEC) provided outstanding support in terms of technical advice, and in providing sources and details regarding recommended test mines. The metallic mine chosen was the M-12 training mine. Nonmetallic test items were fabricated with the same dielectric constant as known threat nonmetallic mines (Figure 1). Both types of mines were approximately 0.3m in diameter.

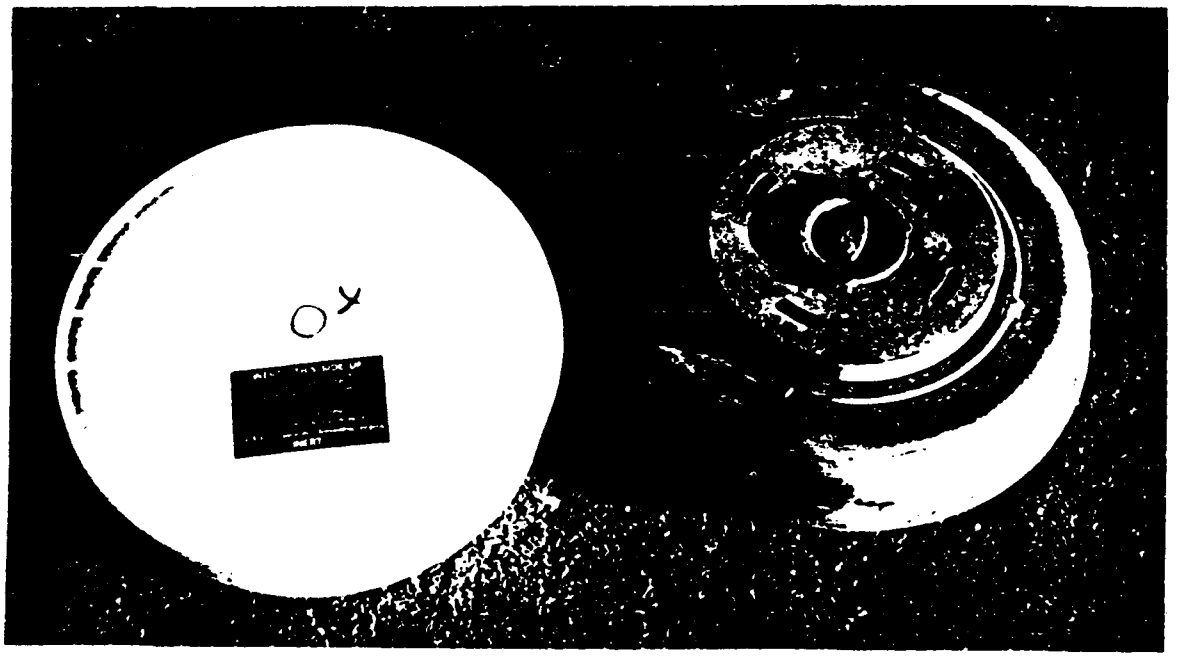


Figure 1 Nonmetallic and Metallic Mines

The test plan was driven primarily by the desire to locate subsurface metallic mines. Nonmetallic mines are also recognized as a potential threat, but the close dielectric constants of such mines and the surrounding sand (approximately 2.5) do not offer much encouragement for detection. They were included in the test plan shown in Figure 2, however, for experimental purposes. Surface mines, both metallic and nonmetallic, were included for several reasons. Although surface mines might not be expected in deliberate minefields, they could be encountered in hasty minefields, or in instances where soil covering shallow mines had been removed by wind, exposing the mines. In addition, they provided a specific indication of the location of the buried mines in the test plan layout. Finally, data were sought on the difference in signals between surface mines and those buried at shallow depths.

A Navy P-3 aircraft carrying X-, C-, and L-band, synthetic aperture radar (SAR) using several polarizations operated by the Naval Air Warfare Center in Warminster, Pennsylvania was used in the demonstration. An airborne radar platform carrying long-wavelength radar was desired because of the relation between radar wavelength and penetration: greater penetration is expected as the wavelength increases, as long as the moisture content is low. Table 1 shows the parameters of this airborne radar system. The resolution shown was considered adequate, given that the mines were spaced at 5m intervals.

Table 1: P-3/SAR Characteristics

System Parameters:			
Bandwidth	60 MHz		120 MHz
Impulse Response Width	3 m		1.5 m
Antenna Parameters:			
	X	C	L
Wavelength (cm)	3.20	5.70	24.00
Frequency (GHz)	9.35	5.30	1.25
Resolution:			
Cell size (m)			
range	1.80	1.80	1.80
azimuth	0.80	0.80	1.10

Phase I

The test site was surveyed and cordoned off with stakes and tapes, and corner reflectors were placed at test site corners. The site was then overflown to obtain images at altitudes of 7,500 and 12,500 feet with X-, C-,

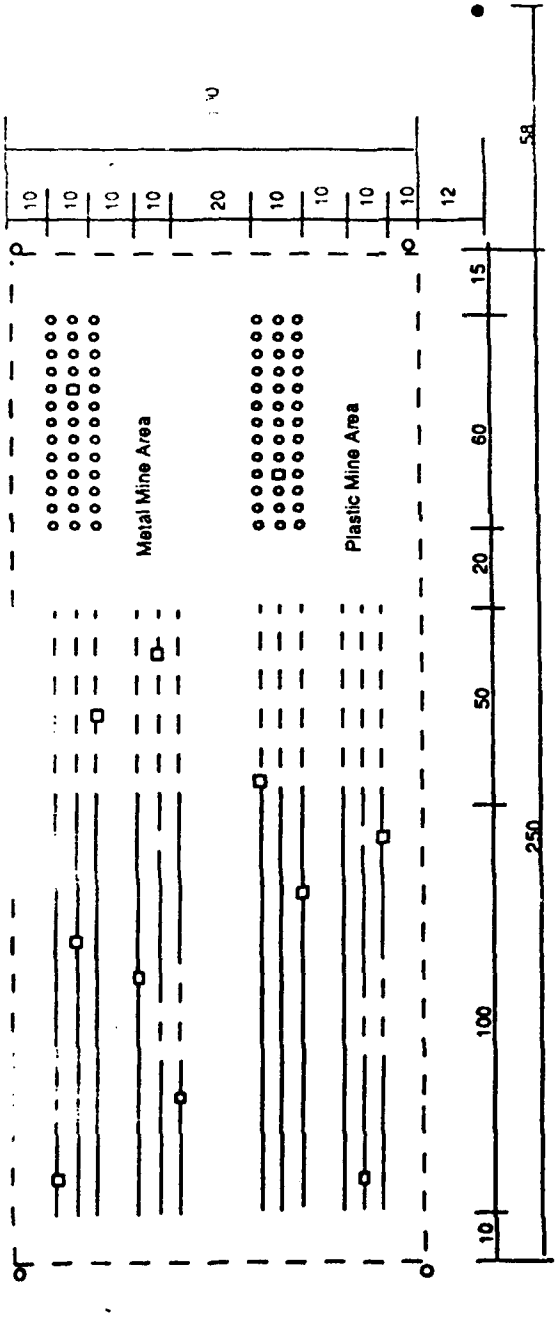


Figure 2 Test Site (Phase I)

and L-band radars. Next, the minefield was constructed using a Small Emplacement Excavator (SEE), which is a militarized backhoe, and a road grader with the blade sharply tilted to produce a v-shaped trench. Mines were surveyed in, with both the intervals between mines and depth being carefully measured (Figure 3). The southern 50 meters of each trench was deliberately left without mines. The first four trenches (on the northeast corner of the site) were dug by the SEE; the remaining trenches were dug by the grader. During site construction, soil samples were gathered from various locations and depths within the trenches.[8] Soil pits were also dug to permit further soil characterization, and surface roughness determinations were made.[9] When all mines were emplaced, the trenches were backfilled. Both the SEE and manual labor were used to backfill the three most easterly trenches. The grader was used to backfill the remaining trenches, covering most of the surface surrounding the trenches with the loose soil removed from the trenches. The area containing surface mines was not significantly disturbed. As soon as the minefield was completed, the site was imaged again. Finally, the minefield was dismantled. The mines were readily located (reflecting the excellent surveying work performed by the Marine Corps personnel) and were removed from the site. The site was then imaged a third time with the radar systems. The imaging overflights were conducted in a three-day period.

Figure 3 Mine Emplacement

Phase II

A month after the first site was dismantled, a second, adjacent site was constructed to obtain additional imagery. The design of this site was the same as the Phase I site, with minor changes (see Figure 4). This site was to be retained for an extended period of time (it is still in existence), and therefore was bordered with barbed wire, with an outer border of concertina wire. Again, a road grader was used to dig the trenches. Care was taken during minelaying operations to confine the area of disturbed soil as much as possible to the trenches in which the mines were laid, in order to simulate the type of disturbance likely to be caused by a minelaying plow. The Phase II site thus contains significantly less disturbed soil than the Phase I site. Soil samples were taken from the same relative locations as in the Phase I site (e.g., samples were taken from both sites at row three, mine 4, etc.). The Phase II site is slightly more moist throughout than the Phase I site (surface soil moisture ranges from 0.47 percent to 1.31 percent), but there were no statistically significant differences between the two sites.[8] Grain size distributions in the soils from the two sites are also comparable.[8]

Radar Imagery

A total of 144 radar phase histories were obtained from overflights of the test site on five separate days. Of these, 95 were taken during Phase I, and 17 were obtained in Phase II. In addition, 32 images were taken three months after the Phase II overflight. HH, VV, HV, and VH polarizations were used in the X, C, and L radar bands. Data were taken at incidence angles of 35 degrees, 50 degrees and 70 degrees. Processing and analysis of the phase histories concentrated on the L-band images, because greater soil penetration was expected with the longer wavelength. The images were processed and analyzed using conventional techniques. New pattern-finding algorithms were also developed during the course of the work. Table 2 summarizes the findings from L-band imagery.

Phase I Results

Unless stated otherwise, both the results described below and the discussion that follows refer to L-band radar.

There is no evidence of buried objects in the imagery. No buried mines or buried corner reflectors were identified by trained Army image analysts in any of the Phase I images. This is true regardless of wavelength, polarization, and incidence angle.

A strong return was obtained in the area of disturbed soil on the test site in some of the images. This signal is strongest with VV polarization, and at the lowest incidence angle (35 degrees) used. Figure 5 shows a strong signal on L-band images where the soil was disturbed.

Pattern-finding algorithms have been developed that can identify/isolate

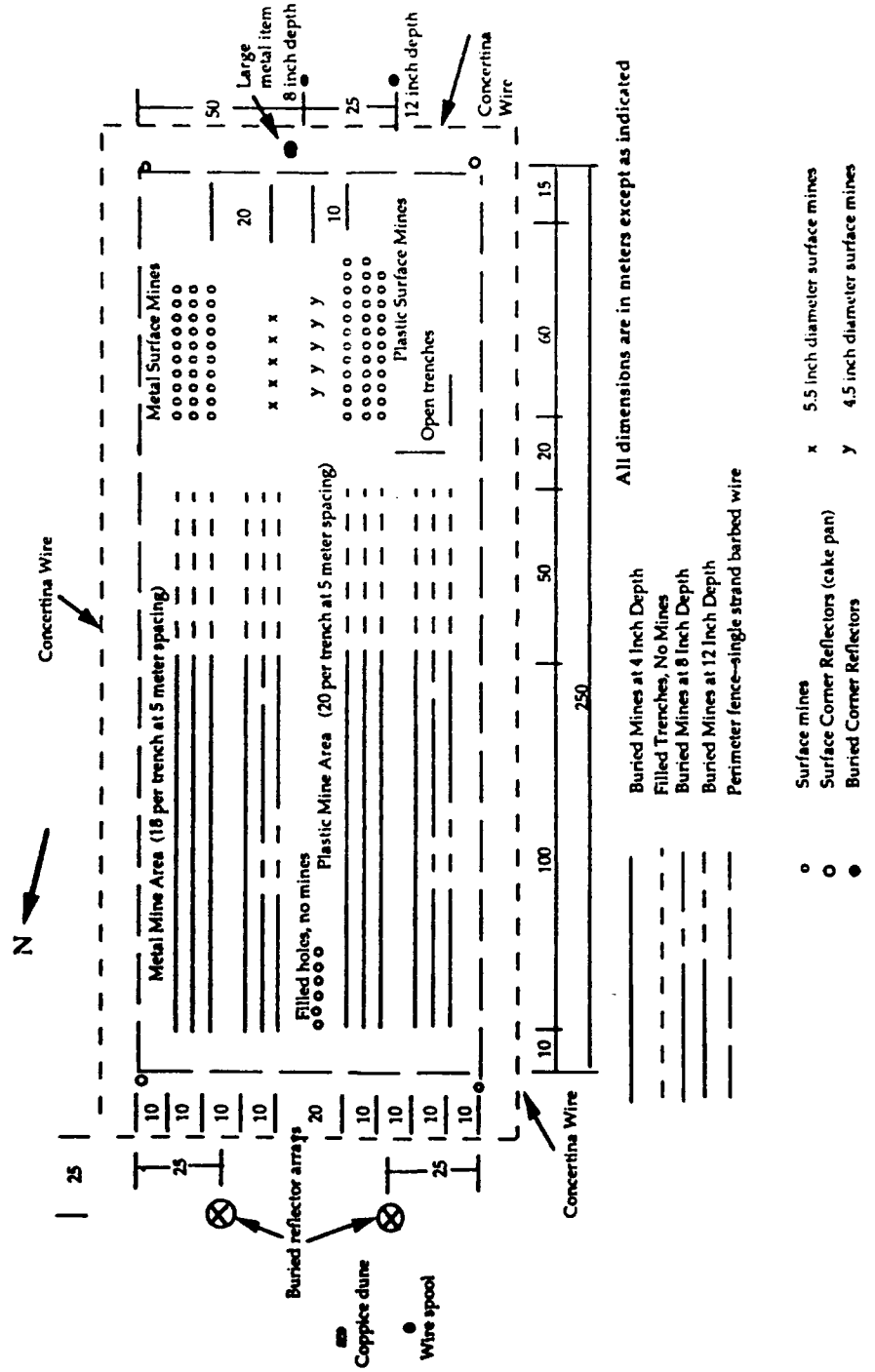


Figure 4 Test Site (Phase II)

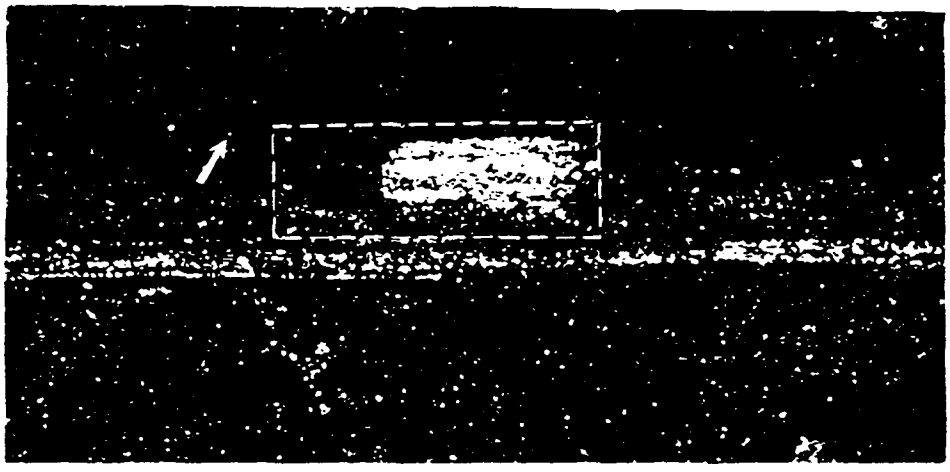


Figure 5 L-band Image Showing Disturbed Soil on Test Site, and Possible Signal from Buried Corner Reflector (arrow)

Table 2: Types of Signals

Incidence Angle	Band/Polarization L_{HH}	L_{HV}	L_{VH}	L_{VV}
Phase I:				
35	disturbed soil no mines	disturbed soil no mines	disturbed soil no mines	disturbed soil weak mines
50	disturbed soil no mines	weak disturbed soil no mines	weak disturbed soil no mines	disturbed soil no mines
70	no disturbed soil no mines	no disturbed soil no mines	no disturbed soil no mines	disturbed soil weak mines
Phase II:				
35	N/A no disturbed soil mines not resolved	N/A	N/A	N/A
50	N/A	N/A	N/A	N/A
70	no disturbed soil mines not resolved	N/A	N/A	no disturbed soil no mines

linear features frequently associated with minefields and the related pattern of disturbed soils. Algorithms have been developed and applied to the linear features associated with the surface metal mines. These algorithms appear robust and may ultimately provide a mechanism to assist image analysts in automatically detecting/isolating linear features associated with minefields.

Strong returns are given by metallic surface objects in some images. Corner reflectors, metal stakes bordering the site, and metal surface mines are clearly visible, and are individually resolved in the C-band imagery. In L-band imagery, the metallic surface mines gave a weak return and were not individually resolved in VV polarization. The surface metal mines were not seen in cross-polarized L-band images nor in L-band images with HH polarization. Fence posts are visible in some L-band images, and on C-band images. Posts in the range direction (East-West) produced stronger images than those in the along-track direction (North-South).

Nonmetallic surface mines are not readily detected. None of the nonmetallic surface mines are visible in any of the L-band images. A few nonmetallic surface mines are weakly visible in C-band images.

Phase II Results

Strong returns were obtained from metal posts, barbed wire, concertina wire, and surface metal mines. In particular, a very bright return was seen from the single roll of concertina wire used to mark the outer boundary of the site. The surface metal mines were highly visible on several of the L-band images, but were not individually resolved. They were individually resolved on one X-band and one C-band image.

There is no evidence of disturbed soil (i.e. trenches, vehicle tracks) on any of the images. L-band imagery shows no evidence of returns that can be attributed to disturbed soil. Figure 6 shows a C-band image of the Phase II site. The barbed wire, and particularly the concertina wire, are clearly visible, but there is no signal from disturbed soil. Note that metallic surface mines are clearly resolved, as was shown in Phase I imagery as well.

There is no evidence of buried mines, either metallic or nonmetallic. Figure 5, however, may show a buried corner reflector. The bright returns at the four corners of the site are corner reflectors located on the surface. A fifth corner reflector is located on the surface near the upper left corner of the site. Halfway between this fifth surface reflector and the reflector at the upper left corner of the image, a corner reflector was buried. This reflector is vertically oriented. Its uppermost tip is located just below the soil surface; its corner is therefore 0.5m below the surface. There is a bright spot in this image in the correct position for a return from this buried reflector; this requires further experimental confirmation.

Discussion

The fact that no signals from buried mines and buried corner reflectors

were readily identified in this experiment using L-band radar indicates that the technology is not as robust as some might expect. The ground penetration achieved by the SIR-A, SEASAT, and SIR-B L-band radars occurred under very special conditions. Analysis of SIR-A images of North Africa followed by intensive field work showed that significant ground penetration occurred,[1,2] and as a result the potential penetrating capabilities of SEASAT were investigated in the Mojave Desert by Blom, et al., who detected subsurface dikes.[3] Subsequently, penetration also occurred in northern Saudi Arabia using SIR-B.[7] The explanation for these results compared with our own at Twentynine Palms undoubtedly lies in the complex nature of microwave ground-penetration physics. According to McCauley, et al.,[1] variations in radar image tone are caused by changes in radar backscatter, which is mainly determined by (1) physical properties such as slope, surface roughness, and characteristics of the soils; (2) subsurface roughness where penetration occurs to a significant degree; (3) the propagation wavelength; (4) the angle of incidence; (5) the polarization of the incident wave; (6) a complex dielectric constant, which in most instances is dominated by moisture content and density; and (7) a complex volume scattering coefficient applicable to random media. Given the complexity and number of variables involved, it is apparent that straightforward explanations for the results described above

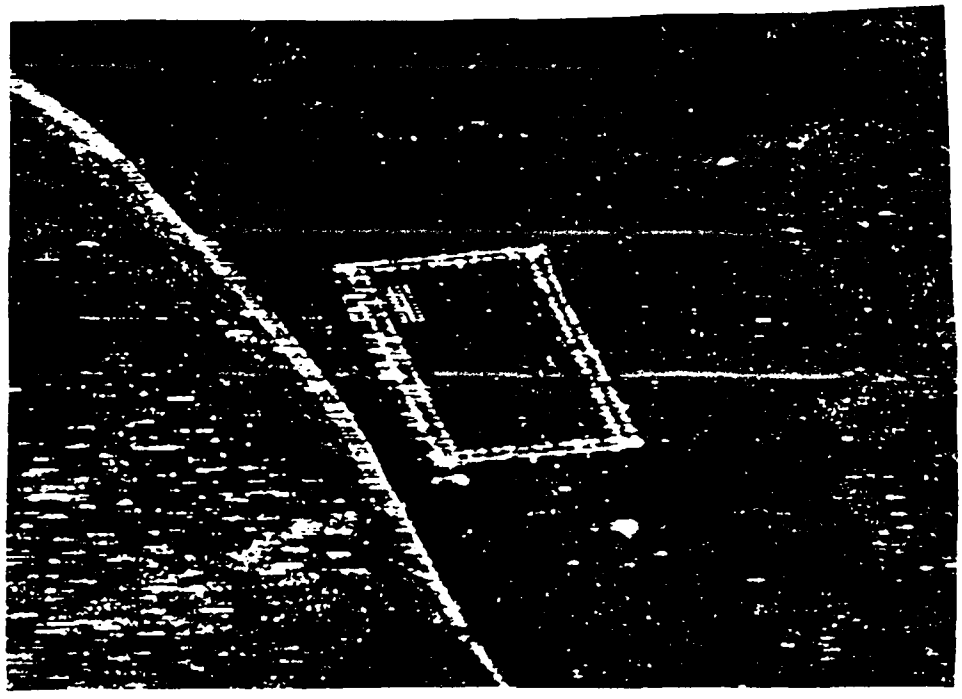


Figure 6 C-band Image Showing Barbed Wire, and Surface Metallic Mines

should not be expected. In addition, the objects sought in this test were metallic mines, approximately 0.3m in diameter. These are small targets when compared with the much larger terrain features in other studies using ground-penetrating radars.[1,3]

Surface roughness, in terms of radar return, deserves a comment. At the Phase I site at Twentynine Palms, surface variations were on the order of one-quarter to one-half wavelength (L band = 24 cm) over a distance of a few meters, and thus approached the criteria for being "radar rough". As stated above, because of the disturbed soil signals obtained in Phase I, efforts were made in Phase II to reduce the amount and area of surface roughness. No disturbed soil patterns were seen on the Phase II imagery. However, to the extent that Phase I was more representative of typical military construction activity, these results suggest that disturbed soil patterns might possibly be used as an indicator of such activity.

The effects of soil moisture warrant comment since there is general agreement that this factor may dominate all others. Many investigators (G.G. Schaber, U.S. Geological Survey, personal communication, 1991) believe that at moisture levels below approximately 2 percent, the soil is adequately dry to achieve penetration when other conditions, such as soil homogeneity, composition (to include salts and clay content), and particle size, are favorable. The effect of soil moisture reducing radar returns was demonstrated by Blom, et al.[3] who were unable to detect buried corner reflectors when soil moisture was in excess of 5 percent. Layering identified in soil pits at the Phase I site is most likely insignificant;[9] clay contents are very low and no salts (such as carbonate or gypsum) were identified; and there were no statistically significant differences in particle size with depth.[8] Surface soil moisture in Phase I ranged from 0.33 percent to 0.50 percent, and in Phase II, from 0.47 percent to 1.31 percent; there were no statistically significant differences in soil moisture with depth or between the two sites. These conditions meet those generally accepted as adequate for penetration to the depths of interest.

The amount and particle size of surface gravel may also have effected the results of this experiment. Blom, et al.[3] report that for L-band radar, surface gravel should be less than 1.5cm to prevent scattering losses; successful penetration has occurred only where sand is predominant and any gravel present is very small and scattered. Gravel in surface samples from Twentynine Palms ranged from 5.2 percent to 34.8 percent, much of which was 1.5 cm or larger in size. Subsurface gravel contents were comparable, as was particle size, so even if surface penetration occurred, scattering losses would be likely from subsurface gravel. Although Roth and Elachi[10] report that only a few particles exceeding the size criteria are needed to prevent detection of subsurface features, it is not clear what amount of gravel is critical. To assess the effect of gravel on these results, some corner reflectors have been buried at a shallow depth at Twentynine Palms; the covering soil consists of dry, sifted sand. A survey with a ground-based radar system is planned when circumstances permit.

The disturbed soil signals on the L-band imagery were strongest at low

angles of incidence. Since the ground surface was "radar rough", return from surfaces at or nearly normal to the incident beam would provide the maximum signal. If the slopes of the soil mounds parallel to the trenches were approximately 35 degrees with respect to the horizontal, these slopes would be roughly normal to the beam of the radar at an incidence angle of 35 degrees as shown in figure 7. This is likely because the angle of repose for granular particles is between 34 and 37 degrees. Signals similar to those noted on the imagery would thus be produced. This suggests that incidence angles equal to the natural angle of repose (i.e. 35 degrees) may be useful in radar systems attempting to identify disturbed soil signals.

Consideration has been given to the possibility that the signals from disturbed soil may result from a change in dielectric constant due to the lower density of the disturbed soil, coupled with the ability of the long-wavelength radar to penetrate and detect this change. Experiments are planned to resolve this issue.

Future work in the area of mine detection and extraction from synthetic aperture radar imagery should emphasize the application of various speckle-noise reduction techniques on the original imagery and also the signal processing of the original phase history information. Several speckle-noise reduction techniques are now available. However, for this effort, only one was used. This was a geometric filter that was applied twice to the original radar image in order to implement the pattern-finding algorithms. Analyzing

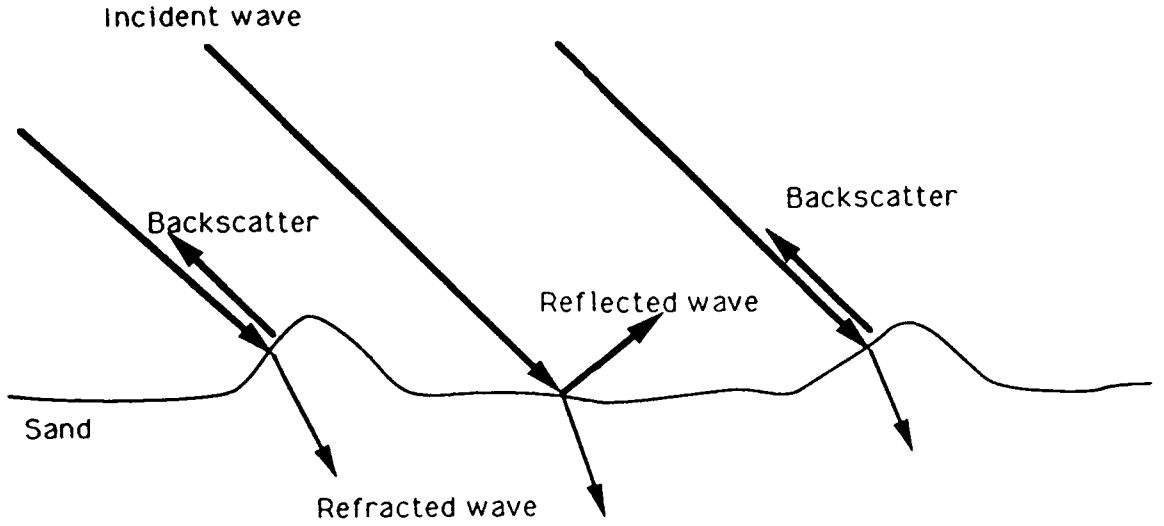


Figure 7 Backscatter Due to "Radar Rough" Surface

the signals in the phase histories deserves attention, as noted. When the signal-to-noise ratio is small as it is for buried mines, it would appear that this would be a good direction for further work.

Future work planned includes the use of systems such as wideband impulse

radars and tomographic radars. Also, a calibrated ground-based system that can record signal return under varying conditions would be very useful in furthering the basic understanding of radar penetration of dry soil.

Conclusions

The results of this project do not provide evidence that a long-wavelength synthetic aperture radar can be used to detect buried mines. However, it is possible that the lessons learned in terms of surface indicators and the progress made in developing pattern-finding algorithms could lead to more robust techniques for locating minefields. It is also likely that further signal processing efforts can contribute to finding minefields and possibly, the buried mines themselves.

Finally, it should be noted that military radar reconnaissance systems typically use short-wavelength radar to identify details of objects and terrain. Long-wavelength systems are not common in operational inventories, and if further research shows that disturbed soil resulting from military operations is preferentially detected by long-wavelength radar, then the addition of such systems to the operational inventory should be considered.

Acknowledgments

We wish to thank Marine Corps personnel from the Combat Development Center (MCCDC) at Quantico, VA and the Air Ground Combat Center (MCAGCC) at Twentynine Palms, CA, and particularly MCAGCC Range Control and the Third Marine Air Wing's 173rd Marine Wing Support Squadron (173 MWSS), for their assistance in creating and operating the test site. We also wish to thank personnel from the Foreign Science and Technology Center (FSTC), Charlottesville, Va; the Navy Air Warfare Command (NAWC), Warminster, PA and particularly Mr. Jim Verdi; and the Belvoir Research Development and Engineering Center (BRDEC), Ft. Belvoir, VA for their advice and assistance. We thank Alan Krusinger, who took thermal measurements, Stuart Allison; who assisted with the preparation of the test site; Ed Simental, who worked on image processing; and Dr. J.N. Rinker, Senior Research Scientist, for his guidance and help; all from TEC. Finally, special thanks are due to John Curtis and Lee Tidwell, U.S.A Waterways Experiment Station, Vicksburg, MS, for their detailed surface roughness and soil characterization effort at the Phase I site.

References

1. McCauley, J.F., G.G. Schaber, C.S. Breed, M.J. Grolier, C.V. Haynes, B. Issawi, C. Elachi, and R. Blom, Science, **218**, 1004-1020 (1982).

2. McCauley, J.F., C.S. Breed, G.G. Schaber, W.P. McHugh, B. Issawi, C.V. Haynes, M.J. Grolier, and A.E. Kilani, IEEE Trans., **GE-24**, 624-648 (1986).

3. Blom, R.G., R.E. Crippen, and C. Elachi, Geology, **12**, 346-349 (1984).

4. Rinker, J.N., Polar Record, **12**, 403-405 (1965).

5. Rinker, J.N., S. Evans, and G. de Q. Robin, Fourth Symp. Rem. Sensing of Environ. Univ. Mich., 793-800 (1966).

6. Schaber, G.G., J.F. McCauley, C.S. Breed, and G.R. Olhoeft, IEEE Trans., **GE-24**, 603-623 (1986).

7. Berlin, G.L., M.A Tarabzouni, A.H. Al-Naser, K.M. Sheikho, and R.W. Larson, IEEE Trans., **GE-24**, 595-602 (1986).

8. Ehlen, J. and J.P. Henley, A Comparison of soils from Twentynine Palms, California and Saudi Arabia, (U.S.A. Topographic Engineering Center, Fort Belvoir, VA, 1990), ETL-0583.

9. Curtis, John O. and L. Tidwell, Twenty-nine Palms, California Test Site Characterization, (U.S.A. Waterways Experiment Station, Vicksburg, MS, 1992), TR EL-92-14.

10. Roth, L.E. and C. Elachi, IEEE Transactions, **AP-23**, 674-675 (1975).

Developmental Research for Designing a Microwave Electrothermal Thruster

Captain Scott S. Haraburda
Chemistry Department, U.S. Military Academy
West Point, NY 10996

Nomenclature

C_p - Heat capacity
E_i - Energy of particle i
$F()$ - Function vector
$F'()$ - Function derivative matrix
h - Planck's constant
H_i - Enthalpy of particle i
k - Boltzmann constant
K - Equilibrium constant
M_i - Molecular weight of particle i
P - Pressure
q_N - Partition function
R - Gas Constant
T - Temperature
V - Volume
ω - Statistical weights
X_i - Mole fraction of particle i
Z - Compressibility factor
β - Inverse product of temperature and the Boltzmann constant
$\Delta\varepsilon_{1i}$ - Electronic energy change from ground state
ρ - Density

Background

Electrothermal Propulsion

In general, there are three major types of rocket thrusters: chemical, nuclear, and electrical. Chemical rocket thrusters are the most commonly used type of thruster. Energy is transferred to the working fluid through chemical reactions (combustion). Nuclear rocket thrusters are practically and

politically difficult to use. Energy is transferred through nuclear energy. Electrical rocket thrusters are not practical in a large force region, such as gravity from large celestial bodies. Energy is transferred via heating coils or EM waves to the propellant fluid.

There are three types of electrical rocket thruster systems. Electrothermal thrusters use electric energy to heat a conventional working fluid. Electrostatic thrusters use ions or colloidal particles as the working fluid. Finally, Electromagnetic (EM) thrusters use EM fields to accelerate the working fluid, usually in the plasma state.

A proposed electrothermal propulsion system uses microwave induced plasmas. Although this system uses an electromagnetic wave, it is classified as an electrothermal thruster because it uses a nozzle (not EM waves) to accelerate the propellant. Schematically shown in Figure 1 is a version of this system[1].

Microwave or millimeter power beamed to a spacecraft from an outside source (such as a space station or planetary base) is focused onto a resonant cavity to sustain a plasma in the working fluid. The hot gas would expand through a nozzle to produce thrust. Alternatively in a self-contained situation, power from solar panels or nuclear reaction could be used to run a microwave frequency oscillator to sustain the plasma.

Microwave Induction

A microwave plasma is very efficient for uses in jet propulsion. Production of these plasmas involves using plasma columns the size of conventional resonance cavities. These cavities are stable, reproducible, and quiescent. These plasmas develop as a result of surface wave propagation and are characterized by ion immobility. The major physical processes governing the discharge are: (a) discharge conditions (such as the nature of the gas), (b) gas pressure, (c) dimensions and material of the vessel, (d) frequency of the EM field, and (e) the power transferred to the plasma.

For microwave plasma electrothermal rocket thrusters, pressures near atmospheric and gas temperatures near 2000 Kelvin are being investigated. In the EM environment, free electrons are accelerated about the heavier and neutral molecules. These electrons collide with other elements or

molecules of the gas to cause them to ionize as previously bound electrons are stripped off. In essence, kinetic energy is transferred from the accelerated electrons to the gas.

Figure 2 illustrates the various discharge properties within a microwave system[2]. The cold propellant receives microwave energy resulting in production of a plasma. This plasma gives off radiation and heat. The excited species flow away from the plasma while the cold species flow towards it. Finally, the plasma excited propellant is recombined outside of the plasma with increased kinetic energy. This thermalized propellant exits through a nozzle as propulsion thrust.

Numerical Methods

Algebraic Equations

Gauss Elimination is one of the more popular methods for solving "X" number of equations with "X" number of unknowns[3]. This method is used to solve for the unknowns involved in calculating the thermodynamic properties of a plasma. The general equation for linear types of these problems is:

$$\underline{\underline{A}}\ \underline{x} = \underline{b}$$

or

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

Here, $\underline{\underline{A}}$ is defined as the matrix of coefficients of the set of linear equations with $\underline{X}$ as the vector of unknowns and $\underline{b}$ as the vector of products. Unfortunately, many algebraic equations are non-linear. This requires another method of solving for the unknowns, which may have several sets of solutions. However, some of these sets have impossible physical meaning.

The Newton Method is one way to obtain solutions to non-linear equations. To solve problems using this method, one needs to rearrange the set of equations (F(x)) such that they all equal zero. This can be written as:

$$F(\underline{X}) = \underline{0}$$

or

$$\begin{bmatrix} f_1(\underline{X}) \\ \vdots \\ f_n(\underline{X}) \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

The Newton Method requires an initial guess of values for the variables. Values of the variables are iteratively modified to quadratically approach the actual solution. The formula for this variable modification is:

$$\underline{X}^{n+1} = \underline{X}^n - \frac{F(\underline{X}^n)}{F'(\underline{X}^n)}$$

Here, $F'(\underline{X}^n)$ is defined as the derivative of $F(\underline{X}^n)$. Because this form is linear, each iteration can be solved using Gauss Elimination. This linear system can be obtained by rearranging this formula.

$$F'(\underline{X}^n)\ (\underline{X}^{n+1} - \underline{X}^n) = - F(\underline{X}^n)$$

Differential Equations

The Finite Difference method may be used to solve Partial Differential Equations[4]. One would approximate each derivative using the Taylor series expansion of a function about a nodal point. The following are examples of approximations for derivatives (use $g = \Delta x$):

(a) forward: $$\frac{df(a)}{dx} = \frac{f(a+g) - f(a)}{g}$$

(b) backward: $$\frac{df(a)}{dx} = \frac{f(a) - f(a-g)}{g}$$

(c) centered: $$\frac{df(a)}{dx} = \frac{f(a+g) - f(a-g)}{2g}$$

For second order derivatives, the following centered difference can be used:

$$\text{(d)}\quad 1 \text{ variable: } \frac{d^2 f(a)}{dx^2} = \frac{f(a+g) + f(a-g) - 2f(a)}{g^2}$$

Solving Partial Differential Equations using this method is similar to that of Galerkin methods in that the Partial Differential Equations are transformed into a linear set of algebraic equations with the functions at the nodal points being unknowns.

Experimental System

Microwave Cavity

An electromagnetic system was used to generate a plasma. The microwave cavity body was made from a 17.8 cm inner diameter brass tube. As seen in Figure 3, the cavity contained a sliding short and a coupling probe (the two major mechanical moving parts of the cavity). The movement of this short allowed the cavity to have a length varying from 6 to 16 cm. The coupling probe acted as an antenna which transmitted the microwave power to the cavity. The sliding short and coupling probe were adjusted (or moved) to obtain the desired resonant mode. A resonant mode represents an eigenvalue of the solution to Maxwell's four electromagnetic equations. Illustrated in Figure 4, two separate resonance modes were used in these experiments: TM 011 (Ls = 7.2 cm) and TM 012 (Ls = 14.4 cm)[2].

Additional features of this cavity included: two copper screen windows located at 90 degree angles from the coupling probe (which allowed photographic and spectral measurements), and two circular holes (in both the base and top plates) to allow propellant and cooling air flow through the cavity.

Plasma Containment

The plasma was generated in quartz tubes placed within the cavity. The inner tube was 33 mm outer diameter and was used for the propellant flow. The outer tube was 50 mm outer diameter and was used for air cooling of the inner tube. Both tubes were about 2 1/2 feet long and were epoxied to aluminum collars. These collars fed the gas and air through the cavity. For additional protection, water cooling was done on the collar downstream of the cavity.

Flow System

Flow of 99.99% pure helium was controlled using a back pressure regulator and a 3/4 inche valve in front of the vacuum pump. Calibrated to zero assuming a complete vacuum from the pump, a Heise gauge with a range from 1-1600 torr was used to measure the pressure of the plasma chamber. Three sets of flow meters were used to measure the gas, water, and air flows. Calibrated using boiling water, type T thermocouples (copper constantan) and an Omega 400B Digicator were used to measure the temperature of the air and water both entering and exiting the cavity.

Microwave Power

A Micro-Now 420B1 (0-500 watts) microwave power oscillator was used to send up to 400 watts of power at a fixed frequency of 2.45 GHz to the cavity. Although rated for 500 watts, energy was lost from the microwave cable, circulator, and bidirectional coaxial coupler.

Connected to the microwave oscillator was a Ferrite 2620 circulator. This circulator provided at least 20 dB of isolation to each the incident and reflected power sensors. The circulator protected the magnetron in the oscillator from reflected signals and increased the accuracy of the power measurements. The reflected power was absorbed by the Termaline 8201 coaxial resistor. The incident and reflected powers were measured using Hewlett-Packard 8481A power sensors and 435A power meters.

Energy Transport

Energy can be transported through conductive, convective, and radiative means.

Radiation

A major loss mechanism in energy transport to the propellant is energy loss to the walls of the discharge chamber. The following parameters were used for this analysis: the cavity material is unpolished brass, the plasma geometry is an oblate ellipsoid, and the wall temperature was 300 Kelvin. Assuming graybody radiation transport within the cavity and knowing the heat rate loss to the cavity wall, one can predict the surface temperature of the plasma (assuming

that it acts as a solid object). Figure 5 shows that temperature as a function of pressure[5].

Conduction / Convection

The remaining loss mechanism in energy transport is energy loss through heat conduction and convection. A way to visually understand this loss is to see the temperature profile within the cavity. Using energy transport equations and known boundary conditions (temperature and velocity profile), one can approximate the profile using numerical methods. For example, figure 6 shows the temperature profile of the air cooling chamber[5]. The major source of error in developing this profile is the error involved in experimentally determining the boundary conditions.

Thermodynamic Properties

Energy, mass, and momentum transport phenomena is involved in this system. To predict these transport processes, one could solve the conservation equations for each. Although it is not the intent of this paper to present these fundamental equations, calculation of some of the parameters involved will be presented. Density is important in determining the amount of particles present in any given location. The energy and enthalpy values are important in determining the amount of energy present. Finally, heat capacity is important in determining the transfer of energy from one location to another. Although diffusion parameters are important in determining the transfer of mass from one location to another, these calculations are not included.

Statistical Mechanics

Thermodynamic properties of the plasma are obtained through use of statistical mechanics and equilibrium properties[6]. This requires the use of partition functions (q_n), which are functions of temperature and volume. For helium, the molecular partition function is a product of its translational (q_{trans}) and electronic (q_{elect}) partition functions. The translational and electronic partition functions are defined as:

$$q_{trans} = \left(\frac{2 p m k T}{h^2} \right)^{3/2} V$$

$$q_{elect} = \sum \omega_{ei} \exp(-\beta \Delta\varepsilon_{1i})$$

Values obtained through these calculations assume ideal gas conditions and local thermodynamic equilibrium.

Equilibrium

Thermodynamic values are calculated assuming equilibrium conditions[7]. The following chemical reactions are used in these calculations:

$$He \overset{K_1}{\rightleftharpoons} He^+ + e^-$$

$$He^+ \overset{K_2}{\rightleftharpoons} He^{++} + e^-$$

The equilibrium constants can be related to the mole fractions of the species as such:

$$K_1 = P \frac{X_{He+} X_{e-}}{X_{He}}$$

$$K_2 = P \frac{X_{He++} X_{e-}}{X_{He+}}$$

The values of these constants can be obtained using statistical mechanics and experimental values for ionization energies ($\Delta\varepsilon_i$)

$$K_1 = \exp(-\beta\Delta\varepsilon_1) \frac{q_{He+} q_{e-}}{q_{He}}$$

$$K_2 = \exp(-\beta\Delta\varepsilon_2) \frac{q_{He++} q_{e-}}{q_{He+}}$$

This provides us with two non-linear algebraic equations and four unknowns - the mole fractions. We can derive two additional equations assuming conservation of particles and electrical neutrality. These equations are:

$$\sum X_i = 1$$

$$X_{e-} - X_{He+} - 2X_{He++} = 0$$

Using numerical methods to solve the four non-linear and linear equations, the mole fractions at various temperatures can be calculated. The values at 1 ATM are provided in Figure 7. Two additional parameters are needed for calculating thermodynamic values. They are the average molecular weight (M) and compressibility factor (Z). Mathematically, they can be calculated as:

$$M = \sum X_i \, M_i$$

$$Z = \frac{M_o}{M}$$

Here, M_o is defined as the molecular weight at 273.15 Kelvin.

Density

The density of the plasma must be known to be able to develop model equations. The functional relationship between density, pressure, temperature, and compressibility is:

$$\rho = \frac{\rho_o \, P \, T_o}{Z \, P_o \, T}$$

Here, ρ_o, T_o, and P_o are the values at standard state conditions of 273.15 Kelvin and 1 ATM.

Energy / Enthalpy

It is essential to know the energy of the plasma so that one can do heat transfer modelling[8]. The energy level of an individual species can be calculated as:

$$E_i = N \, k \, T^2 \left(\frac{d \ln Q_i}{d T}\right)$$

$$= \frac{3}{2} N k T + N \sum \frac{\omega_i \Delta\varepsilon_i \exp(-\beta\Delta\varepsilon_i)}{q_{elect}}$$

Energy may be expressed in terms of enthalpy.

$$H_i = E_i + N k T$$

The total enthalpy can be calculated using the individual enthalpy levels and compressibility.

$$H = Z \sum X_i E_i + Z R T$$

Heat Capacity

To understand the changes in energy levels with temperature, one must know the heat capacity of the fluid[8]. This parameter is defined as the change in enthalpy with respect to temperature at constant pressure.

$$C_p = \left(\frac{d H}{d T}\right)_p$$

The heat capacity of the individual species is calculated as:

$$C_{p,i} = \frac{5}{2} R + N \frac{d}{dT} \left[\frac{\sum \omega_i \Delta\varepsilon_i \exp(-\beta\Delta\varepsilon_i)}{q_{elect}}\right]$$

However, the overall heat capacity must account for the chemical reactions and the compressibility changes. The overall formula for this comes down to:

$$C_p = Z \sum X_i C_{p,i} + Z \sum \left(\frac{d X_i}{d T}\right)_p H_i - \frac{Z H}{M_o} \sum M_i \left(\frac{d X_i}{d T}\right)_p$$

Non-Propulsive Applications

The following is a literature survey listing potential focus for conducting microwave plasma research.

Detoxification of Hazardous Materials

Environmental concerns have increased dramatically over the past few years. One important area of concern is the storage or elimination of hazardous materials[9].

Processing of hazardous wastes can be accomplished through biological, chemical, physical, or thermal means. Biological treatments include activated sludge, anaerobic filters, and waste-stabilization ponds. Chemical treatments include ion-exchange, reduction-oxidation, and neutralization. Physical treatments include distillation, filtration, and centrifugation. And, thermal treatments include incineration.

An emerging technology is that of plasmas. Plasma technology is used in the space and energy industries. The application of the arc / torch electrode technology can be used in the treatment of hazardous waste.

The plasma torch can be used to dissociate pumpable liquid organic wastes into its elements. Additionally, metal recovery from scrap metal can be done using a plasma. The speed of treatment is extremely fast. The plasma can break down toxic compounds within milliseconds, with power ranges near 1 M watt.

As seen in Table 1, many companies today are not only researching the concept of plasma waste treatment, they are developing plant operations using this technology[9]. Furthermore, Westinghouse is working with the U.S. Department of Energy to develop treatment procedures of buried waste drums.

Surface Treatment of Commercial Materials

The surface treatment of materials is very important in industry[10]. Oxide layer growth, plasma deposition, and plasma etching provide many applications in the production of many consumer goods. These three treatments are schematically illustrated in Figure 8.

Plasmas are used to induce and to speed the growth of an ozide layer on materials. For some metals, this oxide layer can be used as a protective coating.

Plasmas are used to deposit specific compounds on surface materials. One widely used application is the generation of a methane plasma to generate a diamond thin-film layer on the surface.

Finally, plasmas are widely used in the electronic industry. Etching of silicon has many applications in the production of integrated circuits. In the past, etching was done using a "wet" technology (the use of liquid chemical reactions). This old technology had problems related to surface wettability and bubble formation. Additionally, the liquid wastes were dangerous and expensive to dispose. Finally, the size was limited to greater than 3 microns. With plasmas, we can obtain sizes less than 1 micron.

Novel Methods in Chemical Reaction Procedures

Plasmas provide a unique environment for chemical reactions with unlimited potential. Using plasmas, energy can easily be transferred to the reaction; thus, allowing us another means to transport energy to an endothermic reaction. Also, the plasma can produce unique radical species which can create unique compounds. This can be very useful in the organic chemistry arena. The list of applications continues.

Conclusions

Although this paper was not intended to provide much scientific and engineering data regarding microwave plasmas, it does provide enough information to suggest the many uses of microwave generated plasmas. Much effort has been invested in researching the propulsive applications of these plasmas. Through this effort, a large source of information can be harnessed towards developing a fundamental understanding in the non-propulsive applications. Using this knowledge, microwave generated plasma technology can be developed which not only benefits the U.S. Army, but the entire country.

References

[1]Hawley, M., Asmussen, J. Filpus, J. Frasch, L., Whitehair,S., Hoekstra, C., Morin, T., and Chapman, R., "A Review of Research and Development on the Microwave-Plasma Electrothermal Rocket," *Journal of Propulsion and Power*, Vol. 5, No. 6, 1989.

[2]Haraburda, S., "Transport Properties of Plasmas in Microwave Electrothermal Thrusters," M.S. Thesis, Michigan State University, 1990.

[3]Johnson, L., and Ross, R., *Numerical Analysis*, 2d Ed., Phillipines, Addison-Wesley Publishing Co., 1982.

[4]Hall, C., and Porsching, T., *Numerical Analysis of Partial Differential Equations*, Prentice Hall, New Jersey, 1990.

[5]Haraburda, S., Hawley, M., and Dinkle, D., "Theoretical Modelling of Diagnostic Evaluations of Microwave Generated Plasma Systems," paper number IEPC-91-032, AIDAA / AIAA / DGLR / JSASS *22d International Electric Propulsion Conference*, 1991.

[6]McQuarrie, D., *Statistical Mechanics*, Harper & Row Publishers, New York, 1976.

[7]Lick, W., and Emmons, H., *Thermodynamic Properties of Helium*, Harvard University Press, Cambridge, Massachusetts, 1962.

[8]Lick, W., and Emmons, H., *Transport Properties of Helium*, Harvard University Press, Cambridge, Massachusetts, 1965.

[9]Ondrey, G., and Fouhy, K., "Plasma Arcs Sputter New Waste Treatment," *Chemical Engineering*, Vol. 98, No. 12, December 1991.

[10]Hopwood, J., "Macroscopic Properties of a Multipolar Electron Cyclotron Resonance Microwave-Cavity Plasma Source for Anisotropic Silicon Etching," Ph.D. Dissertation, Michigan State University, 1990.

Table 1 Current Plasma Detoxification Systems

WASTE	CAPACITY	DESIGNER
Scrap Metal	50 tons/hr	Westinghouse
Toxic Landfill	2.5 tons/hr	Westinghouse
Toxic Material	1.1 tons/hr	Retech, Inc.
Organic Liquids	439 lbs/hr	Aerospatiale

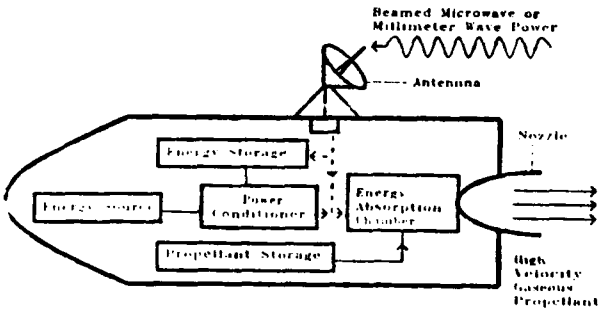


Figure 1. Microwave Electrothermal Thruster.

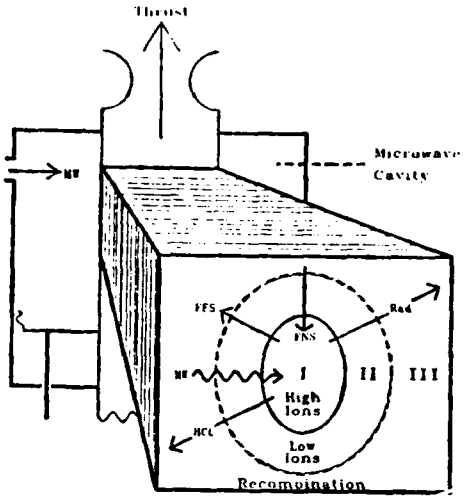


Figure 2. Plasma Discharge Properties.

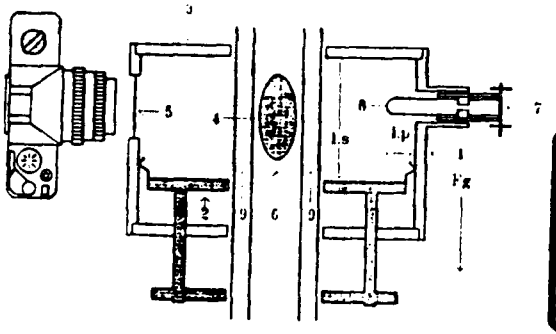


Figure 3. Microwave Cavity

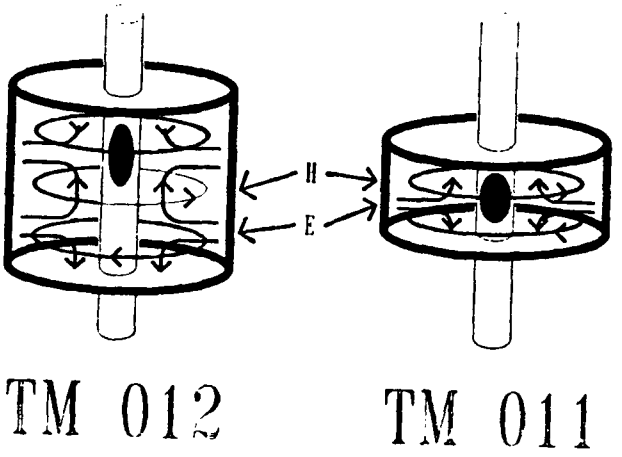


Figure 4. Microwave Electromagnetic Modes.

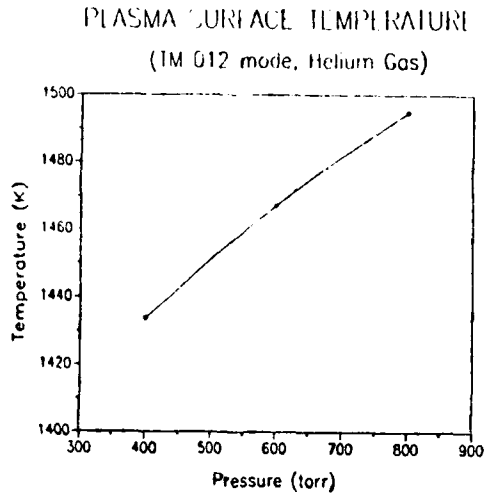


Figure 5. Plasma Surface Temperature Plot.

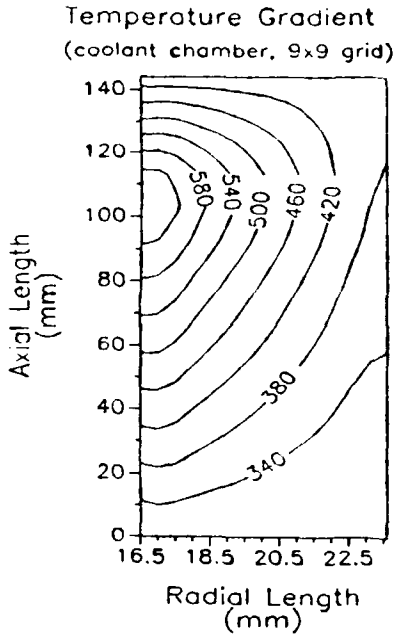


Figure 6. Coolant Chamber Temperature Profile.

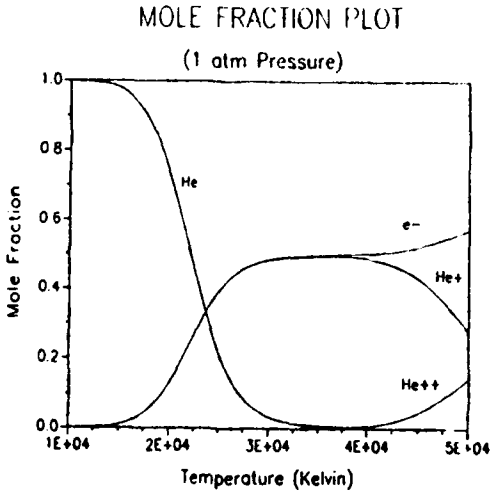


Figure 7. Helium Plasma Mole Fraction Plot.

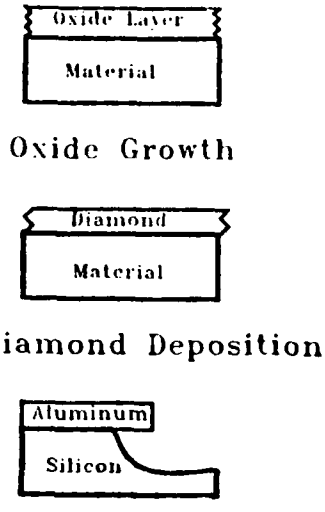


Figure 8. Plasma Surface Treatments.

Artificial Intelligence For Chemical And Biological Detectors

Dr. Alice M. Harper*, Dr. Steven D. Brown, and Mr. Richard Vanderbeek
Department of the Army, CRDEC, SMCCR-RSP-C
Aberdeen Proving Ground, MD 21010

I. Abstract

The artificial intelligence group at CRDEC is currently developing hybrid pattern recognition/neural network systems for transition to instruments and sensors which detect chemical and biological threat agents in real-world, real-time applications. These are presently being tested using pyrolysis/mass spectra from the Chemical and Biological Mass Spectrometer (CBMS).

II. Instrumentation and Artificial Intelligence Design

A. The Chemical/Biological Mass Spectrometer

The Chemical/Biological Mass Spectrometer is being developed to collect aerosols and air-borne vapors from battlefield environments and to analyze and identify the samples for chemical and biological agents. The instrument is consists of two units, a bioprobe for collection and pyrolysis of particles in the range of key biological specimens (1-10 microns) and a miniature mass spectrometer with identification algorithms and expert systems. The CBMS is able to continuously sample the atmosphere and to detect and identify agents in less than two minutes. For chemical agents, the an air stream is continuously transferred through the unit to the mass analyzer. Biological agents are collected in on quartz wool inside an infrared quartz tube pyrolyzer. In order to volatilize the biologicals for analysis in the mass spectrometer, the collected samples are pyrolyzed on a reproducible temperature ramp operating between ambient temperature and 600 degrees Celsius. The fragments produced enter the air stream and are transferred to the mass spectrometer unit.

After entering the mass spectrometer unit samples pass through a silicon membrane to reduce the pressure from atmospheric pressure to

the high vacuum conditions necessary for mass spectrometer operation. A highly focused electron gun is used to charge the vapors for trapping inside the ion-trap quistor. The CBMS is capable of several configurations of mass spectrometry experiments. In the simplest case, ions formed by low energy impact are singly charged and represent a profile of the composition of the original air stream. The mass spectrometer can perform up to 25 scans between 0 and 512 atomic mass units (amu) per second. For biological samples, the time between pyrolysis and the end of appearance of fragments in the mass spectrometer ranges from approximately 30 seconds for fast (ballistic) pyrolysis to, typically, 80 seconds for a slower temperature ramp. In the simplest of configurations the mass spectrometer sums the mass spectral scans over the time envelop of the pyrolysis event. For ballistic conditions, 30 sec x 25 scans per second equals 750 individual mass spectra that can be summed in real time. Figure 1 shows an example of the resulting sum mass spectrum for a sample of aldolase, a representative protein mass profile after pyrolysis. Unless otherwise stated, all spectra used in this paper were collected during a single bioprofiling on the CBMS during an intermediate phase of development. They do not represent the current instrument especially with respect to sensitivity at individual atomic mass units.

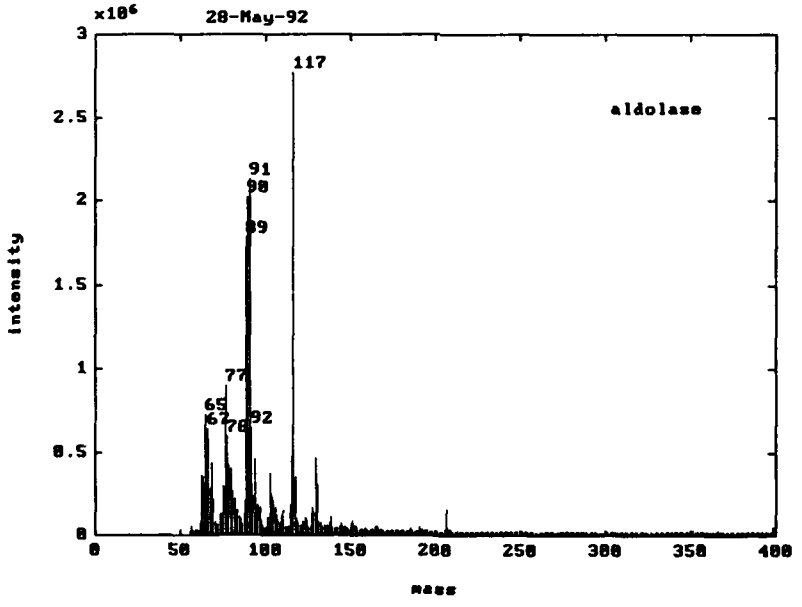


Figure 1: Total mass spectral profile obtained from ballistic pyrolysis.

The ion-trapping capabilities of the instrument enable higher order mass spectrometry experiments to take place. If after ionization of the pyrolysis fragments, all fragments above and below 135 amu are removed from the trap, ions of mass 135 can be collected. This single fragment can be further fragmented by higher energy from the electron gun. This experiment results in a secondary ion scan known as the daughters of mass 135. Figure 2 gives an example of a spectrum produced by this experiment. Daughter ions are scanned at approximately 4 scans per second. Similar isolation of individual ions (singulation) followed by fragmentation produces granddaughter spectra. Presumably singulations such as these could continue until the intensity of the signal is lost or the time required for stepwise singulation becomes prohibitive.

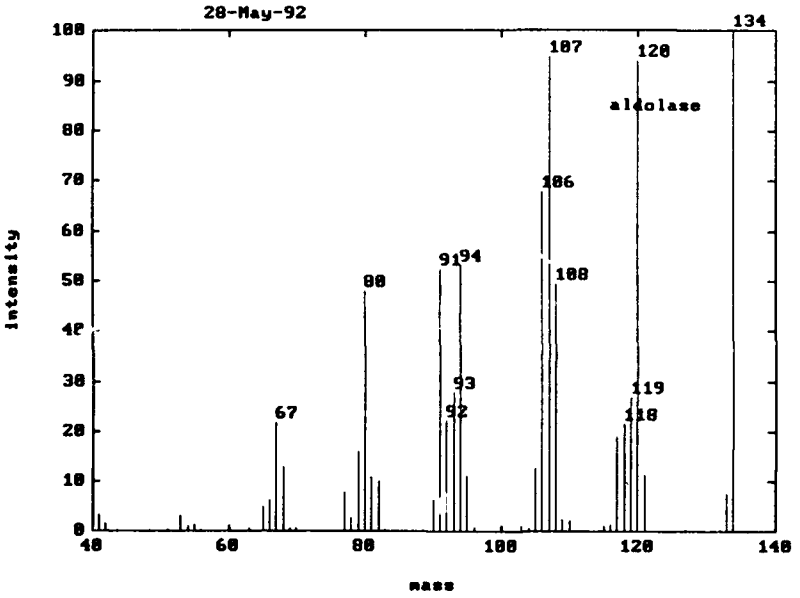


Figure 2: Daughter ion spectrum of parent mass 135.

The rapid scanning capability of the CBMS combined with the 30 to 80 second time period of the scanning gives rises to a rather unique experiment. Figure 3 shows the time profile for the sample in Figure 1 in terms of total ion current. Figure 4 shows this same sample resolved on both the time and mass axes. Figure 3 was formed by summing the mass axis intensities in each one second time window. The time window is presently thought to arise not only from differences in fragment formation in the pyrolyzer as the temperature rises on the ramp, but also from diffusion and chromatography-like effects as the fragments travel to the mass spectrometer and pass through the membrane

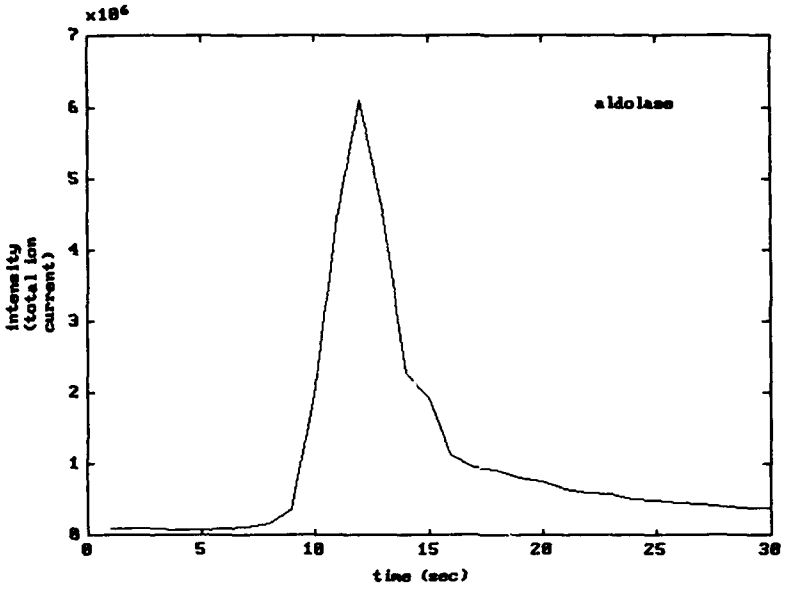


Figure 3: Total ion current as a function of time after pyrolysis.

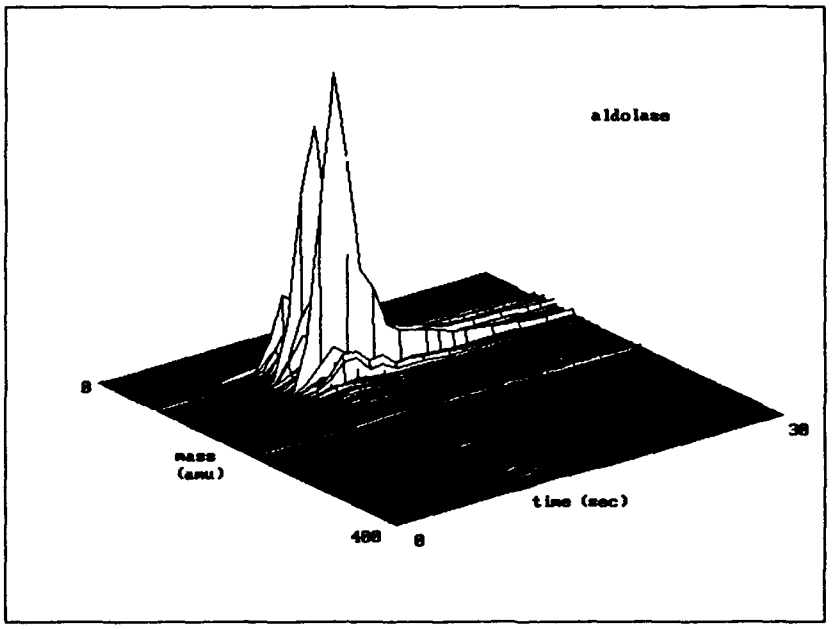


Figure 4: Time evolution of all masses formed by pyrolysis.

barrier. The time resolution produced by these effects presents the possibility that biological fragments can be analyzed in the time-primary ion, time-daughter ion, and higher order domains. The complexity of the pyrolysis signal for biologicals which are mixtures of biochemicals might be reduced if the time axis is resolvable and can be shown to be reproducible.

B. Artificial Intelligence Design for Detection Instrumentation

The rational for the artificial intelligence necessary to accomplish data reduction, pattern recognition and organism identification in complex chemical instrumentation such as the CBMS is outlined in Figure 5.

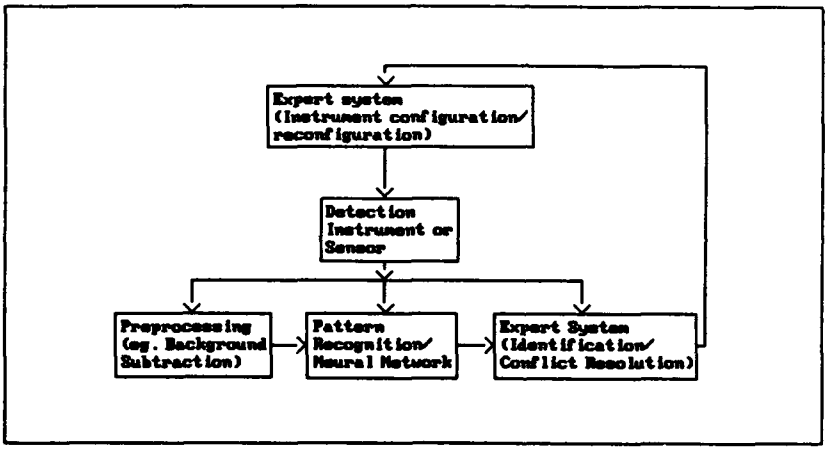


Figure 5. Generic instrumental artificial intelligence design.

Neural networks are presently targeted as the ultimate pattern recognition modeling technique. The expert system must be capable of optimizing rule formulation based on part or all of the mass spectral profile as well as on output from spectra processed in a pattern recognition algorithm or neural network. Several contractors have been designing intelligence algorithms around this design including the TAU Corp. EVAL software and Peter Herrington's FURES[1] (fuzzy logic/pattern recognition). At CRDEC, a concentrated effort to extend and modify these efforts to create a versatile artificial intelligence protocol capable of identification of agents and redesign of experiments which encompasses the variety of configurations this instrument can operate in is underway. The design treats the instrument and its adjustable

parameters as part of the expert system knowledge base and as intermediate decision outputs. Under this configuration, if a detection is unresolvable based on current data, a warning can be issued and the instrument can be reprogrammed to perform a different experiment to resolve any ambiguities. For example, if primary scans are collected in real time and a sudden influx of high levels of biological fragments are detected by the pattern recognition models, a warning is issued. In some cases, identification might proceed without ambiguity. However, if the artificial intelligence cannot identify the organism based on the primary ions, the expert system will trap the nature of the ambiguity and use its rule base to reprogram the instrument to produce, for example, daughter ion scans of one or more ions known to be useful for the resolution of the ambiguity.

The biggest problem expected to arise in real-world real-time sampling is the background at the site of sampling. In battlefield conditions, this background will consist not only of components normally found in the atmosphere such as pollen and air-borne bacteria which pose no threat, but also of the various smokes, diesel exhaust and dust particles created as a result of military operations. For biological identification, any particles in the 1-10 micron region pose a potential problem during identification since these will be pyrolyzed along with the agent threat. In battlefield conditions, such inputs can arise as a sudden influx of materials in the same way the treat is expected to arise. The mass spectra will be a composite of all materials trapped in the pyrolysis unit. The spectra will be mixtures; therefore, any pattern recognition or expert systems models must be able to identify the threat organism in the presence of substantial background signal. Classification algorithms and classification neural networks which do not allow fuzzy decision making will probably fail well below the instrumental sensitivity levels unless all possible background conditions are modeled. CRDEC is considering technology such as associative neural networks, fuzzy backpropagation, statistical neural networks and learning vector quantification as modeling algorithms. The later is also being studied by the TAU team. Discriminant analysis and related techniques are thought to map the data structure too literally with respect to the training set data.

C. Preprocessing

Algorithms for preprocessing are dependent on the instrument and signal. Since the prototype system is being developed in conjunction with an interest in using time resolved mass spectra from the CBMS, the preprocessing tools current used address this particular instrumental setup. Early attempts to eliminate the background arising from the

environment were unsuccessful because this background is convoluted with background arising from hysteresis inside the instrument. The later has components whose contribution to the total signal is temperature and mass flow dependent; hence, this background is time convoluted. Modifications to the instrument are being made to eliminate this effect. Nevertheless, preprocessing of the time resolved spectra in this paper has not included subtraction of the background.

Rank annihilation was used to reduce the size of the time resolved mass spectra for input into the EVAL implementation of the Learning Vector Quantization neural network. The procedure decomposes the time resolved mass spectral profile covariance matrix (X) into two unitary matrices, one (T) containing orthogonal trends in time, the other (M) containing the corresponding trends in the mass spectra and a diagonal matrix S containing the singular values (eigenvalues) of the covariance matrix.

$$X = T * S * M'$$

where the matrix transpose is denoted by '. An automated test using the F ratio determines the number of significant eigenvectors. Once the eigenvectors are extracted a simplex search post rotates the vectors to an "optimal" rotation as defined by a minimum in the entropy (H) where

$$H = -\sum_{i=1}^{n} p_i \log_2(p_i)$$

where

$$p_i = \frac{|s_i|}{\sum_{i=1}^{n} |s_i|}$$

s_i is the intensity of the i-th trial mass spectrum and "n" is the number of masses. A penalty is added to the entropy term if negative masses are encountered. The mass spectrum defined by the optimal rotation (N_1) is removed from the data by a rank annihilation. A residual matrix (N_r) is formed with rank one less than the data matrix (N). This residual is formed as

$$N_r = N - a_1 N_1$$

where a_1 is a scaler. The process is repeated using the residual time resolved mass spectrum and ends when all significant eigenvectors are

eliminated.

D. Neural Network

The neural network used in this study was the TAU implementation of the Learning Vector Quantization (LVQ2) of Kohonen[2]. Briefly, this procedure is a feed forward associative neural network similar to the K-nearest neighbors algorithm used in classical pattern recognition. However, LVQ does not store the training set data as an approximation to the statistical distribution, but uses the training set data to define a set of reference patterns for each class which span the space of the training data, are clustered according to the distribution exhibited by the training set data, and may be influenced by a member of an alternate class if that member's nearest neighbor is one of the class references. During training a penalty function is used when the latter occurs to modify the development of the reference set distribution. In the LVQ2 implementation, the penalty function is implemented to move the discrimination surface toward the intersection of the overlapped groups.

E. Expert systems

The expert systems challenge posed by detection applications is diagramed in Figure 6. The outputs of the neural network (or any

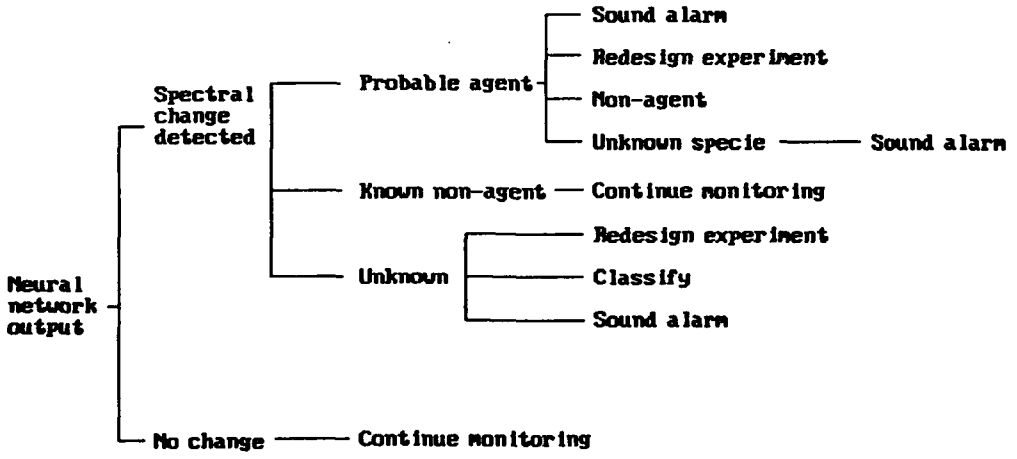


Figure 6: Example of a expert systems scenario arising from neural network results.

classification system) can define a detection situation with identifications that can range from classification of a chemical or biological threat into a known class, a subset of classes, or a non-biological class. Further, the unknown sample may be similar to a known class but also be a new specie.

Although it is possible to use multiple supervised learning neural networks to attempt to resolve all encountered combinations of clustering tendencies, the multi-experiment capability of the CBMS has led to an emphasis on using expert systems to identify the nature of a classification problem. Once the problem is defined the expert system selects a more specific and definitive experiment to resolve the ambiguity. Since, the LVQ2 algorithm computes the distance between an unknown and the reference samples among the classes, if identification does not occur then the distances form a pattern locating the sample in space defined by the reference samples. Figure 7 gives an example of this concept. The ordered distances from the unknown sample form a classification vector [1 2 1 2 1 1 1 2 2 ...]. Subclassification in the expert system is determined by this classification order. In this case the subclassification is the combining of two clusters into one larger cluster. This vector along with any other information found to be relevant form the factor set determining the rule used for redesigning the experiment, subtracting a known interferent profile, or reinspection of the data.

The biggest problem for an instrument which produces large arrays of data is selecting that portion of the information which optimizes the expert systems design. When the daughter ion scans become available, genetic algorithms[3] will be used to optimize the daughter scans used in rule formulation.

III. Results and discussion

Time resolved primary ion scans from the CBMS were obtained by its designers at the Bruker Franzen laboratory in Germany. For each sample five replicates were analyzed. The order of the replicate runs were randomized. Table 1 lists the samples used in this study. Since each set of replicates is the only representation of a class of organisms, the experimental design is clearly inadequate to test the capabilities of the instrument nor can the true class structure be modeled. Therefore, all results reflect the ability to distinguish individual organisms from the remaining classes.

Table 1: Samples used in this study.

1. Spores of bacillus subtilis
2. Coxiela burnetti (heat inactivated)
3. Staphylococcal enterotoxin B
4. DS-2
5. Yersinia enterocolitica (heat inactivated)
6. Aldolase
7. MS-2 Bacteriophage
8. Mixture of pollen and baker's yeast
9. Gasoline
10. Burning vegetation
11. Diesel exhaust

The rank annihilation algorithm discussed the previous section was used to reduce the size of the time resolved mass profiles. Figure 8 shows the result of the first rotation for a typical sample of coxiela burnetti. The time profile carves out an evolution of pyrolysis fragments centered at 9-10 seconds after the sample was pyrolyzed. The mass vector in 8b gives the masses of these fragments. Library searches have failed to reveal the nature or origin of these pyrolysis fragments. Daughter ion scans, when available, may reveal the chemistry and biological origin of the time resolved masses. The time resolved spectrum corresponding to the first rotation is shown in Figure 9a. After the spectrum in 9a was rank annihilated from the original time profile, the residual in 9b resulted. This residual was decomposed to obtain a second time resolved spectrum. Two rotated mass spectral eigenvectors were retained. After elimination of samples which rotated to background vectors because of signal to noise problems, each sample's vectors were reduced to the mass range of 46 to 245 because the instrument exhibited low sensitivity to higher masses. The resulting reduced spectra were input to the LVQ2 algorithm. All samples were correctly classified. Not enough samples were present to test the models on unknown spectra. However, when one sample per class in all classes that contained four or more replicates of the sample were used for a test set under the leave-one-out strategy, the overall prediction fell to 82 per cent.

IV. Conclusions

An artificial intelligence protocol for detection instrumentation has been established. An hybrid expert systems/neural network is capable of handling the identification problem and can also generate the scenario

for additional data acquisition. The techniques were illustrated using the CBMS which is still under development but better experimental designs will be necessary before real-time models can be built.

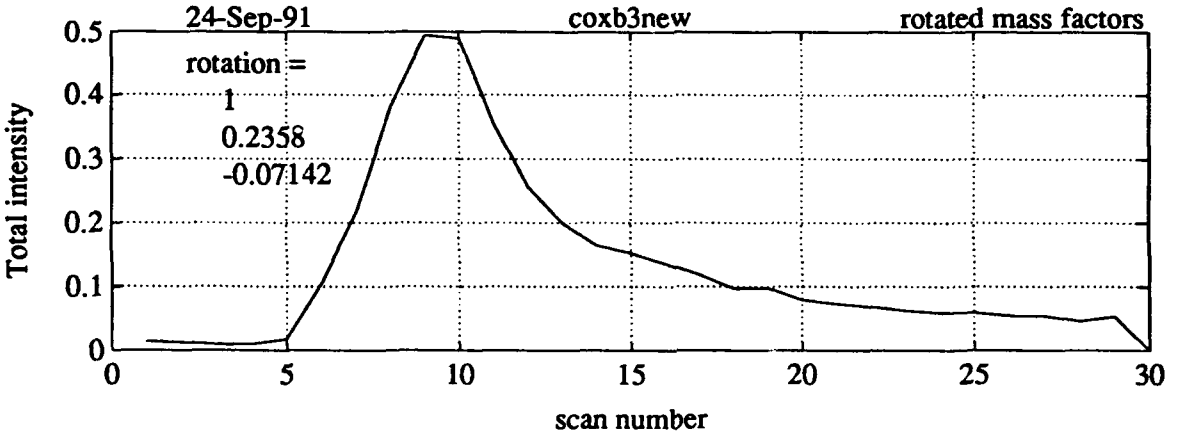
V. References

1. P.B. Harrington and K.J. Voorhees, *Anal. Chem.*62,729(1990).

2. T. Kohonen, *Proceedings of the IEEE*78(9),1464(1990).

3. T.E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*,(Addison-Wesley, Reading, MA, 1989).

a.



b.

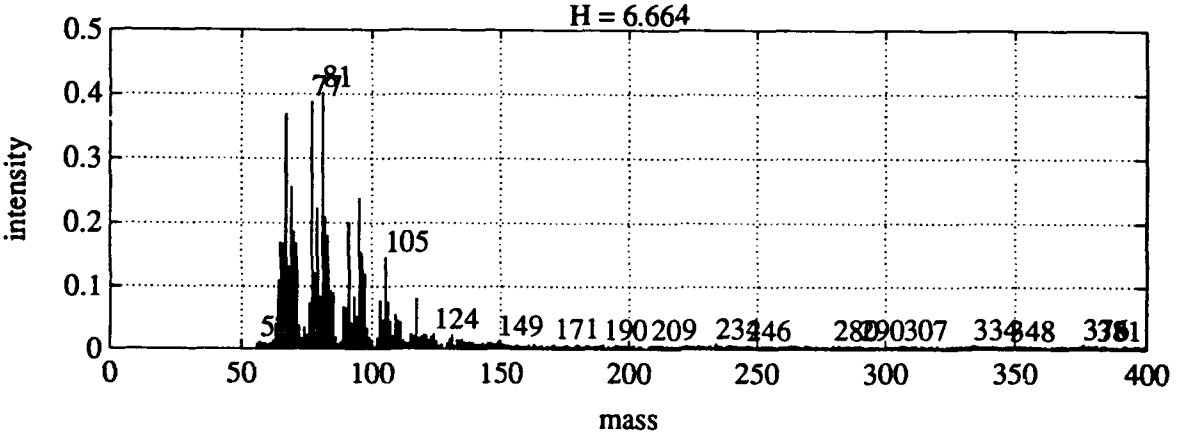


Figure 8: a) Total mass intensity evolving in the first time resolved axis from singular value decomposition. b) The mass spectrum of the fragments which evolve according to the time trend in a)

a.

b.

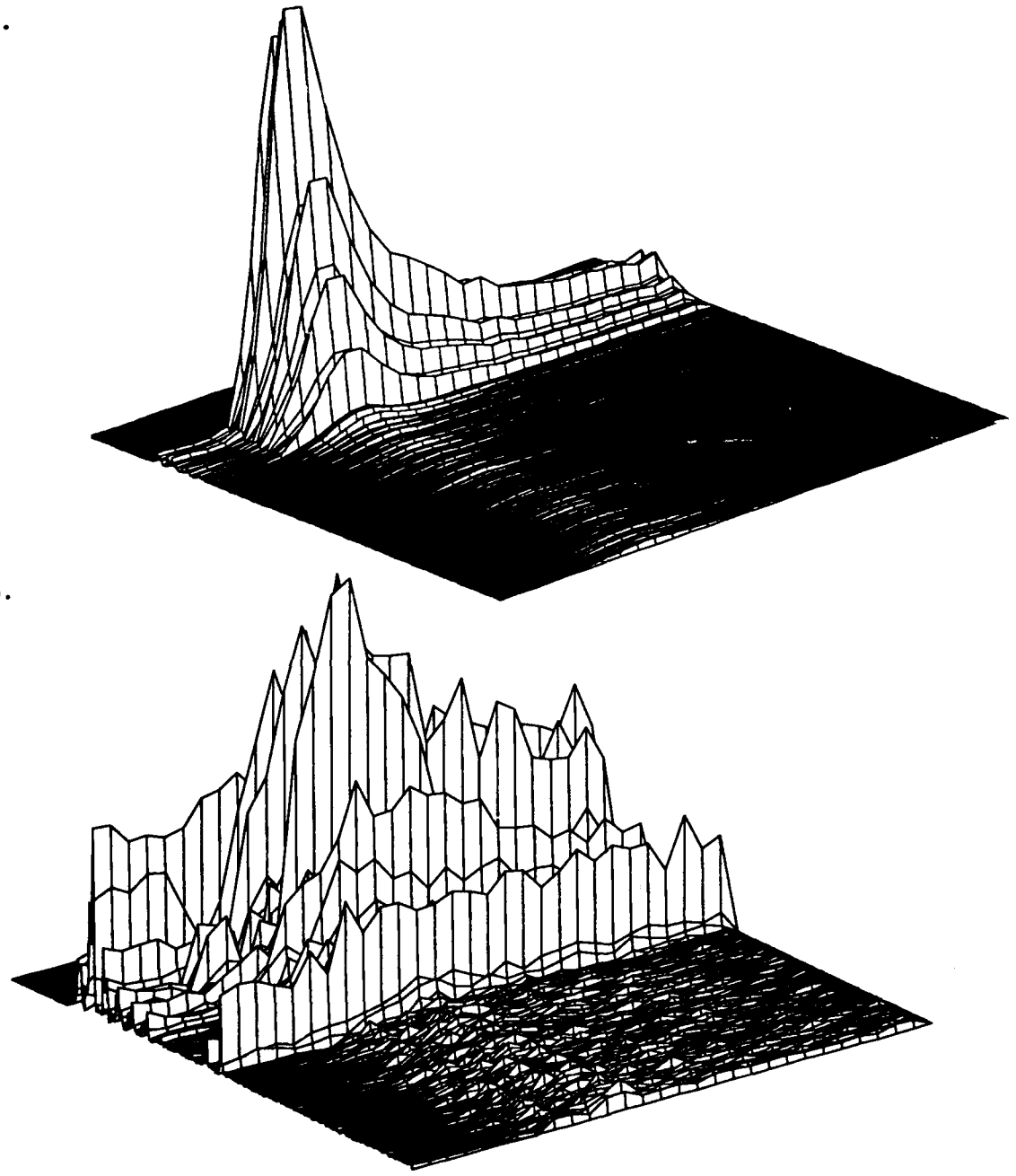


Figure 9: a) The time resolved mass spectal profile define by the first eigenvector of the decomposition of the original data from a coxiela burnetti sample. b) The residual matrix obtained when the profile in a) is annihilated from the original data.

Physics and Device Applications of Silicon Microclusters

Dr. James F. Harvey,*[a] Dr. Hongen Shen,[b] Dr. Robert A. Lux, Dr. Weimin Zhou,[a] Dr. David C. Morton, Ms. Melanie W. Cole, Dr. Mittra Dutta, Dr. Madan Dubey, Mr. Richard C. Piekarz, Dr. Richard T. Lareau, Mr. Armand M. Balekdjian, Dr. Jagadeesh Pamulapati, and Dr. Clarence G. Thornton

US Army Electronics Technology and Devices Laboratory, Ft. Monmouth, NJ 07703

Dr. Charles M. Bowden

US Army Research, Development, and Engineering Center, Redstone Arsenal, AL 35898

Dr. Raphael Tsu[c]

University of North Carolina at Charlotte, Charlotte, NC 28223

1. INTRODUCTION

The recent observations of bright, visible photoluminescence (PL) and resonant tunneling in microparticles of silicon have generated an exciting new field of research, with appealing attributes for novel devices of major impact on Army electronics and optics programs. After reviewing the stimulating discovery of PL from porous silicon (at Malvern, UK)[1] and the observations of resonant tunneling in silicon microcrystallites (at Univ. of NC at Charlotte, supported by the US Army Research Office--ARO)[2] a new program was initiated at the US Army Electronics Technology and Devices Laboratory (ETDL). Its purpose was to exploit the rapidly developing prospects for new materials based on ultra-small crystalline semiconductor clusters or structures formed without special epitaxial or lithographic procedures. It was apparent that these initial experiments opened the door to exciting new classes of devices based on new physics phenomena occurring in these novel materials, promising optoelectronic, optical, and fast tunneling devices fully compatible with silicon IC technology. A coordinating group was

a. National Research Council-ETDL associateship.
b. GEO-Centers, Inc., Lake Hopatcong, NJ 07849.
c. Research sponsored by USARO and ONR.

quickly formed at ETDL to integrate experimental and theoretical approaches to this research across organizational and functional lines. Within several months ETDL researchers had reported the first Raman spectra from porous silicon, the first measurement of pressure dependence, among the first identifications by TEM of the crystalline planes in microcrystallites of silicon with sizes of several nanometers, and had produced extremely small silicon microstructures using two promising alternative techniques. The ETDL program was leveraged by cooperative experiments with the nonlinear optics program at the Research, Development, and Engineering Center of the US Army Missile Command (MICOM) and by capitalizing on ARO sponsored research at universities.

Bulk crystalline silicon has an indirect bandgap and displays only a very faint infrared (IR) photoluminescence (PL) at room temperature. Porous silicon was found to demonstrate very bright visible PL at room temperature. Thus porous Si offered the immediate promise for silicon IC based opto-electronic devices and because of this quickly fostered an extremely active research field. Because of the special promise of porous Si, because of the relatively easy fabrication conditions required to observe its bright PL, and because of the potential for benefiting from the ideas and research of so many other groups, the ETDL program has had the greatest initial success with porous silicon. As a result most of the results reported in this paper deal with porous silicon.

However it should be remembered that this is a project with very general scope and wide implications. The goal of these initial experiments with porous silicon and tunneling in silicon microclusters is to establish the unique transport and optical phenomena governing the behavior of materials composed of these ultrasmall silicon particles. Efforts are in progress in parallel with this research to formulate device concepts utilizing these phenomena. Initially electroluminescent (EL) films and light emitting diode (LED) structures have been examined. Also in parallel, promising new fabrication techniques for silicon microstructures, both in-house and in conjunction with ARO funded university research, are being assessed and developed. These new materials will require new approaches to device design, and the current research has focused on establishing the most advantageous approaches.

2. LUMINESCENCE FROM POROUS SILICON

Porous silicon is made by electrochemically etching a silicon wafer in a solution of HF. As a result many very small pores are formed in the silicon. As the etching process progresses very small particles of crystalline silicon remain between the pores. The exact structure of the porous silicon and the mechanisms responsible for the bright PL are controversial.

HARVEY, SHEN, LUX, ZHOU, MORTON, COLE, DUTTA, DUBEY, PIEKARZ, LAREAU, BALEKDJIAN, PAMULAPATI, THORNTON, BOWDEN, TSU

The strong, visible PL from porous silicon is in sharp contrast to the weak PL in the IR from bulk crystalline Si. Silicon's indirect bandgap of 1.1 eV (Γ point of the valence band to the Δ point of the conduction band) is mainly responsible for the weak IR PL. Optical transitions across this bandgap require a change in momentum of nearly the full Brillouin zone. Three important questions need to be addressed in order to explain the PL from porous Si. First, what happens to the 1.1 eV bandgap of bulk silicon? Canham's original report of PL from porous Si attributed its visible wavelength to a blue-shift in the silicon PL as a result of quantum confinement in microcrystalline silicon structures with diameters on the order of 3 nm.[1] The second question is what mechanism is responsible for the enhancement in the optical transitions across the bandgap? While quantum confinement results in some uncertainty in the momentum (by the Heisenberg principle), this uncertainty is not by itself sufficient to significantly relax the momentum conservation requirement. Third, what is the role of localized states and their relation to the recombination process?

Other investigations have provided support of the quantum confinement model of the PL mechanism for porous Si, in which the silicon bandgap is widened due to the confinement of the carriers in an extremely small particle in the same manner that the energy level spacing is increased for a particle in a box when the box dimensions decrease.[1,3,4] Alternative mechanisms for the visible luminescence have also been raised. Some reports suggest amorphous silicon could be the source of the PL.[5,6] Others have proposed optical transitions in individual molecules present at the silicon cluster surface or in the region between clusters, such as the siloxene molecule.[7,8] So far experimental results have been ambiguous or conflicting with regard to a clear cut determination of the PL mechanism. This portion of this paper discusses the results of structural, analytical, optical, and cathodeluminescence (CL) measurements on porous silicon. Like previous experimental work, these results are not by themselves definitive in identifying the PL mechanism. However they do provide important additional information relevant to the three questions stated above, and they provide important clues to guide efforts at device design.

2.1 STRUCTURAL AND CHEMICAL ANALYSIS

Early reports of PL and optical absorption in porous silicon attributed the results to quantum confinement effects in needle-like crystalline structures with diameters of 2-9 nm.[1,3] TEM micrographs of the porous silicon film taken at ETDL[9] clearly show the crystalline planes in microcrystallites of silicon, as shown in figs. 1 and 2. As shown in fig. 2 the interplanar spacing is 0.31 nm, which corresponds to the spacing for the (111) orientation of bulk crystalline silicon. The microcrystallites seen in these micrographs are not needle-like but rather of an irregular shape. Similar TEM micrographs were independently taken by Lehmann and coworkers.[10] The

lattice planes in the crystallites shown in fig. 1 are rotated by a few degrees from each other in the plane of the micrograph. It is expected that other crystallites may be rotated by a few degrees out of the micrograph plane, in which case their crystal planes would not appear in fig. 1, and they would appear as the amorphous looking region between crystallites in this figure. Initial observations have not seen the typical ring structure in transmission electron diffraction that is the signature of amorphous silicon.

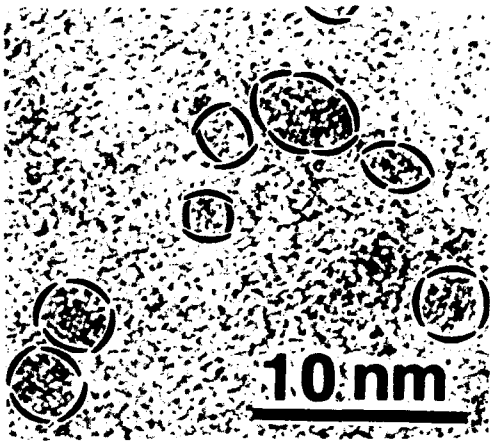


Figure 1. TEM micrograph showing ultra-fine crystallites in porous silicon.

Results of Auger analysis on this sample are shown in fig. 3. Silicon, oxygen, carbon, and nitrogen atoms are seen in this analysis. SIMS analysis on the same sample also shows the presence of hydrogen atoms, which do not have an Auger signature, and a small amount of fluorine.[11] The composition of this film is about 70% (atomic concentration) of silicon, 20% oxygen, and 10% carbon and nitrogen, excluding the hydrogen concentration, which Auger can't measure. The results of independent SIMS analysis elsewhere[12] confirm the qualitative nature of these measurements, and demonstrate a large variation sample-to-sample as the result of different environmental conditions, as well as long time scales for interactions with the environment.

Figure 2. TEM image of a silicon crystallite from porous Si. The marker indicates the interplanar spacing, 0.31 nm.

ESR measurements[13] on this sample indicated that the surface of the silicon microcrystallites were well passivated and free of dangling bonds, which was attributed to bonding with hydrogen or SiO_2. These conclusions were independently confirmed elsewhere by FTIR measurements,[14-16] which noted that the strength of the PL correlated with SiH_2 and that the Si-H bonds were replaced by Si-O bonds over time.

2.2 CORRELATED RAMAN AND PL MEASUREMENTS

For these measurements both Raman and PL spectra were taken from the same spot on the sample and at nearly the same time. The average sizes of the silicon crystallites varied from sample spot to sample spot. At each sample spot three spectra, parallel polarized Raman, cross polarized Raman, and unpolarized PL, were taken. Reference Raman spectra were taken from the polished silicon surface which had not been exposed to the etching process. In contrast to the reference Raman spectra from the crystalline silicon surface, the Raman spectra from spots on the porous Si were unpolarized.[17-19]

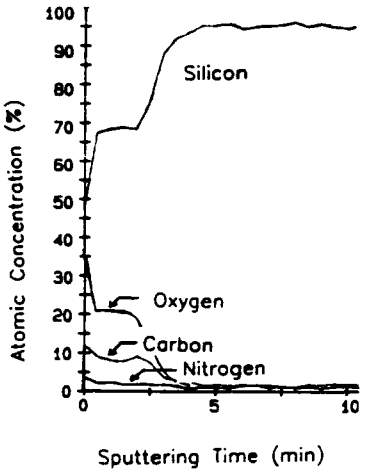


Figure 3. Auger analysis of porous silicon film.

Shown in fig. 4 are Raman and PL spectra from two spots (A and C) on a typical sample. The Raman spectra consist of sharp peaks due to crystalline silicon modes near the peak for the bulk crystalline Si reference, marked "c-Si" in the figure. The line widths of the porous Si peaks are about 5 cm^{-1}, compared with 3 cm^{-1} for the reference c-Si peak. Peak C also displays the splitting of the TO and LO phonon modes. The PL spectra show the broad peaks at energies between 1.55 eV and 2 eV typical of porous Si PL.

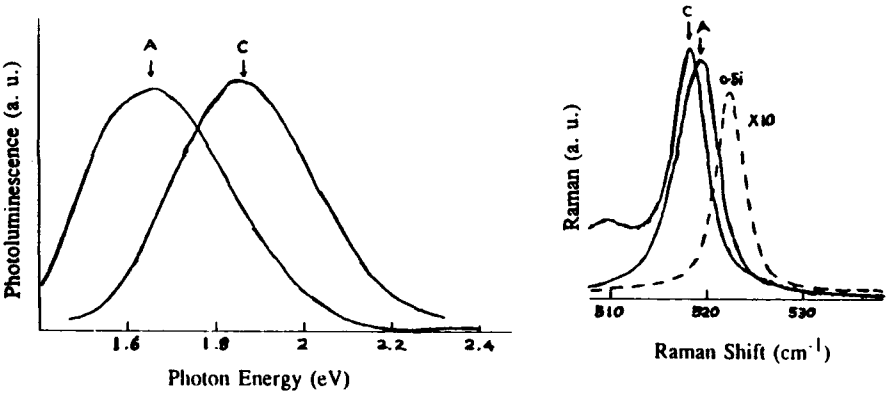


Figure 4. Correlation of Raman and PL spectra from porous Si.

The downshift of the porous Si Raman peaks from the c-Si peak and the LO-TO splitting can be used to estimate the size of the silicon microcrystallites. The electron-phonon interaction in a finite sized particle has a maximum corresponding to a phonon momentum $q=\pi/\ell$, where ℓ is the confinement dimension of the particle. The phonon momentum of the Raman peak from a finite sized particle will therefore be shifted by π/ℓ from the center of the Brillouin zone. From the experimental down-shift in the energy of the Raman peaks, the particle size can thus be estimated using the known phonon dispersion in silicon. The phonon dispersion is different for the LO and TO

phonons, so the experimental phonon splitting can also be used to estimate particle size. Our Raman data results in size estimates between 2 and 3 nm for the porous Si microparticles.

Fig. 4 also illustrates the correlation between PL and Raman spectra taken at different spots on a sample, which is typical of our data. The PL spectra from spots which have the largest Raman down-shifts, and hence the smallest average particle size, have peaks at the largest PL photon energies. This indicates that quantum confinement due to the smaller particle size is increasing the spacing of the energy levels involved in the PL recombination.

2.3. MEASUREMENTS OF THE DIELECTRIC CONSTANT

Angle resolved ellipsometry measurements were made on the porous Si samples, by fitting the reflectance data to known formulas for thin films and extracting the index of refraction and film depth. Our data resulted in values for the index of refraction of the porous silicon film between 1.5 and 1.9, and therefore dielectric constants between 2.3 and 3.6. These values are much smaller than the dielectric constant for bulk crystalline Si at this excitation wavelength, 17.8. The reduction is due to two principal causes. This material is porous, with the voids contribute a dielectric constant of 1 to the mixture. Second, quantum confinement itself will reduce the dielectric constant,[20] by increasing the critical optical transition energies. Based on our estimates of our porosities and on a simple estimation of the dielectric constant of a mixture of silicon particles and voids, the measured dielectric constants appear too small to be the result solely of the voids. Large void fractions, on the order of 90%, are not supported by our structural studies. These measurements demonstrate that the quantum confinement effects can have a significant effect in the reduction of the dielectric constant of porous silicon, and that the dielectric constant within the microcrystallites themselves is reduced.[18,19]

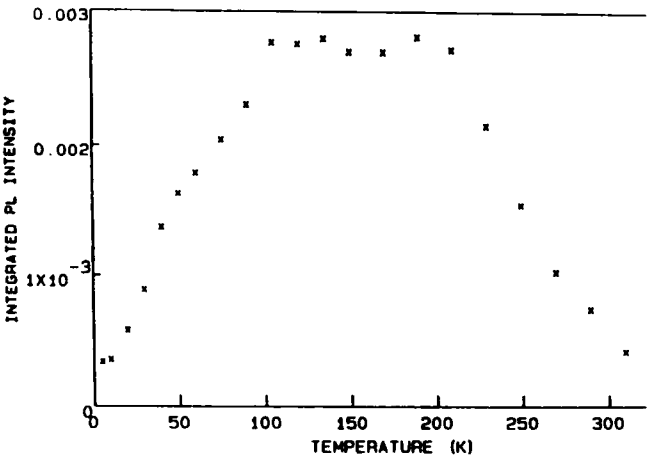


Figure 5. Temperature dependence of the PL intensity from porous silicon.

2.4. TEMPERATURE DEPENDENCE OF THE PL

The temperature dependence of the integrated intensity of the PL is shown in fig. 5.[21] At low temperatures the intensity rises with temperature, with a behavior typical of phonon-assisted recombination in semiconductors. Near room temperature the intensity decreases. The temperature dependence of the porous Si is very similar to that reported for Se doped Si.[22] In this report the falloff in PL intensity near room temperature is attributed to the thermal dissociation of bound excitons. This suggests that the temperature behavior of the porous Si PL near room temperature is due to a similar mechanism. If the temperature dependence of the log PL intensity is plotted against inverse temperature, the relation is nearly linear near room temperature, and a thermal dissociation energy of 100 meV can be extracted. The binding energy of the excitons is inversely proportional to the square of the dielectric constant inside the silicon microparticle. As a result of the quantum confinement on the dielectric constant, as discussed above, relatively shallow defects or impurities in bulk Si would have their binding energies increased when confined in microparticles of these small dimensions. The increased binding energy would then be consistent with this measured dissociation energy. The temperature dependence reported here is not consistent with a recent report of the temperature dependence of the PL of siloxene,[23] which has been proposed as a possible cause of the bright PL from porous Si. Zheng and coworkers reported a steady increase in the PL from siloxene as the temperature decreased down to 10° K, which is contrary to these observations for porous silicon. Our temperature dependence measurements thus raise serious questions about a model based on PL originating from siloxene.

The temperature dependence of the PL time constants for porous Si are shown in fig. 6.[21] The time constants at low temperatures are very long (nearly 1 millisec), decreasing by almost two orders of magnitude at room temperature. The shape of their temperature dependence and their magnitudes are very similar, again, to the PL lifetimes reported by Bradfield for Se doped Si,[22] which again suggests the similarity in PL processes. The PL time constants are due to both radiative and non-radiative channels. At low temperatures there are fewer non-radiative paths, and the overall time constants are longer. As temperature is increased the time constants reflect the temperature dependence of the non-radiative processes. Our results are thus consistent with the temperature dependence of non-radiative transitions involving trapping

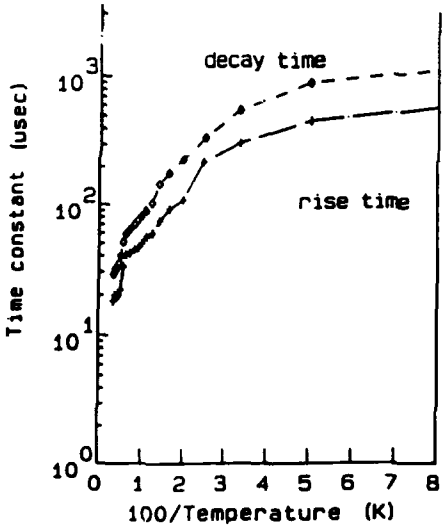


Figure 6. Temperature dependence of the PL time constants for porous silicon.

or localized states.

2.5. EXCITATION POWER DEPENDENCE OF THE PL

For these measurements the PL was detected using a lock-in amplifier system with an acoustic modulator to chop the beam. Use of the acoustic modulator enabled the use of different "duty cycles", in which the peak power and the chopping period would remain the same, but the portion of the period the excitation was on could be varied. A "10/90" duty cycle had the beam on only 10% of the time, and thus had a lower average power than a 90/10 duty cycle at the same peak power.

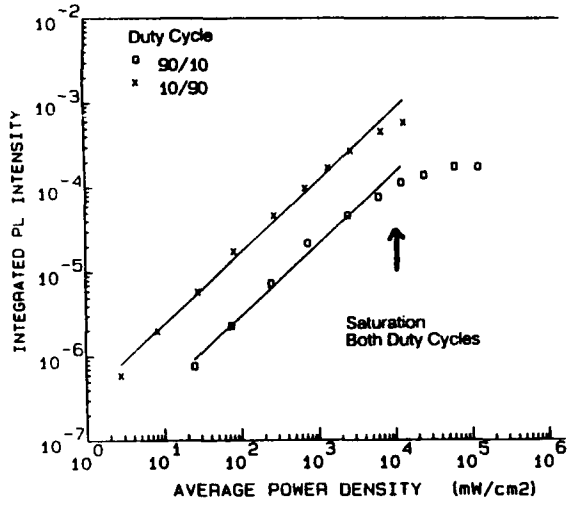


Figure 7. Excitation power dependence of the PL from porous silicon.

The excitation power dependence of the PL is shown in the log-log plot in fig. 7. The slope of 1 on the log-log plot indicates a linear power dependence and that the excitation process is a one photon-one electron process. This linear dependence holds over nearly four orders of magnitude of excitation power, finally beginning to saturate at the higher powers. In order to investigate this saturation effect different duty cycles were used for the excitation. The results for the 10/90 and 90/10 duty cycles are shown in fig. 7 plotted against the average excitation power. Fig. 7 demonstrates that the saturation effect occurs at nearly the same average power for different duty cycles, and therefore for different peak powers. This result indicates that the saturation is due to temperature increases in the illuminated spot on the porous Si sample as a result of increasing absorbed power, an effect which would depend on the average excitation power rather than the peak power. As shown in fig. 5, near room temperature an increase in temperature will cause a decrease in PL intensity. This also suggests that the one photon-one electron excitation process itself will be true at higher excitation powers.[19,24]

2.6 PL EXCITATION SPECTRUM

The PL excitation spectrum for the red peak of porous silicon is shown in fig. 8. This spectrum was taken using a xenon arc lamp and monochrometer. The energy of the PL peak is about 1.55 eV, and this peak energy is quite insensitive to the excitation energy. For the excitation spectrum the PL is monitored at the 1.55 eV peak and the PL intensity recorded as a function of

the excitation wavelength. Fig. 8 shows that the excitation spectrum is a fairly narrow feature peaked near 2.6 eV. The 1.55 eV PL is consistent with the indirect band-gap of Si, widened by quantum confinement. Since the direct band-gap energies of bulk Si are 3.3 eV or greater, the 2.6 eV excitation peak is probably not associated with these transitions. This suggests that it may be due to excitation to an excited level of a quantum confined state.[19,25]

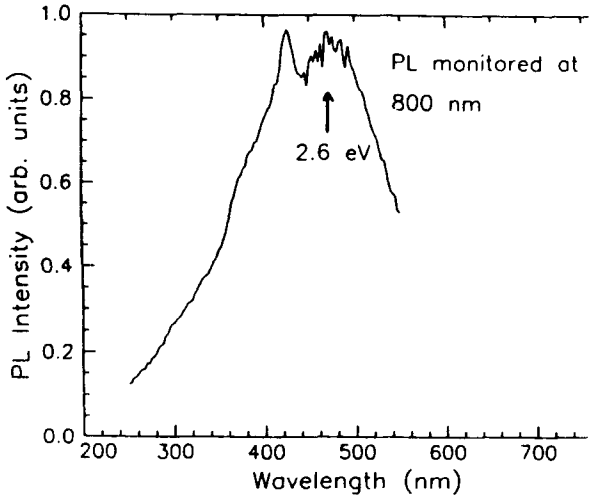


Figure 8. PL excitation spectrum of the red peak of porous Si.

2.7 PRESSURE DEPENDENCE OF THE PL

The first measurements of the pressure dependence of the PL from porous silicon were made at ETDL.[26] The pressure dependence of optical transitions has been extensively used to determine the band structure of semiconductor material. The purpose of these measurements was to determine important information about the electronic structure of porous silicon. Fig. 9 shows the pressure dependence of the peak PL energy over 40 Kbar of pressure. The squares designate data taken as the pressure increased, triangles designate data taken as the pressure decreased. There is a large blue-shift with pressure of nearly 170 meV with 20 Kbars of pressure, followed by a nearly constant peak energy for higher pressure. This leveling off of the PL peak energy at higher pressures is suggestive of the Γ-X band crossing in GaAs at higher pressures. The origin of the apparent hysteresis is not known at this time. The PL intensity is seen to decrease significantly with pressure. The large blue-shift is at least an order of magnitude too large to be due solely to the pressure induced increase in quantum confinement caused by size reduction, as determined from a simple particle-in-a-box calculation using the pressure dependence of the lattice constant. Blue-shifts of the same order of magnitude are seen in samples of porous silicon fabricated by another research group, under different conditions.[26] For porous

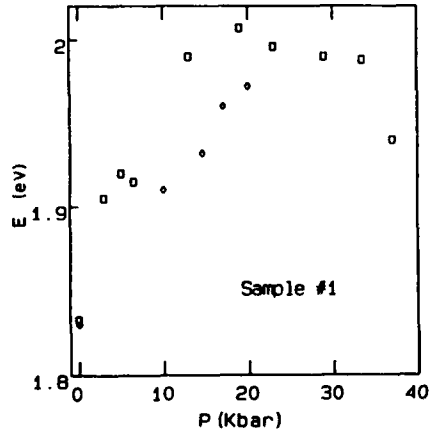


Figure 9. Pressure dependence of the energy of the peak PL from porous silicon.

Si the blue-shift is in wrong direction to be consistent with the pressure induced red-shift in the PL from either amorphous Si[27] or the indirect gap in bulk crystalline silicon.[28] Thus, these results provide strong evidence that the origin of the porous Si PL is not a known phase of amorphous Si and that it is not solely due to quantum confinement of the Γ-X bandgap of bulk silicon.

2.8 CATHODELUMINESCENCE

Under bombardment from high energy electrons, porous silicon initially emits a bright yellow CL. After a short time, on the order of several seconds, the CL dies. Apparently the high energy electrons damage the surface of the microcrystallites. The damage appears to be irreversible. After damage no PL is emitted either. This effect was confirmed independently by a group at IBM.[29] Although studies by other groups have indicated that hydrogen can be driven off the microcrystallite surfaces and replaced, with a corresponding degradation and restoration of the PL, this CL damage appears to be permanent. Once damaged the porous Si apparently does not respond to the usual methods of restoring the PL.

2.9 SUMMARY OF THE RESULTS ON POROUS SILICON AND DEVICE IMPLICATIONS

Our results provide strong support for the *quantum confinement* model of the PL, in which the photon energy of the PL from the silicon microcrystallites is shifted into the visible due to the physical confinement of the quantum mechanical wave functions in the small particles. At the same time our data raise serious questions with the proposals that porous Si PL originates from amorphous Si or from siloxene. However our results also indicate that other processes involving states with higher degrees of localization than provided by the physical confinement of the microparticle, such as those associated with traps, impurities, or defects, may be equally important in the mechanism for the PL. Once the nature of these more localized states is understood, they may be the key to other novel device applications. Our structural studies do not see quantum wire structures, but rather somewhat irregular microcrystallites. These microparticles appear to be separated from each other by hydrogenated silicon or silicon oxides. Electrical conductivity through a porous silicon film will therefore be via sequential tunneling into and out of the microcrystallites or through defects around them. These studies do not appear to support a high void fraction in our samples. Our studies confirm the results reported by others concerning the importance of the passivation and surface chemistry of the microparticles and the variation in time and from sample to sample of the chemistry in the film. Other fabrication techniques for nanoparticles for device application will therefore have to consider surface passivation as an integral part of the fabrication process. The damage caused by high energy electrons indicates that the usual way of exciting EL films may require reexamination

for these materials. Other excitation techniques, such as high frequency excitation using capacitive contacts, should be considered. Efforts at an LED device design have been initiated at ETDL.[30] Encouraging preliminary results have been achieved elsewhere,[31,32] but significant problems must still be solved. Our experimental results indicate that quantum confinement may significantly affect the usual relationships of energy levels expected from bulk silicon. The binding energies of impurities and defects may be greatly altered, drastically changing the procedure of doping materials for device applications.

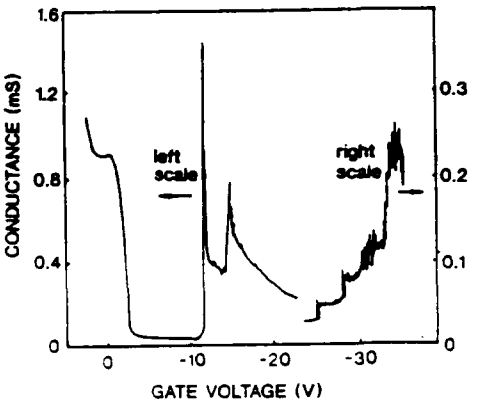


Figure 10. Conductance-gate voltage curve for resonant tunneling through silicon microparticles.

3. RESONANT TUNNELING THROUGH MICROCRYSTALLITES

Fig. 10 shows conductance-voltage curves for tunneling through a very thin film of silicon microcrystallites.[2] The microcrystallites were fabricated by recrystallizing a very thin amorphous silicon layer in an oxygen ambient. The conductance spikes between -10 and -20 volts are due to resonant tunneling through the 3D confinement of the microcrystallite. The steps at higher negative voltages are due to resonant tunneling through 1D confinement, which occurs when the wavefunctions in neighboring crystallites overlap to form an essentially one dimensional state. The large separation between the spikes below -20 volts is due to charge accumulation in the crystallites, in a phenomenon similar to "coulomb blockade." Resonant tunneling devices are inherently resistant to noise, and their basic physical mechanism operates well into the millimeter wave spectrum. These demonstrations of resonant tunneling and coulomb blockade-like effects in a film of silicon microcrystallites open the way for new classes of fast tunneling devices, compatible with silicon VLSI technology.[33]

4. NANOPARTICLE FABRICATION

In addition to the two fabrication methods discussed above for microcluster particles of silicon, two additional novel methods have been under investigation at ETDL. The first method is based on the pulsed power technology which has been developed and perfected here at ETDL. A piece of silicon is subjected to a fast very high current, high voltage pulse. The silicon is vaporized in a manner similar to an "exploding wire" plasma. The silicon atoms recrystallize as they leave the plasma, and the resulting microcrystallites are collected.[34] Fig. 11 is an SEM micrograph of a preliminary shot. In the figure crystallites smaller than 100 nm can be seen

on the collection substrate. A similar exploding wire machine using metal wire targets, operating in Russia, has demonstrated that metal microclusters of 10 nm size can be generated by this method.[35] It is expected that the ETDL technique for silicon can be optimized to produce silicon microparticles of this size. A second method uses a laser ablation technique to deposit a film of microstructures with diameters ranging down to values less than 100 nm in a preliminary experiment.

Figure 11. SEM micrograph showing microcrystallites of Si formed in the pulsed-power fabrication process.

5. NONLINEAR OPTICS EFFECTS

Researchers from MICOM have predicted large nonlinear optics parameters for very small clusters of metals, semiconductors, and superconducting material.[36] Active research is being pursued at MICOM to develop such material for optical device applications, such as harmonic generation and optically bistable switches. Experiments on porous silicon and other silicon microcrystallites produced at ETDL are being coordinated with the MICOM research group to measure the nonlinear optic parameters and to assess the potential for optical devices.

6. SUMMARY

It is abundantly clear that these new materials have tantalizing new properties for device applications. The strong PL from porous silicon may well reflect quantum confinement involving localized defects or impurity states. The sharp conductance-voltage features of silicon microparticles are most readily explained by resonant tunneling. Semiconductor microparticles are predicted to have interesting nonlinear properties of interest in optical devices. Semiconductor device physicists typically use concepts of bandgaps, band offsets, impurity doping energy levels, and carrier transport characteristics in visualizing and formulating device concepts. In these new materials consisting of semiconductor microparticles or microstructures at the nanometer scale, the values of these key quantities can become significantly altered from the expected bulk semiconductor values. Our results make a significant contribution to the expanding body of research defining these new properties. Out of this research will come new procedures and techniques for utilizing these novel materials in devices, and it will open the way for entirely new device concepts and applications.

7. ACKNOWLEDGEMENTS

The authors wish to thank M.J. Wade for assistance with fabrication; L. Calderon and M. Taysing-Lara for assistance with data and data analysis; D.W. Eckart, L. Casas, and T. Monahan for assistance with sample analysis; and E.H. Poindexter for assistance with composition. We especially wish to thank M. Weiner, L. Kingsley, and T. Burke for the use of the photo of the preliminary results of their pulsed power fabrication process and to recognize the novel laser ablation fabrication process being developed by A. Tauber.

8. REFERENCES

1. L.T. Canham, *Appl. Phys. Lett.* 57, 1046 (1990).
2. Q. Ye, R. Tsu, and E.H. Nicollian, *Phys. Rev. B* 44, 1806 (1991).
3. V. Lehmann and U. Goesele, *Appl. Phys. Lett.* 58, 856 (1991).
4. A. Halimaoui, C. Oules, G. Bomchil, A. Bsiesy, F. Gaspard, R. Herino, M. Ligeon, and F. Muller, *Appl. Phys. Lett.* 59, 304 (1991).
5. K.H. Jung, S. Shih, J.C. Campbell, D.L. Kwong, T. George, T.L. Lin, H.Y. Liu, J. Zavada, and S. Novak, to appear in *Light Emission from Silicon*, S.S. Iyer, L.T. Canham, and R.T. Collins, Eds. (Mater. Rsch. Soc., vol. 256, San Francisco, 1992).
6. M.A. Tischler, R.T. Collins, J.C. Tsang, J.H. Stathis, J.L. Batstone, and S. Zollner, to appear in *Light Emission from Silicon*, *op. cit.*
7. T. Motohiro, T. Kachi, F. Miura, Y. Takeda, S. Hyodo, and S. Noda, to appear in *Light Emission from Silicon*, *op. cit.*
8. S.M. Prokes, O.J. Glembocki, V.M. Bermudez, R. Kaplan, L.E. Friedersdorf, and P.C. Searson, to appear in *Light Emission from Silicon*, *op. cit.*
9. M.W. Cole, J.F. Harvey, R.A. Lux, D.W. Eckart, and R. Tsu, to be published in *Appl. Phys. Lett.*
10. V. Lehmann, H. Cerva, and U. Goesele, to appear in *Light Emission from Silicon*, *op. cit.*
11. R.T. Lareau, unpublished data.
12. L.T. Canham, M.R. Houlton, W.Y. Leong, C. Pickering, and J.M. Keen, *J. Appl. Phys.* 70, 422 (1991).
13. F. Rong and E.H. Poindexter, unpublished data.
14. M.B. Robinson, A.C. Dillon, D.R. Haynes, and S.M. George, to be published in *Light Emission from Silicon*, *op. cit.*
15. V. Petrova-Koch, A. Kux, F. Mueller, F. Koch, and V. Lehmann, to be published in *Light Emission from Silicon*, *op. cit.*
16. Y.H. Xie, W.L. Wilson, F.M. Ross, J.A. Mucha, and E.A. Fitzgerald, to be published in *Light Emission from Silicon*, *op. cit.*
17. R. Tsu, H. Shen, and M. Dutta, *Appl. Phys. Lett.* 60, 112 (1992).
18. J.F. Harvey, H. Shen, M. Dutta, and R.A. Lux, *Bull. Am. Phys. Soc.* 37, 565 (1992).
19. J.F. Harvey, H. Shen, D.C. Morton, R.A. Lux, W. Zhou, M. Dutta, and R.

Tsu, to be published in *Quantum Well and Superlattice Physics IV*, G.H. Doehler and E.S. Koteles, Eds. (SPIE vol. 1675, Bellingham, WA, 1992).
20. R. Tsu and L. Ioriatti, *Superlatt. Microstruct.* 1, 295 (1985).
21. H. Shen, L. Calderon, W. Zhou, M. Dutta, J.F. Harvey, and R. Tsu, *Bull. Am. Phys. Soc.* 37, 669 (1992).
22. P.L. Bradfield, T.G. Brown, and D.G. Hall, *Phys. Rev. B* 38, 3533 (1988).
23. X.L. Zheng, H.C. Chen, W. Wang, and K.N. Manjularani, *Bull. Am. Phys. Soc.* 37, 670 (1992).
24. M. Taysing-Lara, L. Calderon, H. Shen, J. Pamulapati, J.F. Harvey, M. Dutta, and R. Tsu, *Bull. Am. Phys. Soc.* 37, 670 (1992).
25. D.C. Morton, J.F. Harvey, H. Shen, R.A. Lux, and R. Tsu, *Bull. Am. Phys. Soc.* 37, 669 (1992).
26. W. Zhou, H. Shen, J.F. Harvey, R.A. Lux, M. Dutta, F. Lu, C.H. Perry, R. Tsu, N.M. Kalkhoran, and F. Namaver, submitted to *Appl. Phys. Lett.*
27. B.A. Weinstein, *Phys. Rev. B* 23, 787 (1981).
28. B. Welber, C.K. Kim, M. Cardona, and S. Rodriguez, *Solid State Commun.* 29, 1021 (1975).
29. J.L. Batstone, M.A. Tischler, R.T. Collins, and J.C. Tsang, *Bull. Am. Phys. Soc.* 37, 564 (1992).
30. M.F. Tompsett and R. Tsu, patent disclosure, CECOM docket no. 4728.
31. N. Koshida and H. Koyama, *Appl. Phys. Lett.* 60, 347 (1992).
32. N.M. Kalkoran and F. Namavar, to be published in *Light Emission from Silicon, op. cit.*
33. J.F. Harvey, R.A. Lux, and R. Tsu, patent disclosure, CECOM docket no. 4741.
34. J.F. Harvey, R.A. Lux, C.G. Thornton, H. Singh, R.J. Zeto, M. Weiner, L. Kingsley, and T. Burke, patent disclosure submitted to CECOM.
35. C.G. Thornton, *Trip Report to USSR, 24 June-8 July 1990.*
36. Y.Q. Li, C.C. Sung, R. Inguva, and C.M. Bowden, *J. Opt. Soc. Am. B* 6, 814 (1989).

BIODEGRADATION OF CHEMICAL WARFARE AGENTS: DEMILITARIZATION APPLICATIONS

Steven P. Harvey and Joseph J. DeFrank
Chemical Research, Development and Engineering Center
Aberdeen Proving Ground, MD 21010-5423

INTRODUCTION

By treaty agreements, the chemical stockpiles of both the U.S. and the former Soviet Union are currently slated for destruction. However, chemical demilitarization technology now centers on incineration which has encountered public opposition due to environmental concerns and excessive cost, currently estimated at 50 times the cost of producing the chemicals *(1,4)*. There are several reasons why chemical demilitarization would be an ideal application of biodegradation. First, unlike other bioremediation scenarios where contaminants are found in soil or groundwater, in this case, the chemicals to be degraded are in an exactly known location and a relatively pure form (typically distilled). Consequently, it is possible to "feed" the chemicals to the bacteria in a precisely controlled manner. Second, biodegradation occurs at room temperature in aqueous solution, thus precluding any possibility, however remote, of a catastrophic vaporization of agent. Third, a major objection has been that so much money would be spent on incinerator construction that it would be necessary to subsequently use chem for disposal of other hazardous wastes to amortize expenditures. This has brought criticism of the Army from surrounding communities. The fermentors which would be used for biological demilitarization on the other hand, could subsequently be very useful in the production of enzymes, antibodies or other biological products. These fermentation processes have no adverse environmental impact.

The chemical stockpile at APG may be uniquely appropriate for biological demilitarization. Mustard is the only agent stockpiled there and it is stored in bulk.

Consequently, there are no explosive components of munitions to be disposed of. A major criticism of previous demilitarization alternatives has been that they would only dispose of the chemicals, not the other equipment. The empty one-ton mustard storage containers can be decontaminated to the 3X level with sodium hydroxide, the contaminated wash from which could be fed directly into the fermentor medium. Decontamination to the 5X level can then be achieved in the decon/detox incinerator which is currently used for that purpose.

RESULTS

Mustard hydrolysis has been studied extensively *(5-9)*. Briefly, the reaction proceeds through a series of sulfonium ion intermediates to thiodiglycol (Figure 1). The first (rate-limiting) step is the formation of a cyclic sulfonium ion which then reacts with water to form the alcohol. Hydrolysis occurs most rapidly at dilute mustard concentration, high pH and low chloride ion concentration.

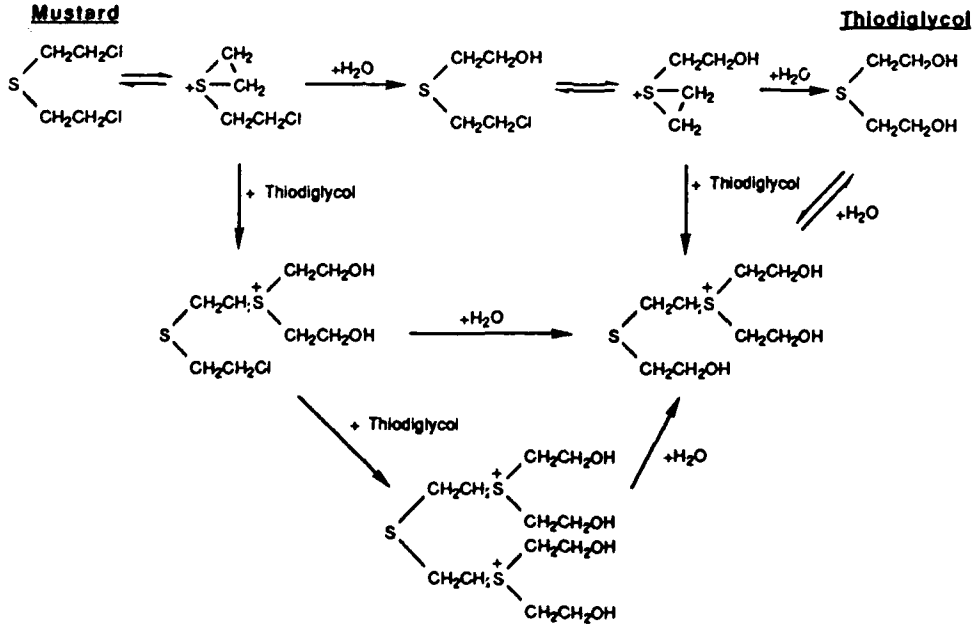


Figure 1. Hydrolysis of Mustard. Adapted from *(9)*. Mustard hydrolyzes through a series of sulfonium ion intermediates, finally to thiodiglycol. The first step, the formation of the cyclic intermediate, is rate-limiting. Thiodiglycol acts as a nucleophile to produce the branched intermediates.

Isolation and identification of thiodiglycol-degrading bacteria. Literature searches found no previous reports of thiodiglycol biodegradation. Enrichment cultures were set up in mineral salts medium using thiodiglycol as the sole carbon source. Soil samples containing indigenous bacteria were added and cultures were incubated up to two weeks. Tubes which became turbid were diluted 1:500 in 10 mM thiodiglycol minimal medium. After 10-15 such dilutions, bacteria were streaked on nutrient plates at least three times for single colonies and inoculated back into thiodiglycol minimal medium. Two different strains were isolated (SH18 and SH42) which grew on 10 mM thiodiglycol as sole carbon source. Neither organism grew without an exogenously added carbon source, indicating CO_2 was not being utilized as a carbon source. Disappearance of thiodiglycol with corresponding increase in $O.D._{595}$ of the culture was monitored by gas chromatography (GC) *(11)*. Bacteria were identified by gas chromatographic fatty acid analysis (Microbial I.D. Inc., Newark, DE) and Biolog carbon source utilization plates with the corresponding Microstation System software. SH18 showed a fair (37.5%) identity with *P. pickettii* by GC fatty acid analysis and was not identifiable with the Biolog plates. SH42 showed a very high (74.1%) identity with *A. xylosoxidans* ssp. *xylosoxidans* by GC fatty acid analysis. When tested with the Biolog system, SH42 corresponded exactly (95 out of 95 carbon sources) with the results most commonly seen with *A. xylosoxidans* ssp. *xylosoxidans*. Both bacteria grew optimally at 28-30°C and more slowly at temperatures up to 34°C. SH18 grew optimally on thiodiglycol between pH 7.0 and pH 8.6. Optimum growth of SH42 on thiodiglycol occurred between pH 8.4 and pH 9.4. All subsequent experiments with SH18 were conducted at pH 7.5 and with SH42, at pH 9.2.

Growth of SH18 and SH42 on hydrolyzed mustard as sole source of carbon. Bacteria were grown in hydrolyzed mustard (16 mM) as sole carbon source . Resulting cultures were analyzed by ^{13}C nuclear magnetic resonance (NMR) to determine their post-growth organic profiles (Figure 2). In the case of the SH18 culture, analysis detected only one remaining product, 2,2′-dihydroxydiethyl sulfoxide (thiodiglycol sulfoxide), which comprised approximately 16% of the original carbon from mustard. Thiodiglycol sulfoxide was not detected in a sterile control containing hydrolyzed 16 mM mustard, consistent with a metabolic origin for the compound. Since SH18 will not utilize thiodiglycol sulfoxide as its sole source of carbon for growth, it may represent a metabolic

Strain	% Original Carbon Remaining	Compounds Detected
SH18	16	Thiodiglycol sulfoxide
SH42	2 - 3	Thiodiglycol, thiodiglycol sulfoxide, small amounts of ethers or thioethers

Figure 2. Results of ^{13}C NMR analysis of SH18 and SH42 grown on hydrolyzed mustard.

dead-end product. In the case of the SH42 culture, a one hour analysis detected no carbon-containing products in the culture medium. A subsequent 72 hour analysis of this same culture improved the signal-to-noise ratio and it was seen that a maximum of approximately 3% of the carbon from mustard remained in the medium in the form of thiodiglycol, thiodiglycol sulfoxide and small amounts of several compounds tentatively identified as ethers or thioethers.

Toxicity of hydrolyzed and biodegraded mustard. Because thiodiglycol is essentially nontoxic, the expectation was that the medium would be nontoxic after hydrolysis prior to inoculation. However, since the pathway of degradation is unknown, it is possible that toxic metabolites are formed during growth. In order to determine the toxicity, if any, of the culture, two commercial bacteria-based toxicity testing kits were used. Toxi-Chromotest (Orgenics, Inc., Columbia, MD) uses the chromosomally-encoded ß-galactosidase enzyme of *Escherichia coli* test bacteria to catalyze the cleavage of a chromogenic substrate. The intensity of color is inversely proportional to the degree of cell killing by the compound to which the bacteria are exposed. Results show no detectable toxicity from either the sterile medium containing hydrolyzed mustard only, or the post-growth culture medium with either SH18 or SH42. In all cases the $O.D._{405}$ is greater than that of the blank containing only *E. coli* cells and buffer (Figure 3).

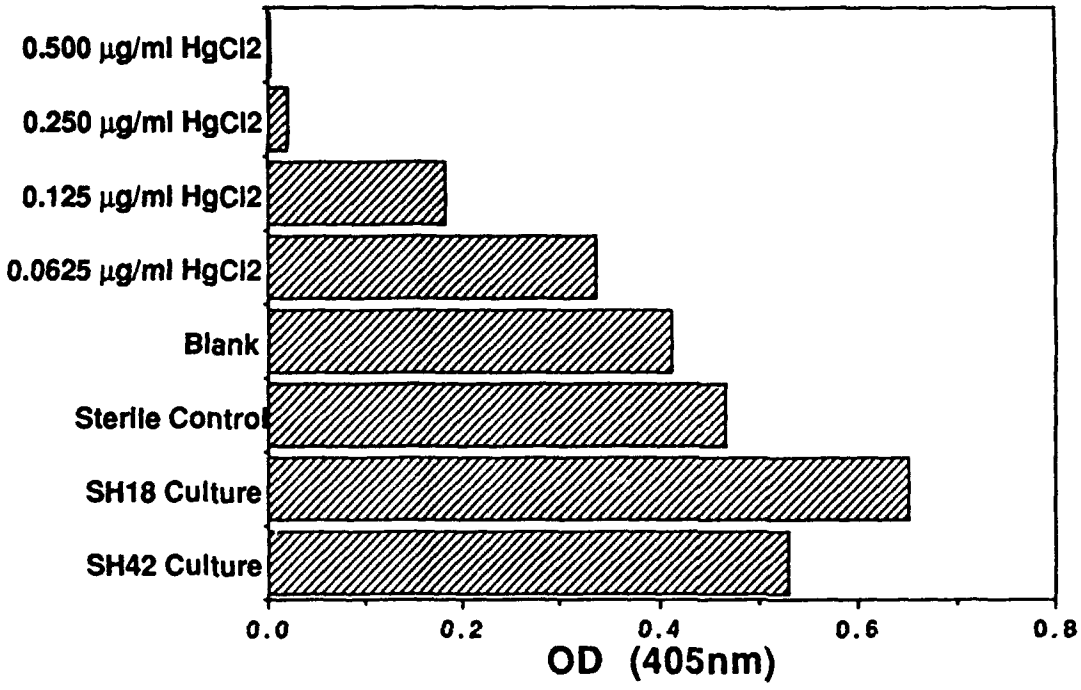


Figure 3. Results of Toxi-Chromotest. O.D.$_{405}$ is inversely proportional to toxicity. $HgCl_2$ is the standard toxic compound. The blank consists of test bacteria and buffer alone (no toxicity). Bacteria in wells of a microtiter plate are incubated with 100 ul of the indicated concentration of $HgCl_2$, hydrolyzed 16 mM mustard or post-growth cultures of SH18 or SH42 bacteria grown on hydrolyzed mustard. $HgCl_2$ results are averages of 2 determinations, others are averages of 8 determinations each.

The second test, SOS-Chromotest (Orgenics, Inc.), specifically detects genotoxicity with *E. coli* which have an SOS promoter cloned upstream of the ß-galactosidase gene. Genetic damage activates the SOS promoter stimulating synthesis of ß-galactosidase. The chromogenic substrate is therefore cleaved in proportion to the degree of genotoxicity. This test was of particular interest because of mustard's alkylating effects on DNA. Results (Figure 4) show that the sterile control, SH18 and SH42 cultures all exhibit no detectable genotoxicity (no increase in O.D.$_{405}$).

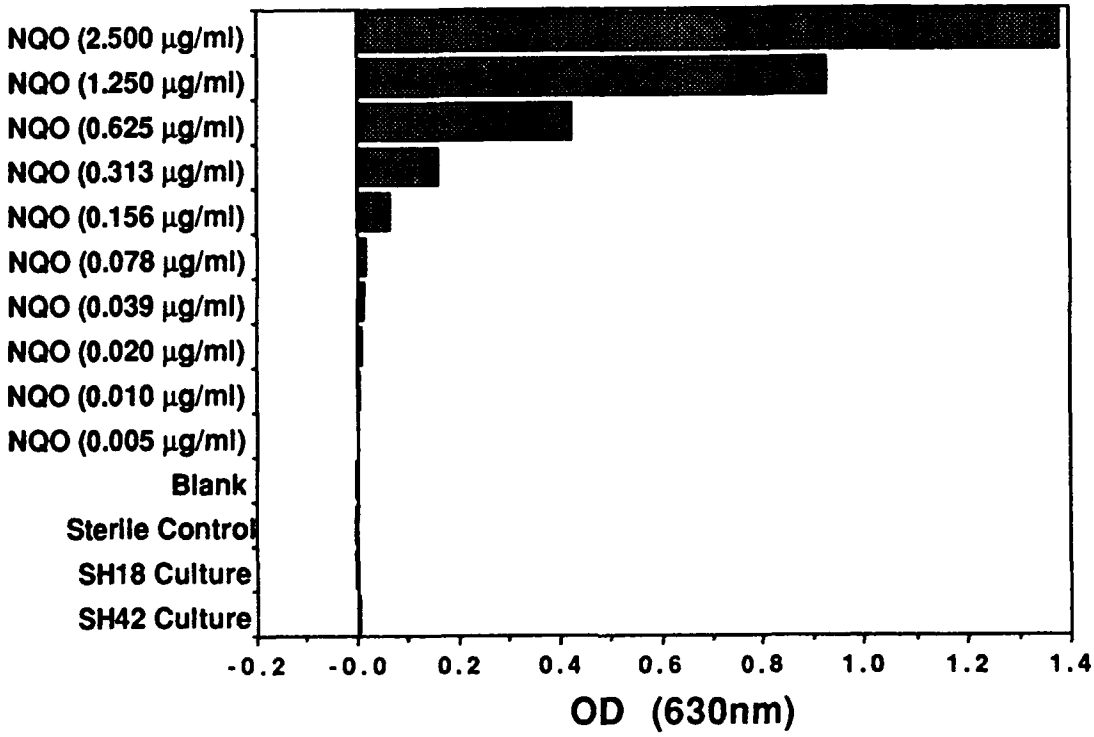


Figure 4. Results of SOS-Chromotest. $O.D._{405}$ is proportional to toxicity. 4-Nitroquinoline is the standard genotoxic solution. Procedures are the same as described for Figure 3.

Fermentation study with munitions-grade mustard from the U.S. chemical stockpile. Although the initial studies demonstrated that a very large percentage of the carbon in the mustard hydrolysis medium could be consumed (at least 97%), it was also of interest to attempt the degradation on a larger scale. For that purpose, a fermentation study using munitions-grade mustard from the chemical stockpile was conducted. A sample was obtained from the circa-1940's chemical stockpile at APG. The mustard was dark brown to black in color (highly purified mustard is colorless) and contained some insoluble particulate material. Presumably, the color and particulate material were remnants of the one-ton steel container in which it had been stored for approximately 50 years. 125 ml of this mustard was hydrolyzed in a slight molar excess of NaOH. Mustard was measured by volume and no impurities were accounted for in the stoichiometry. The solution was stirred overnight and undissolved particulate material was allowed to settle prior to decanting.

For fermentation, this solution was diluted 1:1 in dH_2O and 3 g/L KH_2PO_4, 1 g/L $(NH_4)_2SO_4$, 0.1 g/L yeast extract and 10 ml/L Wolin Salts solution were added. pH was adjusted to 9.2 with a few drops of H_2SO_4. Assuming pure mustard and complete hydrolysis, the thiodiglycol concentration was 50 mM. This medium was filter-sterilized and added to a sterile 1.25 L fermentation vessel on a New Brunswick BioFlo III fermentor. 100 ml of a midlog culture of SH42 growing on thiodiglycol as sole carbon source was added as seed culture. Fermentation was conducted at pH 9.2 at 30°C with constant aeration. Change in the concentration of thiodiglycol and possible appearance of metabolites in the medium was monitored by ^{13}C NMR at various times during growth. Results are shown in figure 5. During the log phase of growth, mustard consumption was approximately 50 ul/hr/L of culture.

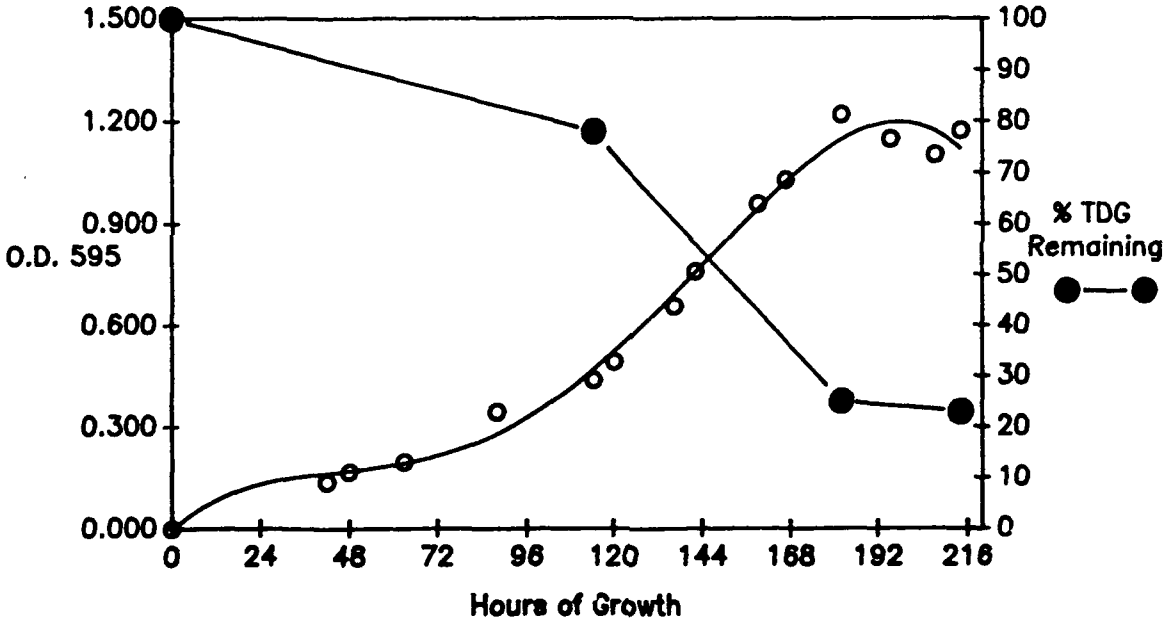


Figure 5. Growth of SH42 on hydrolyzed mustard and ^{13}C NMR analysis.

13C NMR Results from fermentation study. The NaOH hydrolysis step yielded a relatively pure thiodiglycol product. There are several minor impurity peaks at 33 ppm, 34 ppm, and 72 ppm. Approximately 77% of the thiodiglycol was consumed during the fermentation. An apparent minor degradation product peak also developed around 39.5 ppm.

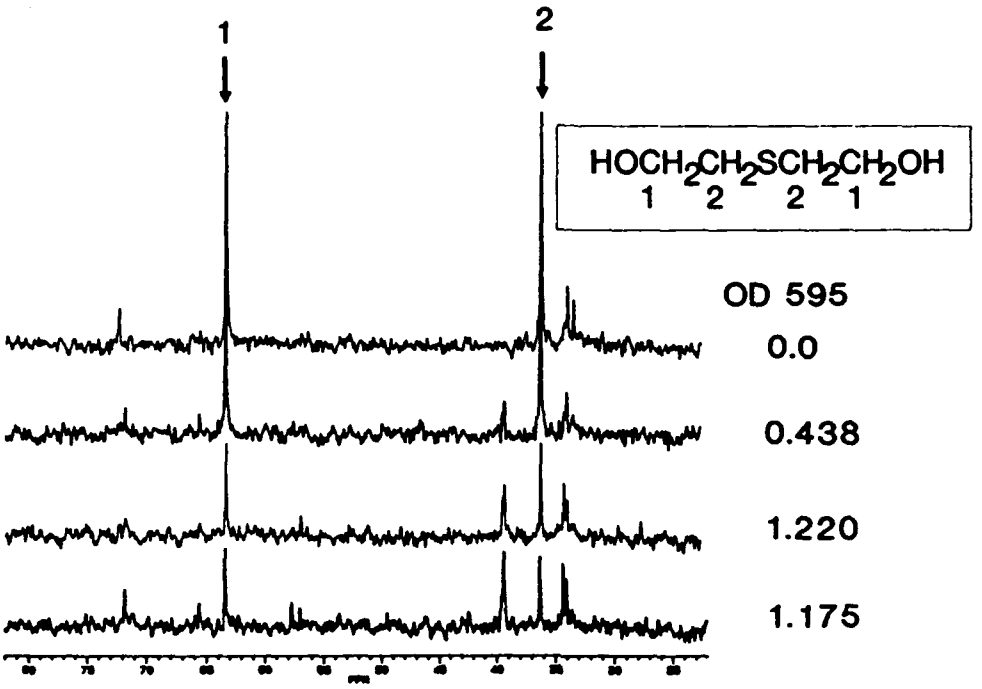


Figure 6. ^{13}C NMR results from fermentation study.

CONCLUSIONS

In summary, these are the first data to show that mustard can be alkaline hydrolyzed and biodegraded; utilized as the sole source of carbon by two different strains of bacteria. NMR analysis of the resulting cultures shows, in the case of SH18, approximately 16% of the original carbon from mustard remaining in the form of thiodiglycol sulfoxide. In the case of SH42, a maximum of only 3% of the original carbon from mustard remains in the form of (primarily) thiodiglycol, thiodiglycol sulfoxide and small amounts of several possible ether or thioether compounds.

Two different microtoxicological tests detected no toxicity from either culture. The tests show that hydrolysis alone will detoxify the aqueous mustard solutions (microtoxicology results from sterile controls). However, it is an important practical consideration that although thiodiglycol is essentially nontoxic, it is also a mustard precursor (chlorination yields mustard) *(9)*. This is the rationale for U.S. export controls on thiodiglycol. Therefore, hydrolysis alone will not convert mustard to an acceptable product for disposal.

Munitions-grade mustard from the U.S. chemical stockpile was also degraded by SH42 in a one liter fermentation. The rate of degradation during the logarithmic growth phase was approximately 50 ul/hour/liter of culture. Refinements of the growth conditions would be expected to yield significant improvements in this figure. Even at this level, the degradation rate is in the range required to be able to degrade the entire APG stockpile in approximately three to four years (the time required would depend on the final degradation rate achieved and the total fermentation capacity).

Proposed fermentation facility. The possible configuration of a mustard demilitarization facility (partial) based on the chemical/biological technology described in this paper is shown in Figure 7. A relatively small agent containment facility would be the site for the removal of the mustard from the bulk container and its hydrolysis to thiodiglycol. It is estimated that each ton container of mustard would be diluted into 15-30,000 L of alkaline water. Once complete hydrolysis is confirmed, the thiodiglycol would be pumped to a holding tank within the fermentation facility. At this point, the pH of the solution would be adjusted to the optimum for growth of the microorganisms and any trace nutrients added. This solution would be fed into the fermentors in a continuous or sequential manner. It is anticipated that the fermentation plant will have 4-8 fermentors running at any one time and at different stages in the process. Upon the completion of the mineralization process, the bulk of the culture would be harvested by separation of the biomass and removal of salts and other soluble materials by ultrafiltration/reverse osmosis. If appropriate, 1-2,000 L of culture would remain in the fermentor to act as the inoculum for the next fermentation run. The bulk of the water used in this process would be recycled, thus greatly reducing the potential impact on waste water treatment facilities. Although an exact cost for this process is difficult to determine at this time, a rough estimate based on conventional fermentation facilities is as follows:

Capitol Equipment (fermentors, tanks, pumps, purification equipment, etc.)	$25,000,000
Building	5,000,000
Operating Expenses (4 years) (salaries, materials/supplies, utilities)	20,000,000
TOTAL	$50,000,000

Once the facility has been completed (2-3 years), the actual demilitarization process would take 3-4 years.

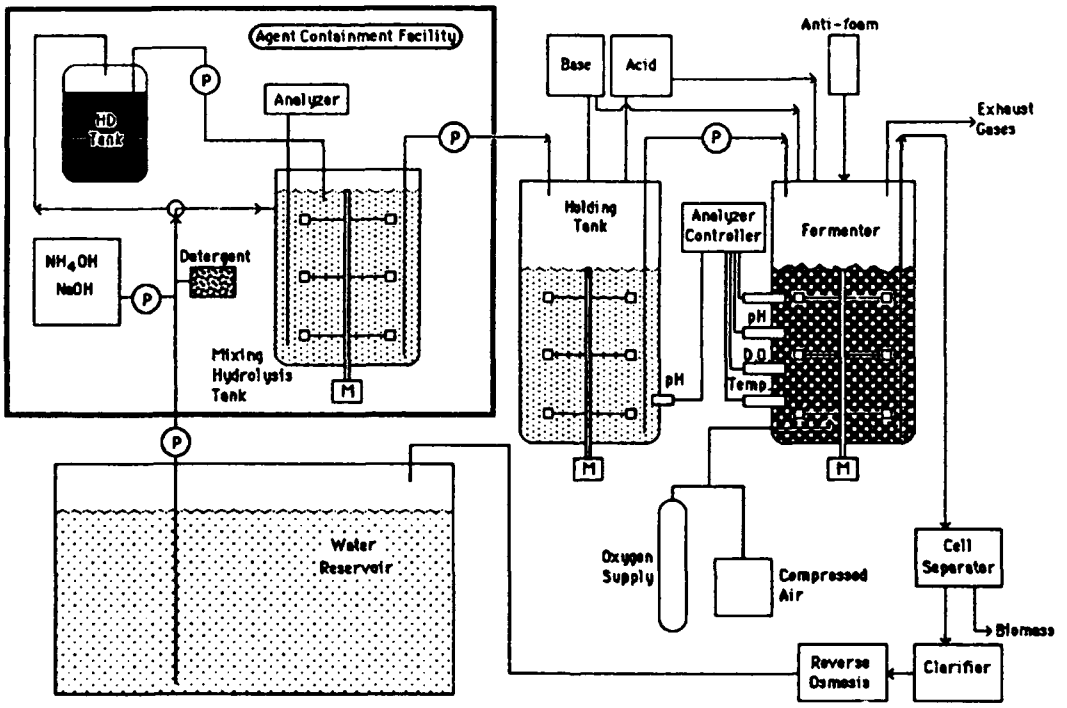


Figure 7. Proposed fermentation facility for mustard demilitarization.

REFERENCES AND NOTES

1. L.R. Ember, *Chemical and Engineering News* **68**:9 (1990).

2. F.P. Ward, in *Biotechnology and Biodegradation*, D. Kamely, A. Chakrabarty and G.S. Omenn, Eds. (Gulf Publishing Co., Houston, 1990), pp. 147-154.

3. A. Picardi, P. Johnston, R. Stringer, *Alternative Technologies for the Detoxification of Chemical Weapons: An Information Document*, (Greenpeace International, Washington D.C., 1991), pp. 20-39

4. T.M. Beardsley, *Sci. Am.* **263**, 48 (1990).

5. P.D. Bartlett and C.G. Swain. *J. Am. Chem. Soc.* **71**, 1406 (1949).

6. Y.C. Yang, J.R. Ward, R.B. Wilson, W. Burrows, J.S. Winterle. *Thermochim. Acta* **114**, 313 (1987).

7. S.P. McManus, N. Neamati-Mazraeh, B.A. Hovanes, M.S. Paley, J.M. Harris. *J. Am. Chem. Soc.* **107**, 3393 (1985).

8. Y.C. Yang, J.R. Ward, T. Luteran. *J. Org. Chem.* **51**, 2756. (1986).

9. Y.C. Yang, L.L. Szafraniec, W.T. Beaudry and J.R. Ward. 1988. *J. Org. Chem.* **53**, 3293 (1988).

10. Biolog plates were purchased from Biolog, Inc., Hayward, CA. Thiodiglycol was purchased from Sigma and had a purity of greater than 99%. Vacuum-distilled mustard from our stock was used. **Warning:** Mustard is a dangerous vessicant and must be handled in a closed system or in a hood with a minimum velocity of 100 ft/min.

11. Gas chromatography with flame photometric detector was performed as follows: One microliter of aqueous culture was directly injected into a Hewlett-Packard 5840A gas chromatograph equipped with a flame photometric detector with a sulfur filter to which was connected a Hewlett-Packard 5840A GC terminal. The GC was equipped with a 6 foot coiled glass column (I.D. 2 mm) packed with 100/120 mesh Gas Chrome Q support phase which was coated with a 3% load of OV-101 as

the stationary phase. Injection port temperature was set at 230° C while detector temperature was set at 240° C. Oven temperature was programed between 60° C and 230° C at a rate of 30° C/minute. The final temperature was held for 15 minutes prior to integration. Helium flowing at 90 ml/min was used as carrier gas. Detector gases were hydrogen, air and oxygen flowing at 220, 110 and 60 ml/min respectively. Slope sensitivity was set at 0.04 while chart speed was set at 0.5 cm/min.

12. NMR analysis was according to the following procedure: Samples were monitored using ^{13}C NMR spectroscopy. The spectra were recorded using a Varian VXR-400S Superconducting FTNMR System operating at 100 MHz for ^{13}C. In all cases, spectra were obtained at a probe temperature of 21°C in an unlocked mode. Tetramethylsilane (TMS) in chloroform was used as the external standard. For each sample analyzed, 592-1192 transients were acquired in absolute intensity mode. The number of transients was dependent on the signal-to-noise required or desired. The spectra were obtained using a pulse width of 8.8 usec (57°), a sweep width of 25 KHz, an acquisition time of 1.0 sec, and a pulse delay of 2.0 sec. WALTZ decoupling was used for full proton decoupling. Quantitative data were obtained by digital integration of the peak areas.

A sterile control was run under identical conditions prior to each sample analyzed. The area of the SCH_2 resonance (delta 32.5) in the sample was divided by the area of the SCH_2 resonance in the control sample to obtain a value for the amount of thiodiglycol consumed. Similarly, the area of the OCH_2 resonance (delta 59.5) in the sample was divided by the area of the OCH_2 resonance in the control sample. The two values were averaged to obtain the percentage of thiodiglycol consumed at time, t. GC-mass spectroscopy was used for confirmation of NMR results.

13. We thank William T. Beaudry, Linda L. Szafraniec (CRDEC Chemical Division) and Adam N. Stroup (CRDEC Biotechnology Division) for ^{13}C NMR analysis of many samples, Jude J. Height (CRDEC Analytical Division) for assistance with the mustard work, and James J. Valdes (CRDEC) and Ananda M. Chakrabarty (University of Illinois, Chicago) for advice on the project.

Vulnerability of Protective Structures to Artillery Fire

Ms. Pamela G. Hayes* and Mr. William L. Huff
U.S. Army Engineer Waterways Experiment Station
3909 Halls Ferry Road, Vicksburg, Mississippi 39180-6199

Introduction

The Structural Mechanics Division, Structures Laboratory, U.S. Army Engineer Waterways Experiment Station (WES), has been conducting research, sponsored by the Office, Chief of Engineers, to develop improved techniques and weapons effects criteria for the rapid emplacement of protective structures and fighting positions for ground-based forces. The goal of the research program is to increase the battlefield survivability of personnel, weapon systems, and equipment from the effects of conventional and saboteur weapons using materials and construction techniques compatible with existing and future logistics systems. The improved techniques and criteria are then used for the modification or design of more survivable protective structures or fighting positions.

One aspect of the design or modification of protective positions is to evaluate the vulnerability of the current positions to a particular threat and quantify the effects of modifications or redesign. In the past, the only way to make this type of determination was to construct the positions and conduct repetitive weapons effects tests. In a time of rising construction and test costs and ready availability of personal computers, it has become much more cost effective to begin evaluating a position's vulnerability and subsequent increases in survivability due to modifications and redesign using computerized probabilistic methods.

A computer code, Computer-Based Resistance Analysis (COBRA)[1], was developed at WES for use as a tool to use in the design and modification of protective positions. The code evaluates the probabilistic vulnerability of protective structures, fighting positions, some equipment, and personnel to the effects of indirect fire weapons. Modifications have been made to the code to increase the type of material properties and weapon threats that can be modeled in the code. This is not a difficult procedure since the code was written in modular form; therefore, modifications for the code occur as new problems are encountered. Using the computer code, it is possible to identify areas of the protective positions that are the most

vulnerable to the selected threat and direct the redesign or modification of the position in that area. It also allows an evaluation of the survivability gained by providing additional protective measures to positions. This type of evaluation may show that in some cases the increase in survivability is not proportional to the increased manpower and monetary resources, which must be spent for an additional level of protection. This paper will describe the procedure used to evaluate vulnerability of a protective structure and fighting positions and show the increase in survivability for a Tactical Operations Center when additional protective measures are employed.

Vulnerability Analysis

The vulnerability analysis of protective positions evaluated using computerized probabilistic methods depends on the correct simulation of the target geometry, failure criteria, and attack scenario. WES participated in a live-fire artillery field exercise where it was possible to evaluate the vulnerability of a sand-grid command post bunker and several dragon fighting positions. The target descriptions, failure criteria, and attack scenario for this exercise were simulated using the COBRA code.

The COBRA code represents protective positions as a combination of critical and noncritical components, which define the structures, personnel, and equipment that are relevant to the vulnerability analysis. It is possible to model all components of the protective position in great detail; however, components that are not critical to the function of the position are often crudely modeled. The crudely modeled components often represent elements of the position that provide shielding to critical elements. Each component of the target is assigned a component number and is defined in space as a series of adjacent triangles that have vertices in x, y, and z coordinates. The target description process can be simplified by using component-generating programs that prepare the required triangular description data for simple shapes.

The dragon fighting positions were modeled based on a description in the U.S. Army Engineer Survivability Manual[2]. The overall dimensions of the modeled dragon position were 18 ft long and 10 ft wide. The major components of the position consisted of a trench, side berms, timber roof, overhead soil cover, and two men. The position consisted of 4 in. by 4 in. timbers spanning the trench and providing support for the overhead soil cover. The width of the overhead cover was 4 ft, the height 1.5 ft, and the length 10 ft. The berms were positioned on each end of the 14-ft-long trench and were 10 ft long, 2 ft wide, and 1.5 ft high. In the live-fire exercise, three dragon positions were constructed. It is important to note that the different dragon positions in the test varied in construction details from position to position and in each iteration of the test so that

Figure 1. Dragon position constructed for a live-fire field exercise.

the modeled position will not truly represent the test positions. It should, however, be possible to make a general comparison of model predictions and test results. One of the dragon positions is shown in Figure 1. One of the men was modeled standing to the side of the overhead cover and was approximately 6 ft tall. The other man was sitting under the overhead cover and had a height of 3.5 ft. Figure 2 shows the components of the dragon position.

The vulnerability analysis requires that the failure criteria for a position be defined in terms of its components. This is accomplished by assigning each component material properties and damage mechanisms. The relationships of the individual component vulnerabilities to the vulnerability of the entire position are defined in the system definition using Boolean logic. The critical components that make up the position are grouped according to function into critical subsystems. The subsystems can then be grouped to form systems essential for the function of the position. Using subsystem and system definitions in the failure criteria make it possible to identify the areas of the position receiving the most damage and allow a quantification of the increases when the position has been redesigned or modified in some way. The components designated as being critical to the dragon position were the men, overhead soil cover, roof,

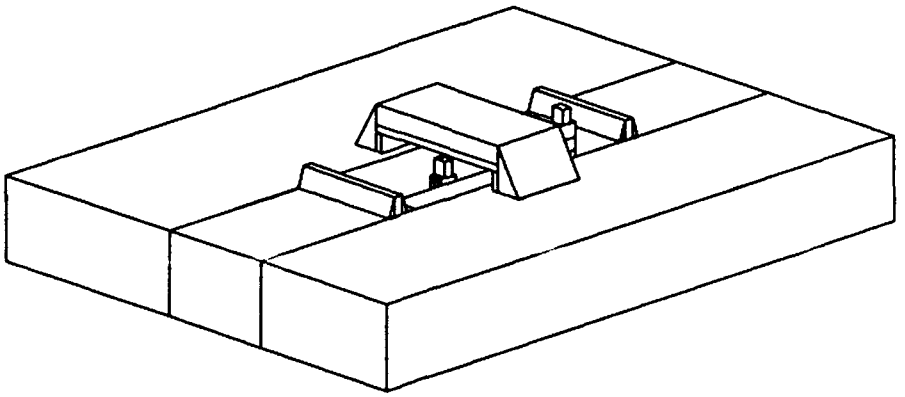


Figure 2. Dragon position model for COBRA vulnerability analysis.

and the berms. Subsystems were set up to define the damage to the berms, overhead soil cover, and roof. These components were not designated as critical; however, they were essential to the vulnerability analysis. If the amount of damage to these components is known, it may be possible to design positions that would provide better protection. The vulnerability of the position would then be defined by the probability of the men in the position being incapacitated.

A sand-grid command post bunker was also modeled for the live-fire exercise. Sand-grid is a polyethylene plastic honeycomb grid material that is shipped collapsed and can then be expanded and filled with dirt to construct revetments or load-bearing walls. The bunker was constructed from two load-bearing, sand-grid walls supporting a timber roof covered by 1.5 ft of soil cover. Another sand-grid wall was placed approximately 5 ft from the ends of both the load-bearing walls. This provided access to the bunker from each end on both sides and still provided excellent fragmentation protection. A typical sand-grid command post bunker is shown in Figure 3. The bunker walls were 20 ft long, 6.5 ft high, and 3 ft wide. In order to locate localized damage to the sand-grid walls, they were modeled in 5-ft-long sections. The timber roof was 20 ft long and 12 ft wide, and soil cover with a height of 1.5 ft was modeled. The bunker was modeled with two men inside.

Figure 3. Sand-grid command post bunker.

The failure criteria for the sand-grid bunker were based on the protection provided to the two men in the bunker. It was also important to determine the damage to the load-bearing, sand-grid walls supporting the roof and to determine if the end walls were going to be damaged to the point that they no longer provided fragmentation protection. The walls in the field test were constructed with locally available fill dirt, and the modeled walls were assigned material properties for similar soil. Previous experiments have shown that it is possible for a section of the load-bearing wall to be significantly damaged, and the remaining sections of the wall will continue to support the bunker roof and soil cover. In the failure criteria for the bunker, it was assumed that 5 ft sections of the load-bearing walls were redundant support elements. An expression was also written to determine the amount of damage to the bunker roof and overhead soil cover. Figure 4 shows the components of the sand-grid command post bunker.

After the target description has been defined and the failure criteria established, it is necessary to define the attack scenario. The attack scenario is defined by projectile parameters including fuzing, nose shape

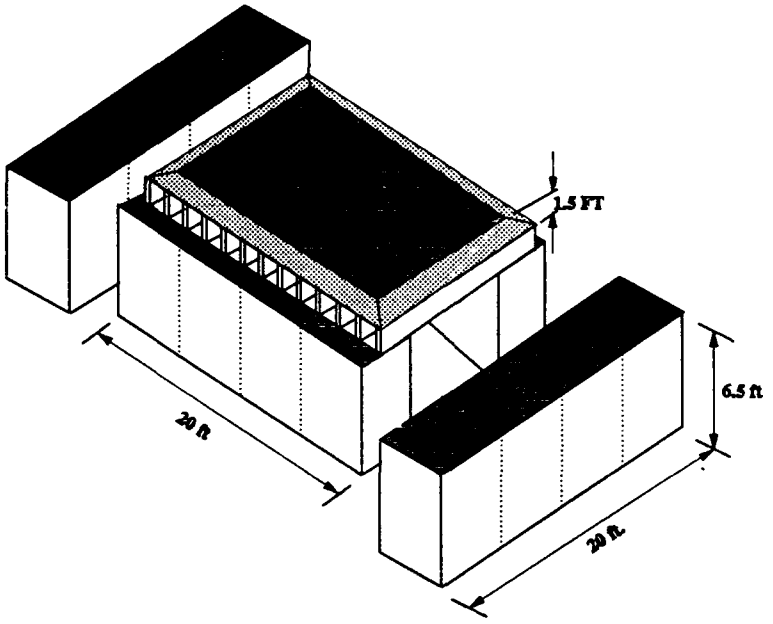


Figure 4. Components of sand-grid bunker.

factor, explosive weight, and fragment descriptions and distribution data. Additional parameters describing the attack scenario are the weapon system errors such as precision, target location, and mean point of impact. The aim points are input as a single point or multiple points. It is also necessary to describe the incoming azimuth of the projectiles, the angle of fall, impact velocity, and the number of rounds in an attack. The number of attacks is input to specify the number of times the scenario will be run to get a probabilistic determination of the position vulnerability.

The attack simulation begins by calculating the path of the projectile to the target. The code uses the aim point coordinates and the weapon system errors to calculate the impact points which are normally distributed inside the weapons error ellipse. The projectile is delivered to the burst point along an entry path with a specified azimuth and incoming angle. It is traced in short increments from the edge of a user specified target hit-miss box along the entry path until it encounters a component or completely misses the target. The fuzing condition of the projectile determines which calculations are conducted when the projectile encounters a component. Projectiles that have contact fuzing are detonated at the point where the projectile encounters the component. Proximity-fuzed projectiles are detonated at the proximity distance from the point where the component is

encountered. Delay-fuzed projectiles are assumed to operate only in soil or in air after encountering soil.

After the location of the burst point has been determined, it is necessary to determine which components are affected by the projectile detonation. This is accomplished by generating rays that project from the burst point in a spherical pattern. The program calculates the intersection of the ray with each component it encounters. The program can then subject the nearest components to the blast and fragmentation associated with the projectile detonation. If the components nearest the detonation are defeated, the blast and fragmentation effects are then applied to the next component on the ray.

Blast propagation models for air and soil are used to determine the loading on various structural elements. The most commonly used methods for calculating the blast damage to components are to treat them as materials that deform elastically or with tabular data compiled for damage to concrete components. If a component cannot be evaluated by either of these means, it is possible to enter a vulnerability equation directly in the input file. The damage to the components is calculated with regard to the possible blast and fragmentation damage from previous rounds. The cumulative damage to the structural components from the effects of blast is represented by the cumulative probability of kill for the component.

The potential for damage from fragmentation to the components of the target is determined by assuming that fragments travel along the rays from the burst point to the components on the ray. The nose band of fragments which the ray passes through determines the fragment weight and velocity pairs that will be used in the damage calculations for the components on that ray. The fragment data input to the program is typically determined from static tests; therefore, the program adjusts the input fragment velocity and direction to reflect the projectile velocity and direction. If the distance that the fragments must travel along the ray before intersection with the first component is greater than 2 ft, the velocity of the fragment is adjusted for the effects of air drag. Once the fragment traveling on the ray intersects with a component, it is necessary to calculate the depth of penetration and determine if the fragment perforated the component. Fragments that perforate a component continue to travel along the ray with a reduced velocity until intersection with the next component. Fragment damage is evaluated using one of three different methods. The first method is the volume method where the probability of failure of a component corresponds to the amount of cumulative volume that has been removed from the component. The second method is the use of algebraic equations to determine a probability of kill based on the fragment weight and velocity. The third method is the use of fragment

velocity versus probability of kill curves that are entered for the component type at each fragment weight.

After all the damage calculations have been completed, the program compiles the probability of kill statistics for the subsystems and systems in the vulnerability criteria for the position. The process of calculating burst points and the associated blast and fragmentation damage is repeated until the specified number of projectiles have been examined or the defined system vulnerability reaches a cumulative probability of kill of 1.0, which signifies the completion of one sortie against the target. To obtain a statistically significant measure of the vulnerability of the position, it is necessary for the program to rebuild the position to its original undamaged configuration and repeat the sortie until the value for the vulnerability of the position falls within a specified confidence interval.

The results of the vulnerability analysis for the dragon positions and the sand-grid command post bunker were compared to the results of a live-fire artillery exercise. The results of the vulnerability analysis showed that the sand-grid bunker should provide excellent protection for personnel from the blast and fragmentation effects of near-miss artillery rounds. These rounds caused no significant damage to the load-bearing walls, the bunker roof, or the end walls of the bunker. The vulnerability analysis did, however, show that the personnel inside the bunker may be injured if there was a direct hit on the bunker roof or the end walls of the bunker. The results of the live-fire exercise confirmed these predictions and showed that the bunker was capable of continuing to provide protection to personnel against subsequent rounds after a direct hit on one of the end walls of the bunker.

The vulnerability analysis of the dragon positions quantified the predicted damage to the position itself and the personnel within. It was possible to compare the predicted damage to the overhead cover and berms to the actual results of the live-fire exercise. The U.S. Army Engineer School evaluated structural damage to the three dragon positions and quantified the results by determining the maximum half-buried TNT charge that the overhead cover could withstand. The results were reported as a percent reduction in strength of the overhead cover. It was determined that an average reduction in strength for the overhead cover was approximately 38 percent. The reduction in strength of the overhead cover predicted from the vulnerability analysis was 40 percent. The reduction in size of the berms for the dragon positions in the actual test was not quantitatively measured. After the test WES approximated the reduction in size of the berms using descriptions of the berms from the Engineer School and photographs of the positions. The estimated value for reduction in size of the berms was an average of 23 percent. The average predicted value of reduction in size for the berms was 35 percent. The comparison of

the predicted and the actual reduction in strength of the overhead cover shows very good agreement. There is not as good an agreement between predicted and actual results for the reduction in size of the berms. It is believed, however, that the agreement is satisfactory because no quantitative measurements were actually made of the berms after the test and that the range of assessed values encompassed the predicted value. A comparison of the predicted probability of kill for the personnel in the positions calculated using known impact points showed good agreement with the test results.

Protective Measures for Increased Survivability

The vulnerability analysis of protective positions can be determined using the COBRA code; however, the main advantage of the computer code as a design tool is the ability to quantify the increases in survivability to positions as additional levels of protection are employed. An example of this can be shown by determining the vulnerability of a Mobile Tactical Operations Center (TOC) with varying levels of protection.

The TOC consisted of four M577 command post carriers with extensions. The unprotected position consisted of the command post carriers placed in a cross configuration with each vehicle backed into one leg of the cross. In this position the command post carriers had little protection from the blast and fragmentation of near-miss artillery rounds. To provide the TOC with more protection from an artillery attack, the vehicles were placed in trenches using the same configuration. Each perpendicular trench was 94 ft long and 12 ft wide. Soil berms were placed around the edges of the trenches to provide additional protection. This position was designed to provide protection to the vehicles and personnel from the near-miss artillery rounds. The next level of protection for the TOC was to provide a roof structure with a 48-in. overhead soil cover to the dug-in position. This provided additional protection to the working area between the vehicles. Both the dug-in and the overhead cover positions are shown in Figure 5. The vulnerability criteria for the position stated that at least two of the M577's should be operational and have a crew of at least three people. The criteria for killing the M577 was a direct-hit artillery round.

The analysis for this position was used to determine the decrease in survivability as more manpower and resources were spent for additional levels of protection. The vulnerability of the unprotected TOC was determined to be 60 percent. When additional time and resources were spent to construct the dug-in position the vulnerability of the position dropped to 41 percent. The additional resources spent to then construct the overhead cover on the dug-in position further decreased the probability of

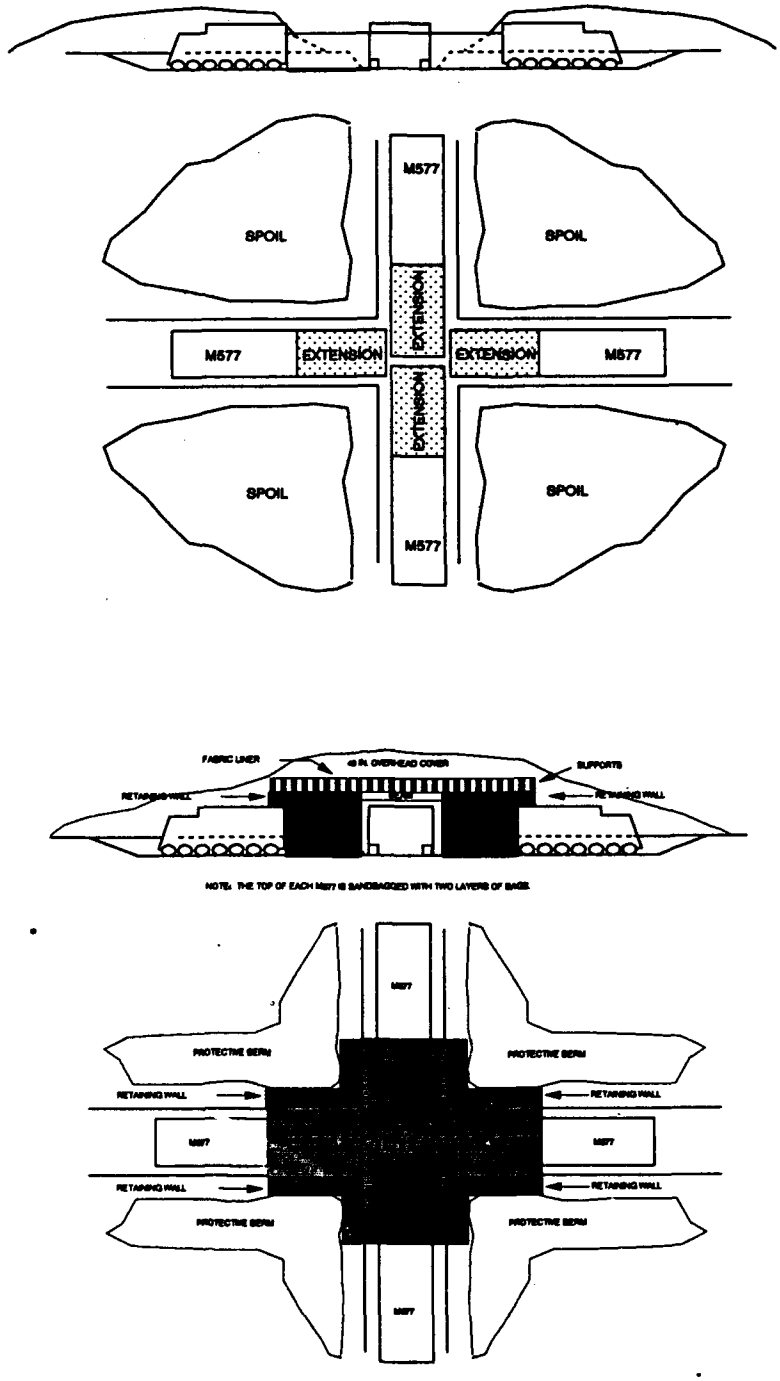


Figure 5. Dug-In and Overhead Cover positions for the TOC.

kill for the position to 19 percent. A plot of the vulnerability of the TOC is shown in Figure 6. The vulnerability analysis has shown that there is approximately a 20-percent increase in survivability when the position is dug-in and another 20-percent increase when overhead cover is employed. This measure of increase in survivability makes it possible to determine how much protection should be implemented in relation to the resources available.

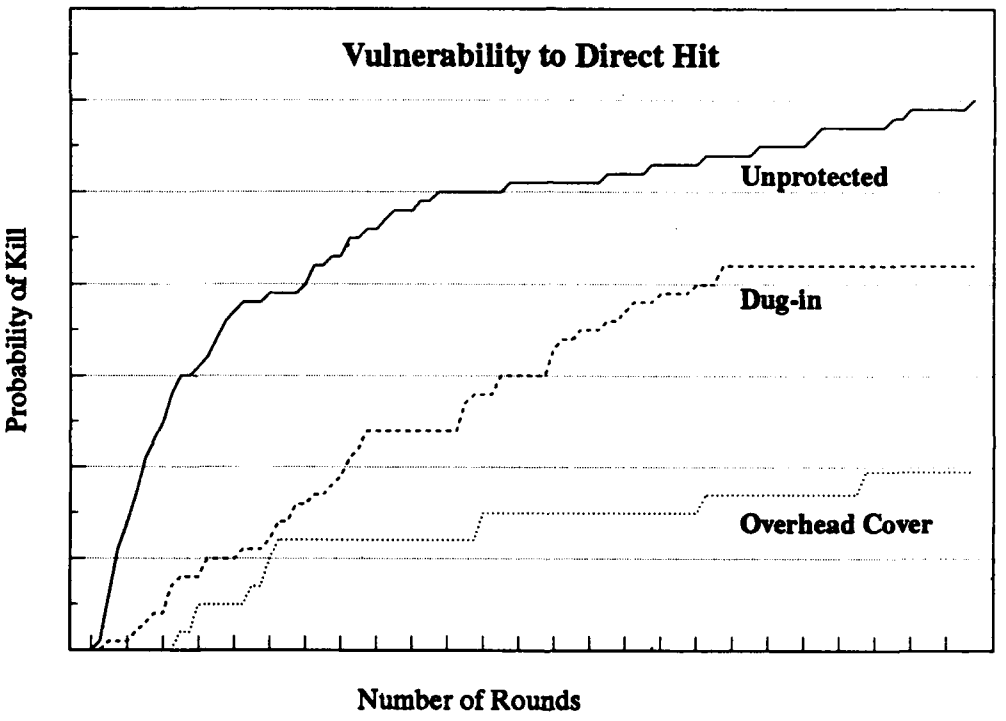


Figure 6. Vulnerability analysis for the TOC.

Conclusions

Using the COBRA code for vulnerability analysis of protective positions has significant advantages over a total test and evaluation type analysis. The probabilistic methods for vulnerability analysis allow several different designs to be evaluated before actual field evaluation begins. It also shows which components of the position are causing the failure and which components could possibly be redesigned or modified to increase protection. The vulnerability analysis of the dragon positions

quantified the reduction in strength of the overhead cover and reduction in size of the berms, as well as predicting a probability of kill for the personnel. The analysis of the command post bunker also showed that the damage to bunker walls could be predicted along with the probability of kill for the personnel. Good correlation was obtained in a comparison of the data for both of these analyses and the results of the live-fire artillery exercise.

The vulnerability analysis of the TOC shows the type of data that could be available to a commander and his staff as he decides how to delegate the resources at his disposal. The unprotected probability of kill can be compared with the decreased probability of kill resulting from committing resources to constructing additional protective measures. By including operational factors with this data, the battlefield commander will be provided with a model for optimizing his survivability resources.

Acknowledgments

This paper is based on work sponsored by the Office, Chief of Engineers, U.S. Army, Washington, D.C. The cooperation of the authorities at WES that permitted the preparation and publication of this paper is appreciated.

References

1. Cheek, James B., Jr., "Development of a Program for Computer-Based Resistance Analysis (COBRA)," Technical Report SL-80-5, (U.S. Army Engineer Waterways Experiment Station, Vicksburg, Mississippi, 1980).

2. Headquarters, Department of the Army, "Field Manual 5-103, Survivability," (Washington, D.C., 1985), pp. C4-C5.

Laser-Induced Luminescence: Soldier Vulnerability

Lisa B. Hepfinger, Ms
U. S. Army Natick Research, Development and Engineering Center
ATTN: STRNC-ITT
Natick; MA 01760-5019

Introduction

The potential vulnerability of the individual combat soldier to detection from laser-induced luminescence was first identified to this Conference 20 years ago[1]. In the ensuing years, the Woodland Battle Dress system of uniforms has been fielded to provide the individual combat soldier with improved camouflage protection. The electronics field has changed dramatically, and what was possible then only with specialized instrumentation can now be done much more simply and effectively. The proliferation of lasers on the battlefield, in the form of target designators, range finders and aiming lights, in combination with the routine use of night vision devices, requires that the phenomena of laser-induced luminescence of military materials be revisited.

When light impinges upon a material, it can be transmitted, reflected, or absorbed. If light is absorbed, the incident light may, in some cases, cause light of a longer wavelength to be emitted from the material. This emitted light is called luminescence, a radiative transition caused by the absorption of radiation. Figure 1 depicts a Jablonski energy level diagram[2,3] illustrating the various photophysical processes that a molecule can undergo following excitation.

When radiation of the required frequency is absorbed by a molecule capable of luminescence, an electron is excited to a higher energy level. When the electron returns to its ground state, the energy emitted can be separated from the light reflected by the excitation source with the proper use of filters[4]. Both the absorption, or excitation, and emission phenomena occur over a range of wavelengths known as the excitation and emission regions.

When lasers are used on the battlefield, they may induce luminescence of natural and man-made materials that can be observed during periods of

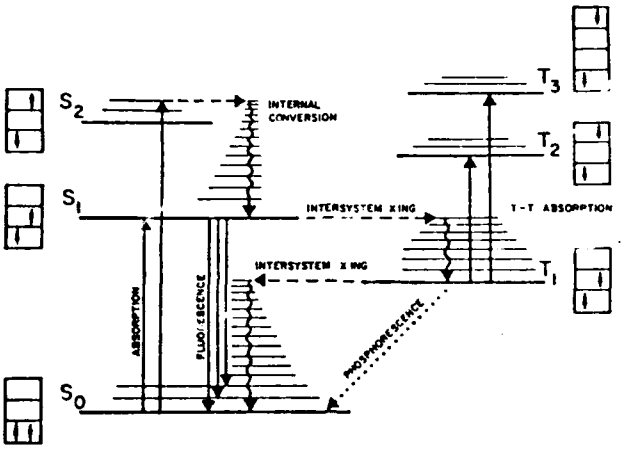


Figure 1. Energy levels and energy transfer in a complex molecule. S_0 is the ground state. S_1 and S_2 are excited singlet states. T_1, T_2 and T_3 are excited triplet states.

darkness with a night vision device or during the day with other detectors. If the laser light causes strong luminescence of a target that differs from that of the background, the signal may be exploited as a means of detection and identification. Previous work[5,6,7] has shown that luminescence can be detected at distances of up to several kilometers.

One of the methods used in the original work for detecting the luminescence of military materials was to utilize a starlight scope as the detector. Figure 2 shows the typical spectral response of the starlight scope available in the 1970s[8,9]. Image intensification technology has progressed since then, and the sensitivities of typical second (GEN II+) and third generation (GEN III) scopes available now are also shown in Figure 2 (personnel communication with C. Hollish, Litton Systems, Inc., 2 January 1992). Using third generation devices, detection ranges can be increased over those possible with second generation devices, or the devices can be operated at lower ambient light levels using the same detection range.

Figure 3 shows the response curves of the two scopes plotted with the emission observed from a typical uniform fabric illuminated with a xenon light source set for excitation at 467.1 nm (Natick Technical Report in preparation, L. Hepfinger). (Similar emission results are seen when the excitation source is a helium-neon (632.8 nm) laser.) The spectral range of luminescence excited in the fabric overlaps quite nicely with the

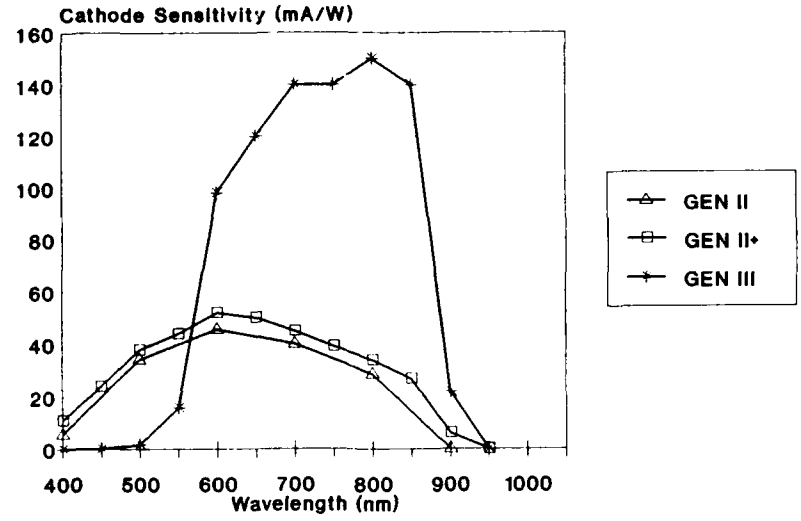


Figure 2. Spectral Response of Image Intensifiers.

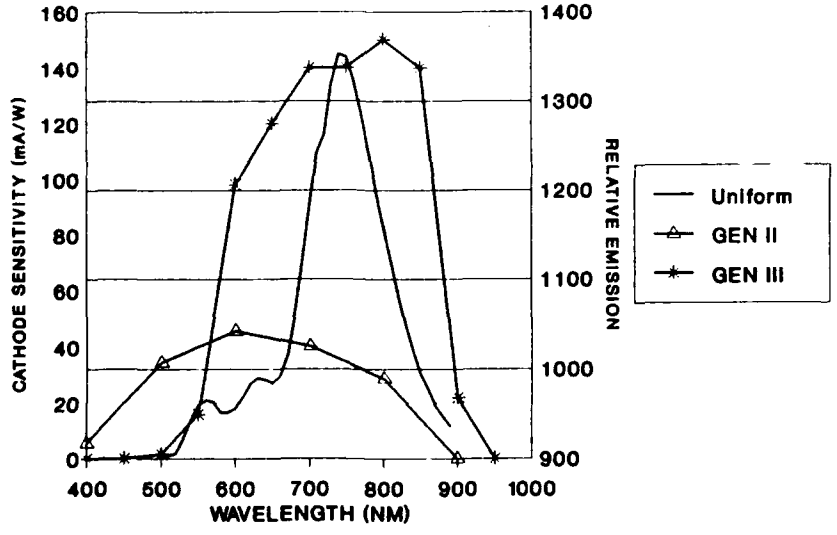


Figure 3. Spectral Response of Scopes (left axis) and Relative Emission of Uniform Excited with 467.1 nm Irradiation (right axis).

spectral response range of the third generation night vision device.

A study was undertaken to characterize the wavelengths of excitation, reflection and emission for typical fielded and experimental materials currently used for camouflaged clothing, individual equipment and tentage. The materials investigated are listed in Table 1, and were all new, unwashed samples. These materials were used to determine which items are most susceptible to detection by laser-induced luminescence and therefore require the reduction or elimination of a signature which has potential for exploitation by an adversary. There is presently no means of differentiating a friend from foe using laser-induced luminescence, since both are expected to exhibit some signature. However, it is in the Army's best interest to reduce the signature created by this phenomena.

Experimental

An SLM Aminco 8000C spectrophotometer was used to determine the emission spectra of the materials shown in Table 1. The samples were excited with a xenon light source set at three different wavelengths, 362.4 nm, 467.1 nm and 529.2 nm. These wavelengths were chosen because of their fairly strong xenon emissions, and they allowed the entire spectral range of interest, 400 to 900 nm, to be scanned for emittance. Reflectance spectra from 400 to 900 nm were obtained using a HunterLab UltraScan Spectrocolorimeter.

Results and Discussion

The Woodland-patterned materials were divided into four groups based on their fiber content, as shown in Table 2. The materials colored for a desert environment are considered separately. All of the materials could be caused to luminesce using 467.1 nm excitation, so only those results are discussed. Figure 4 shows the spectra typical of Light Green 354 emission on Types I, II and III. Emission on Type I is much stronger than the other two and is plotted against the axis on the right.

Table 2. Woodland-Patterned Materials Grouped by Fiber Content

Type	Material	Sample ID
I	Nylon/cotton blends	1,4,7,11
II	100% Nylon	5,8,9,10,14
III	100% Cotton	2,6
IV	Aramid	12,13

Table 1. Materials Included in Study

	Specification	Title	Type	Typical Usage
1	MIL-C-44031	Cloth, Camouflage Pattern: Woodland, Cotton and Nylon, Class 1	W	Battle Dress Uniform (BDU)
2	MIL-C-43468	Cloth, Camouflage Pattern; Wind Resistant Poplin, Cotton, Type III	W	Hot Weather BDU
3	MIL-C-44034	Cloth, Twill, Camouflage Pattern, Cotton and Nylon for Desert Uniform, Class 3	D	Desert BDU
4	MIL-C-3924	Cloth, Oxford, Cotton Warp and Nylon Filling, Quarpel Treated, Class 3	W	BDU Cap
5	MIL-C-43473	Cloth, Coated, Nylon, Polyurethane Coated, Type III	W	Ponchos
6	MIL-C-43468	Cloth, Camouflage Pattern; Wind Resistant Poplin, Cotton, Type IV with Quarpel	W	
7	MIL-C-44031	Cloth, Camouflage Pattern: Woodland, Cotton and Nylon, Class 2 with Quarpel	W	Battle Dress Overgarment (BDO)
8	MIL-C-43734	Cloth, Duck, Textured Nylon, Class 2	W	Body Armor

NOTE: W = 4-color Woodland camouflage print, D = 3-color Desert camouflage print and M = Desert Tan 459

Table 1. Materials Included in Study (continued)

	Specification	Title	Type	Typical Usage
9	MIL-C-44043	Cloth, Ballistic, Nylon, Light-weight, Water-repellent Treated, Class 3	W	Body armor, helmets, armored clothing
10	MIL-C-43734	Cloth, Duck, Textured Nylon, Class 5	W	Individual Tactical Load-Bearing Vest
11	MIL-C-43191	Cloth, Wind Resistent Sateen, Cotton and Nylon, Class 3 with Quarpel	W	Field Coat
12	MIL-C-83429	Cloth, Plain and Basket Weave Aramid, Class 5	W	Aircrew BDU
13	MIL-C-43842	Cloth, Oxford, Aramid	W	Fire retardant clothing
14	MIL-C-44187	Cloth, Laminated, Waterproof and Moisture Vapor Permeable	W	Extended Cold Weather Cloth-ing System
15	MIL-C-44103	Cloth, Duck, Polyester, Fire Water and Weather Resistant, Class 2	M	Tentage

NOTE: W = 4-color Woodland camouflage print, D = 3-color Desert camouflage print and M = Desert Tan 459

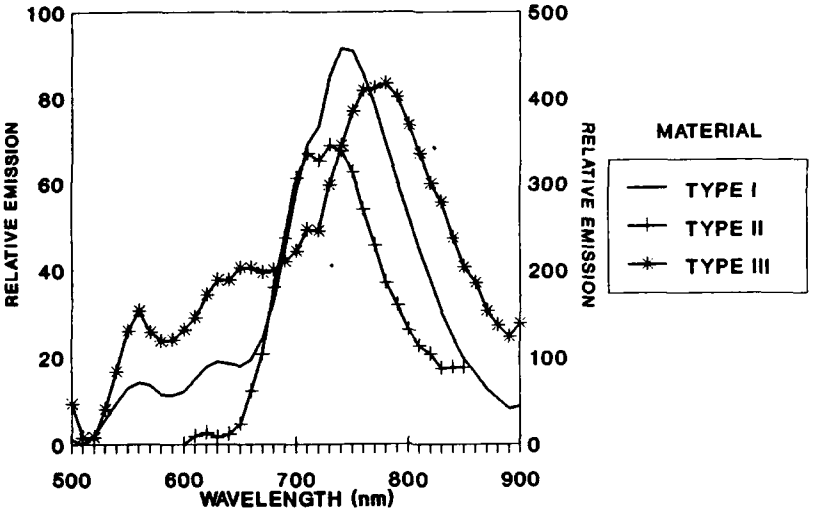


Figure 4. Relative Emission of Light Green 354 with Excitation at 467.1 nm. Type I is Plotted Against Right Axis.

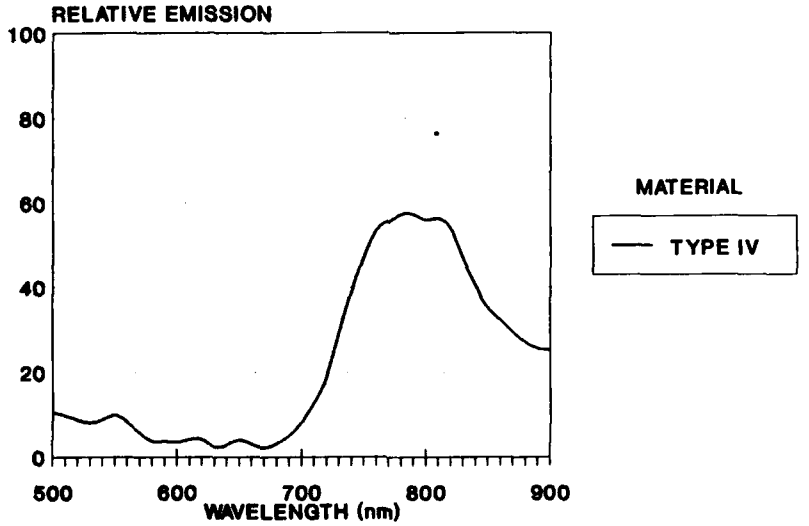


Figure 5. Relative Emission on Dark Green 355 with Excitation at 467.1 nm.

Emission spectra variation with different fiber types and blends is not unexpected. Different classes of dyes must be used on different fibers to produce similar colors, due to the nature of the dye-fabric interactions. The fabric used in the Type I materials is a 50/50 blend of nylon and cotton, and the contribution to the luminescence of the two different classes of dyes used to print the two different fibers in the fabric can be seen. In the Type I materials, Light Green 354 is expected to luminesce the most, since this is usually the ground shade of the fabric. The ground shade for these materials is generally dyed first to be a close match to Light Green 354, and then the remaining colors of the pattern are printed over the Light Green 354. One of the dyes commonly used to match the ground shade in fabrics containing nylon is Acid Blue 258[10], which luminesces in the same range as chlorophyll. When the remaining three colors of the pattern are printed, this luminescence is diminished for Dark Green 355, Brown 356 and Black 357.

In the Type IV materials, the luminescence observed was weak in comparison to Types I, II and III. Figure 5 shows the strongest Type IV emission. In these materials, the Dark Green 355 was found to luminesce the strongest. These fabrics are difficult to print and require a specialized printing process to provide camouflage protection, so it is not unexpected to see a change in the luminescence pattern for these materials.

Figure 6 illustrates the typical reflectance curves seen for the Woodland-patterned materials. The sharp rise in reflectance between 650 nm and 700 nm imitates the spectrum of chlorophyll and provides broadband camouflage protection. Figure 7 shows the reflectance of Light Green 354 plotted with its emission.

Figures 8 and 9 show the emission spectra for the two desert-colored materials. Sample #3 is luminescing fairly strongly from 500 - 700 nm, and Sample #15 is showing some weak luminescence between 600 and 700 nm.

Luminescence excited in the Type I materials, sample 3 and chlorophyll was easily seen in an indoor, simulated woodland environment at a distance of 40 feet, using a 0.95 mW helium neon laser and a third generation night vision device filtered to remove reflection by the laser. Emission from the other Woodland materials, except for Type IV, could also be discerned.

Conclusions

All of the samples tested did exhibit luminescence in a range that can be detected with night vision devices when excited at 467.1 nm. A desert environment without foliage would make Sample #3 more susceptible to detection than an area with some vegetation. Dyestuffs need to be chosen that either completely quench the fabric's luminescence or cause

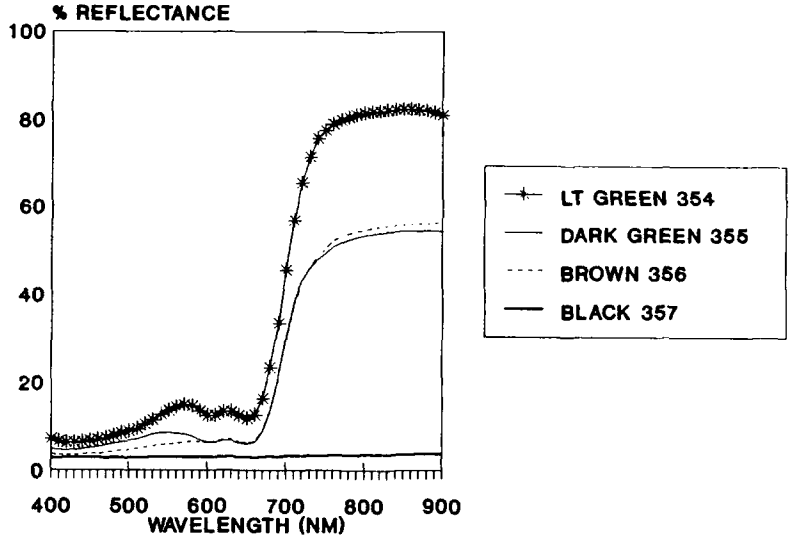


Figure 6. Reflectance spectra typical of Woodland-patterned camouflage.

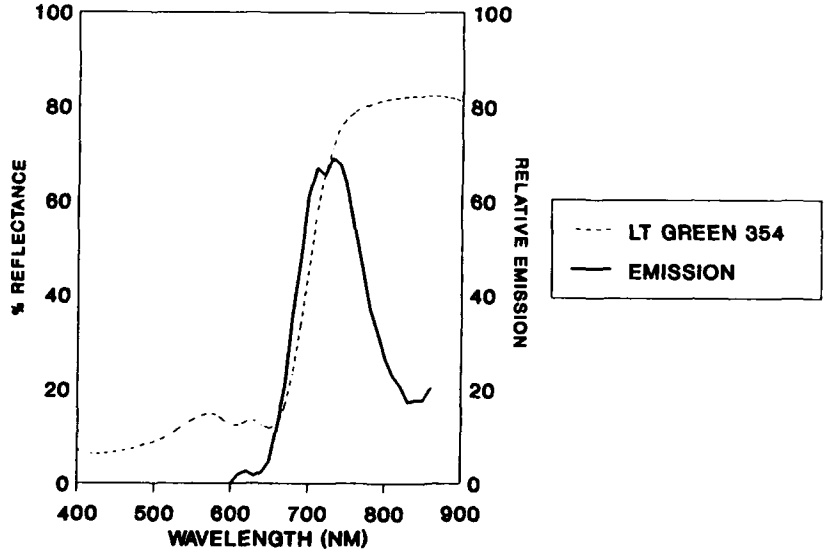


Figure 7. Typical reflectance of Light Green 354 (left axis) and emission spectra (right axis) with excitation at 467.1 nm.

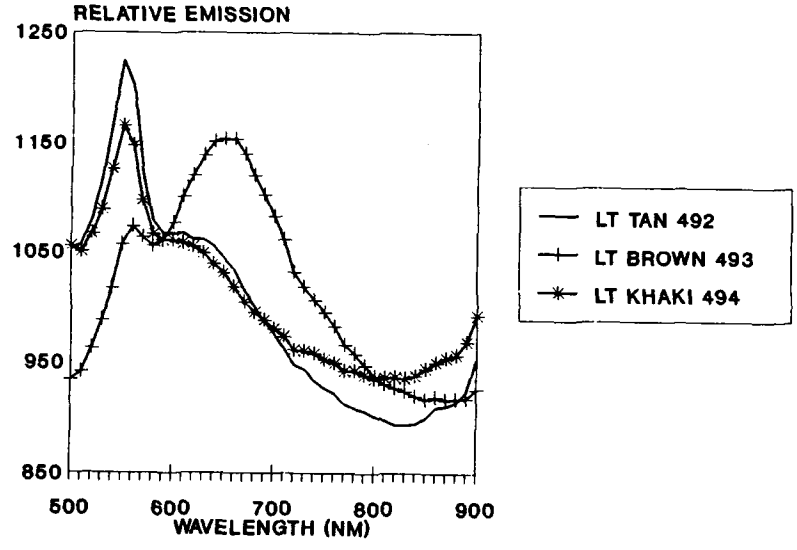


Figure 8. Relative Emission of Sample #3 with Excitation at 467.1 nm.

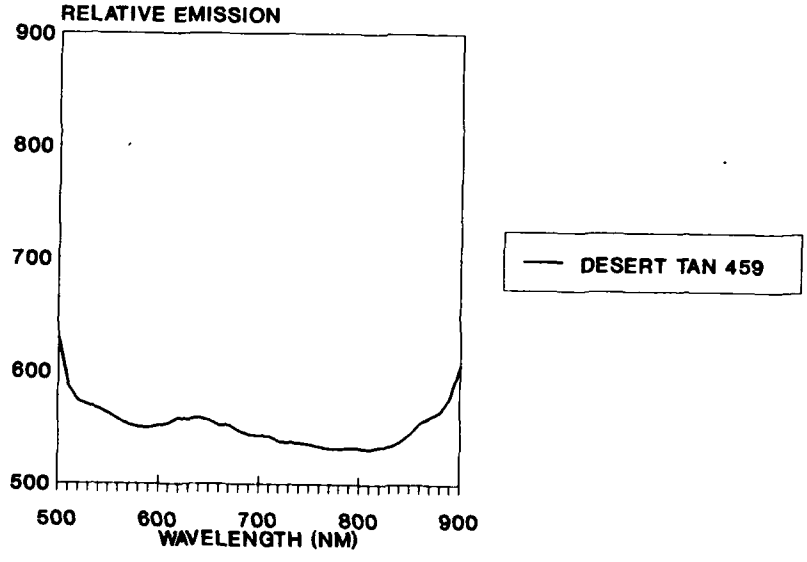


Figure 9. Relative Emission of Sample #15 with Excitation at 467.1 nm.

luminescence that more closely matches the intensity and spectral range of that of chlorophyll, without adversely affecting the near-infrared camouflage protection provided by the fabric. Additional materials, including foreign uniform materials, need to be included in follow-on work. Laser-induced luminescence can be converted to an advantage, since it has potential for use as a means of identifying friendly forces. This concept will be investigated further.

Acknowledgments

The author would like to acknowledge the contributions of Maurice N. Larrivee and the following individuals from the Textile Chemistry and Camouflage Branch of the Individual Protection Directorate at Natick for their help in acquiring emission and reflectance spectra: Tracy Lucciarini, Michael Nolan and Richard Quynn.

References

1. Spectroscopic Studies of Luminescence Emitted by Dyes Irradiated by Laser Light, J. A. Sousa and J. F. Roach, Army Science Conference Proceedings, Volume III, pages 360-369, 29-23 June 1972. (AD 750 377)

2. M. Orchin and H. H. Jaffe, Symmetry, Orbitals and Spectra (Wiley Interscience, New York, 1971), page 297.

3. Geo-Centers, Inc, Final Technical Report, MANPRINT Support for Soldier Identification Concepts using Laser-Induced Luminescence, Contract No. DAAK60-90-D-0002-0088, (Geo-Centers, Inc., Newton Centre, MA 02159) December 1991 (unpublished).

4. F. W. Billmeyer, Jr. and L. B. Hepfinger, Color Res. Appl. 8, 12-16 (1981).

5. J. A. Sousa, J. F. Roach, E. F. Wyner, M. Nakashima and R. Grojean, Laser-Induced Luminescence Studies (I) Remote Detectability under CW Radiation, US Army Natick Laboratories, Natick, Massachusetts (February 1973).

6. J. A. Sousa, J. F. Roach, M. Nakashima, E. F. Wyner, J. Weinstein, R. C. Clapp and R. Grojean, Laser-Induced Luminescence Studies, US Army Natick Laboratories, Natick, Massachusetts (June 1974).

7. Richard E. Grojean, John A. Sousa, Joseph F. Roach, Elliot F. Wyner and Masato Nakashima, Remote Pollution Probing by Laser-Induced Luminescence Techniques, Optical Engineering, Vol 17, No 2 (1978) 139-142.

8. A. O. Ramsley and W. G. Yeomans, Psychophysics of Modern Camouflage, Army Science Conference Proceedings, Volume III, Jun 1982, pages 79-93.

9. H. K. Pollehn, Performance and Reliability in Third-Generation Image Intensifiers, in Advances in Electrons and Electron Physics, Vol. 64A, Academic Press, Inc, London, 1985, pages 61-69.

10. A. O. Ramsley and W. B. Bushnell, Development of the U.S. Woodland Battle Dress Uniform, NATICK/TR-81/008, U. S. Army Natick R&D Laboratories, Natick, Massachusetts (January 1981) (AD A096 884).

An Approach to Mitigating Atmospheric Effects on Image-Based Pattern Recognition by Neural Networks

Dr. Donald W. Hoock Jr.
U. S. Army Atmospheric Sciences Laboratory
White Sands Missile Range, NM 88002-5501

1. INTRODUCTION

Automated classification of man-made objects on the battlefield is an important military application of pattern recognition. Ideal performance under laboratory conditions can be degraded under field conditions, however, by weather, terrain, illumination and both natural and man-made countermeasures such as smoke, dust, chaff and camouflage. This paper presents research results for a technique developed by the U. S. Army Atmospheric Sciences Laboratory, Atmospheric Mitigation and Exploitation Division, to reduce the impact of propagation losses and battlefield smoke and dust obscuration on classifier performance. While developed primarily to mitigate effects of low target contrast conditions, the approach should also have a wider application to reducing the effects of terrain clutter and target camouflage countermeasures.

The following sections describe the method mainly as it applies to image-based target classification using feed-forward, single-layer and multi-layer neural networks trained through back propagation. Originally the goal of the research was to assess or predict classifier performance changes as a function of target contrast losses through the battlefield atmosphere. However, this research led naturally to examining classifier designs for which the impact of propagation losses would be minimized. The result is a novel training strategy in which the classifier is not only designed to recognize whether a target belongs to a predetermined class, but also to actively "ignore" the impact of other undesirable patterns, including obscurants, in the process. The mitigation technique is first described qualitatively. A theoretical justification is then given for the approach based on the combined response of each output neuron in a single neural network processing layer to both a target and an obscurant cloud. Finally, recent successful tests of the technique based on actual field measurements are presented for a truck partially obscured by both black absorptive smoke and white reflective smoke at obscuration transmissions as low as 20%.

As one part of the testing, the classifier was subjected to the effect of terrain clutter surrounding the target as well as to smoke covering the scene. Improved performance in these cases suggests a wider range of application of the technique, including improved discrimination against terrain clutter and camouflage patterns using the mitigation strategy. A concluding section summarizes the research and its benefits.

2. BACKGROUND: STATISTICAL PATTERN CLASSIFICATION

A generic statistical or matched-filter classifier employs a numerical algorithm. It acts on a stream of input data to produce an output value in each of several predefined object "classes" (for example, wheeled vehicle, tracked vehicle, fixed wing aircraft, etc.). The inputs are generally one or more types of data including: image radiance sampled across a scene; a range map across an image; a spatial data transform; a selection of computed image features; or processed spectra. The classifier output in each class spans the extremes of "the object definitely belongs to this class" (which we will arbitrarily assign a numerical value of one) to "the object definitely does not belong to this class" (which we will arbitrarily assign a value of zero). A smooth (perhaps non-linear) change between these extremes of zero and one permits the classifier to be trained and to cope somewhat with small perturbations such as noisy input data, small misregistration errors, moderate image blur, and other factors.

The classifier produces its outputs by forming weighted (perhaps non-linear) combinations of the input data. Different weights are determined for each output class through a "training phase". During training the classifier weights are adjusted to converge on the correct answer (a one or zero) in each output class and over a robust set of input examples of the objects or data sets representative of the classes.

When presented with a novel test image or input data set that was not part of the training set, the classifier will generally respond with values that are neither precisely one in the correct class nor zero in the other classes. A decision space is therefore defined to aid in assigning the object to its proper class. One simple representation of the decision space is shown in the figures later in this paper. There the output in the correct class is plotted along the horizontal axis, and the largest output from all the other classes is plotted on the vertical axis. In this two dimensional plot the classifier performs well if the output data lie in the lower right corner of the plot. If the output data lie in the upper left corner of the plot then there is high probability of misclassification. Outputs that lie in the region of the diagonal from the lower left corner to the upper right corner of the plot represent objects that either: do not fall into any of the predefined classes; contain a mixture of dominant features spread over more than one class; or simply represent such a poor match or such noisy data that no attempt should be made at classification.

3. A QUALITATIVE DESCRIPTION OF A MITIGATION APPROACH

This last situation is precisely what occurs when a target object that would otherwise be well-classified is now partially obscured by screening smoke, dust, chaff or similar countermeasures. A simple, but adequate, model describing the process of obscuration of a target object of image L_t to produce a partially obscured image L_o is given by:

$$L_o = T L_t + (1 - T) L_s \qquad (1)$$

where T is the propagation transmittance, L_s is the image of a totally obscuring cloud, and L_o is the resulting image that is now available for input to the classifier.

The transmittance T ranges from one (totally clear conditions) to zero (totally obscured conditions). It thus acts to interpolate L_o between a totally unobscured target L_t and a totally obscuring cloud L_s. Actually L_s itself has a relatively weak dependence on T for obscurant transmittances above 0.05. At transmittances below 0.05 an object's remaining contrast is usually so low that there is little practical chance for object classification. (An exception could be a target with very intense "hot spots" which could still show through the obscurant even if 95% of its energy is removed). The magnitude of L_s, which is technically known as the "limiting path radiance", does significantly depend on the scattering, emission and absorption properties of the obscurant as well as on the state of natural or man-made illumination energy at the wavelength of interest.

The transmittance T can also vary over the scene. Under certain conditions T can be quite uniform (for example, haze; large area obscuring clouds; long paths through rain, fog or water vapor; etc.). Under other conditions both T and L_o can be quite variable (for example, turbulent diffusion within obscurant plumes; foreground thermal clutter; etc.). The statistics for T and L_s in smooth and structured aerosol clouds have been investigated and reported elsewhere[1,2,3].

Although arguably not as well founded in theory as for obscuration processes, Eq. 1 could also be viewed as the simplest model to introduce parametrically the increasing effects of camouflage patterns, terrain clutter, and additive noise L_s into clear target images, feature sets or spectral inputs L_t. In analogy with Eq. 1, T and (1-T) are then the fractional mixtures of pristine object data L_t and countermeasure pattern L_s respectively. This latter pattern could be, for example, a camouflage scheme, a terrain texture, a spectral interferant, noise, etc.

One can consider strengths and weaknesses of different approaches to mitigating the degrading effects of L_s and classifying the input L_o as an object belonging to class L_t. Three such approaches are:

a. The classifier output could be corrected "after the fact". The effect of T and L_s on classifier outputs can be computed once the classifier is trained. Corrections could thus be made based on prevailing conditions in the field. The weakness of this approach is that it requires that T and L_s be measured or modeled in the field and then applied to classifier outputs. Its strength is that it can be applied to an existing classifier already designed and trained on pristine target object data.

b. The classifier could be trained on a wider set of input data including some for the target objects partially degraded by T and L_s levels. Perhaps a few different sets of weights could be generated for very different degradation conditions. The best or appropriate set would later be selected in the field as needed, based on the current or expected environment. The major weakness of this approach is that it dilutes the power of the classifier. It will be looking to classify a "target in smoke" or "a target in rain" rather than concentrating on its primary purpose: to classify the target itself. The strength of the method is that mitigation measures can be taken before going to the field.

c. The classifier could be trained to recognize the path radiance L_s (or equivalent contaminating pattern) as one or more separate new classes to be discriminated from the target objects by the classifier itself. However, since one wants the classifier to "ignore" rather than "recognize" the presence of the countermeasure, the training scheme must be modified. As shown in the next section, the classifier can be trained to output an intermediate value (for example, one-half) in the target classes when shown pure images or data sets of L_s with no target present. Thus, outputs in the target object classes are trained to neither be biased toward nor away from classifying the target object when both L_s and L_t are present in the input data L_o. This method has the disadvantages of increasing the number of weights stored in the classifier and requiring longer training computation. However, its strength is that nothing additional needs to be measured or adjusted in the field. Only data for the pristine target objects L_t need be collected for the training sets. (The latter can be an important consideration since most existing training data sets are taken at short ranges in chambers or under clear conditions).

We believe the last of these approaches is the most logical. The following section shows the theoretical basis for this training strategy.

4. THEORY BEHIND TRAINING ON OBSCURANTS AS A SEPARATE CLASS

Figure 1 shows the layers of a generic multi-layer neural network classifier. Inputs are image pixel intensities or other types of data as discussed above. There are N inputs $X_i^{(0)}$, i= 1, 2, ..., N in layer p= 0. A weight matrix $A_{ij}^{(p)}$ and an additive constant factor $\theta_j^{(p)}$ linearly combine neuron outputs from the previous (p-1) layer into a vector $Y_j^{(p)}$ as given by

Eq. 2. These are then transformed by a nonlinear sigmoidal function $X_j^{(p)} = S[Y_j^{(p)}]$ given in Eq. 3. (Different equations with a similar effect of sigmoidal output behavior are sometimes used). Sigmoidal output $X_j^{(p)}$ varies between zero and one as $Y_j^{(p)}$ varies between minus and plus infinity. Internal layers, called "hidden" layers, usually range from one to three in number. The number of neurons typically decreases by a factor of three to five in each successive layer. The final layer outputs M decision values, one for each of the M classes to which the image might belong. Subsequent processing (not shown) applies a decision algorithm to select the most likely class for the object.

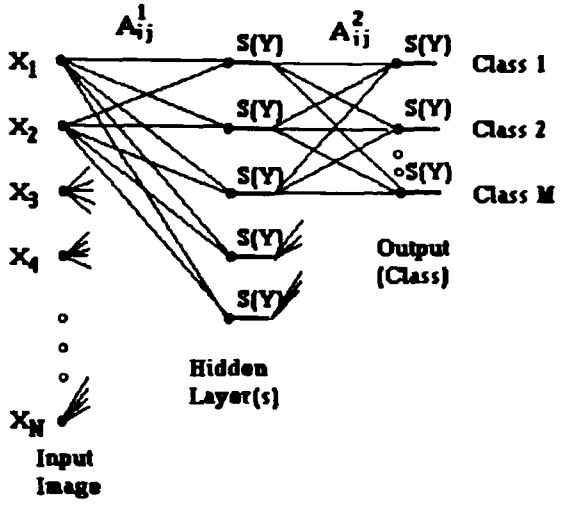


Figure 1. Generic Classifier as a Multi-layer Neural Network.

$$Y_j^{(p)} = \sum_{i=1}^{N^{(p-1)}} X_i^{(p-1)} \; A_{ij}^{(p)} + \theta_j^{(p)} \qquad (2)$$

$$X_j^{(p)} = S\left[Y_j^{(p)} \right] = \frac{1}{1 + e^{-Y_j^{(p)}}} \qquad (3)$$

The network used in the examples of this paper inputs image pixel intensities. Weights are initially given small random values. The well known method of back propagation[4,5,6] is then used to iteratively modify the weights as the network is shown training data sets in random order. Using the traditional training strategy, $X_j^{(p)}$ in the output layer is trained toward a value of one if the input image belongs to class j and to zero otherwise. Training halts when all outputs are within a small error (typically at most 1% or less) of their correct output class values.

Suppose a distribution T_i of transmittance values over pixels i is redefined as a mean T and a small fluctuation f_i about the mean:

$$T_i = (1 + f_i)\, T \qquad (4)$$

The response Y_j^{obsc} of a single layer network in output neuron j to the obscuration process of Eq. 1 is (suppressing the layer variable p):

$$\begin{aligned} Y_j^{obsc} &= \sum_i A_{ij} X_i^k T_i + \sum_i A_{ij} L_{si} (1 - T_i) + \theta_j \\ &= T \left(\sum_i A_{ij} X_i^k + \theta_j \right) + (1 - T) \left(\sum_i A_{ij} L_{si} + \theta_j \right) + \eta_j \\ &= T Y_j^k + (1 - T) Y_j^o + \eta_j \end{aligned} \tag{5}$$

where L_{si} is the path radiance limit for pixel i, Y_j^k is response to the unobscured target object (clear training image X_i^k) from class k, Y_j^0 is the response to the fully obscured image alone (no target), and η_j is a residual fluctuation term defined by:

$$\eta_j = T \sum_i A_{ij} (X_i^k - L_{si}) f_i - \theta_j \tag{6}$$

From Eq. 3, define $S_j = S[Y_j]$ and solve for the exponential term:

$$\begin{aligned} \left(\frac{1}{S_j^{obsc}} - 1 \right) &= e^{-Y_j^{obsc}} \\ &= e^{-T Y_j^k} \; e^{-(1-T) Y_j^0} \; e^{-\eta_j} \end{aligned} \tag{7}$$

so that the sigmoidal output response S_j^{obsc} in terms of the combination of responses to the clear object S_j^k and the obscuration alone S_j^0 is:

$$S_j^{obsc} = \frac{1}{1 + \left(\frac{1}{S_j^k} - 1 \right)^T \left(\frac{1}{S_j^0} - 1 \right)^{(1-T)} e^{-\eta_j}} \tag{8}$$

The residual fluctuation term η_j will typically be small. This is partly because f_i is a relatively small fluctuation about zero. The sum usually involves many A_{ij} when many target details contribute to the classification. This will smooth out some of the fluctuations. Also, the obscurant cloud is statistically independent of X_i^k (the clear target object) since the cloud fluctuations are due to atmospheric turbulence which is independent of the scene background. Since f_i is related to the transmittance, it will be relatively independent of the path radiance L_{si} as noted earlier. Finally, it is common to scale all training images to

the same average intensity (in this case it would include the totally obscured images) to prevent bias toward brightness rather than pattern as the network performs its classification. Thus, the mean of X_i^k will generally equal the mean of L_{si} during training.

Equation 8 is the key to the mitigation strategy. We would like for the response to the obscuration S_j^{obsc} to be as similar to that for the unobscured object S_j^k as possible. By inspection, if S_j^0 is trained to a value very close to 1/2, then $(1/S_j^0 - 1) \approx 1$. This term, even when raised to the (1-T) power, will still only multiply by a value close to one. While the term $(1/S_j^k - 1)$, with S_j^k trained very close to one or zero, will be affected by being raised to the T power, the sigmoidal output S_j^{obsc} will still be relatively close to one or zero as long as T is not too small (say above 0.2). Finally, we note that the residual obscuration fluctuation term η will still contribute some noise to the output, but its effect should typically be small for the various reasons cited above.

This training strategy is summarized in schematic form in Fig. 2. Note that there would be no benefit from this type of strategy if one trained the neural net on object images that were already degraded by examples of obscuration. Then one could not separate the response to the target from the response to the obscurant. Note also that the training value of 1/2 in the response of the target classes to training images of obscurant patterns means that we are training the network to explicitly "ignore" obscuration in the target classes rather than to "recognize" that the scene contains both target and obscurant. In fact, once trained, the output from the obscurant class neuron is simply discarded when the classifier is presented with test images. It is not used in the decision making process to assign the object in the test image to the most probable class.

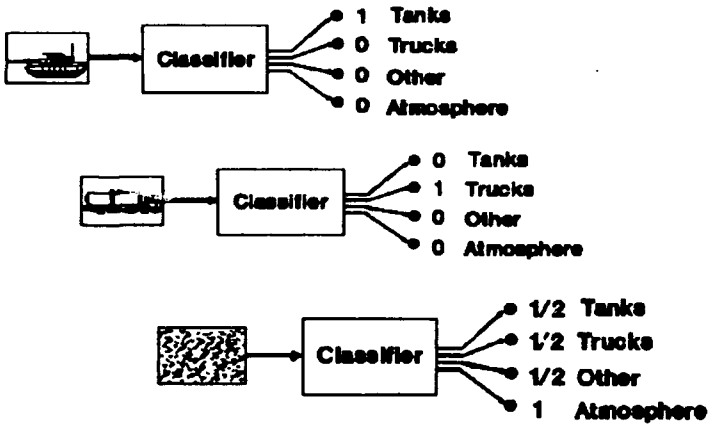


Figure 2. Atmospheric Mitigation Training Strategy.

Analogous arguments are possible when the contaminating pattern is one of camouflage or terrain clutter instead of obscuration. The only requirements are: (1) It must be meaningful to have a range of values from pristine (uncontaminated) state to a fully contaminated state; and (2) The

patterns of the input training images in the class to be "ignored" must be sufficiently different from the patterns of the objects to be "recognized" so that the network training will converge.

During the early phases of this research, the method was tested on synthetic images and synthetic obscuration conditions, including noise. Both linear classifiers and neural networks were examined[6,7,8]. These comparisons consistently show that incorporating a separate class for obscuration improves performance over the traditional training strategy. The next section presents results using real images and actual obscuration examples from field measurements to test the technique.

5. A COMPARISON FOR EXAMPLES OF OBSCURED OBJECTS

The method was tested on experimentally measured, visible wavelength band images of a truck at 1 km that was obscured by white fog oil screening smoke and by black graphite smoke. (Tests on kaolin dust as the obscurant yielded similar results). The experiment was designed as follows.

A single-layer, feed-forward neural net with 800 inputs (pixel intensities) and a multi-layer, feed-forward neural net with 200 pixel inputs, 80 neurons in the first hidden layer, and 20 neurons in the second hidden layer were each trained using back propagation. Both classifiers were tested on the same input images and were tested separately using both the "traditional" training technique (trained on clear target images only) and the mitigation technique (including the "atmosphere" as a separate class along with clear targets).

The networks were first trained on three classes of target images (truck, tank, and aircraft). Results from this classifier were also compared to those found by retraining the networks on four classes of target images (truck, tank at two aspects, and aircraft). The networks were also trained separately on these target combinations plus an additional class represented by synthetic images of "obscuration" which contained patterns of slowly varying fluctuations across an otherwise uniform modeled smoke, consistent with smoke obscuration statistics[1,2]. While actual smoke-obscured images were used in testing, actual smoke images were not used to train the classifier. Training included no information about whether the test smoke would be white or black. All training images were scaled to the same mean brightness (including the obscuration class) to prevent bias toward brightness rather than pattern discrimination. Similarly, training targets had the same relative size.

The goal of the experiment was to determine performance changes only as the level of obscuration at the test site was changed. No tests of performance with changing target aspect or target range were performed. Only the truck was used in testing.

The test procedure consisted of taking clear air images of the truck just before the obscurant was released. These clear images were used in training so that illumination did not become a variable. Smoke was then released in a plume approximately 20 to 30 m wide and 500 m in front of the truck for a period of 10 minutes. Sixteen images each of the truck obscured by white smoke and the truck obscured by black smoke were selected from the test imagery. These test images were chosen so that the average transmittance to the truck was reduced to the range of 20 to 70%, as determined from contrast measurements on objects near the truck and by transmissometer measurements several meters downwind. Fluctuations in obscuration over the target were present in many of the test images.

A rectangular region around the truck was cut from each image and sampled at fixed gain on a gray scale of 0 to 255 as a 40 horizontal pixel by 20 vertical pixel array for input to the single-layer network and again as a 20 pixel by 10 pixel array for input to the multi-layer network. The truck was registered by hand into the center of the image with an estimated error of less than ± 1 pixel, although boundaries of the truck became difficult to see near the lower transmittance levels. (Registration error of up to 3 pixels was the major contributor to the somewhat noisier performance shown in the original report of these results[6].)

Figure 3 shows the digitized image of the truck before reduction to 40 by 20 and 20 by 10 pixel images for input to the classifier. Note that the slight 1 km long slant path introduces desert scrub background all around the truck. Figure 4 is a still camera close-up of the truck seen at ground level to help the reader find the target in Fig. 3.

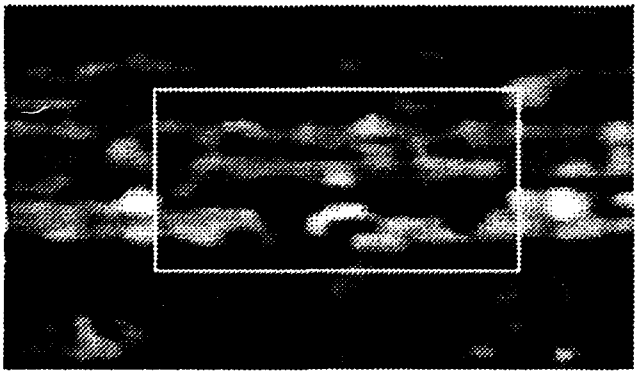


Figure 3. Digital Image of the Truck Target over the Slant Path.

Figure 4. Ground-Level Close-up of the Truck.

Performance comparisons were made for the two extremes of a perfectly segmented target and a completely unsegmented target. In the former case, only those pixels actually belonging to the target were used by applying

the same mask of target pixels in both training and testing. Any pixels in the image that lay outside the target were given the same uniform grey level used in training. However, recent techniques in classification are moving toward the use of unsegmented objects. Thus comparisons were also made of performance using the entire (unsegmented) rectangular image. Background as well as actual obscuration in front of the target both become part of the test image that the classifier has not seen during training.

Figures 5 and 6 show the impacts of fog oil and graphite smokes on outputs from the single-layer neural network using 16 unsegmented truck images. These contain desert terrain surrounding the truck in each image as well as smoke obscuration of truck and background. The classifier was trained on three object classes (set #1): truck, front view tank, and fixed wing aircraft. Classification of the truck using the traditional strategy is obviously poor. Excellent improvement is provided, however, when the atmospheric (smoke) training strategy is used instead.

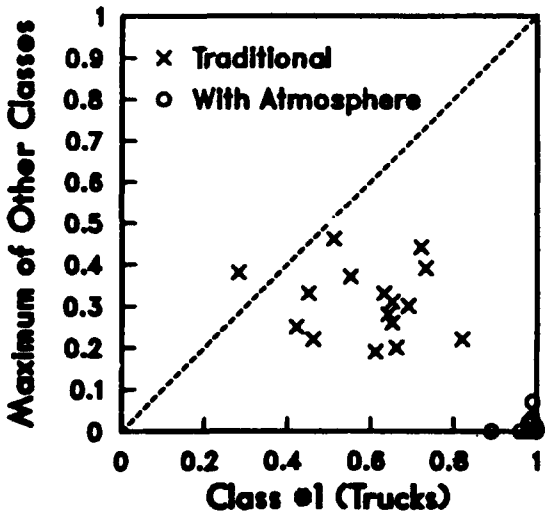


Figure 5. Unsegmented Image of Truck in Fog Oil Smoke; Single-Layer Net Trained on 3 Classes (set #1).

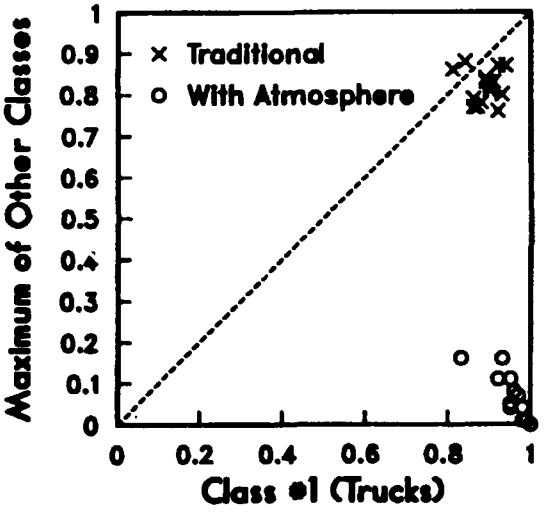


Figure 6. Unsegmented Image of Truck in Graphite Smoke; Single-Layer Net Trained on 3 Classes (set #1).

Figures 7 and 8 show the responses to perfectly segmented truck images (containing no background) of the single-layer neural network. The three training classes (set #2) were: truck, side view tank, and fixed wing aircraft. Performance using the traditional strategy in Fig. 7 would be good even in the presence of the fog-oil smoke. In Fig. 8, however, performance of the traditional training strategy in graphite smoke remains poor. Again, improvement is shown in using the atmospheric mitigation training method. Spread in data is due both to different transmittance and to fluctuations in transmittance in the different images.

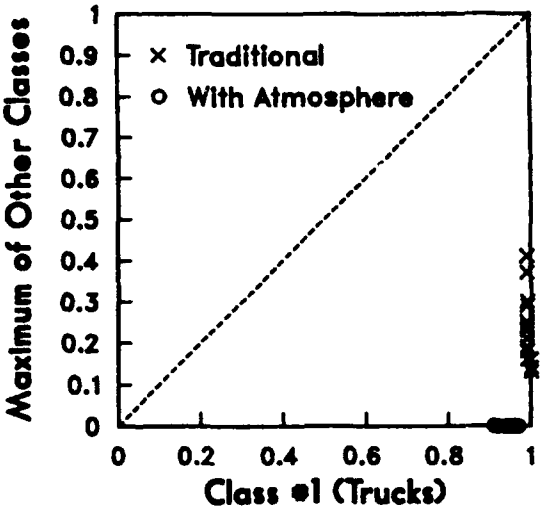


Figure 7. Segmented Truck from Image in Fog Oil Smoke; Single-Layer Net Trained on 3 Classes (set #2).

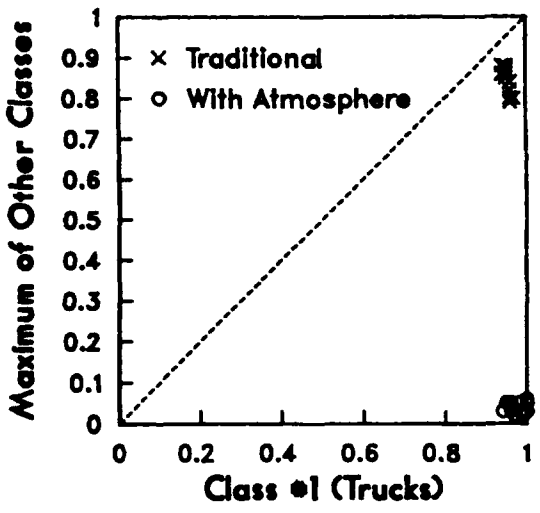


Figure 8. Segmented Truck from Image in Graphite Smoke; Single-Layer Net Trained on 3 Classes (set #2).

Figures 9 and 10 show responses of the multi-layer neural network to unsegmented truck images in fog oil and graphite smokes for the same target set (#2) as in Figs. 7 and 8. Traditional training still results in poor performance, partly due to the fewer pixels (one-fourth) on target.

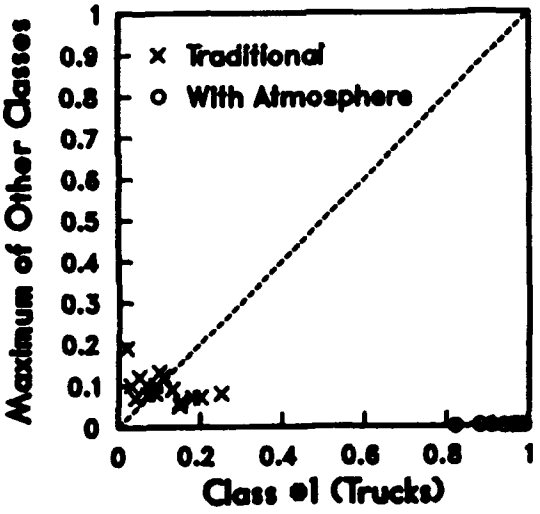


Figure 9. Unsegmented Image of Truck in Fog Oil Smoke; Multi-Layer Net Trained on 3 Classes (set #2).

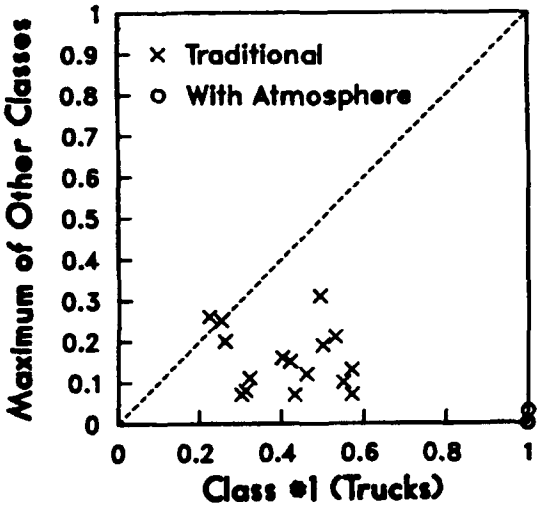


Figure 10. Unsegmented Image of Truck in Graphite Smoke; Multi-layer Net Trained on 3 Classes (set #2).

The mitigation strategy is obviously superior. Figure 11 shows the responses of the multi-layer neural net to perfectly segmented truck images in graphite smoke using the three class training set #2. Performance can be directly compared to that from the single-layer network in Fig. 8 which uses the same images. Figure 12 shows the multi-layer network response to unsegmented truck images in graphite using the same training set #1 and images as the single-layer network of Fig. 6.

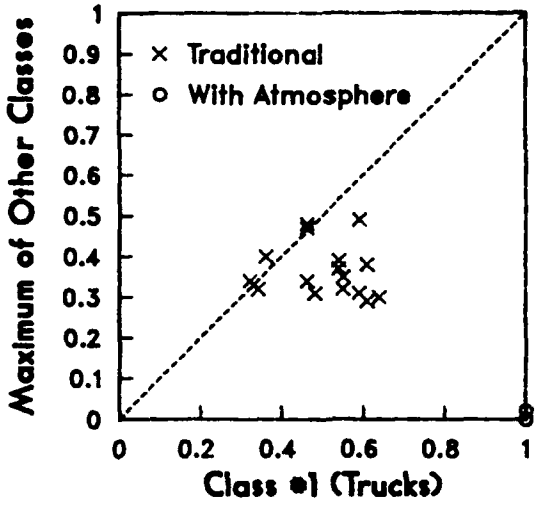


Figure 11. Segmented Truck from Image in Graphite Smoke; Multi-Layer Net Trained on 3 Classes (set #2).

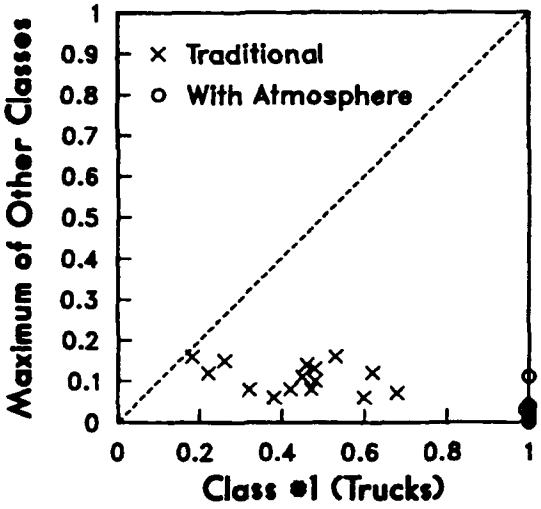


Figure 12. Unsegmented Image of Truck in Graphite; Multi-layer Net Trained on 3 Classes (set #1).

Figures 13 through 16 test if the increase in number of classes affects conclusions by training on all four classes (truck, front view tank, side view tank, and fixed wing aircraft). Figures 13 and 14 show single-layer network responses to unsegmented truck images obscured by fog oil and graphite. Figures 15 and 16 show multi-layer network responses to these same images, but with one-fourth as many pixels in the image. While performance is somewhat worse with this increase in number of classes, improvement using the mitigation technique is still excellent.

6. CONCLUSIONS

We have shown that image-based neural net classifier performance in obscurants and other adverse propagation conditions can be improved by incorporating the atmosphere into the classifier training. The approach has been described both qualitatively and theoretically to require a training strategy to "ignore" rather than "recognize" the obscuration in the scene. This novel strategy has been demonstrated to improve performance in image-based classification for both single-layer and

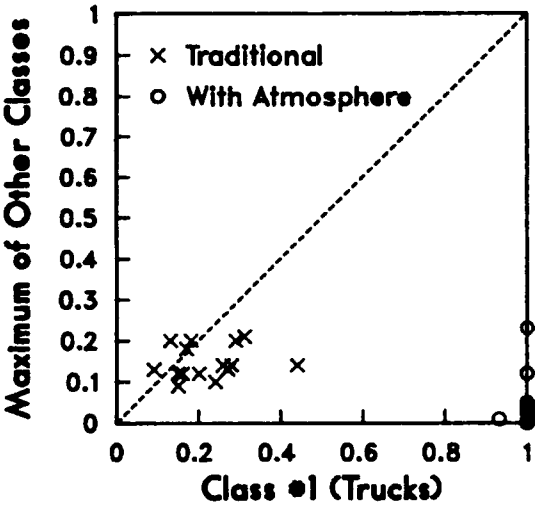


Figure 13. Unsegmented Image of Truck in Fog Oil Smoke; Single-Layer Net Trained on 4 Target Classes.

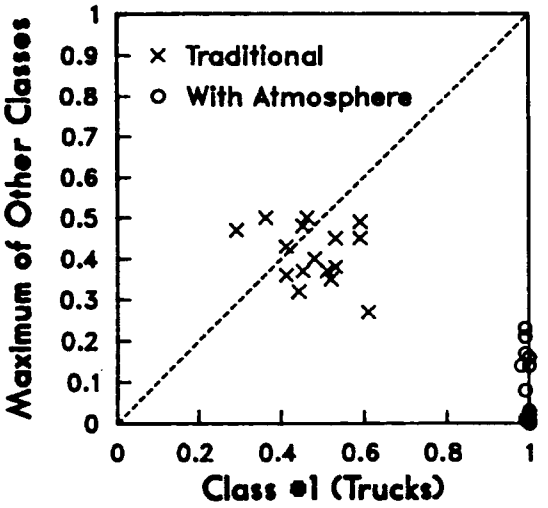


Figure 14. Unsegmented Image of Truck in Graphite Smoke; Single-Layer Net Trained on 4 Classes.

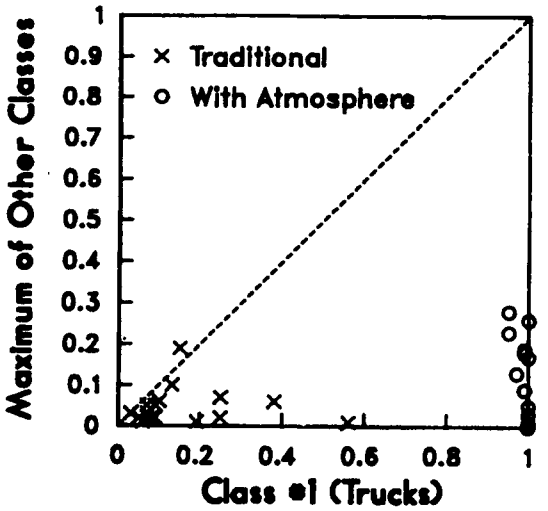


Figure 15. Unsegmented Image of Truck in Fog Oil Smoke; Multi-layer Net Trained on 4 Target Classes.

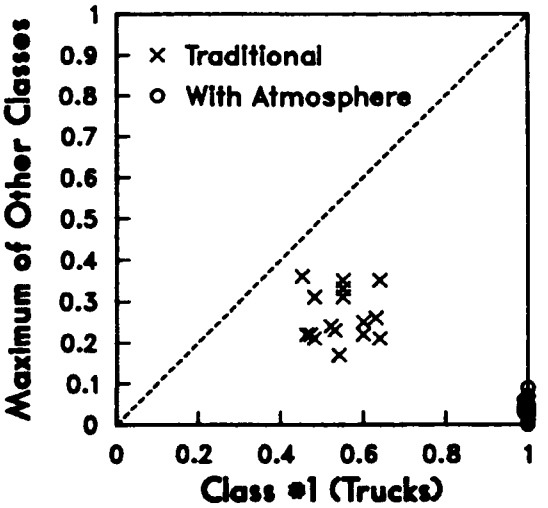


Figure 16. Unsegmented Image of Truck in Graphite; Multi-Layer Net Trained on 4 Target Classes.

multi-layer neural networks using actual images of a truck obscured by both highly scattering white smoke and highly absorptive black smoke at transmittances as low as 20%. On unsegmented targets containing background clutter the method also showed considerable improvement. Thus, the mitigation technique should have application to reducing effects of camouflage, clutter and similar countermeasures on military use of pattern recognition and target acquisition on the battlefield. It is the author's belief that if more techniques to mitigate effects of the battlefield environment and countermeasures were incorporated into basic system designs, then the environment could be made more a part of the solution and less often a part of the problem to applications of emerging technologies on the battlefield.

ACKNOWLEDGEMENT

The author acknowledges the work of Mr. John Giever of the Physical Sciences Laboratory, New Mexico State University under contract DAAD07-91-C-0139 who programmed the neural networks and processed imagery for analysis in this project.

REFERENCES

1. D. W. Hoock, *Modeling Time Dependent Obscuration for Simulated Imaging of Smoke and Dust Clouds*; *Proc. of the SPIE*, 1486, 164-175 (1991).
2. D. W. Hoock, *Theoretical and Measured Fractal Dimensions for Battlefield Aerosol Cloud Visualization and Transmission*, in *Proc. of the Battlefield Atmospherics Conference, Ft. Bliss Texas, Dec. 1991*, U. S. Army Atmospheric Sciences Laboratory Report, White Sands Missile Range, NM, (1992, in press).
3. D. W. Hoock and J. C. Giever, *Detectability and Simulation of Smoke and Dust Clouds in Images*, in *Proc. of the SPIE*, 1312, 298-309 (1990).
4. A. Lapedes and R. Farber, *How Neural Networks Work* in *Neural Information Processing Systems*, D. Anderson, ed., (American Institute of Physics, NY, 1988).
5. R. Lippman, *An Introduction to Computing with Neural Nets, IEEE Acoustics, Speech and Signal Processing*, April, 4-22, (1987).
6. J. C. Giever and D. W. Hoock, *Atmospheric Propagation Effects on Pattern Recognition by Neural Networks*, in *Proc. of the SPIE*, 1486, 302-313, (1991).
7. D. W. Hoock and J. C. Giever, *Propagation Effects on Image Moments, Edges and Linear Classifiers*, in *Proc. of the Tenth Annual EOSAEL/TWI Conference*, (U.S. Army Atmospheric Sciences Laboratory Report, White Sands Missile Range, NM 88002-5501, 1990).
8. J. C. Giever and D. W. Hoock, *Analyzing Atmospheric Propagation Effects on Neural Network Pattern Recognition*, in *Proc. of the Eleventh Annual EOSAEL/TWI Conference*, (U. S. Army Atmospheric Sciences Laboratory Report, White Sands Missile Range, NM 88002-5501, 1991).

Optimization of Particle Model for Large-Scale Computing

Mr. David A. Horner*, Dr. John F. Peters, Mr. Alejandro R. Carrillo
US Army Engineer Waterways Experiment Station
3909 Halls Ferry Road, Vicksburg, MS 39180-6199

BACKGROUND

Many military applications for soil-structure interaction and soil mechanics involve large discontinuous deformations. Such problems include sinkage of vehicle tracks and tires, projectile penetration, ground shock, catering and explosions of shallow mines. The traditional use of continuum mechanics is limited to small strains with continuous deformation. Although discontinuities can exist within a continuum as a special condition, it is difficult, with a continuum description, to capture the spontaneous formation and destruction of slip and separation surfaces that are observed in particulate masses. The principal difficulty comes from the mathematical description of the kinematics that describe the movement of material "particles" within the continuum. In a continuum, movement must obey compatibility relationships that preclude formation of slip planes and separations. However, in real materials, large deformations can occur that violate compatibility in the strict sense imposed by continuum formulations. To model discontinuous deformations, asis done in finite element analysis of fracture propagation or at interfaces in soil-structure interaction problems, special joint elements are often introduced to model slip or separation.[1] Use of these elements complicates analysis of small deformation problems and becomes excessively complicated for large deformation problems. More importantly, the joint elements require *a priori* assumptions to be made on the location and propagation of the discontinuities.

An alternative to the continuum description for soil and rock mechanics problems is the Discrete Element Method (DEM)[2] which models the material as a collection of individual particles which interact only at inter-particle contact points. The DEM is a particular type of "particle model," a generic term for the class of simulation models where the discrete representation of a physical phenomenon involves the use of interacting particles. Particle models have successfully been used to treat a wide class of problems arising in plasma physics, solid state

physics, astrophysics, and fluid dynamics. In a particle model, each particle has a set of conserved attributes (e.g., mass), and momentum variable attributes (e.g., position, velocity) and some constitutive relationships describing the interactions among particles. The variable attributes evolve according to the equations of motion. The equations of motion are integrated to determine the velocity of the particles. The particle position are obtained from the velocity by integrating the equation $dx/dt = v$.

The highest resolution particle model is one where there is a one-to-one correspondence between the physical and computer particles. For example, in a molecular dynamic simulation each atom is represented by a computer particle and the particle contains the attributes of the atom (e.g., mass, charge, position). Rock and soil mechanics problems modeled using DEM have been one-to-one particle models.

Several authors have used DEM to model granular soil assemblies.[2,3,4] DEM has been used to investigate how mechanical stress waves propagate in granular material and how they are influenced by medium microstructure.[5] DEM has been used to model the results of tunnel failure resulting from a nuclear explosion.[6]

Sophisticated algorithms have been developed to describe the evolution of the particulate system including the formation and breaking of inter-element contacts.[7] Computing forces between elements requires relationships to describe normal and shear interaction at the contacts. In some models the individual particles can "break" when stress conditions within the particle reach some critical level.[8] Typically, soil particles have been modeled as two-dimensional circular rigid disks. Ting[9] has developed an ellipse-based two-dimensional particle to represent contact flatness and particle angularity. Six-sided solid shapes have been used to model granular material in three dimensions.[10]

The predominant disadvantage of DEM is its enormous computational requirement. At present, the maximum number of particles that can be feasibly handled in DEM computations is no more than a few tens of thousands of particles. This is primarily because of the large computational requirement of keeping track of all particle contact locations. The goal of the present research is to increase the feasibility limit by at least a factor of ten, a value that at best, permits study of particulate systems at laboratory scales. For example, a laboratory direct shear soil test using the smallest allowed soil specimen will contain over 200,000 particles of sand and triaxial test soil specimens can contain well over a million particles.

Use of one-to-one modeling of a particulate medium for a full scale problem is clearly beyond present computational capabilities. To model

full scale problems, a model is needed that somehow treats the aggregate behavior of particles yet retains the particle model's ability to depict large non-affin· motions. The present research fits into this more general goal in two important ways. First, whatever averaging scheme is ultimately developed, it is likely that the general integration scheme will remain unchanged. In the averaged scheme each particle will represent the aggregate behavior of many particles. The motion of the aggregated particles, however, would still be computed by an explicit integration scheme. The explicit scheme is necessitated by the changing geometry of the system. In addition, the number of aggregated particles needed to model real problems in three dimensions will undoubtedly be large. Thus, real problems will involve particle systems on the order of 200,000 to 1,000,000 particles.

The second need for a large particle model is to model laboratory aggregated on a one-to-one basis. The principal requirement of the averaged particle model is that it converges to the behavior of the one-to-one model as the analysis mesh is refined. Therefore, for the aggregated model to be useful, the one-to-one model must provide an accurate representation of the real particulate medium. To provide the link between the particle model representation and actual soil behavior, particle models must, as a minimum, be able to model simple experiments on soil. Thus, a critical element of the research is to improve computational techniques to increase the particle content to significantly above the current few tens of thousands particles.

MECHANICS FOR SOIL PARTICLE MODEL

Figure 1 illustrates the interaction between two particles. An interactive force, acting as a repulsive force between the two particles, is computed if the two particles are in contact. Two particles are considered to be in contact only if the distance, D, between their centers is less than the sum of their effective radii.

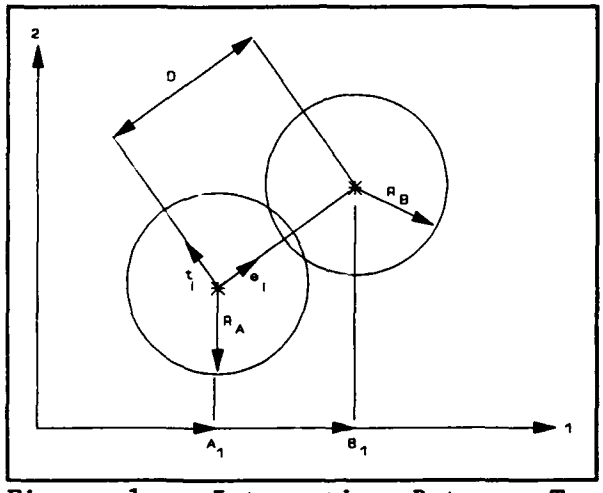


Figure 1. Interaction Between Two Particles

If two particles are in contact, the normal unit vector, e_i, and the shear unit vector, t_i, are determined. The normal unit vector points from the center of one particle to the center of the other particle. For the two-dimensional case, the shear unit vector is determine by rotating the normal unit vector 90 degrees. For the three-dimensional case, the shear unit vector lies in a plane perpendicular to the normal unit vector and its direction is determined by the relative velocity of the two particles in contact.

The normal force between the two particles, which is a repulsive force aligned with the normal unit vector, is computed as:

$$f_{(n)}^{AB} = \Psi K_{(n)eff}^{AB} \tag{1}$$

Where

$K_{(n)eff}^{AB}$ is the normal effective stiffness constant for the two particles and Ψ is a penalty function design to prevent particle penetrations.

To determine the shear force component of the particle-particle interaction, the tangential relative velocity is computed:

$$V^{AB}_{(t)\,1} = (V_A - V_B)\, t_1 \tag{4}$$

The tangential relative velocity is integrated with respect to time to provide the relative tangential displacement for the time step:

$$\Delta \delta^{AB}_t = V^{AB}_{(t)} \Delta t \tag{5}$$

The shear force increment for the time step is then computed:

$$\Delta f^{AB}_{(s)} = K^{AB}_{(t)\,eff} \Delta \delta^{AB}_t \tag{6}$$

Where $K^{AB}_{(t)\,eff}$ is the tangential effective stiffness constant.

While the two particles remain in contact, the shear force increment is added to the total shear force from the previous time step, i.e.,

$$f^{AB}_{(s)\,N} = f^{AB}_{(s)\,N-1} + \Delta f^{AB}_{(s)} \tag{8}$$

Where the indices N and N-1 refer to times t_N and t_{N-1}, and $\Delta t = t_N - t_{N-1}$.

The magnitude of the shear force is also compared to a maximum shear force that is based on a Coulomb-type friction law. If the shear force exceeds that maximum shear force, the shear force is set equal to the maximum shear force:

$$f^{AB}_{(s)} \leq f_{(s)\,max} = C^{AB}_{eff} + f^{AB}_{(n)} \tan\phi^{AB}_{eff} \tag{9}$$

An artificial viscous damping force is also introduced to the problem. The use of damping other than by friction is necessary so that the assemblies reach a state of equilibrium for all conditions.[2] A viscous damping force is computed for each particle. Its magnitude and direction are based on the current absolute velocity of the particle.

Once forces have been determined for each particle, the instantaneous acceleration of each particle over the current time step is determined as the resultant force acting on the particle divided by the particle mass. The updated velocity and location of each particle are then determined by integrating the equations of motion.

It should be noted that two types of forces are involved in the simulation. The normal and viscous forces depend only on the current positions and relative velocities; these forces can be computed from the

current state of the assemblage. In contrast, the shear forces depend on the relative movement of the contact surfaces which requires knowledge of previous particle positions. The shear forces, therefore, depend on the history of particle motions. Much of the literature on very large particle systems (e.g. Hockney and Eastwood, 1981) involves the computation of inter-particle forces from potentials and thus the forces are not history dependant. An issue in the present work is the extent that history dependence might influence the optimization efforts.

OPTIMIZATION

The goal of the optimization effort was to significantly increase the number of particles that could be simulated. For the original unoptimized code, runtime was the controlling factor. A four-thousand-particle simulation required over two hours of runtime for one thousand integration steps. However, as capability increased, computer memory also became a consideration.

The original unoptimized code was an implementation of the general algorithm as given by Hockney and Eastwood.[11] The algorithm is essentially equivalent to a short-range, variable-neighbor molecular dynamics algorithm. This algorithm consists of three main processes: (1) find particles in contact, (2) calculate forces between particles, and (3) integrate the equations of motion (move particles). The only distinguishing feature is the introduction of shear forces, requiring an additional element dealing with previous contacts. This additional element forced some variations to the standard molecular dynamics algorithms.

The first step in the optimization process was to restructure the existing code to improve machine performance. The original code performed at 5.5 MFLOPS (million floating point operations per second) on the CRAY Y-MP, which has a peak speed of 333 MFLOPS per processor. By improving vectorization, use of functional units, memory traffic, and input/output, performance was increased to over 180 MFLOPS, depending on the number of particles. Overall, runtime was reduced by an order of magnitude. In practice, the maximum time reduction obtainable by vectorization is a factor of around ten on the CRAY Y-MP. Thus further improvement from restructuring is not likely.

Although performance increases due to restructuring were significant, further improvements were required to meet project goals. This forced algorithm evaluation. The most time-consuming process was finding the particles in contact. The original code used the simplest of methods: computing distances between all particles. This method, though easily implemented, scales computationally as $O(N^2)$ for a N-particle problem

(i.e., the number of computational operations to perform a simulation is proportional to the square of the number of particles). This is inefficient and impractical for large scale problems containing only localized forces. A more appropriate technique is to use a chaining mesh (Hockney and Eastwood, 1981) method to produce a "neighbors-list" for each particle. This method breaks the physical space of the simulation down into a regular grid of cells. Cell size is set equal to the maximum particle diameter. The list of neighbors is thus restricted to a particle's own cell and the 26 surrounding cells. An implementation using the nearest neighbor list scales as O(N).

If previous contact were not involved, the general practice would be to sequence through the cells to evaluate a particle for particle-particle interaction with the particles contained in its own cell and the lower 13 adjacent cells. If all 26 surrounding cells were evaluated, it would duplicate calculations for inverse reactions. However, by adding a particle to its own cell and all of its 26 surrounding cells, and then limiting interactions to those particles put into a cell after the current particle, the sequence of interactions is maintained. This cut the search space for previous contact and previous contact storage in half.

Figure 2 shows relative timings for the original code, the restructured code, and the neighbors list code for a 1000 time step simulation of a free fall test. The free fall test consisted of allowing the particles to fall from their initial position onto a rigid floor. The restructured code was clearly an improvement over the original code but, as noted above, retained the characteristic $O(N^2)$ increase in computations as the number of particles increased. The algorithm based on nearest-neighbor lists greatly reduced the CPU time and, more important to achieving progress, scaled computationally as O(N).

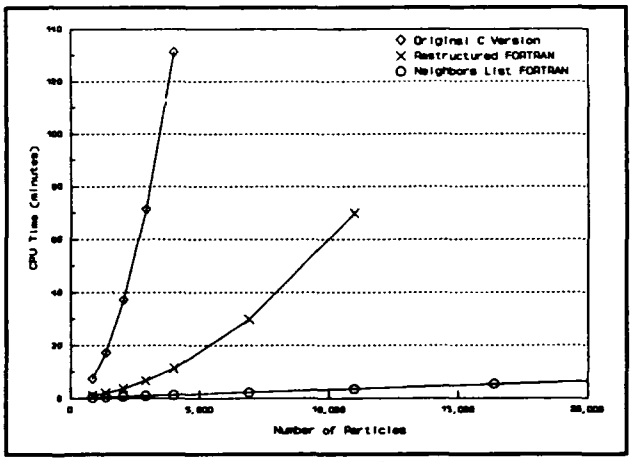


Figure 2. CPU Timings for Three Stages of Code

Figure 3 shows variations in runtimes for the free fall test and confining test simulations for the neighbors-list code. The confining test consisted of placing the particles into a rigid box and compressing the particles from the top with a rigid plate. The timings for the confined test were greater than those of the free fall test for the same number of

particles because the average number of particle contacts was larger. The timings for both test conditions remained linearly related to the number of particles.

Simulations were performed for 10 time steps using 600,000 particles for the free fall problem and using 219,488 particles for the confined problem. The results are tabulated below:

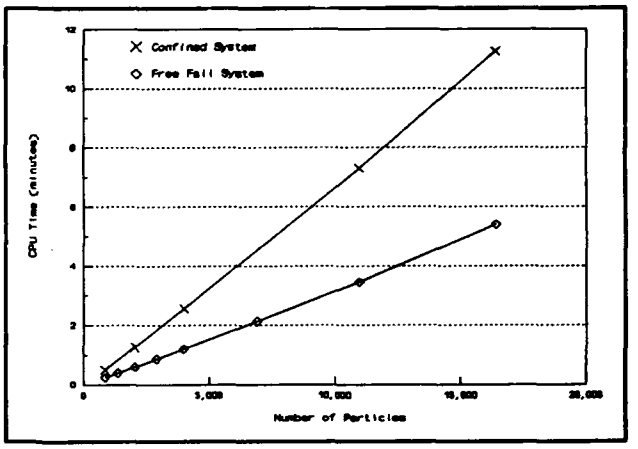


Figure 3. Variations in Runtimes

N	Problem	CPU time (min.)	Memory (MW)
600,000	Free Fall	90.5 (905.)*	108.64
219,488	Confined	2.6 (263.)*	42.23

* Estimate for 1000 time steps

The estimated times to execute 1000 time steps are 905 min. for the free fall problem and 263 min. for the confined loading problem. The corresponding times extrapolated from Figure 3 are 200 min. for the free fall problem and 150 min. for the confined loading problem. The linearly-extrapolated times thus underestimate those estimated from actual computations by factors of 4.5 for the free fall problem and 1.75 for the confined loading problem. Note that the estimates for the free fall problem based on timings for the early time steps may be too large because the execution cycles become shorter as the particle mass expands and contacts become fewer. In any case, the increase in execution times scale as O(N) as N is increased to the target range of N = 1,000,000.

The principal restriction to performing practical simulations for large N is memory capacity. The memory required to perform the free fall simulation with 600,000 particles constitutes nearly the entire capacity of the CRAY Y-MP at the U.S. Army Engineer Waterways Experiment Station (WES) (with five of the six processors idle). Considering the run times involved, it would not be economically feasible to capture such a large percentage of the computer's capacity for the extended run times required. Thus, to make numerical simulations practical for large N, it will be

necessary to develop "out-of-core" computational algorithms and employ multitasking to make full use of all six processors.

CONCLUSIONS

Overall, dropping the scaling of the problem from $O(N^2)$ to $O(N)$, combined with the restructuring for machine performance, greatly enhanced the capability of the model. Simulation of soil tests that contain the number of particles found in laboratory soil tests (200,000 particles) can now be performed. In-core solutions can reach magnitudes of as many as 600,000 particles. For future consideration, running 600,000 particles on the CRAY Y-MP at WES in core uses almost all of the memory, leaving five processors idle. For problems of this magnitude, the use of multitasking to make use of these processors is of notable concern. Other considerations are the development of an out-of-core method for simulations greater that 600,000 particles and the use of massively scalable machines for even greater performance.

ACKNOWLEDGMENT

The tests described and the resulting data presented herein, unless otherwise noted were obtained from research conducted under the MILITARY RESEARCH DEVELOPMENT TEST AND EVALUATION PROGRAM of the United States Army Corps of Engineers by the U.S. Army Engineer Waterways Experiment Station and the Army's High Performance Computing Research Center (AHPCRC). AHPCRC is operated by the Unversity of Minnesota and funded by the Army Research Office, Contract Number DAAL03-89-C-0038. Permission was granted by the Chief of Engineers to publish this information.

REFERENCES

1. R. N. Yong and A. W. Hanna, *Journal of Terramechanics*, Vol. 14, No. 3, (1977).

2. P. A. Cundall and O. D. L. Strack, *Geotechnique*, Vol. 29, No. 1, (1979).

3. J. Christoffersen, M. Mehrabadi, and S. Nemai-Nasser, *Journal of Applied Mechanics*, Vol. 48, (1981).

4. R. J. Bathurst and L. Rothenburg, *Journal of Applied Mechanics*, Vol. 55, (1988).

5. A. Shukla and M. Sadd, *Wave Propagation and Dynamic Load Transfer due to Explosive Loading in Heterogenous Granular Media with Microstructure*, (U.S. Air Force Office of Scientific Research Under Contract No. F49620-89-C-0091, Bolling Air Force Base, 1990).

6. F. Heuze, O. Walton, D. Maddix, R. Schaffer, and T. Butkovich, *Proceedings of Mechanics Computing in 1990's and Beyond*, American Society of Civil Engineers, Columbus, Ohio, May 20-22, Vol. 2, (1991).

7. J. M. Ting, B. T. Corkum, C. R. Kauffman, and C. Greco, *Journal of Geotechnical Engineering*, Vol. 115, No. 3, (1989).

8. P. A. Cundall and R. D. Hart, *Development of Generalized 2-D and 3-D Distinct Element Programs for Modeling Jointed Rock*, Miscellaneous Paper SL-85-1, U.S Army Engineer Waterways Experiment Station, (1985).

9. J. M. Ting, *Proceedings of Mechanics Computing in 1990's and Beyond*, American Society of Civil Engineers, Columbus, Ohio, May 20-22, Vol. 2, (1991).

10. J. Ghaboussi and R. Barbosa, *International Journal for Numerical and Analytical Methods in Geomechanics*, Vol. 14, (1990).

11. R. W. Hockney and J. W. East, *Computer Simulations Using Particles*," (MacGraw-Hill, New York 1981).

Host Respiratory Protection For Lethal Staphylococcal Enterotoxin B

Robert E. Hunt*, LTC, VC
Anthony J. Johnson, COL, VC
Marti Jett, Ph.D.
Jeenan Tseng, Ph.D.
Jack L. Komisar, Ph.D.
Catherine L. Wilhelmsen, LTC, VC
Walter Reed Army Institute of Research
Washington, DC 20307-5100

Louise Pitt, Ph.D.
David Ruble, CPT, VC
United States Army Medical Research Institute of Infectious Diseases
Fort Detrick, Frederick, Maryland 21702-5012

Army scientists during the nineteen sixties prepared, purified, assayed and determined the effects of Staphylococcal Enterotoxin B (SEB) in primates[1-3]. A significant observation was a pattern of acute intoxication that progressed to a severe shock-like syndrome with high mortality following the administration of microgram (μg) quantities of purified SEB. Because of the toxin's overall biological and physical properties, attention was directed to its potential usefulness as a biological warfare agent in overt, covert, terrorist, and conventional warfare activities. It was the threat agent potential that led the U.S. Army Medical Research and Development Command to conduct biomedical research for development of defenses against the toxin's incapacitating and lethal effects.

Staphylococcal enterotoxin B (SEB), along with six structurally and biologically related serotyped classes of enterotoxins (SEA, SEB, SEC1,

SEC2, SEC3, SED, and SEE) have been major causes of food poisoning worldwide[4,5]. These enterotoxins are extracellular proteins produced by the ubiquitous bacterium *Staphylococcus aureus*. It was the incidence and misery of staphylococcal food poisoning that led scientists to develop vaccine candidates and attempt to immunize humans as a means of reducing the incidence of staphylococcal enterotoxin food-borne illness. Early attempts were hampered by unpurified enterotoxin and questionable serological assays[6]. It was only after the enterotoxins were properly identified serologically and purified as distinct molecules that specific immunization regimens could be addressed. Bergdoll, as early as 1966, investigated the immunological properties of purified SEB in rhesus monkeys after toxoiding the molecule with formalin[7]. After alum-adsorption of the formalinized SEB toxoid, he was able to induce antibody formation in rabbits and increase the resistance of rhesus monkeys to oral SEB challenge about 100 fold. The results of the immunization experiment indicated that 3-intramuscular injections of 62.5 μg for the initial and first booster doses of SEB toxoid and 250 μg for the second booster, all at five-week intervals, were needed for this protection. His results also indicated that the resistance to SEB was due, at least in part, to the presence of anti-toxin antibodies in the plasma of resistant animals. His research and that of others have contributed valuable information on immunoprophylaxis through vaccination against staphylococcal enterotoxins and toward development of an effective toxoid vaccine against SEB. Yet, there remains no effective vaccine that fully protects against these staphylococcal enterotoxins.

Beginning in February 1991, scientists from the Division of Pathology, Walter Reed Army Institute of Research and the U.S. Army Medical Research Institute of Infectious Diseases have been testing the degree of protection afforded rhesus monkeys immunized with microspheres of a biocompatible, biodegradable poly (DL-lactide-co-glycolide) polymer (DL-PLG) that encapsulates the SEB toxoid. The decision to encapsulate the SEB toxoid was made following failures of soluble and alum-adjuvanted SEB toxoids to provide appropriate protection against SEB respiratory exposure in rhesus monkeys[8,9].

Severe immediate hypersensitivity was an additional adverse response observed when immunized and SEB exposed monkeys were subsequently challenged by the respiratory route. Major factors for microsphere encapsulation of the SEB toxoid were to develop an oral vaccine against SEB that especially stimulated the production of secretory immunoglobulin A (sIgA) and to provide protection in immunized rhesus monkeys against SEB aerosol challenge. Additionally, it was theorized that effective production of anti-SEB sIgA antibodies would suppress the immediate hypersensitivity response encountered following immunization with soluble or alum adsorbed SEB toxoid.

The microsphere formulation, SEB toxoid encapsulation, and biologic testing in mice had been previously accomplished through contract arrangements with a commercial firm[10]. After optimizing the formulation of microspheres composed of DL-PLG and SEB toxoid, immunization studies were conducted in mice[11]. It was shown that (a) three oral immunizations, (b) an initial parenteral immunization followed by an oral, intramuscular or intratracheal booster immunization, or (c) an initial and booster immunizations administered intratracheally provided antibody responses in the plasma and bronchoalveolar lavage (BAL) fluid. Also, a strong IgA antibody response was induced in the bronchoalveolar fluid following the intratracheal administrations. This led to questions and the decision to immunize rhesus monkeys by various routes to determine the most effective method to induce protection against lethal SEB aerosol exposure. It is the unique results of these early SEB protection studies accomplished in rhesus monkeys that form the basis for this report.

The study consisted of 40 young, healthy rhesus monkeys; 4 controls and 36 immunized. Nine different immunization regimens employing three different routes of immunization were accomplished (table 1).

TABLE 1. IMMUNIZATION REGIMENS

OR/OR	OR/IM	OR/IT
IM/OR	IM/IM	IM/IT
IT/OR	IT/IM	IT/IT

Primary/Secondary

OR=Oral; IM=Intramuscular; IT=Intratracheal

The routes of immunization were oral (OR), intramuscular (IM), and intratracheal (IT). Each immunization regimen consisted of a primary immunization (OR, IM, or IT) followed 49 days later by a secondary booster immunization (OR, IM, or IT). The microsphere vaccine was administered at a dose of 100 μg in 0.5 milliliters (ml) volume for the intramuscular and intratracheal immunizations and 1.0 milligram (mg) in 0.5 ml volume for the oral immunizations. Operational and resource constraints allowed only ten monkeys, one from each immunization regimen and one control, to be studied as a group.

The aerosol SEB challenge was performed 80 days after the booster immunization. A standard lethal SEB challenge dose of 15 LD_{50} was scheduled for all rhesus monkeys; however, the actual challenge doses varied among study groups. The SEB challenge dose received by each rhesus is listed (tables 2 and 3). The actual exposure doses were determined from samples taken from the exposure chamber during challenge.

To gain more information on the role of anti-SEB antibodies in protection and to obtain additional data on the reported incidence of immediate hypersensitivity, the immunoglobulin (Ig) isotype IgM, IgG, and IgA titers were measured in the plasma and bronchoalveolar lavage fluids (BAL) of all monkeys using an ELISA test. The isotype titers for IgG and IgA are listed for each monkey (tables 2 and 3). It was determined that IgG and IgA in the plasma and bronchoalveolar lavage

fluid contributed to protection as the majority of the survivors had high levels of both isotype antibodies. The isotype IgM response was inconsistent and had no noted relationship to protection.

All four rhesus monkeys (100%) administered the IM/IT immunization regimen were protected against the aerosol lethal dose challenge. In addition, two monkeys in the IT/IT and two in the OR/IT immunization regimens were protected. In contrast, no other immunization regimen protected more than one monkey and there were no protected monkeys in the OR/OR, OR/IM, and IM/OR regimens. These data emphasize the importance of the intratracheal route as a booster immunization for protection against aerosol exposure.

There are several implications in this study that are relevant to protection against biological threat agents delivered as an aerosol. First, since the toxin was directed toward the respiratory tract, it was hypothesized that protection would correlate with the level of local IgA anti-SEB antibodies in bronchial fluid. As shown in the data above, this was indeed the case. Monkeys with the highest IgA and IgG levels in BAL were protected whereas those with the lowest levels were more susceptible to the lethal challenge. Second, it was hypothesized that the route of immunization would influence the level of respiratory antibodies and hence be a prime determinant of protection. Indeed, the question of which route of immunization best induces secretory immunity was central to this study. While researchers have long suspected the importance of local immunization, there is a notable paucity of protection data comparing various immunization regimens, especially in primates. For example, based on what is known about lymphocyte trafficking following oral immunization, it may have been thought that oral immunization would be superior to respiratory (and certainly IM) immunization regimens. On the other hand, if systemic immunity alone were important for protection against the lethal event induced by this toxin, IM/IM immunization regimen may have been the most effective regimen. The success of intratracheal immunization, especially as a booster in the IM/IT regimen, in generating protection against the highest lethal toxin challenges attempted in this study

demonstrates that immunization via the respiratory tract may be the most promising regimen to induce antibody-mediated immunity against aerosolized biologic warfare agents such as staphylococcal enterotoxins.

In summary, the lives of rhesus monkeys were completely protected from a high dose (15 LD50) SEB aerosol lethal challenge by an initial intramuscular injection of 100 micrograms of formalinized SEB toxoid encapsulated in biodegradable and biocompatible DL-PLG copolymer microspheres and a 49th day booster 100 microgram dose administered intratracheally. No evidence of hypersensitivity was noted following the aerosol challenge. Additionally, the data indicates that protection against SEB is provided, at least partially, by the production of IgG and IgA anti-SEB isotype antibodies in the plasma and bronchoalveolar lavage fluid. The immunization regimen consisting of an initial intramuscular injection followed by a booster vaccine dose administered intratracheally, the latter being the mucosal site of SEB challenge, proved to be an important factor in protection.

Future endeavors will include aerosol rechallenge of the surviving monkeys to evaluate the duration of protection and to further assess the potential for immediate hypersensitivity reactions to occur in immunized monkeys. Additionally, following the technical development and evaluation of DL-PLG microsphere SEB toxoid booster immunization by the intranasal (IN) route in preliminary mouse studies, a scientific study is planned to aerosol challenge and assess a group of monkeys immunized with an IM/IN immunization regimen. Complete characterization of respiratory route immunization to include safety test analysis of empty DL-PLG microspheres administered to the lungs of monkeys is planned. Formulations of DL-PLG copolymers constructed as resorbable sutures, resorbable suture clips, and as microsphere control release implants are approved for use in humans and currently have a safe history[11]. Continued success with novel selective immunization of the respiratory tract against lethal toxins and potentially other agents can have a major positive impact in protecting the health and fighting strength of soldiers.

Table 2. **Protection: Monkeys Immunized with Microspheres**

DAY 98 PREBLEED/PREBAL ANTIBODY TITERS (SRI)
End Point Titration-ELISA

SURVIVORS MONKEY	IMMU*	IgG PLASMA	IgG BAL	IgA PLASMA	IgA BAL	Chall LD50
89B161	IM/IT	800000	320	45	40	13.8
90B003	IM/IT	1000000	1280	1800	320	18.1
90B031	IM/IT	500000	160	250	20	7.0
90B078	IM/IT	1800000	1280	1000	320	22.2
90B009	IT/IT	1000000	320	380	160	17.6
90B055	IT/IT	900000	2560	1800	1280	7.4
89B064	OR/IT	200000	80	25	20	12.9
90B066	OR/IT	5000	<10	60	20	7.8
90B025	IT/IM	20000	10	25	20	7.1
90B030	IM/IM	1800000	80	280	<10	6.7
90B062**	OR/IM	900000	640	30000	<10	18.4
90B067	IT/OR	80000	40	400	20	7.4

Primary/Booster immunization regimen
OR=Oral; IM=Intramuscular; IT=Intratracheal
IMMU*=Immunization regimen; BAL=Bronchoalveolar lavage;
Chal=Challenge
** Deleted from study due to illness and treatment prior to challenge.

Table 3. **Protection: Monkeys Immunized with Microspheres**

DAY 98 PREBLEED/PREBAL ANTIBODY TITERS (SRI)
End Point Titration-ELISA

NON-SURVIVORS MONKEY/IMMU*		IgG PLASMA	IgG BAL	IgA PLASMA	IgA BAL	CHAL LD50
89B088	CONTROL	25	<10	25	<10	13.0
90B001	CONTROL	50	<10	50	<10	18.1
90B029	CONTROL	25	<10	25	<10	6.8
90B050	CONTROL	25	<10	25	<10	18.4
89B001	IT/IT	189000	<10	25	<10	13.1
90B075	IT/IT	700000	1280	400	640	21.8
90B008	OR/IT	28000	<10	50	<10	17.1
90B098	OR/IT	20000	<10	25	<10	21.9
89B116	IT/IM	200000	20	25	<10	13.8
90B02'	IT/IM	22000	320	50	<10	18.2
90B103	IT/IM	40000	40	25	<10	20.9
89B210	IM/IM	850000	40	25	<10	13.2
90B010	IM/IM	450000	<10	50	<10	18.2
90B060	IM/IM	20000	<10	25	<10	18.6
89B034	OR/IM	100000	<10	25	<10	13.3
90B018	OR/IM	28000	<10	50	<10	18.1
90B043	OR/IM	10000	10	25	<10	7.6
89B060	OR/OR	25	<10	25	<10	13.1
90B014	OR/OR	875	<10	50	<10	18.5
90B027	OR/OR	2000	<10	170	<10	7.3
90B079	OR/OR	1000	<10	700	<10	22.4
89B096	IM/OR	30000	<10	50	<10	13.0
90B002	IM/OR	13000	<10	50	20	17.4
90B016	IM/OR	40000	<10	25	<10	7.2
90B110	IM/OR	20000	<10	25	<10	21.
89B079	IT/OR	25	<10	25	<10	13.2
90B023	IT/OR	110000	20	50	<10	18.5
90B092	IT/OR	9000	20	25	<10	18.7

Primary/Booster immunization regimen
OR=Oral; IM=Intramuscular; IT=Intratracheal
IMMU*=Immunization regimen; BAL=Bronchoalveolar lavage;
Chal=Challenge

References

1. W.R. Beisel, *Toxicon* **10**, 433-440 (1972).
2. G.J. Crawley, I Gray, W.A LeBlanc, and J.W. Blanchard, *J. Infect. Diseases* **116**, 48-56 (1966).
3. L.F. Hodoval, E.L. Morris, G.J. Crawley, and W.R. Beisel, *Appl. Microbiol.* **15**, 187-192 (1968).
4. M.S. Berdgoll, Staphylococcal Intoxications, in *Foodborne Infections and Intoxications*, H. Riemann and F.L. Bryan, Eds. (New York: Academic Press, Inc; 1979), pp. 443-
5. M.S. Bergdoll, Enterotoxins, in *Microbial Toxins, III, Bacterial Protein Toxins*, T.C. Montie, S. Kadis, and S.J. Ajl, Eds. (New York: Academic Press, Inc; 1970), pp. 265-326.
6. G.M. Dack, E.O. Jordan, and O. Woolpert, *J. Preventive Med.* **5**, 151-159 (1931).
7. M.S. Bergdoll, *Infect. & Immun.* **116**, 191-196, (1966).
8. US Army Medical Unit. Annual Progress Report, Fort Detrick, MD pp. 101-102 and 243-250 (1968).
9. US Army Medical Unit. Annual Progress Report, Fort Detrick, MD pp 227-233 (1969).
10. Southern Research Institute, Annual Report, Project 6160, Contract DAMD17-C-6162, 1991.
11. J.H. Eldridge, J.K. Staas, J.A. Muelbroek, T.R. Tice, and R.M. Gilley, *Infect. & Immun.* **59** (9), 2978-2986 (1992).

Modeling Human Performance to Predict Unit Effectiveness

Steven R. Hursh*, LTC(P)
Walter Reed Army Institute of Research
Washington, D.C. 20307-5100

Richard E. McNally
Science Applications International Corporation
626 Town Center Drive, Suite 205
Joppa, MD 21085

A variety of Army standard models are used to analyze and predict the impact of new equipment and operational doctrine on the combat effectiveness of military units fighting hypothetical battles. Absent from most models is any consideration of human performance in modulating the effectiveness of units. In 1989, the Medical Research and Development Command initiated a program to incorporate into an Army standard model consideration of medical products, both materiel and concepts, that improve the health and performance of soldiers. The purpose was two fold: first, accurately portray the impact of soldier health and performance on unit effectiveness and, second, document the extent to which medical products enhance the ability of military units to carry out their warfighting mission.

The Medical Research and Development Command, through cooperation with the Ballistic Research Laboratory, selected the Army Unit Resiliency Analysis (AURA) Model as the first test bed for these concepts. The model was improved to incorporate the latest medical data on soldier health and performance across a range of variables: infectious disease, biological and chemical weapons effects, heat stress, and sleep deprivation. Medical interventions, both current and future, were modeled for their ability to attenuate these threats and enhance soldier performance and unit effectiveness.

In this paper the impact of products that have evolved from the neurosciences will be discussed. First, modeling has been applied to estimate the battlefield toxicity of potential nerve agent attacks and to predict the likelihood and severity of casualties. A variety of medical protective measures have been assessed and shown to improve the resiliency of typical combat units. Second, modeling has been applied to estimate the degradation of individual and unit performance under conditions of continuous operations that permit only limited amounts of sleep. A series of sleep disciplines were modeled and the net benefits of adequate sleep for overall unit productivity were demonstrated.

OVERVIEW OF MODELING STRATEGY

The purpose of medical modeling is to demonstrate the impact of medical research and development products, both concepts and materiel, on the Army's warfighting capabilities using accepted models of battlefield threats and combat effectiveness. There are four goals to this program:

1) demonstrate the impact of medical products for improved unit effectiveness, 2) assist the transfer of medical technology and information for operational planning, 3) generate timely information to medical research directors and developers as decision aid for planning and milestone decisions, and 4) provide feedback to researchers concerning data gaps and model refinement.

To meet these goals, a common modeling strategy has been followed and is diagramed in Figure 1. First, a specific battlefield threat forms the focus of each modeling task. These threats encompass a broad range of medical hazards to the soldier such as biological and chemical weapons, infectious disease, stress and fatigue, environmental extremes (heat, dehydration, etc.), and traumatic or accidental injury. The threat is then specified within the context of a battlefield scenario that includes the geographic location, meterological conditions, weapons characteristics, combat units and tasks, uniform and protective devices, and operational doctrine.

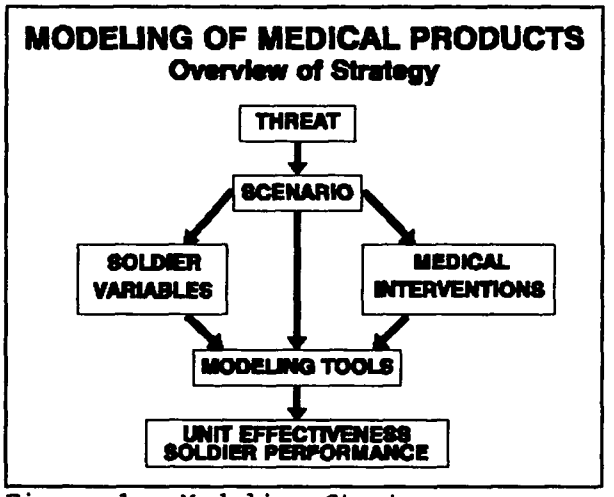


Figure 1. Modeling Strategy

For each threat and scenario, appropriate soldier variables must be modeled that represent the likely impact of the threat in that scenario. These human factors include death or incapacitation, rate of return to duty after injury, and performance degradation. Degradation factors may be subdivided further into cognitive factors, physical work capacity, physiological limits such as water requirements and thermal load, uniform restrictions and encapsulation, and individual variability. The model of unit effectivenss must have the capacity to simulate the impact of all these factors, individually and in combination. A the same time, medical interventions are specified that can reduce degradation of some or all of these performance factors. The medical interventions may include the following: preventative drugs or vaccines, physical conditioning, therapy drugs, physical barriers, and operational doctrine such as sleep discipline, work schedules, acclimatization, and water discipline. The next step is to select a modeling system that can portray the impact of each medical intervention on these individual soldier variables and then simulate the effects of those soldier variables on unit effectiveness. The modeling system may be a single simulation program or a series of modules that compute different aspects of the problem and feed into a final estimate of unit effectiveness. Ultimately, the final product is a representation of the impact of the medical interventions on casualties, individual performance and unit effectiveness.

To date a central modeling system for estimating unit effectiveness has been the *Army Unit Resiliency Analysis (AURA)* developed by the Ballistic Research Laboratory (BRL). This model was selected because it contained rudimentary algorithms to represent many soldier performance factors, was an accepted Army modeling system, contained representations for a variety of military units that were coordinated with the appropriate

schools, and was open to further improvement by BRL to incorporate the latest medical data and performance representations. The current version of AURA represents the cooperative activities of the proponent laboratory, the medical laboratories of the Medical Research and Development Command, and the primary contractor, Science Applications International.

Other modeling tools have been used as required. The first case to be discussed below illustrates the use of the Modular Chemical Assessment Structure (MCAS) for estimating chemical and conventional casualties as an input factor to AURA. Other modeling tools have been used to represent biological weapons effects, heat strain, and non-auditory blast injury.

MODELING OF MEDICAL THERAPY OPTIONS FOR NERVE AGENT DEFENSE

Introduction: The 1980s were characterized by a reexamination and reemphasis on the development of measures to increase the ability of U.S. forces to survive and fight on a chemically contaminated battlefield. Ironically, although the Warsaw Pact is now dissolved and the likelihood of facing a Soviet adversary on the battlefield has evaporated, the probability of U.S. forces facing a CW threat nonetheless has increased. During the recent Persian Gulf conflict (Operation Desert Shield/Storm), the risk that U.S. forces would be subjected to chemical attacks was greater than at any time since World War I.

The chemical defense research and development programs of the 1970s and 1980s have yielded many improved capabilities for detection of chemical agent and protection against and treatment of injuries caused by CW agents--and improvements are in progress. In addition, the medical community has made substantial contributions to these efforts--and is also on the brink of significant improvements. In this section, some of the improved products of medical chemical defense research will be examined. During the past two decades, we acquired the capability to treat nerve agent casualties with oximes, such as pralidoxime chloride (2-PAM). In the 1980s, preexposure drugs for nerve agents were fielded, and several drugs that showed great promise are now nearing transition from the laboratory bench and animal studies to development programs that will place them in the hands of our fighting men and women and their medical providers. Enhanced treatments, including the use of anticonvulsant drugs, reduce damage to the nervous system caused by nerve agents, thereby shortening recovery time. Developments this decade may include improved preexposure compounds, and future improvements might be nerve agent antibodies--a product of biotechnology--that will increase the levels of nerve agent that soldiers can tolerate.

Objective: This paper presents an abbreviated discussion of the comparative effectiveness of current, enhanced, near-term, and far-term therapies for nerve agent intoxication. In addition, a "no therapy" approach is presented for comparison. A full report is available.[1]

Approach: The effectiveness of several nerve agent treatment alternatives are addressed. We evaluate the effectiveness of these treatment alternatives with two levels of physical protection from agent effects: no protection (MOPP 0) and full mission-oriented protective posture (MOPP IV).

Figure 2 illustrates the general approach used for these studies. First, CW attack profiles were established (weather, terrain, type of weapon, agent deployed, etc.). Then, the MCAS (Modular Chemical Assessment Structure) model was used to develop agent dosage and deposition, chemical

casualty data, and return-to-duty distributions. The principal MCAS outputs are numbers of casualties resulting from the chemical attack, given the protection and treatment options applied, and profiles for return of casualties to duty. These data are input into AURA (Army Unit Resiliency Analysis) to determine unit effectiveness over time. AURA also accounts for decrements in effectiveness resulting from the burdens of protection.

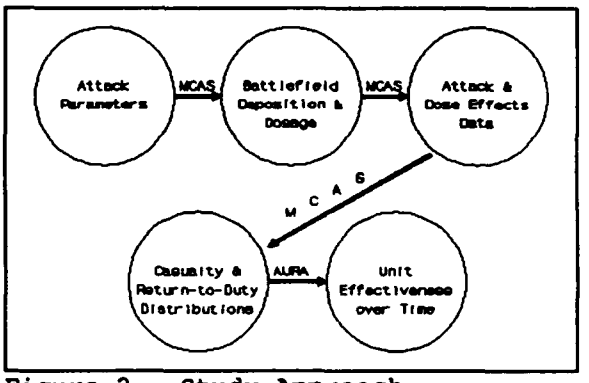


Figure 2. Study Approach

Chemical attack parameters were based on methods developed for the U.S. Army Medical Research Institute of Chemical Defense (USAMRICD) study "Battlefield Levels of Soman 1989" and have been used to predict threat agent dose data for the current study.[2] The former study developed munition expenditure requirements, resulting agent concentration and dose to protected and unprotected personnel, and casualty data for artillery and air-delivered attacks under both summer and winter conditions. The munitions used were 152-mm artillery projectiles with unthickened soman and bulk fill bombs with thickened soman.

Two hundred fifty-six rounds of artillery, fired in one volley, were required to achieve the effects studied. Forty bombs, delivered simultaneously from four aircraft, were required. We studied each fire plan in a Central Europe summer day, a Central Europe winter dawn, and a Southwest Asia summer day environment, against both an air-delivered bulk chemical bomb and an artillery-delivered chemical shell scenario. The bulk bomb provided multiple line sources over the target. The artillery scenario provided multiple point sources over the target. As an example, the deposition pattern for the artillery attack on a SWA summer day is shown in Figure 3.

Therapies Considered: Nerve agents are a group of structurally similar organophosphorus compounds with differing inorganic fragments (cyanide, fluorine, or sulfur) and organic fragments. They share a common phosphorus-oxygen structure. They exhibit the common physiological effect of interfering with the enzyme acetylcholinesterase (AChE); consequently, the balance of the neurotransmitter acetylcholine (ACh) is drastically altered. ACh is an essential neurotransmitter for the parasympathetic nervous system, involved in breathing, perspiration, and other essential functions, and necessary for the innervation of skeletal muscles, as well as for other central nervous system functions. ACh is released by nerve endings to carry the nerve impulse to the end organs. AChE neutralizes the released ACh after the firing of the nerve impulse. When nerve agents block the action of AChE, the resulting excess of ACh causes an uncontrolled overstimulation of the skeletal muscles, parasympathetic end organs, and other central nervous system end effectors. This, in turn, causes the symptoms that are characteristic of nerve agent poisoning (pinpointing pupils of the eye and resulting dimness in vision, difficulty breathing, drooling, excessive sweating, nausea, vomiting, cramps, involuntary defecation and urination, twitching, jerking, staggering, headache, confusion, drowsiness, coma, convulsions, and death).[3]

The therapies addressed are listed in Table 1. The Mark I Nerve Agent Antidote Kit administers 600 mg of the oxime *N*-methylpyridinium -2-aldoxime chloride (2-PAM) and 2 mg of atropine sulfate. Every soldier carries three Mark 1 kits for self- or buddy aid. *Atropine* is one in a class of drugs which block the action of ACh, thereby mitigating the toxic effects at the neurons that results from nerve agent poisoning. *2-PAM* is one of several oximes widely believed to reactivate AChE bond by nerve agents. *Pyridostigmine pretreatment* is a compound that reacts with AChE, thereby protecting the enzyme from destruction by nerve agents.[4] *Anticonvulsant therapy* counters convulsions which are a major effect of nerve agent poisoning which can lead to brain damage and probably contributes to lethality.[5] For purposes of this study, diazepam was the anticonvulsant drug modeled. *Nerve agent antibodies* are proposed as a future protection. It seems likely that molecular biology will provide an effective nerve agent pretreatment that will reduce agent toxicity in man manyfold--probably by early next century.[6]

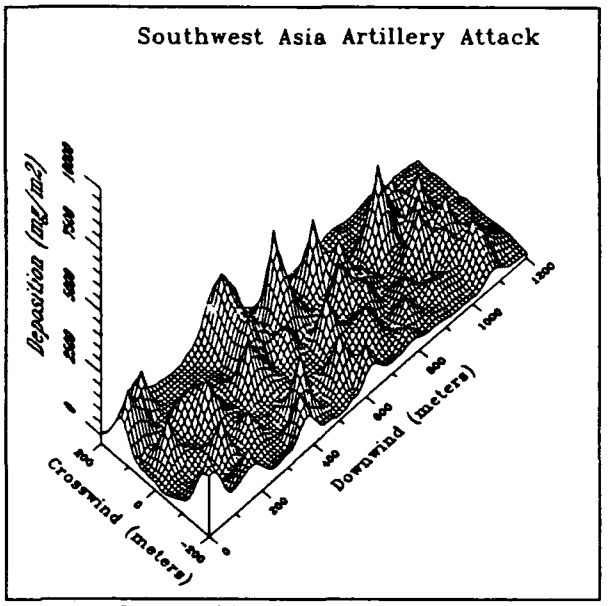


Figure 3. Artillery Attack

Table 1. Nerve agent treatment alternatives.

Option	Drugs Utilized
No therapy	None
Current therapy	Atropine & 2-PAM (Mark I Antidote Kit)
Enhanced current therapy	Pyridostigmine pretreatment, with atropine and 2-PAM
Near-term therapy	Pyridostigmine pretreatment, with atropine, 2-PAM, and anticonvulsant
Far-term therapy	Biotechnology product (nerve agent antibody)

Mission-oriented protective posture (MOPP) provides for utilization of the chemical protective ensemble in a manner tailored to the threat of chemical attack. "MOPP 0" is the standard battle dress uniform. MOPP I through IV progressively add or close components until the individual is completely sealed in the full ensemble. MOPP IV protection is at a cost of restrictions to range of motion, tactile sensitivity, and field of vision;

also, the additional weight of the ensemble imposes limitations. If the individual has a heavy physical workload, there is yet another burden: increased heat stress as the temperature and relative humidity within the ensemble components increase. In hot climates, this burden is significant for even light-work scenarios. Table 2 describes the protection options considered.

Table 2. Protection options.

Option	Equipment
None	Battle dress uniform (BDU) with normal load-bearing equipment (LBE) and individual weapon
MOPP IV	BDU and all MOPP gear, with LBE and weapon; MOPP donned at time of attack

In general, R&D observations have indicated that the alternate treatment regimens described in Table 1, in comparison to the no-treatment or the current-treatment alternatives, will result in improvements to such factors as resistance to injury (threshold, visual incapacitation, incapacitation, death) and time to return to duty. In the current study, we looked at the relative effectiveness of the five medical alternatives in terms of unit effectiveness over time for both of the physical protection options.

Modeling Tools:

MCAS: Modular Chemical Assessment Structure (MCAS) is a chemical agent dose and effects simulation. MCAS develops the vapor and liquid agent challenges on the target, as a function of time, which would result from an attack scenario, given specified meteorological and terrain conditions. The agent challenges described above were developed with MCAS. MCAS permits varying the response of the troop population in response to the challenge, based on the protection available and the effectiveness of treatment. Thus, numbers of casualties resulting from the attack scenarios and numbers of soldiers who return to duty after medical intervention can be evaluated with various treatment alternatives. Protective ratios derived for lethal effects under each treatment alternative were developed in coordination with the U.S. Army Medical Research Institute for Chemical Defense and are contained in the full report.[1] Note that anticonvulsant with a treatment regimen, in comparison to the equivalent treatment without anticonvulsant, does not change the protective ratio, rather it hastens recovery and return-to-duty.

AURA: The principal tool employed to evaluate unit effectiveness with various medical product alternatives is the Army Unit Resiliency Analysis (AURA) model.[7] AURA is a one-sided, company/ battalion-level methodology that uses inputs from a variety of accepted models and other sources to assess the effects of many types of attack or conditions on unit performance. AURA describes the unit physically and functionally, and may be used to determine the extent to which it can perform its prescribed functions (missions), given the availability of physical assets (persons and equipment). AURA describes units in terms of the combinations of

persons and equipment required to perform mission functions. AURA also considers whether alternate persons or equipment would be available to perform mission functions in the event that specific equipment or persons were lost, the effectiveness of those alternatives in relation to the principal means, and similar types of issues. AURA can accept input from other analytical models to establish personnel and equipment losses and, in turn, to determine unit effectiveness. The units examined during this study are listed in Table 4.

Table 4. Units evaluated.

Unit	Subunit[a]
Headquarters Unit	Headquarters and Headquarters Battery, Field Artillery Battalion, Heavy Division (TOE 06-366J)
Artillery Unit	155-mm Self-Propelled Battery, M109 (TOE 06-367J)
Attack Helicopter Unit	Attack Helicopter Company, AH-1 Cobra (TOE 17-187J)
Infantry Anti-Armor	Infantry Anti-Armor Company (TOE Unit 07-248J)
Ammunition Supply	Combat Service Support Company (TOE Point 09-64)

[a] *From U.S. Army Table of Organization and Equipment (TOE).*

AURA was used to develop unit effectiveness for these units, based on their AURA descriptions, degradation due to MOPP, and chemical casualties-developed by MCAS. No conventional casualties or other causes of degradation to unit effectiveness were played. MCAS outputs the percentage of the target population in four casualty categories for input to AURA: threshold effects, vision impairment, incapacitation, and death. MCAS also provides the number of days from the attack until incapacitated casualties return to duty. AURA inputs from MCAS reflect the effects of medical protective or therapy drugs that reduce casualties or hasten return to duty. Therapy side-effects are also represented, such as the visual effects of atropine. No replacements were played in the AURA simulations. AURA results are sensitive to the ability of a unit to use alternate unit members to accomplish critical tasks. The units studied in these experiments showed a range of sensitivity to personnel losses with the Ammunition Supply Point showing the least sensitivity and the Headquarters unit showing the most sensitivity to casualties.

Results and Analysis: We examined unit effectiveness after soman attack by two weapon systems, against five types of units, in two levels of protection, applying five nerve agent treatment alternates, under three sets of climatic conditions. Altogether, 300 sets of conditions were represented. For the sake of this abbreviated report, we will examine only the results achieved after an artillery attack against the artillery unit and the attack helicopter unit, under Southwest Asia day conditions. We selected these units because, from among the units studied, they fall roughly midway in sensitivity to loss of specific skills. Southwest Asia

day conditions were chosen because the region is of particular interest as of this writing. We used the artillery attack because the results appear to be driven by a fairly uniform distribution of agent over the target area.

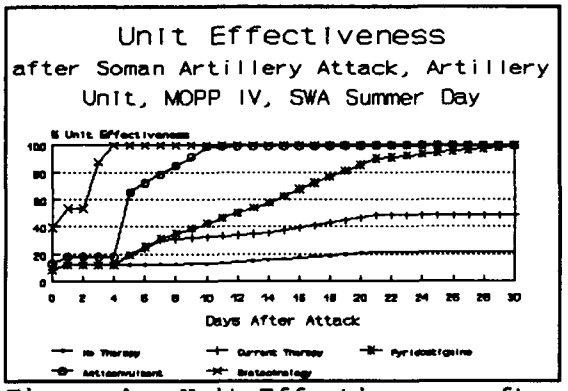


Figure 4. Unit Effectiveness after Soman Artillery Attack, Artillery Unit, MOPP IV, SWA Summer Day

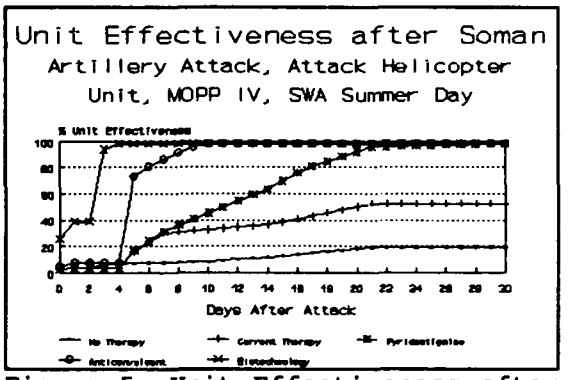


Figure 5. Unit Effectiveness after Soman Artillery Attack, Attack Helicopter Unit, MOPP IV, SWA Summer Day

We will emphasize a comparison of the current doctrine for treating nerve agent casualties in the field (atropine and 2-PAM), with the enhanced regimen (adding pyridostigmine pretreatment), the near-term improvement (adding pyridostigmine and anticonvulsant therapy), and the far-term treatment (nerve agent antibody). A key point of the analysis is the time to recover combat effectiveness after the attack. For purposes of this analysis, we consider units to be combat-effective if they are ≥80% effective.

Figure 4 shows the results when an Artillery unit in MOPP IV suffers the hypothetical artillery attack. The figure plots unit effectiveness as a function of time for all five treatments. With the current therapy, the unit never regains significant capability. Pyridostigmine pretreatment enables the unit to regain significant effectiveness after 17 days. Anticonvulsant therapy has a clearly beneficial effect, restoring the unit to combat effectiveness after 7 days--unit capability improves by more than 50% after day 4, as unit members who do not experience convulsions return to duty. Pretreatment with nerve agent antibodies rapidly improves capability after day 2, as visually-impaired personnel recover; at day 3, the unit reaches 80% effectiveness; and, after day 1, at least 50% effectiveness is shown.

Figures 5 shows the results of an artillery attack against the Attack Helicopter unit. The Helicopter unit shows about a 10% additional reduction in initial post-attack effectiveness, which might be due to this unit's higher-than-average sensitivity to vision effects. Current therapy fails to restore the Attack Helicopter unit. With pyridostigmine pretreatment, the unit is restored to near-full-combat effectiveness. However, they are less than 80% effective for 18 days. Anticonvulsive therapy restores the unit significantly by the 5th day. Nerve agent antibodies notably enhance the performance of the unit. Initial loss of effectiveness (25-35%) is less than with other therapies. The unit is

restored to greater than 80% effectiveness in 3 days or less and to more than 90% effectiveness in less than 4 days.

Conclusions: Current therapy is only marginally effective under the attack conditions studied, and fails to return to duty the number of casualties required to restore combat effectiveness. Current therapy augmented with pyridostigmine pretreatment will eventually restore combat effectiveness. Without personnel replacements, however, unit recovery will take several weeks. This latter result can probably be accounted for as follows: while atropine plus 2-PAM with pyridostigmine is effective in restoring AChE levels, other effects of nerve agents extend the recovery period of some casualties. Even so, over a 30-day period, this therapy is effective in ultimately restoring combat effectiveness. Current therapy augmented with pyridostigmine pretreatment and anticonvulsive therapy significantly improves units' time of recovery to combat effectiveness. Only one-third to one-half the time required with current therapy enhanced with pyridostigmine pretreatment is required for recovery. We can reasonably expect that use of nerve agent antibodies will further shorten unit recovery times. Other comparisons in the study not detailed here indicated that no therapy option is sufficiently effective to permit troops to operate in a high chemical threat environment without physical protection. No unit studied under MOPP 0 conditions was restored to combat effectiveness during the study period.

EVALUATION OF SLEEP DISCIPLINE TO SUSTAIN UNIT PRODUCTIVITY

"Continuous operations," as defined in U.S. Army Field Manual 22-9, is continuous land combat with some opportunity for sleep, although this sleep may be brief or fragmented.[8] The successful development and application of a doctrine for managing sleep and alertness in continuous operations is essential to success on the battlefield. The potential technological advantages of night vision and electro-optical devices for all weather operations increase the potential for continuous operations. With limited combat personnel available for future contingency operations, the U.S. Army cannot expect to conduct continuous operations by rotating personnel in shifts or by fighting soldiers and units to exhaustion and replacing them with fresh personnel (WRAIR Technical Report No. BB-87-1).

To be successful, the combat unit must maintain an efficient balance of sleep and work in order to sustain high levels of productivity in continuous operations. Ideally, every individual would receive adequate sleep to maintain effective performance on a daily basis; this performance must be maintained in all activities during the day. Nominally, the amount of sleep necessary to do this depends on the individual and the tasks to be performed. Continuous operations are routinely conducted by U.S. Navy, U.S. Air Force, and U.S. Army aviation units with 6-8 hours of sleep in every 24-hour period. It is difficult and potentially dangerous to conduct unit level exercises to evaluate the long-term impact of the continuous operations with different sleep disciplines. Laboratory measurements of effectiveness on experimental tasks do not readily translate into performance of tasks necessary to maintain operational effectiveness. As a step to examine the implication of different sleep disciplines, an effort was undertaken to define, develop, test, and evaluate a methodology to incorporate sleep into the AURA model. This effort was completed in October 1989.[9] Since that time other improvements have been incorporated

in the sleep modeling process and results are summarized here.

Objective: The objective of this study was to determine the impact of alternative sleep regimens on two Army units, one performing a mixture of physical and cognitive tasks, the other performing primarily cognitive tasks. The units selected were the M109 Artillery Battery and the Headquaters Unit, see Table 4. The greatest impact of sleep deprivation is evident in soldiers performing cognitive tasks as opposed to those performing strictly labor-intensive tasks. The results from the Headquaters Unit, which arranges primarily cognitive tasks, served as a "worst case" view of the effects of sleep deprivation.

Approach: The AURA methodology was used to measure the impact of sleep deprivation on each unit's performance. The AURA model determines, for a given unit under a given set of conditions, unit effectiveness in performing a pre-specified mission. This effectiveness value was used as a measure of performance for the current study and a basis of comparison for the cases studied.

Sleep Methodology: The sleep methodology within AURA has been documented in an earlier report.[9] A discussion of this methodology and of the modifications made to the methodology during the course of this study will be included in this report.

In the sleep methodology used for the previously referenced studies, there were six key variables to identify the sleep cycle within AURA.[10] These variables included the definition of the unit of effective sleep, the maximum allowable accumulation of these effective sleep units, the level of sleep units which are equivalent to 100% effective performance, and the use rate of effective sleep units during work activities. The final variable used in the methodology was the size of the reservoir of effective sleep units that can be saved for use at a later time. Three additional variables were developed for use during this analysis. These new variables included a parameter to specify a required work period. This variable allowed us to control sleep periods more easily and directly than did the previous methods of adjusting the minimum sleep units needed to remain awake. Two other variables were added to the AURA model to allow the representation of individual variations in required sleep. These will be discussed in detail in the section on the AURA methodology.

Sleep Accumulation Function: The basic sleep accumulation function is constructed in four segments (see Figure 6). For the first 10 minutes, no "minutes of effective sleep" are generated; for minutes 10 through 30, a person is given half-credit for the amount of time asleep (i.e., 10 minutes). So a person at 30 minutes into their sleep period will have accumulated 10 "minutes of effective sleep," reflecting the fact that individuals do not start deriving full benefit from sleep until 30 minutes after they have fallen asleep. From 30 minutes until the point where they have made up 80% of the difference between where they started at the beginning of the sleep period and maximum extent of their reservoir (that level defining full restedness), a person is able to accumulate effective sleep on a one-to-one basis one minute of effective sleep for each minute asleep. After replenishing 80% of the deficit, a person only accumulates half of a minute of effective sleep for every minute the person sleeps. The benefit to be derived from any period of rest is calculated by solving a set of simultaneous equations described in the main report. For example,

if a person is in balance working 18 hours and sleeping 6 hours, during that 6-hour sleep period he would accumulate 283 1/3 effective minutes of sleep and be at maximum sleep accumulation at the end of the sleep period. If a person was very tired, having worked more than 18 hours, then a 6-hour sleep period could potentially yield 340 "minutes of effective sleep" in a 6-hour period because the reservoir remained below the 80% threshold of diminishing returns. Finally, once a person sleeps sufficiently long to reach the maximum of their reservoir, further sleep accrues no further benefit to the sleep accumulation function.

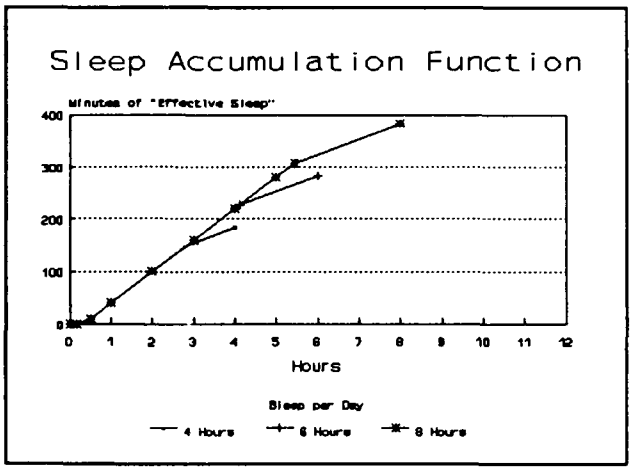


Figure 6. Sleep Accumulation Function

Work Degradation Function: The rate at which sleep units are used is based on the observations of degradations in performance in individuals deprived of sleep over extended periods. The Department of Behavioral Biology of the Walter Reed Army Institute of Research has observed that the maximum time that a person can work without effective rest is 4 days. Measures of a variety of mental (cognitive) tasks indicate that the rate of correct performance declines an average of 25% per day.

It has also been observed that a person in their normal sleep cycle (where they are in balance) will be able to retain 100% performance in their every day of work in that environment. A person that has a nominal sleep requirement of 6 hours of sleep/18 hours of work will still perform at 100% of their capability at the end of the 18-hour work period. With these limits in mind, assuming work is to continue, we modeled a linear drop in the effectiveness at which they were able to perform their job. For this analysis, all jobs in the Headquaters Unit were considered to be cognitive and, therefore, subject to a performance degradation of 25% per day without effective rest. The Artillery Unit was represented by a range of degradation functions corresponding to the mix of both physical and cognitive tasks; maximum degradation for any task was 25% per day.

Sleep Requirements Distribution: In order to model individual variations in sleep requirements, the differences in sleep balance points for individuals were assigned according to an approximate normal distribution of requirements. The distribution of sleep requirements was categorized into four classes:

- 10% of the unit was assumed to require 7.5 hours of sleep
- 40% of the unit required 6.5 hours of sleep
- another 40% of the unit required 5.5 hours of sleep
- 10% of the unit required 4.5 hours of sleep

The above factors resulted in an approximation to the normal distribution with a mean of 6 hours and a standard deviation of 45 minutes. The differences in sleep balance point required an adjustment to the size of the reservoir for individuals performing different jobs. The absolute rate at which "minutes of effective sleep" were used was individualized to reflect both the rate of use imposed by the job and the adjusted size of the reservoir for the individual.

Sleep Disciplines Considered: We evaluated a series of sleep regimens or disciplines that represent different doctrinal prescriptions of sleep per day. For the Headquaters Unit we examined 4, 6, 8, and 10 hours of sleep per day; for the Artillery Unit we tested regimens of 4, 5, 6, 7, 8 and 9 hours per day.

Unit Effectiveness and Productivity: The prime output of AURA for this study was unit effectiveness at 4 hr intervals across a 20 day continuous operations scenario. When the unit was at work AURA estimated the effectiveness of the unit given the current level of fatigue computed as described above. When the unit was taking sleep according to the prescribed discipline, AURA portrayed the unit at zero effectiveness. For any given sleep regimen, AURA depicted the effectiveness of the unit when awake and the duration of zero effectivenss when the unit was asleep. Each unit under each regimen showed a daily cycle of work and sleep for 20 days. We calculated for each day an average hourly productivity that reflected both the quality of performance when awake (degraded perhaps by fatigue) and the duration of scheduled sleep. Note that increases in the length of sleep imposed a trade-off; the increased sleep improved performance when awake but lengthened the period of zero productivity when asleep. By comparing average productivity across the range of sleep regimens, AURA provided an estimate of the *optimal sleep discipline* that maximized productivity.

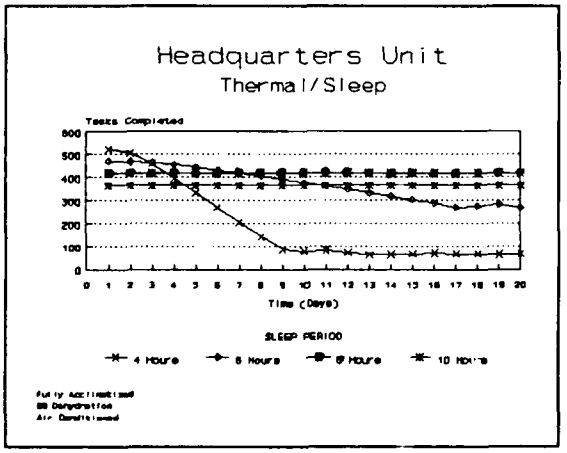


Figure 7. Tasks Completed, Headquarters Unit

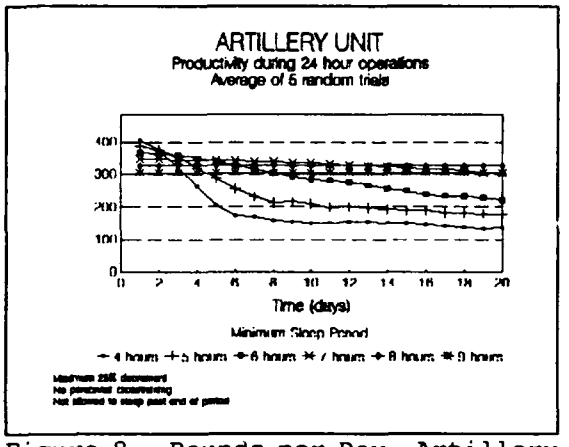


Figure 8. Rounds per Day, Artillery Unit

Results: Figures 7 and 8 show the impact of sleep deprivation for a twenty day period. The Headquaters unit (Figure 7) is required to perform 26 fire

support tasks per hour. Thus, the Y-axis illustrates the number of fire support tasks completed per day for each of the different sleep regimens. In all cases, the effects of sleep deprivation clearly drive the unit's ability to complete its mission over the 20 day period. Of the disciplines studied, 8 hrs of sleep produced teh greatest sustained productivity. Six hrs of sleep produced greatest productivity for scenarios lasting less than 10 days. Four hrs of sleep sustained performance for only 3-4 days.

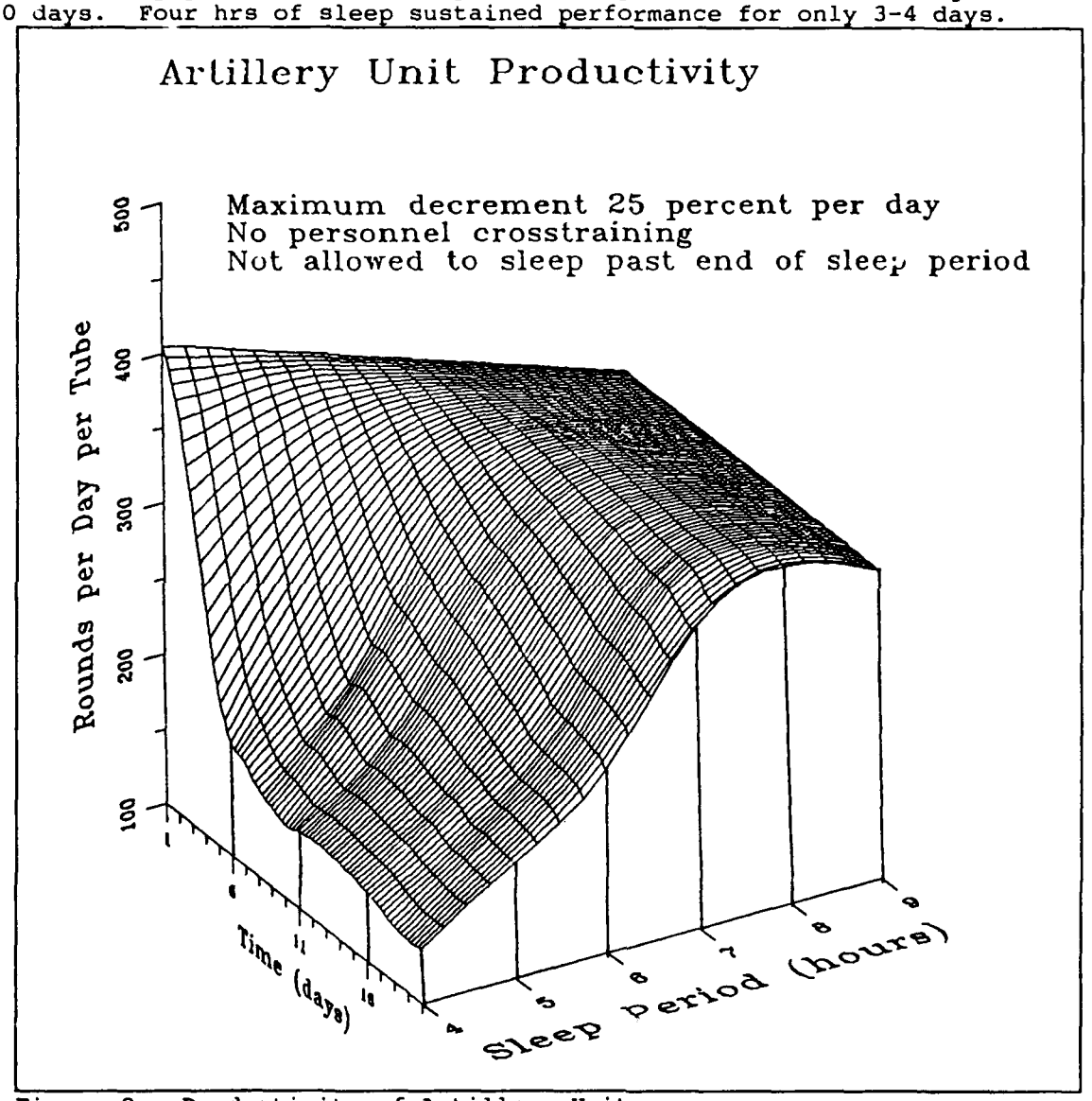


Figure 9. Productivity of Artillery Unit

The performance of the Artillery Unit (Figure 8) was less affected by marginal levels of sleep (eg. 4 hours) compared to the Headquaters Unit, primarily because of the generally less cognitive nature of the tasks and

the unit's ability to perform its mission with several alternative work strategies. Nevertheless, performance became unacceptable after 4 days of continuous operations. Seven hrs of sleep produced optimal performance for two weeks, while eight hours of sleep sustained the unit indefinitely. For both units studied, sleep regimens greater than eight hrs were actually detrimental to overall unit productivity by scheduling more sleep than necessary.

Figures 9 graphically illustrates the impact of various sleep disciplines on the Artillery Unit. These results show that the 4, 5 and 6 hour sleep regimens are not sufficient to prevent sleep deprivation effects on unit performance. While the 9 hour sleep regimen provides optimal performance during work periods, a 7 to 8 hour regimen provides the greatest overall productivity over the 20 day period.

Conclusions: Seven and 8 hours of sleep provide sustained performance levels for long term operations. Six hours of sleep provides high performance for up to 10 days, but is subject to accumulated degradation in the level of performance over time. Four hours of sleep provides highest levels of performance for short bursts of activity, but after a few days the unit will lose combat effectiveness and require a respite to recover.

REFERENCES

1. *New Medical Therapy Options for Nerve Agent Defense*: J.P. Bulger, R.E. McNally, D.E. Lee, and A.M. Machovec (U.S. Army Medical Research and Development Command Task Order Report, Contract No. DAAL02-90-C-0071, T.O.# 0001AE, Science Applications International Corporation, Joppa, Maryland, 1991).

2. *Battlefield Levels of Soman*: R.E. McNally and D.L. Bareis (U.S. Army Medical Institute of Chemical Defense, Aberdeen Proving Ground, Maryland, Contract No. DACMD17-88-C-8155, 1989) p. 3.

3. *U.S. Army Field Manual No. FM 3-9 Military Chemistry and Chemical Compounds.* (U.S. Government Printing Office, Washington, D.C., 1975), pp. 3-9.

4. *Nerve Agent Pretreatment Transition Information Paper: Physostigmine Plus Adjunct*: D.L. Bareis (U.S. Army Medical Institute of Chemical Defense, Ft. Dietrick, Maryland. Contract No. DAMD17-88-C-8155, 1989).

5. *Proceedings of the Workshop on Convulsions and Related Brain Damage Induced by Organophosphorus Agents, 1-2 February 1990, Aberdeen Proving Ground, Maryland*: G.M. Kiebzak and A.J. Feister (U.S. Army Medical Research Institute of Chemical Defense, Aberdeen Proving Ground, Maryland, 1990) p. 12, 179.

6. *Nerve Agent Scavengers*: Medical Chemical Defense: A Bulletin on Research and Development. (U.S. Army Medical Research and Development Command, **3(1)**, 1989), p 2.

7. *Input Manual for the Army Unit Resiliency Analysis (AURA) Computer Simulation Model: 1990 Update, (UNCLASSIFIED)*: J.T. Klopcic, R.M.

Sheroke, and S.K. Price (Ballistic Research Laboratory Technical Report BRL-TR-3187, 1990).

8. *U.S. Army Field Manual No. FM 22-9, Soldier Performance in Continuous Operations.* (U.S. Government Printing Office, Washington, D.C., 1983)

9. *Evaluation of Sleep Discipline in Sustaining Unit Performance*: R.E. McNally, A.M. Machovec, and D.T. Ellzy (U.S. Army Medical Research and Development Command Task Order Report, Contract No. MDA903-88-D-1000, Science Applications International Corporation, Joppa, Maryland, 1989).

10. *The AURA Fatigue and Heat Stress Algorithms*, (UNCLASSIFIED): J.T. Klopcic, (Ballistic Research Laboratory Technical Report BRL-MR-3802, 1989).

High Bandwidth, Wide Field of View Lasercom Demonstration

Dr. John C. Johnson
USASDC, Advanced Technology Directorate,
Huntsville, AL 35807-3801

Abstract

Laser satellite networking is a key element of effective communications operations to support both strategic and tactical missions. Lasercom offers a number of important advantages over conventional RF satellite communications. The shorter wavelength available using lasers provides higher data rates at less power and smaller apertures, both resulting in lower weight requirements. On the other hand, lasercom entails more difficulty in acquisition and tracking because of the narrow beams used. Technology problems to be overcome before intersatellite laser communications can reliably outperform RF communications include acquisition in the presence of significant background light from the earth, tracking to resolutions of a few microradians, high speed modulation of semiconductor lasers with close to one watt of power in a diffraction limited beam, high bandwidth low noise detector response, and demonstrated long term performance. We have developed critical technologies to solve some of these problems, and demonstrated them in a laboratory testbed which also supports development and testing of network protocols and algorithms. Our hardware provides new capability in background light rejection by using innovative atomic line filter technology, improved tracking accuracy by using innovative zero backlash Roto-Lok drive gimbal telescopes, and increased communications bandwidth by incorporating multi-link networking protocols.

Introduction

For communications between satellites, higher data rates (> 1 Giga bits/sec) can be achieved using optical laser communications than is possible using conventional Radio Frequency (RF) transceivers (approximately 20 mega bits/sec). In addition, these data rates can be achieved by using less power, and smaller mirrors (and thus at less weight) by using optical frequencies instead of RF carriers for the signal. The reason for this is because of the divergence of the beams which is proportional to the wavelength in the diffraction limit, and can be made

much smaller for the laser transmitter. The laser has a wavelength on the order of one ten thousandth that of the RF signal. The disadvantage of using narrow beams is that acquisition and tracking become more difficult since only very small pointing errors can be tolerated.

Small laser divergence results in smaller size, weight, and power as well as resulting in a lower probability of intercept and a high resistance to jamming.

A typical RF system requires a 3 meter dish antennae with 150 watts of transmit power. The overall RF system would weigh approximately 340 pounds and require on the order of 350 watts. A typical laser system would require a relatively small telescope (0.15 meter) and a relatively low transmit power (1 watt) with a data rate capability of 1000 mega bits/sec. The overall laser system would weigh approximately 150 pounds and require on the order of 100 watts.

The overall program goals and achievements are summarized in figure 1. The primary objectives were to demonstrate several critical technologies which would enhance the feasibility of fielding a laser satellite communications network. The first innovation is acquisition in background light using an Atomic Line Filter. The second innovation is the use of high data rate simultaneous video and digital data. The third innovation is the use of a lightweight, low cost Roto-Lok drive gimbal. The fourth innovation is the development of a high bandwidth, high accuracy tracking system.

Following the demonstration of these critical technologies an additional goal was to integrate all the critical technology components into a laboratory testbed for the purpose of demonstrating the simultaneous operation of all necessary components. This testbed was designed and laid out such that upgrades to the system are easily accommodated. The testbed was also designed to accommodate networking protocols and algorithms. The testbed as shown in figure 2 consisted of three nodes with laser transceivers and two additional nodes which were hard-wired to simulate laser transceivers for the purpose of evaluating network protocols and algorithms.

The optical path of the laboratory laser transceiver is shown in figure 3. A beacon for acquisition and tracking diverges from a small aperture and has a wavelength of 852 nano meters matching an atomic line filter. The tracking signal is received through a four inch aperture and locked on to with wide and narrow field of view cameras. The four inch aperture is shared with the transmitted and received communications beams at 780 nano meters, which are separated from the 852 nano meter tracking signal using a dichroic mirror. The communications rate versus range for the laboratory transceiver is shown in figure 4. Increased data rates

GOAL	PARAMETER	SOW SPECS	PROGRAM ACHIEVEMENT
• Demonstrate Critical Technologies	Field of View	10 mrad	14 x 19 mrad
– Acquisition in background light using Atomic Line Filter	Δλ Bandwidth	0.002 nm	0.0024 nm
– High data rate optical communications (diode laser, APD) (U of D scheme for simultaneous video and digital)	Efficiency	---	0.25
	Video Carrier Frequency	1.2 GHz	1.206, 1.212 GHz
	Simultaneous Data	50 MBPS	50 MBPS
	Separate Digital Data	---	1 GBPS
– Lightweight, low cost Roto-Lok drive gimbal	Open Loop Pointing	100 μrad	80 μrad
– High bandwidth, high accuracy tracking	Closed Loop Pointing	10 μrad	7 μrad
	Bandwidth	---	> 6 Hz
• Integrate Laboratory Test Bed			
– Demonstrate simultaneous operation of necessary components	# of Transceivers	3	3
	# of Nodes	5	5
– Provide facility to incorporate upgrades	# of Links	10	10
• Support networking protocols and algorithms	U of F Tested Algorithms	---	3

Figure 1: Program Goals and Achievements

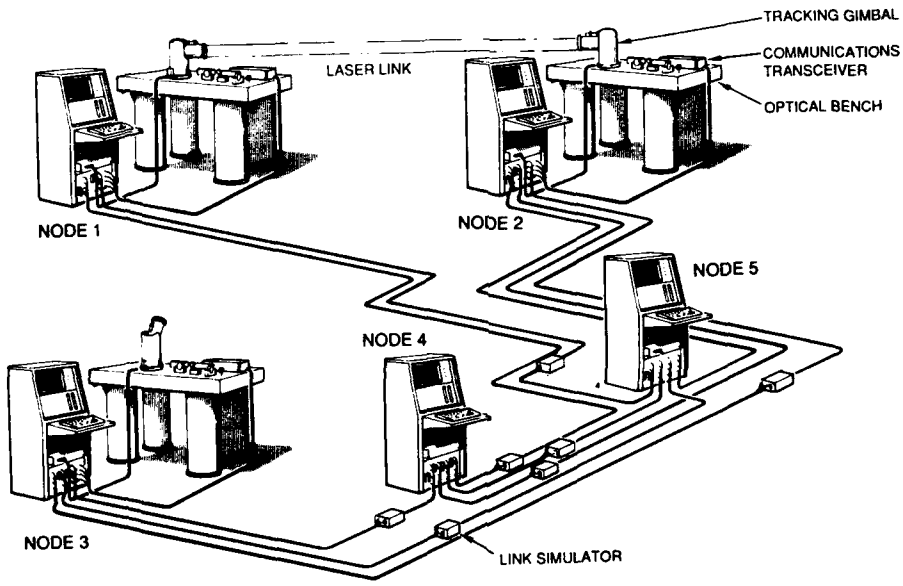


Figure 2: Interactive Laboratory Demonstration Testbed

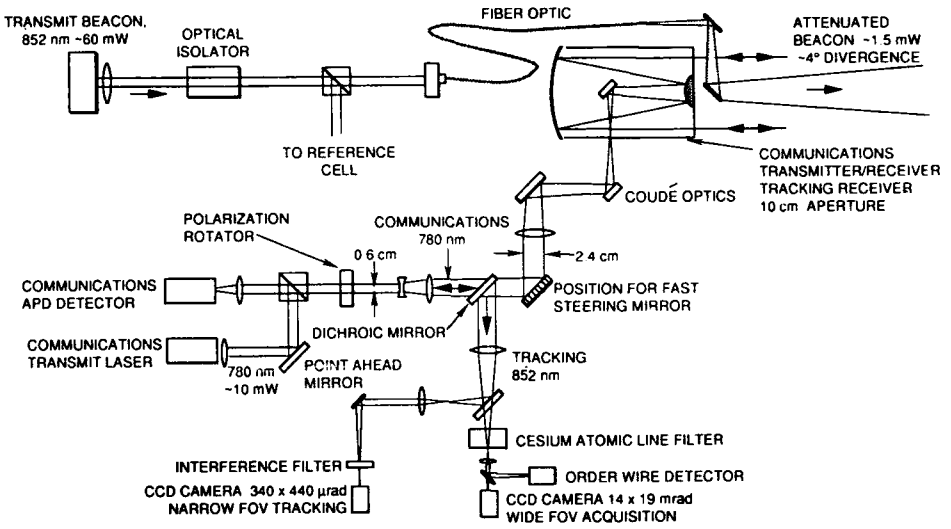


Figure 3: Laboratory Transceiver Optical Path

ASSUMPTIONS:
10 cm telescope, area A = 0.006 m^2, diffraction limited divergence θ = 14 µrad
Transmitter peak power P_o = 200 mW [currently realistic, 850 nm]
Transmit efficiency η_o = 0.5, receive efficiency η_r = 0.5
Mispointing efficiency η_p = 0.5 (5 µrad mispoint)
Avalanche photodiode detector with 60 photons/bit for 10^{-8} BER [currently realistic at 10-50 MBPS]

EQUATION:

$$\text{Data rate} = \frac{P_o \eta_o \eta_r \eta_p A}{\pi(\theta/2)^2 R^2} \times \frac{4.2 \times 10^{18} \text{ photon}}{\text{W} \cdot \text{sec}} \times \frac{1 \text{ MBit}}{60 \times 10^6 \text{ photons}} = \frac{6.82 \times 10^{10} \text{ MBPS km}^2}{R^2}$$

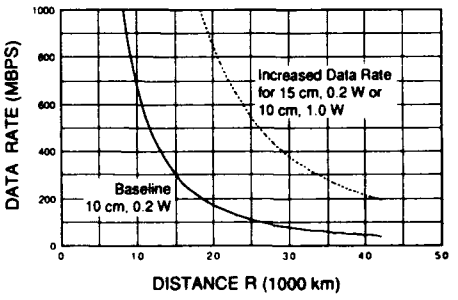


Figure 4: Communications Rate versus Range for Laboratory Transceiver (solid line) upgraded to currently available 200 mW peak power transmit diode laser. Increased data rates (dashed line) are possible with a higher power laser or larger aperture.

available are shown in figure 4, dashed line, if a higher power laser is used or a larger aperture is used. Data rates of one giga bits per second are possible at ranges of greater than 10,000 kilometers.

Technical Approach

The three major technical requirements for a free space laser communications system are the ability to do fast acquisition, high accuracy tracking, and high data rate, high reliability communications. Fast acquisition is necessary in a network because links must be broken and reformed as satellites travel in different orbits around the earth, and acquisition should occupy a small fraction of the total link time. Our approach to acquisition is to open up the receiver field of view to about 1° so it is large enough to view the acquisition beacon from the other satellite based on initial ephemeris information without excessive scanning. To have such a wide field of view with the earth in the background requires the use of an extremely narrow bandwidth filter, such as the atomic line filter demonstrated in this program. For high data rate, high reliability communications, we choose direct detection of a modulated laser diode using a low noise avalanche photodiode. Finally, for high accuracy pointing and tracking, we have incorporated the zero backlash Roto-Lok drive.

The background light problem to be overcome in the acquisition stage is illustrated in Figure 5. One satellite (for instance in geosynchronous orbit) is trying to acquire the tracking beacon of another satellite (in low earth orbit) with the sunlit earth in the background. The instantaneous field of view (FOV) which can be stared at with the acquisition CCD camera depends on the amount of light from the beacon, compared to spatial variations in the amount of light from the earth, reflected off of bright objects such as clouds or dark objects such as the ocean. By minimizing the amount of this background reflected light getting to the receiver, the field of view can be maximized. Signal to noise ratios for a given camera pixel are compared for three different light rejection filters. A background earth radiance (clouds) of $0.2 W/m^2$ nm sr is assumed.[1] For sub-pixel accuracy pointing, a signal to noise ratio (s/n) of better than seven is desired. For the baseline parameters given in the figure, the interference filter (bandwidths ~ 2nm) gives a s/n of 0.5 at the desired 1°FOV. Atomic line filters[2] can reduce the light transmission bandwidth by a factor of 1000 to 0.002 nm. One image preserving line filter developed under this program achieved this bandwidth[3]. Disadvantages of this filter include the necessity for an extra pump laser, and a low efficiency of 0.5%. Ideally, the filter would have a high throughput efficiency, be simple, and have a bandwidth close to 0.05nm so that the transmit beacon wavelength would not have to be adjusted for relative Doppler shifts between satellites. We have identified a filter which meets these goals and gives a s/n of 7 out to a field of view of 2°. It is called a Faraday Atomic line filter[4].

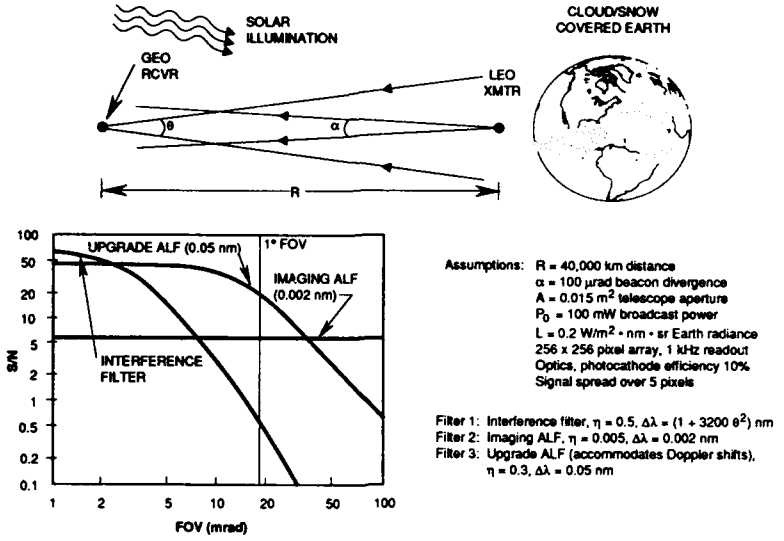


Figure 5: Increased Acquisition Field of View with Atomic Line Filter. Background light from the sunlit earth is blocked 100 times better than with an interference filter, allowing a ten times wider instantaneous field of view.

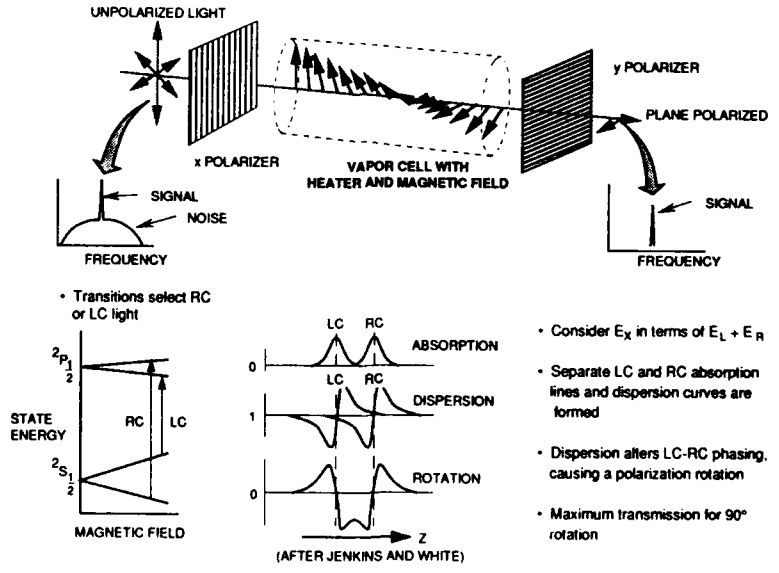


Figure 6: Principles of Faraday Atomic Line Filter Operation. Crossed polarizers block background light, while the polarization of the signal wavelength is rotated, allowing it to pass through the filter.

Operational principles of the Faraday Filter can be understood by reference to Figure 6. Basically, it consists of a sandwich comprised of two crossed polarizers on the outside and a cesium vapor cell in a magnetic field on the inside. Except at a specific wavelength (852 nm) corresponding to an absorption in the cesium vapor, background light is unaffected by the vapor cell, and effectively blocked by the crossed polarizers. For wavelengths very near the 852 nm absorption, however, the polarization of the light can be rotated in the vapor cell, thus allowing it to pass through the second polarizer. The path of the light is unaffected, so spatial and temporal information is maintained. The basis for the polarization rotation is the resonant Faraday effect. Because of the applied magnetic field, the index of refraction (and thus propagation velocity) is different for right and left circularly polarized light. A phase delay of 180° is equivalent to a linear polarization rotation of 90°.

Measured and calculated transmission spectra for a Faraday Filter are shown in Figure 7. We measured a transmission efficiency of 48% for polarized light. (Because of Fresnel reflections, for which AR coatings could be, but were not applied, the optics themselves only had a throughput efficiency of 54%). The total integrated bandwidth for background light was 0.007 nm for temperature T=80° C, magnetic field B=100 Gauss, and path length L=1.2 cm. As shown for 120° C and 200 Gauss, the transmission spectrum is affected by the temperature and magnetic field. Our experimental results and theoretical calculations agree in detail. Our calculations indicate that a central flat passband about 6 GHz wide can be made under appropriate conditions. Intersatellite Doppler compensation requires about $\pm$6 GHz, so the transmit beacon wavelength would only need to toggle between a "moving away" and "coming toward" position. For space applications, permanent magnets would be used, as for the 60°C, 250 Gauss spectrum in the upper right hand corner.

For efficient high data rate communications (minimizing transmitter power requirements), we plan to use digital modulation of the transmit laser, and direct detection using an avalanche photodiode. An atomic line filter is unnecessary for the communications channel because the transmitter is more collimated (than for the acquisition beacon) and the receiver field of view is small. We were able to reliably transmit and receive a repeating digital waveform at 1 GBPS in the laboratory to demonstrate laser modulation and detection at these speeds. For demonstration purposes, we simultaneously transmitted digital communications at 10 MBPS (used to test networking algorithms) and video signals upconverted to a 1.2 GHz carrier frequency as shown in Figure 8. By monitoring the video channel, we could see pointing, acquisition, tracking, and establishment of high data rate communications in a time scale of about two seconds.

For very high data rates over long ranges, the transmitted

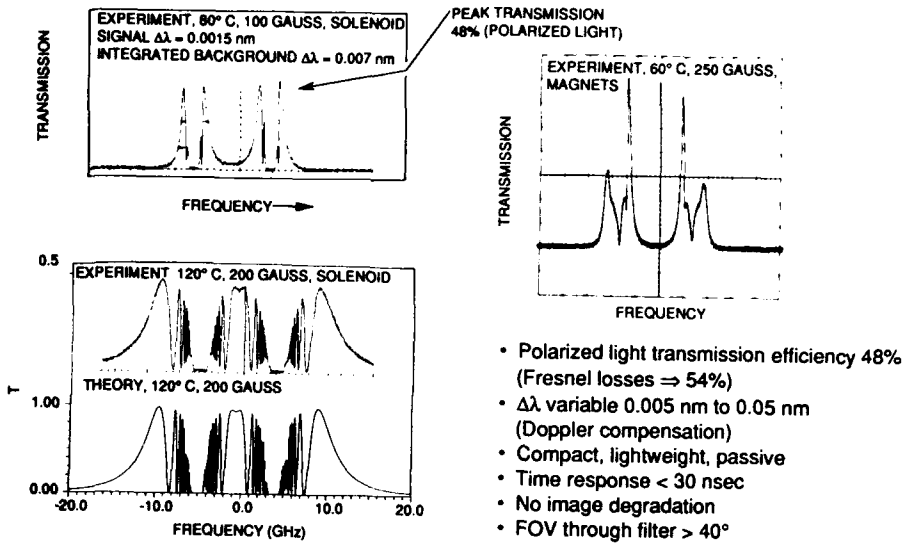


Figure 7: Performance results of Faraday Filter. The transmission spectrum is a function of cesium temperature and magnetic field. Our experimental and theoretical spectra agree in detail.

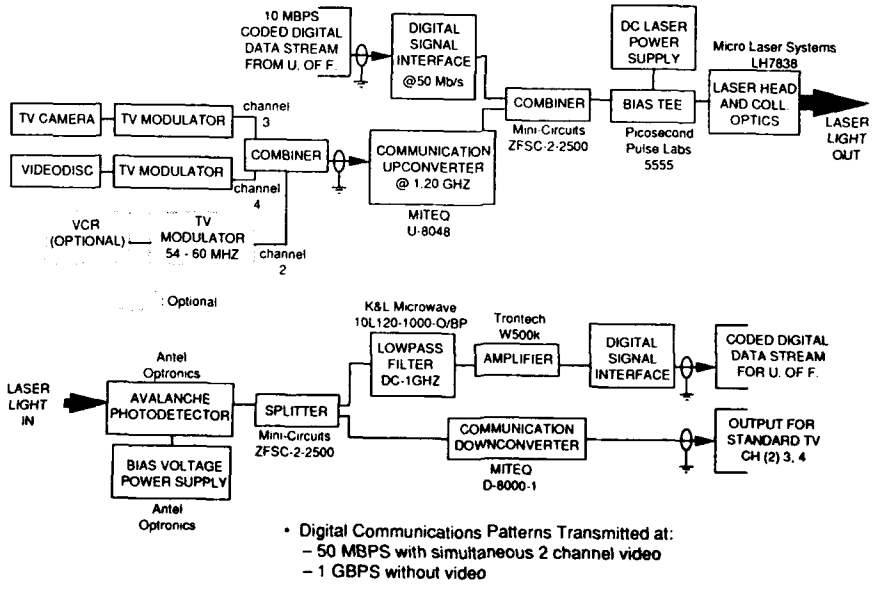
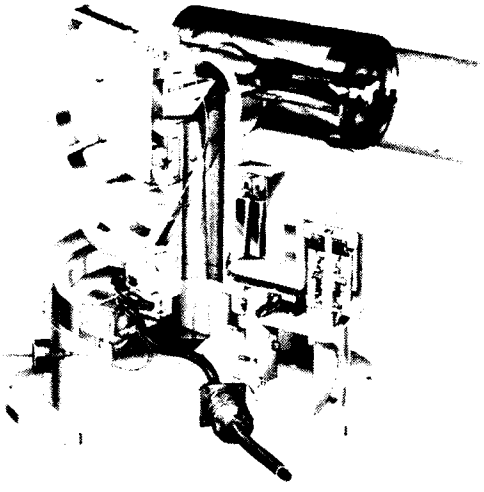


Figure 8: Communications System Block Diagram. Digital data and upconverted video signals are transmitted simultaneously.

communications laser divergence must be minimized, and the transceiver telescope must be pointed accurately within a few microradians. For accurate closed loop tracking in the presence of satellite jitter, a very low backlash, very high stiffness steering mechanism must be used. Possible choices include a direct drive motor, an offset gear drive, and the Roto-Lok cable drive, which we have chosen. A direct drive requires very high torque (and therefore weight and power) and imparts considerable angular momentum to the host structure. A gear drive is very difficult to make backlash free and will impart jitter to the platform. The Roto-Lok drive overcomes the problems inherent in the other drive systems.

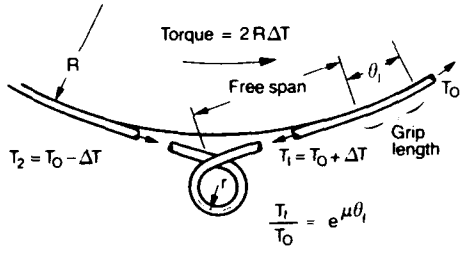
Figure 9 shows a cutaway view of the Roto-Lok drive gimbal developed under this program, and shows the principles of its operation. A drive motor shaft (or Capstan) is coupled to a large moving circular platform (drum) with many cables (eight for our transceivers). The cables are attached to the drum at their ends, wrapped around the drum, and once around the capstan. Friction keeps the cable from slipping on the drum or capstan, so that as the capstan turns, the drum turns with it. Since there are many cables, any imperfections on individual cables are averaged out, minimizing backlash well below that of gear driven systems with no weight penalty. Cable stretch is minimized because the free span between friction points is very small. Finally, the capstan and drum turn in opposite directions at different speeds determined by the turndown ratio of their relative diameters. By designing the relative moment of inertia to be in this same ratio, a reactionless drive system imparting no angular momentum to the host satellite platform can be built.

The tracking control system used in the laboratory testbed is described in Figure 10. Initial pointing was done with reference to a shaft position encoder on the drive motor, with an angular position resolution of 65 micro rad (later upgraded to 1.25 micro rad). After acquiring the beacon laser (in the presence of simulated background light) closed loop tracking was accomplished by centroiding on a CCD camera with a 1 degree field of view behind an atomic line filter. Fine tracking was done with a second camera with a 400 micro rad field of view. Tracking performance is shown in Figure 11. With the high resolution shaft position encoder, a tracking accuracy of 7 micro rad at a bandwidth of about 6 Hz was achieved, compensating for atmospheric fluctuations in the room. The bandwidth was probably limited by the 60 Hz camera readout rather than the Roto-Lok drive. In further tests incorporating a fast steering mirror and a quadrature avalanche photodiode detector, tracking accuracies of about 1 micro rad were achieved at bandwidths of about 100 Hz. Such a system may be needed for 1 GBPS communications at 42,000 km (geosynchronous) ranges. However, as described later, useful data rates and ranges should be achievable with much less complexity and at lower weight using the Roto-Lok drive gimbal alone.



- Low backlash, high stiffness

- The capstan is placed near the drum to increase stiffness by reducing the free span of cable



- Beyond the angle θ_1, the cable is unaffected by the load T_1

Figure 9: Roto-Lok Drive Gimbal Telescope Principles. Multiple cables wound around the motor shaft (radius r) and the drum (radius R) couple them together in a near zero backlash configuration.

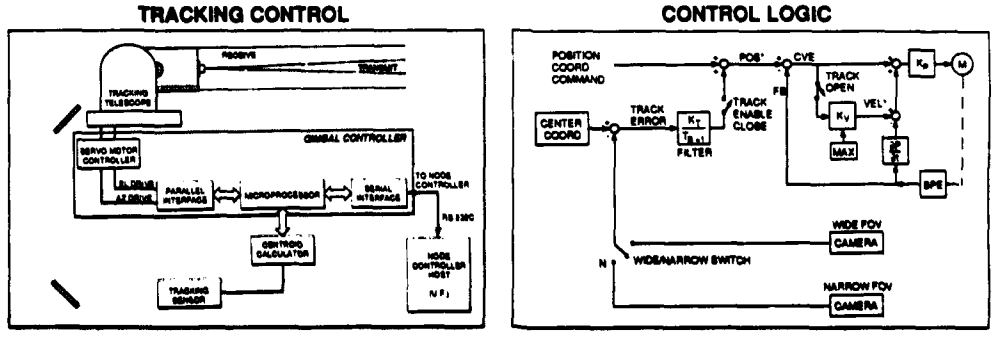
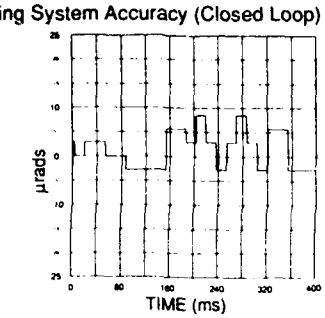
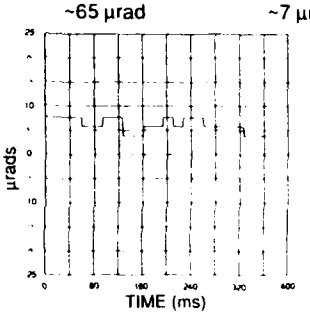


Figure 10: Tracking Control System. The control loop uses centroiding offsets from the wide or narrow field of view cameras to null the gimbal to the appropriate shaft position encoder count.

PARAMETER	S.O.W.	TRANSCEIVER 1 &2	HIGH RESOLUTION TRANSCEIVER 3
Wide Acquisition Field Of View	10 mrad	14 x 19 mrad	14 x 19 mrad
Narrow Tracking Field Of View	–	400 µrad	900 µrad
Gimbal Pointing Accuracy (Open Loop)	100 µrad	~400 µrad	~80 µrad
Gimbal Pointing Jitter (Open Loop)	–	± 65 µrad	± 1.25 µrad
Tracking Bandwidth	–	~6 Hz	~6 Hz *UPGRADE TO 100 Hz*
Tracking System Accuracy (Closed Loop)	10 µrad	~65 µrad	~7 µrad

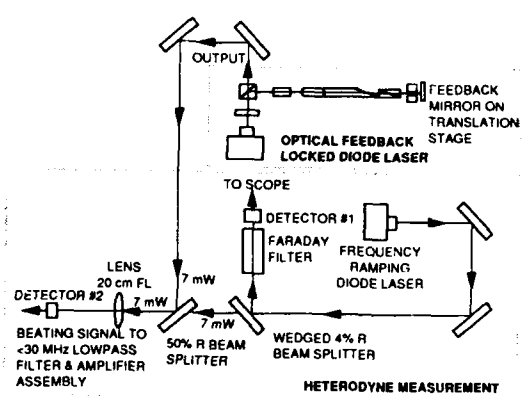


OPEN LOOP ATMOSPHERIC JITTER



CLOSED LOOP TRACKING

Figure 11: Tracking System Performance. With the narrow field of view camera and 1.25 micro rad resolution shaft position encoder better than 7 micro rad pointing accuracy was achieved at 6 Hz.



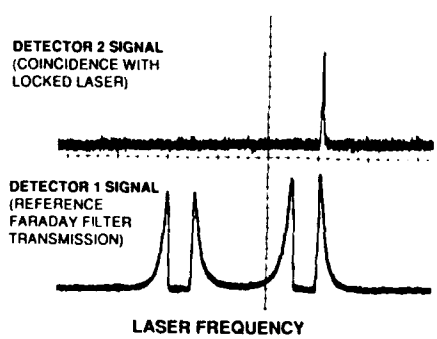


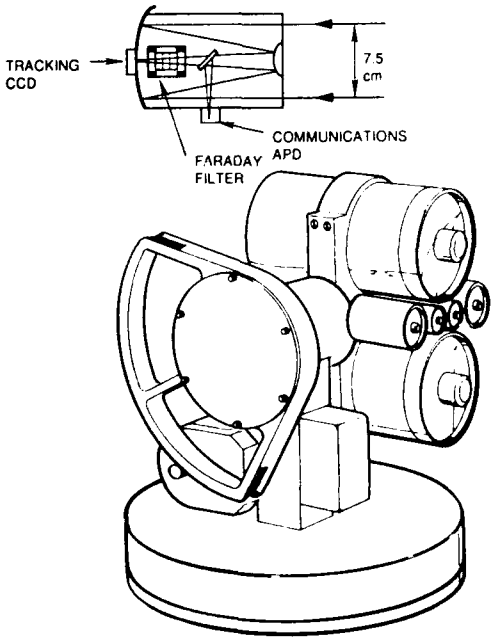
Figure 12: Diode laser wavelength stablization with optical feedback. The laser wavelength is matched to the atomic line filter passband using optical feedback through a second filter which is part of the laser cavity.

A final technology issue which needed to be solved for our atomic line filter based system was stabilization of the diode laser acquisition beacon to the wavelength passed by the filter. This can be accomplished by tuning the laser by changing its temperature and current, and closing an electronic control loop based on transmission through a reference filter. However, the temperature and current resolutions necessary are much less severe if optical feedback, as shown in Figure 12, is used. If some of the laser light is diverted through a Faraday filter (matching the receiver of interest) and reflected back into the diode laser it can automatically stabilize the laser at the wavelength of interest. The figure shows results of a heterodyne experiment showing such locking taking place on one of the Faraday Filter Transmission Peaks.

Conclusions

Laser communications shows great promise for achieving high data rate, low probability of intercept, links between satellites. A network of surveillance and communications satellites incorporating lasercom could reliably and rapidly transmit critical information around the world, as well as to and from high flying aircraft with short range radio links to the ground. Lasercom promises data rates above 100 MBPS at significantly lower power and weight (and therefore launch cost) than is possible with radio frequencies. In this program we have demonstrated innovative technologies which address some of the technical issues associated with intersatellite lasercom, including rapid acquisition, even in the presence of background light, high accuracy pointing and tracking, and high data rate communications. Because of these accomplishments, free space laser communications is ready to come out of the laboratory for implementation, first in aircraft to aircraft links where high data rates are needed, and then in space.

To demonstrate the feasibility of lasercom in comparison to radio frequencies, reliability will be of critical importance. Therefore, a prototype deployable system should be simple and redundant, demonstrating significant data rates (20 MBPS) at significant range (5000 km) in a compact, light weight package. Figure 13 shows a functional design for such a system. First, optics are not shared between transmitter and receiver, eliminating complicated systems to avoid cross-talk. Second, the doubly redundant communications and acquisition transmitters have increased divergences (100 micro rad and 1 milli rad respectively) to reduce pointing accuracy requirements to 40 micro rad. This can be accomplished with a 1° FOV, 256 x 256 pixel, 200 Hz readout CCD coupled to the Roto-Lok drive, eliminating the need for a complicated tracking system. (Sub pixel pointing accuracy is easily achieved). Because of increased laser divergences, small transmitter optics can be used. All lasers and detectors can be mounted directly on the gimbal telescope, increasing efficiency, and reducing interface complexity. Heat can be pumped to the



- Wide field of regard gimbal
- Doubly redundant
- No shared optics between transmit and receive, no Coude optics
- Easy to integrate
- 1.2 cm tracking beacon, 50 mW, 1 mrad divergence
- 2.5 cm communications laser, 100 mW, 100 μrad divergence
- 7.5 cm receiver
  - Tracking through 0.05 nm ALF
    256 x 256 pixel, 200 Hz CCD
    20 mrad IFOV, 40 μrad tracking
  - Communications through APD
    200 μrad FOV
  - Point ahead done with CCD centroiding
- Data rate 20 MBPS at 5000 km
- Weight ~ 40 lbs
  Power ~ 40 W

Figure 13: High reliability, light weight transceiver prototype. Doubly redundant system for reliable lasercom demonstration, with greatly simplified optical path (no shared transmit/receive optics).

ASSUMPTIONS:
Double redundancy, calculate for single functioning components
7.5 cm receive telescope, area A = 0.004 m^2
Transmit Divergence θ = 100 μrad
Transmitter peak power P_o = 100 mW [currently realistic, 850 nm]
Transmit efficiency η_o = 0.5, receive efficiency η_r = 0.5
Mispointing efficiency η_p = 0.5 (42 μrad mispoint)
Avalanche photodiode detector with 60 photons/bit for 10^{-8} BER [currently realistic at 10-50 MBPS]

EQUATION:

$$\text{Data rate} = \frac{P_o \eta_o \eta_r \eta_p A}{\pi(\theta/2)^2 R^2} \times \frac{4.2 \times 10^{18} \text{ photon}}{\text{W} \cdot \text{sec}} \times \frac{1 \text{ MBit}}{60 \times 10^6 \text{ photons}} = \frac{4.46 \times 10^8 \text{ MBPS km}^2}{R^2}$$

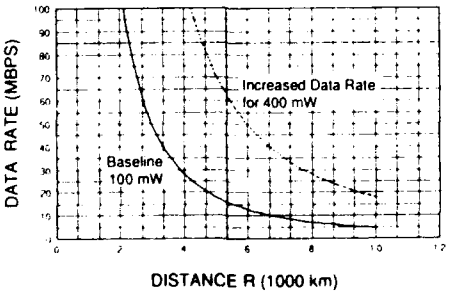


Figure 14: Data Rate versus Range for lightweight transceiver prototype.

cold side of the gimbal using heat pipe technology. Light weight permanent magnet Faraday atomic line filters will be used for background light rejection for acquisition and tracking. Finally, the gimbal telescope system will be made reactionless using the Roto-Lok drive technology, so that minimum angular momentum jitter is transferred to the host satellite. Data rate versus range for the transceiver prototype described in Figure 13 is shown in Figure 14. All components are currently available commercially, with the rated noise figures. Although the proposed system is doubly redundant, the data rate is based on only one set of components being functional, and would be higher if all components were active.

Acknowledgements: Funding for this program was provided by the Strategic Defense Initiative Organization (SDIO), Innovative Science and Technology Directorate. We would like to thank Dr. Kepi Wu for his encouragement and support. The atomic line filter development and testbed integration was done by Thermo Electron Technologies Corporation under the direction of Dr. C.S. Liu and Dr. Eric Korevaar. The high data rate communications scheme was provided by Dr. Charles Ih of the University of Delaware. The Roto-Lok drive gimbal was developed by Don Carson of Sagebrush Technology. The networking algorithms used on the testbed were developed at the University of Florida under the direction of Dr. Randy Chow and Dr. Richard Newman-Wolfe.

References:

1. W.L. Wolfe and G.J. Zissis, *The Infrared Handbook*, (Environmental Research Institute, Michigan, 1978), pp. 3-38, 3-44.
2. J.A. Gelbwachs, *IEEE J. Quantum Electronics* **24**, 1266 (1988).
3. E. Korevaar, M. Rivers, C.S. Liu, *SPIE vol. 1059 - Space Sensing, Communications, and Networking*, 111 (1989).
4. J. Menders et al., *Optics Letters* **16**, 846 (1991)

SOLID PROPELLANT ELECTROTHERMAL GUN PROPULSION

Dr. Arpad JUHASZ *
U.S. Army Ballistic Research Laboratory
Aberdeen Proving Ground, MD 21005, U.S.A.

Mr. Stan SMITH
U.S. Army Strategic Defense Command
Huntsville, AL 35807, U.S.A.

Dr. Zvi KAPLAN, Dr. D. MELNIK, Dr. D. SAPHIER
Soreq Nuclear Research Center
Yavne 70600, ISRAEL

LTC Maxi BLUM
Ministry of Defense, State of Israel
Tel-Aviv 61909, ISRAEL

1. BACKGROUND

Hypervelocity gun launcher (HVL) technology requirements have been identified both for future ground-based terminal defense (TD) as well as tactical kinetic energy (KE) applications.[1,2] The common element of an HVL launch for both types of systems is a high projectile velocity leading to short times of flight and increased kinetic energy on target. In the TD application, short time of flight offers quick reaction intercept capabilities against leakers, targets which have penetrated longer range defense architecture. HVL could permit several guided projectiles to be fired from a single launcher or from multiple launchers during a single engagement. Compared to conventional guided missile systems, the absence of rocket motors and warheads would make projectiles less expensive, and offers the potential of reducing the cost of an engagement, even when several projectiles are launched to increase the probability of success. In the tactical area, increased velocity is important in defeating state-of-the-art armor protections at extended ranges and increasing hit and kill probabilities against maneuvering targets. Both electromagnetic (EM) and electrothermal-chemical (ETC) devices have been under study as candidates for HVL. Though ultimately limited in performance as compared with EM, ETC propulsion is generally accepted as the near-term electric solution to the HVL problem[2].

Most of the developments in ETC propulsion have taken place since 1985. Initial efforts focused on plasma controlled bulk loaded liquid or gelled propellant systems. Recent work has yielded interesting results with a solid propellant variant of ETC (SPETC). This paper describes SPETC results which may form the technical basis for a future point-defense element in a terminal defense architecture for theater or national missile defense. While the technology described is focused on the mission area defined by the Corps Surface to Air Missile (CORPS-AM) operational requirement document, it is equally applicable to the anti-armor and air defense roles.

2. INTRODUCTION

One attractive feature of ETC propulsion is the possibility of using existing gun tube and mount technology to attain greater muzzle energies than obtainable from conventional guns. A schematic description of a solid propellant ETC gun system appears in Figure 1.

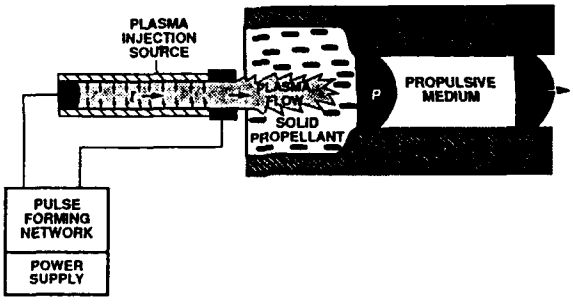


Figure 1. Schematic Description of a Solid Propellant ETC Gun.

In addition to the gun tube and mount, a SPETC gun includes an electric pulsed power source, usually in the form of a pulse forming network (PFN), and a plasma injector linked appropriately to the ammunition cartridge to assure ignition of the propellant charge and control of the interior ballistic (IB) process.

The rationale of the ETC concept is that electric control of the IB process can lead to optimum ballistic progressivity, usefully burning greater amounts of propellant by plasma injection at key points in the IB process.[3] Calculations at the Ballistic Research Laboratory indicate a potential muzzle energy increase of 25–45% for SPETC vs. optimized conventional solid propellant guns. The data presented below were obtained as part of a joint U.S.-Israeli R&D effort on SPETC guns. The following main issues will be addressed:

1. Electrical plasma sources.
2. The 60-mm experimental gun.
3. Plasma jet - solid propellant interaction.
4. Methodology for improving IB acceleration processes.
5. Experimental results showing improved SPETC guns performance.

3. ELECTRICAL PLASMA SOURCE

It has been known that an electrical discharge in a dielectric tube can operate as a plasma jet generator in a pre-determined, controlled manner.[4] The electrical energy is converted into thermal energy by establishing an electrical discharge in a dielectric tube. This discharge develops a high pressure and high temperature plasma which emerges in the form of a plasma jet through a hollow electrode, into the gun chamber. A schematic view of the device is shown in Figure 2.

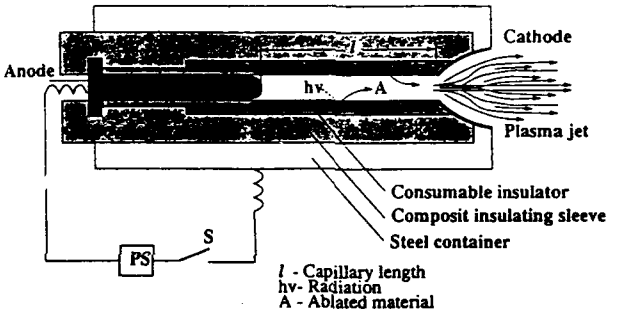


Figure 2. A Schematic View of High Pressure Injector and Plasma Discharge.

A simple model which describes the discharge phenomenology and establishes basic analytical relations between its main parameters has been developed.[5] It permits estimation of plasma temperature, pressure, density, and discharge resistance as well as their dependence on parameters, such as tube radius, length, and current density. These relations are used as practical design tools for plasma jet formation systems matched to the driving power source.

4. EXPERIMENTAL SET-UP

The experimental data presented below were obtained using the 60-mm SPETC gun. The main parameters of this gun are presented in Table 1. Figure 3 shows the basic construction of the device.

Table 1. Specifications of 60-mm SPETC Gun

Breech volume	2,250 cm^3
Projectile travel	3.60 m
Nominal projectile weight	1.35 kg
Nominal breech pressure	470 MPa (4,700 Bar)

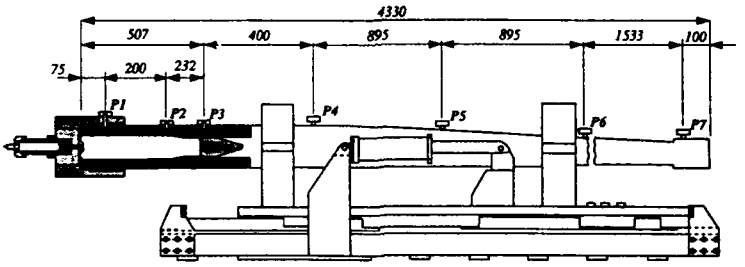


Figure 3. The 60-mm SPETC Gun.

The gun is machined to incorporate an ammunition cartridge. It is mounted on a cradle containing a recoil mechanism. A threaded breech housing allows the insertion of a breech block. The electrothermal plasma injector is contained in the breech block. The plasma generator is located at the base of the ammunition cartridge permitting axial plasma injection. A flexible electrical connection links the plasma injector and power supply. The power supply itself is comprised of several PFNs. Each PFN can be discharged into the load independently as long as impedance matching rules are used to guarantee a high transfer efficiency of electrical energy to the plasma injector load. A maximum pulse length of four milliseconds is possible.

Design rules have been developed to generate various pulse shapes as needed for different experiments. One common pulse shape begins with a low power, 100–300 MW, ignition pulse lasting for 2.5 ms, followed by an energy boost at a 1–2 GW power. An example of such a "staircase" pulsed power shape is shown in Figure 4.

In order to provide an efficient energy transfer, adequate impedance matching between the plasma injection unit and network is required. Impedance matching is done based on the plasma model[5] using network simulation programs and appropriate discharge parameters. When properly done, ninety-five percent of the stored electrical energy can be transferred to the discharge load.

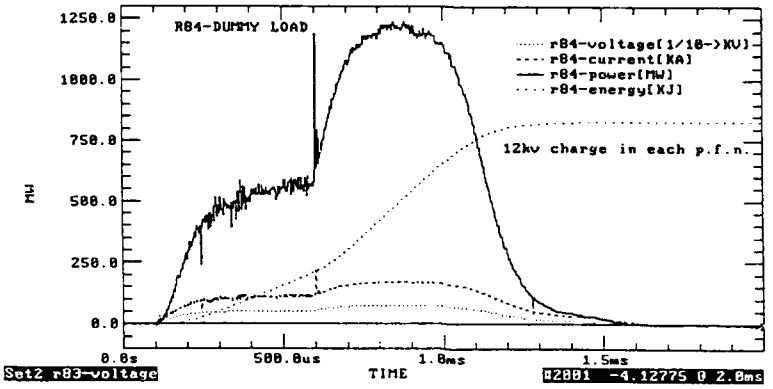


Figure 4. Discharge of Two PFNs to Generate a Staircase Pulse Shape Into a Restive Load.

The following diagnostics are used to characterize SPETC experiments. The electrical system is characterized for each PFN by monitoring the current, charging voltage and voltage across the load. Internal ballistics and projectile kinematics are mainly characterized by piezoelectric pressure gauges mounted along the chamber and barrel. These sensors provide the history of the pressure behavior at each location and the arrival times of the projectile at each pressure sensor in the barrel. External ballistic measurements are made with a Doppler microwave system.

5. THE INTERACTION BETWEEN PLASMA JETS AND SOLID PROPELLANTS

A variety of tests involving propellant type, mass and geometry, projectile mass, free volume in the chamber, and injected plasma characteristics were performed. Initial efforts concentrated on the development of controlled ignition and smooth ballistic pressure-time histories under plasma injection conditions. Subsequent efforts involved a study of the effects of plasma injection on the interior ballistic process. The data obtained have permitted a tailoring of conditions to obtain enhanced plasma-augmented gun performance.

Early experiments with plasma injected into a cartridge containing a granular propellant charge with no pre-arranged passage for plasma flow through the bed resulted in uncontrolled burning and resultant pressure waves. Figure 5 shows the results of such an experiment. A maximum breech pressure of 500 MPa, vs. the 200 MPa expected, was observed. This undesirable behavior is similar to conventional ballistic situations involving localized charge ignition.

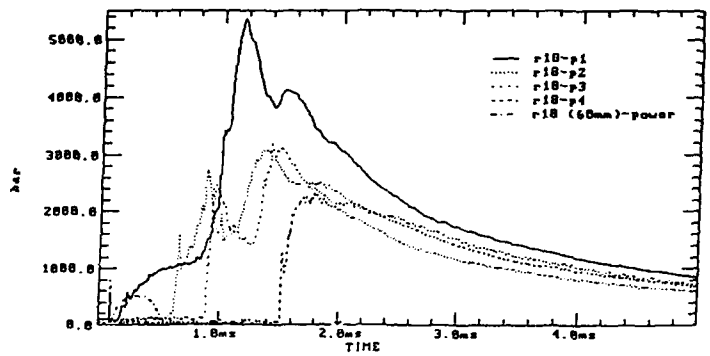


Figure 5. Uncontrolled Burning Developed When Plasma was Injected With no Passage for its Flow.

To avoid the uncontrolled burning phenomenon, plasma injectors were designed to imitate conventional center core igniter geometry. This yielded far smoother, more controlled ballistic processes. Ultimately, a low mass cellulose acetate (CA) tube, axially located in the cartridge to maintain space for the plasma to spread homogeneously through the charge, was adopted. This configuration yielded the smoothest most reproducible IB pressure-time behavior. The importance of these experiments is the proof that essential design concepts used in conventional solid propellant charges, i.e. avoiding localized ignition and providing for distributed energy transfer in the charge, are applicable to the plasma-ignited situation. This greatly facilitates control of the SPETC process.

Once these charge configuration effects on the plasma augmented ballistic process were understood, a series of experiments were conducted to learn about the influence of the amount of energy as well as the power of the injected plasma on the ballistic process. In line with intuitive expectations, both pressurization rate and maximum pressure attained were found to be functions of the power, duration, and total plasma energy injected into the charge. Moreover, the process proved to be well behaved, in fact quite large amounts of plasma energy, on the order of 1.6 MJ, have been safely injected into the 60-mm cartridges filled with solid propellant.

Effects of plasma energy on the IB process obtained using two types of igniter tubes are illustrated in Figure 6. The breech pressure curves of four different firings using the same amount of propellant are presented. Plasma injections of 300 and 500 kJ, respectively, were done using both metal-polycarbonate and low mass cellulose acetate igniter tubes. An increase in pressure and rate of pressurization is observed as the amount of injected plasma is increased, indicating

enhanced propellant burning as a function of increased plasma energy. The CA tube also led to faster pressurization. The effects of plasma power level on the interior ballistic process were found to be similar. That is to say, the rate of pressure rise as well as the maximum pressure attained in a given experiment significantly increased as the power level of the injected plasma increased.

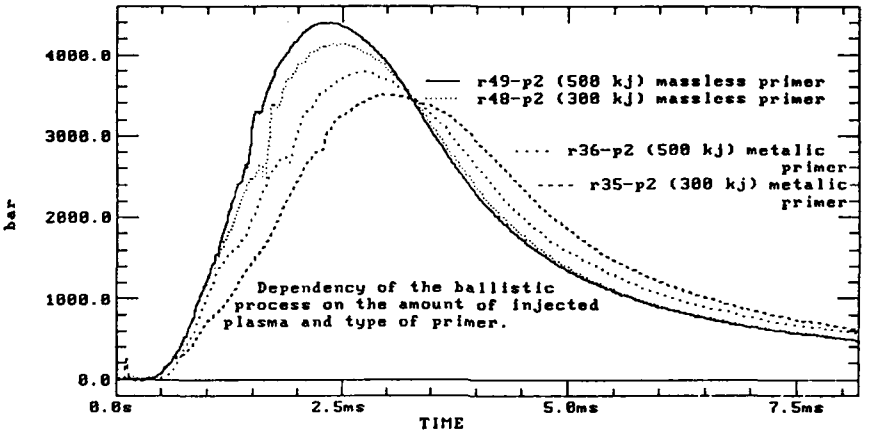


Figure 6. Effects of Primer Tube and Plasma Energy on the Ignition and Interior Ballistic Process.

An interesting series of observations were made concerning ignition delay times, that is, the time difference between the initiating pulse and the onset of pressurization of the plasma augmented ballistic processes. Under plasma injection conditions, ignition delays were found to be significantly shorter and more repeatable than under conventional black powder ignition conditions. In addition, as the amount of injected plasma increased, the delay between the trigger pulse and pressure onset decreased. In otherwise identical experiments involving injection of 34 and 150 kJ plasma energy, respectively, ignition delay times of 2 and 0.5 ms were noted. This is in sharp contrast to the 10-ms ignition delay time obtained using a black powder igniter.

In summary, the effects of plasma on the ignition and combustion properties of solid propellants seem to be well behaved. Plasma ignition appears to offer shorter, more precise ignition delay times than conventional ignition. In addition, increased power and energy levels of the injected plasma result in increases of both the pressurization rate and the maximum pressure attained in the gun. The significance of these findings is the experimental demonstrations of a new quantifiable mechanism for accelerating the IB process. In conventional IB this could only

have been done using either smaller webs or a faster burning propellant composition.

6. METHODOLOGIES FOR IMPROVED LAUNCH USING THE SPETC TECHNIQUE

Insights gained from the experimental data discussed above were used to develop the strategy for obtaining improved SPETC gun performance. The approach involved the delivery of plasma energy to the propellant charge at various stages of the interior ballistic cycle. Two factors of plasma jet technology, the high energy density of the plasma and the possibility of adding this energy in a precisely timed and controlled manner, provide the opportunity of influencing the IB process. In principle, the electrothermal energy can be added at the ignition phase, before maximum breech pressure is reached, synchronously with maximum pressure, or after maximum pressure is past. The basic idea governing the addition of plasma energy is to "widen" the interior ballistic pressure-time profile as close to the maximum pressure as possible, maximizing the area under the curve, hence the work that can be done on the projectile.

Several strategies could be used for injecting a boost of electrothermal energy to extend the maximum value of the pressure-time curve during the IB process. This injection mode may last as long as the projectile senses the resultant pressure rise. Most simply, the injection may be used to add energy to the gas in the chamber to maintain the maximum pressure value for a longer time, thereby enhancing its ability to do work on the projectile. Energy injected in this way would indeed improve performance, though somewhat less efficiently than the chemical energy of the propellant. Due to the size of PFNs, however, such an energy injection strategy may not be ultimately desirable to the user community. If the plasma energy is used to increase chamber pressure in later portions of the IB cycle to help burn propellant which would otherwise not have been consumed, however, the plasma energy could be used more efficiently to do work on the projectile. Such would be the case in using plasma augmentation to burn high density charges with "oversize" web propellant in the gun. Similar arguments may be made for energy addition according to a variety of injection profiles at other stages of the IB process.

Given the complexities of the interior ballistic cycle, mathematical tools simulating the SPETC process are necessary to work out the potential quantitative advantages to be gained from the various plasma injection strategies discussed.

Accordingly, the experimental efforts were conducted in cooperation with a theoretical study modeling the SPETC process. The modeling is based on a lumped parameter approach as in Corner.[6] The basic IB modeling addresses the mass and energy balance of the propellant and the gun; the propellant form function behavior, i.e. surface area history, and its relation to the rate of gas production; and the projectile dynamics and pressure gradient along the gun. In addition, other factors such as heat loss, friction, and chambrage effects are addressed. As a first approximation, it has been assumed here that the hot plasma jet injected into the propellant bed interacts with the burning propellant indirectly, by affecting the overall energetics of the system. Thus, in this model plasma injection appears as an additional term in the gas energy balance equation. Computations performed on this basis were sufficient to predict the basic behavior of the SPETC ballistics process.

7. EXPERIMENTAL RESULTS AND DISCUSSION

With the empirically based insights into the effects of plasma on the ignition and burning of propellant in hand and the theoretical IB model mentioned above, it became possible to address the issue of tailoring the plasma injection to enhance performance from a solid propellant ETC gun. Before discussing the data obtained, however, it is important to define a meaningful baseline for purposes of comparison. We begin with the premise that SPETC needs to provide tangibly improved performance over current conventional capabilities. Given the interest in using existing gun tubes for enhanced performance systems, the benchmark adopted was the greatest muzzle energy that could be delivered by the same gun, using the same projectile, at the same maximum pressure, using conventional ignition and solid propellant charges. This is referred to as optimized conventional performance. Based on these considerations, a muzzle velocity of 1,620 m/s was established as the baseline for the 60-mm gun parameters described in Table 1. The value was computed theoretically and verified against experimental firings.

The following set of three experiments presents experimental proof for the feasibility of controlling the IB pressurization process and enhancing muzzle energy by injecting a high power ET pulse at the time of maximum pressure.

Figure 7 is considered an electrical reference shot where a rectangular pulse of plasma at a power of 200 MW was injected to cause ignition. The pulse width is about four milliseconds, i.e. about one millisecond longer than the time to maximum pressure. This large pulse width was chosen purposely to keep the capillary

discharge "alive" during the IB process so as to allow injecting additional ET energy in subsequent experiments. The same propellant was used as in the firing in Figure 8. Due to the influence of the ET pulse on the combustion rate and maximum pressure, however, the propellant charge mass was reduced by 130 g in this firing.

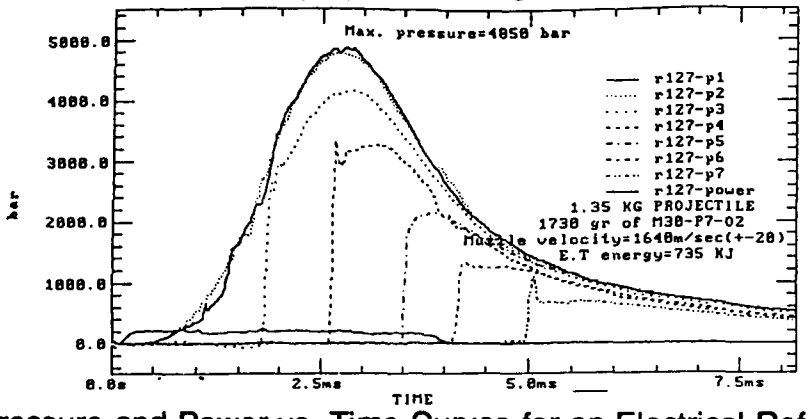


Figure 7. Pressure and Power vs. Time Curves for an Electrical Reference Shot.

Figure 8 describes the results of a firing in which a second boost of plasma energy, during the interval between 2.5 and 3.5 ms, was injected into the gun. The same propellant was used as in Figure 7 above. The muzzle energy for this particular shot was about four percent higher than for the reference shot. The influence of the ET boosted pulse on the breech pressure is shown in Figure 9. The broadening of the pressure-time curve after maximum pressure is seen to match, with a slight time offset, the stepwise injected plasma boost.

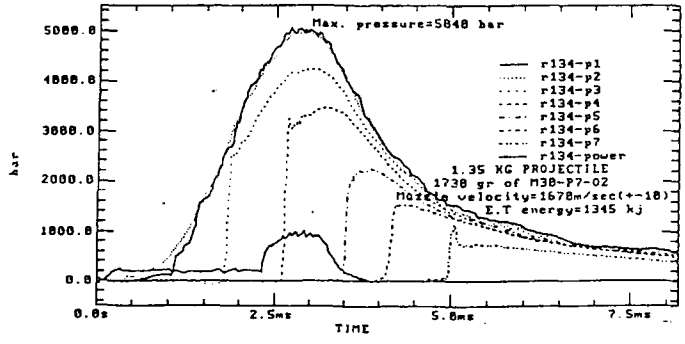


Figure 8. Pressure and Power vs. Time Profiles for an ET Boosted Shot.

Figure 10 demonstrates the best results achieved so far using the stepwise plasma injection synchronized to extend the maximum value of the pressure-time curve. The muzzle energy enhancement was 6% over optimized conventional.

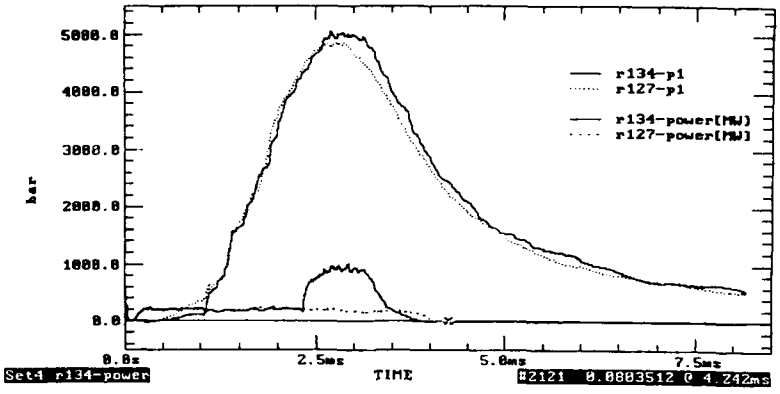


Figure 9. A Comparison of Breech Pressures and Plasma Power Curves for Rectangular vs. Stepwise Plasma Pulse.

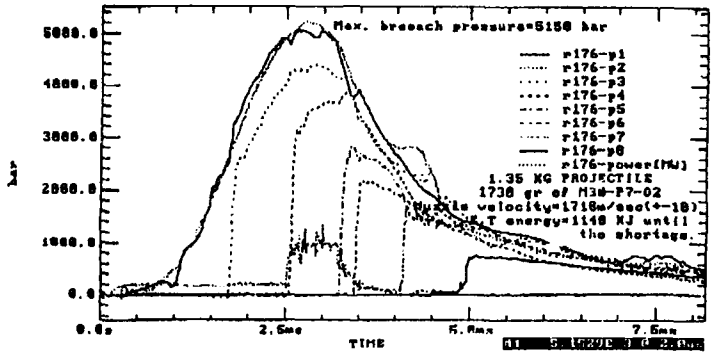


Figure 10. Pressure and Power vs. Time for Shot No. 176.

Another feasible methodology for achieving higher muzzle energy is based on using different propellant geometries, i.e. granulations and webs. Generally, larger web propellants would allow injecting ET plasma at earlier as well as later stages of the ballistic cycle. They may also permit the use of larger charge masses without exceeding permissible pressures in the gun.

To evaluate firing data in which the maximum pressure did not exactly match the baseline value discussed earlier, conventional IB computations were used to calculate a muzzle energy vs. maximum pressure envelope for optimum conventional performance from the 60-mm gun. This tool (the solid curve in Figure 11) could then be used to determine whether the performance from any individual

SPETC firing exceeded the performance potential of conventional solid propellant technology, independently of the precise pressure attained in a given test. The results of a number of SPETC firings are plotted with the optimum baseline curve in Figure 11. As can be seen, a number of shots fall clearly above the line, indicating that superior performance was achieved. At present, the best unequivocal performance increase documented is 6.8%.

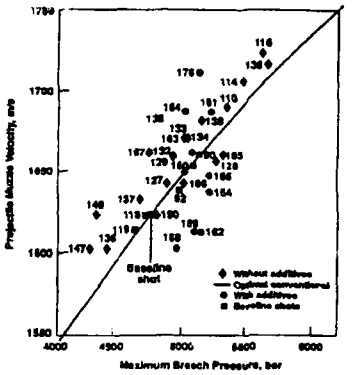


Figure 11. Muzzle Energy vs. Maximum Pressure, Comparing Theoretical Optimum Baseline Curve with Experimental Results.

8. SUMMARY AND CONCLUSIONS

The data presented above are highly encouraging. It is also gratifying that the key features of SPETC, i.e. the ability to safely inject large amounts of plasma energy into a solid propellant bed, the controlled positive effects of plasma power and injected energy on the pressurization process, and the effects of the plasma on reducing the ignition delay time of ballistic firings have been independently confirmed in the U.S.[7] The controlled results from both the present as well as the U.S. study strongly indicate that solid propellants can retain their physical integrity, hence their form function behavior, while burning under the influence of plasma injection. Results unique to the present study indicate that increased plasma energy injected during the latter portion of the ballistic cycle, using a stepwise profile, can cause desired changes in the shape of the IB pressure-time curves and attendant enhancements in ballistic performance. Moreover, theoretical results closely match experimental data. Current efforts are focused on increasing the amount of plasma energy injected into the gun and on evaluating a wider range of propellant grain types and webs. Given success with these efforts, a near term muzzle kinetic energy enhancement of 25–35% over the optimized conventional baseline may be possible.

ACKNOWLEDGMENTS

Technical discussions with Dr. K. White, Mr. W. Oberle, and Ms. G. Wren of BRL relating to experiment and theory are gratefully acknowledged. Funding was provided by the U.S. Strategic Defense Initiative Office, U.S. Army Laboratory Command, and the Israeli Ministry of Defense.

REFERENCES

1. LTG R. Hammond, "U.S. Army Strategic Defense Command Electrothermal (ET) Program," IAT Workshop on Electromagnetic Power Sources and Launchers, Austin, TX, 9 January 1991.

2. MAJ R. Lundberg, "ARDEC/FSAC Electric Armaments Program Information Briefing," IAT Workshop on Electromagnetic Power Sources and Launchers, Austin, TX, 9 January 1991.

3. A. Juhasz, J. Knapton, and K. White, "Process Control Issues in ETC Propulsion," CPIA Publication 557, Johns Hopkins University, G.W.C. Whiting School of Engineering, Columbia, MD, November 1990.

4. S. A. Goldstein, D. A. Tidman, R. L. Burton, D. W. Massey and N. K. Winsor, "Cartridge Containing Plasma Sources for Accelerating a Projectile," U.S. Patent 4,715, 261, December 29, 1987.

5. A. Loeb and Z. Kaplan, "A Theoretical Model for the Physical Processes in the Confined High Pressure Discharges Electrothermal Launchers," *IEEE Transactions on Magnetics*, vol. 25, p. 342, 1989.

6. J. Corner, *Theory of Interior Ballistics of Guns*," Wiley, NY 1950.

7. H. McElroy, J. Greig, and A. Juhasz, "Plasma Augmented Ball Propellant Ballistics," Proceedings of the 5th International Propellants and Propulsion Conference, Picatinny Arsenal, Dover, NJ, November 1991 (in press).

KATECHIS, WALKER

DORMANCY FOR THE GROUND BASED INTERCEPTOR

James C. Katechis
Project Manager, Ground Based Interceptor Program
U.S. Army Strategic Development Command

John J. Walker
SPARTA, Inc., Huntsville, Alabama

The Ground Based Interceptor, known by its short name of "GBI", is one of the key Elements in the National Missile Defense (NMD) Segment of the Global Protection Against Limited Strikes (GPALS) System. Deeply involved in research and development activities for the US Army Strategic Defense Command, GBI is making great strides in becoming a weapon system that will be able to defend the United States against Intercontinental Ballistic Missiles (ICBMs) and Submarine Launched Ballistic Missiles (SLBMs) by intercepting hostile reentry vehicles (RV) in the midcourse of their flight. Pictured below is an illustration of a NMD battle scenario.

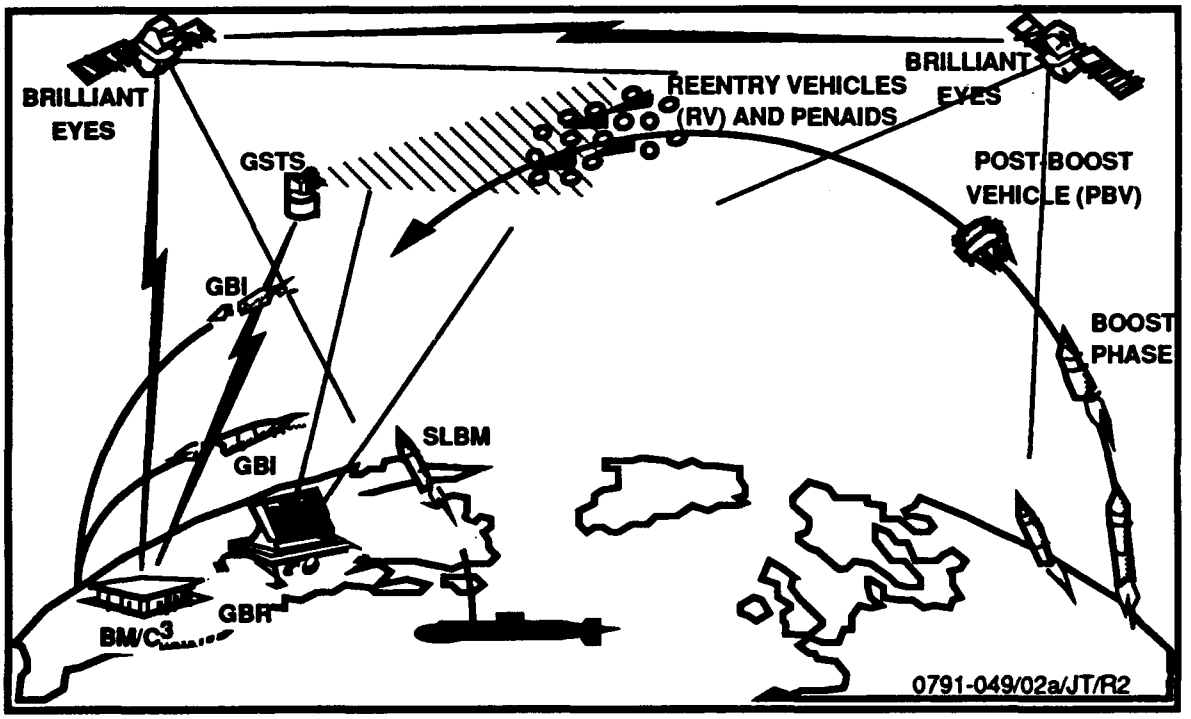


Figure 1. Notional NMD Battle Scenario

KATECHIS, WALKER

OPERATIONAL OVERVIEW

GBI is the interceptor of the National Missile Defense System concept and will consist of an air vehicle, launch site, ground support equipment, personnel and support services. All subsystems of the defense system will be linked together through the Command and Control Element (CCE) to efficiently coordinate the use of sensor and weapon assets.

The GBI operational engagement scenario is based on early warning sensor subsystems detecting, identifying, and designating hostile RVs to the maximum extent possible and transmitting the tracking data to the CCE. The CCE determines the path of the target based on this data from the surveillance and tracking systems and plans a GBI engagement. Launch commit parameters are sent from the CCE to the GBI base, and prior to launch, trajectory parameters are down-loaded from the launch site to a specific interceptor. After launch, the first stage booster separates (approximately 30 seconds into flight) and initiates the second stage booster. The kill vehicle (KV) separates as the second stage burns out (approximately another 30 seconds) and then repositions, pointing the seeker to the predicted target position.

During the flight, divert maneuvers are performed to adjust the trajectory for boost errors and to respond to in-flight updates of the target's track. As the seeker is uncapped, the target and any associated objects will be acquired. The target is designated using pattern matching algorithms or onboard target selection capability. Once the target has been designated, the KV transitions to track mode and performs its final trajectory adjustments. In the terminal homing mode the KV selects an aimpoint on the target body in order to maximize its probability of kill. Just before impact, a kill enhancement device may be deployed for added insurance of collision and lethality. Destruction of the target is accomplished by physical impact with the KV. The intercept event, which can occur from one to ten minutes of flyout, can be monitored by radars and tracking systems for final kill assessment action if required.

The tempo is picking up for GBI. Currently in the final stages of a Functional Technology Validation (FTV) program, GBI is preparing for deployment as an element of our Nation's strategic missile defensive system. With the passage of the Missile Defense Act of 1991 by Congress, the GBI has been tasked to be ready to deploy to an initial NMD site by the 1996-97 timeframe. Leveraging off its predecessor program, the Exoatmospheric Reentry Vehicle Interceptor Subsystem, the latest mature technologies are being assembled to develop an interceptor that is highly reliable and totally mission responsive. In that regard, GBI is looking to "dormancy" as an interceptor concept to meet weapon system high reliability objectives while at the same time driving down life cycle costs.

DORMANCY

A design goal since the beginning of the GBI program, dormancy is being implemented as a means to achieve high reliability standards, to assure that the missile will function as required when needed. Dormancy is also being used as a basis to drive down total life cycle costs by minimizing operations and support requirements needed to sustain GBI in the field. Defined under the dormancy

concept as a maintenance free missile, GBI will not be operated, tested or have power applied in the field except at launch. The interceptor when deployed, will be used and stored like a round of ammunition and incorporated into a certification program (to assure weapon performance high reliability). The interceptor's operability will be validated through stockpile reliability testing which includes comprehensive component testing and live annual service practice firings (figure 2).

By incorporating technical advancements into the interceptor design, GBI is becoming a missile that will not require to be kept up-and-running in order to meet immediate launch requirements. Essentially, the interceptor remains in an off condition until committed by the operations center for battle. When committed and activated, the interceptor goes through quick preparatory procedures--conducting a series of health checks, loading flight data, aligning--and launches, all within a matter of seconds. By remaining dormant until launch, parts are not used, batteries are not activated, components are not incurring failures, and consequently, the interceptor reliability is enhanced and supportability requirements are minimized.

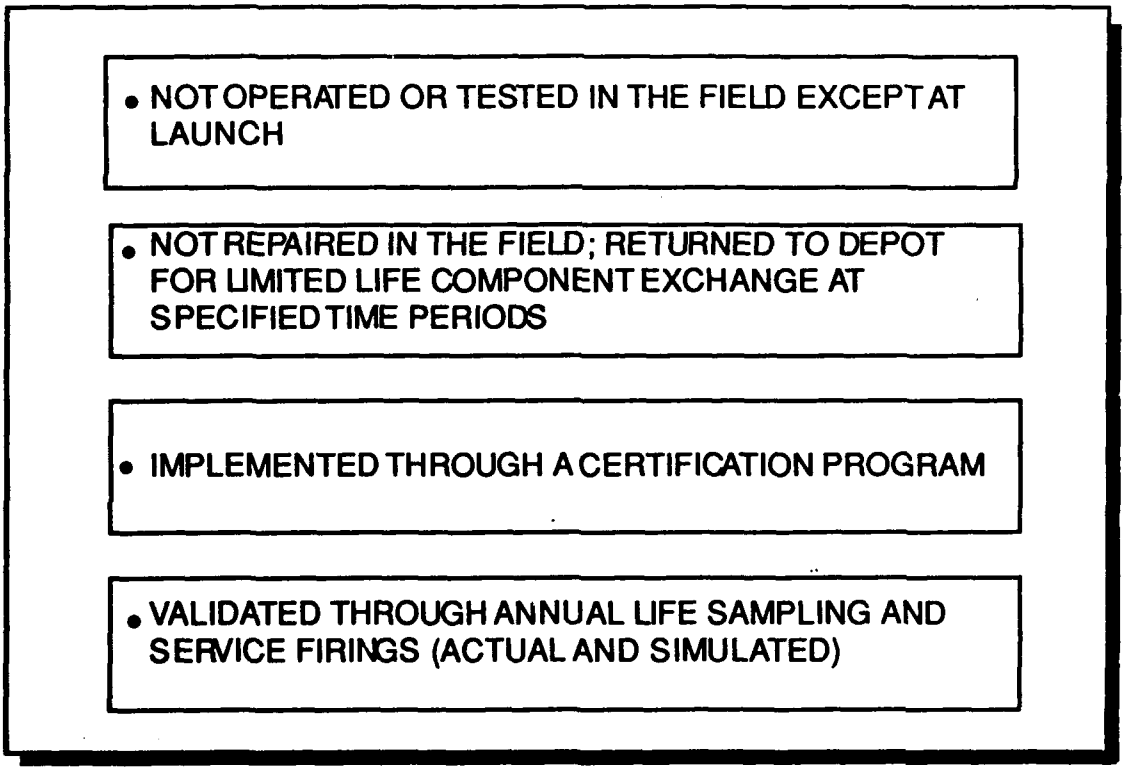


Figure 2. GBI Dormancy

Dormancy offers realistic and immediate potential for high payoffs. From the standpoint of operational readiness, dormancy through the certified round program, facilitates upfront investment of highly reliable parts and components, successive testing and burn-in processes, and a responsive data base which can yield accurate component degradation forecasts. As an O&S cost enhancement, dormancy eliminates the need for spares and repair parts, test sets and facilities to perform highly technical supportability requirements in the field. Perhaps the biggest GBI O&S cost saver will be realized in reducing the requirement for skilled technicians to perform the direct and general support logistics requirements at the deployed site. It is envisioned that GBI once deployed, will be not be stressed from the tactical sense (ground mobile), but will be battle ready in a dormant fixed site environment.

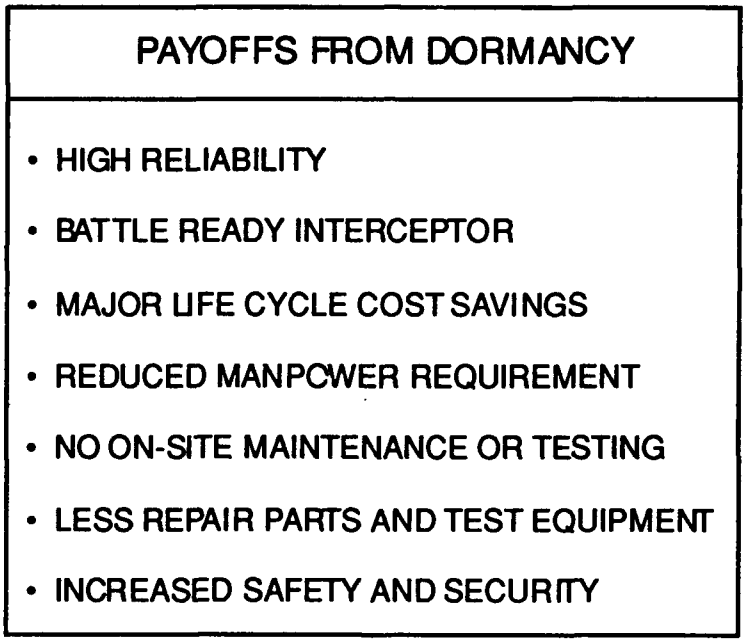


Figure 3. Dormancy Benefits

DORMANCY TECHNICAL CHALLENGE

Dormancy is being implemented through a number of technological advancements being developed and tested in the GBI program. One of the most significant achievements has been the ability to cool the seeker, which will search for hostile reentry vehicles in space, to extremely low temperatures while on the rise during the boost phase of flight. As the interceptor begins its launch sequence, gaseous nitrogen is circulated to bring the seeker components down to a very low temperature. By the time the interceptor finishes its multi-stage boost sequence, the components are cooled and ready for target acquisition functions. Other technical enhancements are being implemented -- quick and accurate alignment, long life and high reliability components, upfront reliability testing, and others -- to meet GBI's dormancy objective.

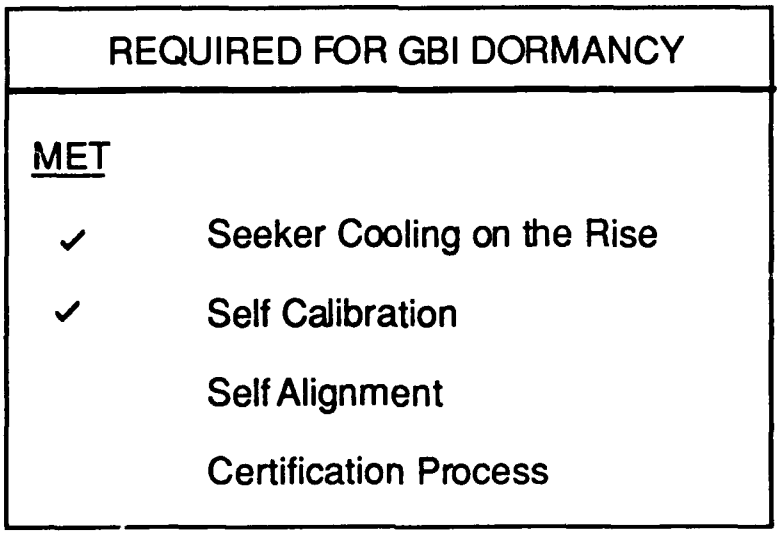


Figure 4. Dormancy Technical Requirements

The dormancy task is not an easy one to implement for the interceptor. The major technical challenge undertaken was to be able to cool an onboard seeker to very low temperatures in order to meet mission requirements. Although technology was available to solve this problem by hooking up to a cryogenic cooler (refrigeration type) unit as a part of the ground launch support equipment, this defeated the cost-saving dormancy concept.

With the standard refrigeration type cooling process, one can vividly see the never ending logistics requirements growing: resupply of the cooling material, purging of the cooling lines, maintaining an adequate vacuum, requirements for a clean room, high failure component replacement and repair, interceptor disassembly, on-site direct and general support maintenance, calibration, quality testing, special test equipment, and on and on. And most importantly, the woes of having an interceptor that would require constant and dedicated maintenance attention in a field environment and have the potential for breeding other evils that would degrade the reliability and ability of the interceptor.

The fix was a new technology requirement. After much thought and design interaction, an inflight cooling capability was placed on the drawing board, designed, fabricated and was validated during a test flight last year. The design specification for a cooled seeker is still there, but this process to cool it has been changed. Cooling will be done inflight, negating the needs for a continuously operating cryogenic unit. As the interceptor begins its launch sequence, gaseous nitrogen will be circulated to bring the seeker components down to the very low Kelvin temperature. By the time the interceptor finishes its multistage boost sequence, the components will be cooled and ready for target acquisition functions.

Other technical design remedies are also being developed to maintain the dormant interceptor objective. The seeker is being produced under design

conditions that will facilitate long calibration shelf life. In the case of seekers produced for the FTV experiments, the seekers have remained in accurate calibration without having to be recalibrated (over 24 months). Another case is the *interceptor inertial measuring unit (IMU). Initial research* concluded that the interceptor internal gyro would have to be continuously spun-up for the missile to be ready for a quick reaction launch. A solution to this dilemma was found by incorporating plans for a cost effective fiber optic gyro that will meet GBI's dormancy and critical response requirements. Here, the IMU will be in an "off" state, yet ready to perform the critical alignment mission. Unwanted noise generated by transition to its operative state will be minimized by being housed in climatically controlled launch stations. And case after case, design recommendations continue to be brought forward that resolve technology issues -- calibration, reliability, integration -- in order to support the dormancy goal.

The GBI approach to dormancy is to minimize stress. A single power "on/off" can be as much as 1000 times more stressful or looking from another viewpoint, effective in causing failures than one hour of dormancy. The key is selecting parts

- Select Parts that are Less Prone to Failure During Dormancy
- Reduce Sensitivity and Stress Environments
- Implement Strict Contamination Control
- Reduce Out-of-Control Manufacturing Process

Figure 5. Approach to GBI Interceptor Dormancy

that are less prone to failure during dormancy, reducing their inherent sensitivity to stress environments, implementing strict contamination control processes and reducing out-of-control manufacturing processes such as contamination, lifted bonds, dielectric breakdown, fluid leakage, moisture, inadequate cleaning, loose conduct particles, inadequate burn-in, etc.

EVOLUTION OF THE DORMANCY CONCEPT

The dormancy concept began back in the late 1970s when the US Army was experiencing poor system readiness levels with the HAWK and Shillelagh Air Defense missile systems. A task force was formed consisting of Government product assurance technical experts from the US Army Missile Command HAWK Project Office who were charged to evaluate the poor readiness levels and recommend a quick, but lasting fix of the problem. What the task force found were a number of significant influencing environments.

First, that extended missile run cycles during increased readiness alerts induced failures. Commanders were bringing up their missiles to full battle status and conducting checks to insure operability. These actions were causing parts to

wear and ultimately, inflicting failures in the missiles. Secondly, the missiles were being subjected to numerous tests -- annual training tests, annual General Inspections, maintenance tests, Brigade, Group and Battalion operational readiness exercises, and just routine practice tests -- causing much stress and strain to activate and practice maintaining the missiles, taking their toll on system operability and availability. Third, by being maintained in the field under less than desirable repair environments, faults were being induced by repairmen who were ironically charged with bringing the missiles back to an operable status. Lastly, the entire maintenance concept required a significant investment of repair parts and test equipment at unit, direct support, general support and depot levels of repair as well as the skilled technicians to conduct the maintenance.

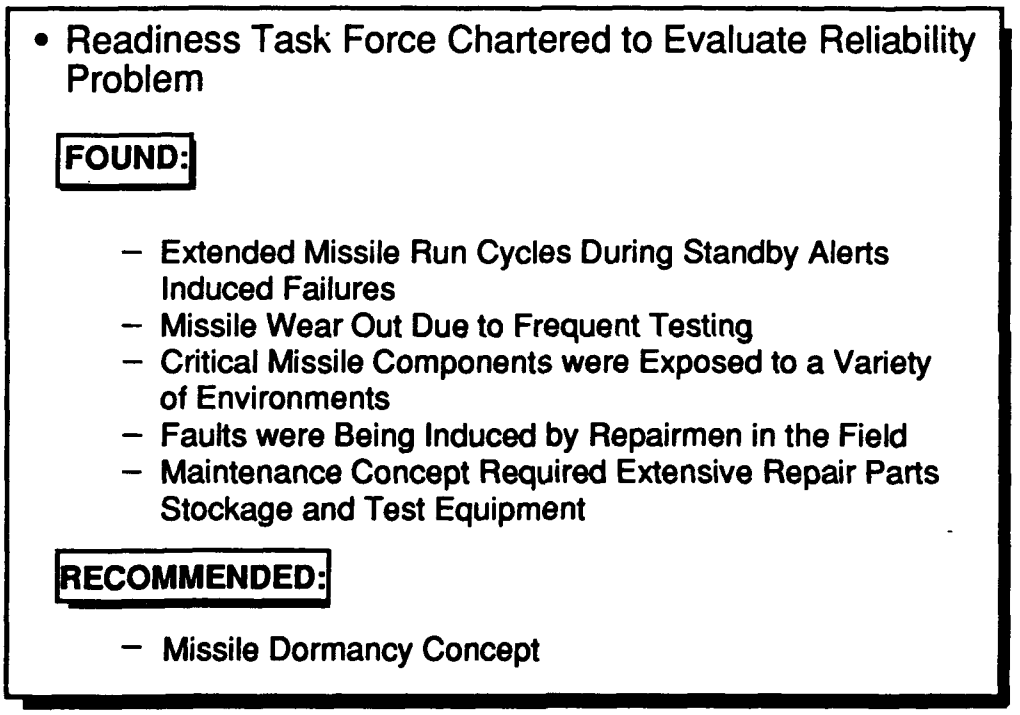


Figure 6. Evolution of the Dormancy Concept

KATECHIS, WALKER

The task force recommended a number of actions which essentially led to the dormancy concept for an "Improved HAWK":

- Designed to Remove Undesirable Characteristics
- New Characteristics
  - Solid State
  - Quality of Components
  - Longer Life Components
  - Stability
  - Reduced Operating/Testing Time
  - Eliminate Field Testing
  - Curtail Missile Disassembly and Modification in an Uncontrolled Field Environment

The results of the changes to the HAWK program were immediate. Readiness rates began to soar, less maintenance was being required and personnel reductions could be made. Also, tangible results were found in the live annual service practice firing tests. Without conducting missile preparation before flight tests, HAWK intercepts (kills) reached unprecedented highs.

But there is more to the success story than just significant readiness improvements. The dormancy concept resulted in significant cost savings. Figure 7 illustrates the savings realized in military personnel, repair parts, facilities and test equipment which cumulatively reduced O&S costs by about half of the original commitments.

BASIC HAWK AND IMPROVED HAWK COST COMPARISONS

AVERAGE COST/MISSILE INVENTORY/YEAR

	BASIC HAWK	IMPROVED HAWK	
Military Support Personnel	$2,316	$1,044	For TRME and CONUS Mgt Personnel
Repair Parts Support	446	418	Mgt and Hardware
OME/FME Amortization	347	106	(TRME) Equipment
DS/GS Building Amortization (20 Yrs)	17	106	(TRME)
Overhaul by CONUS Depot	1,500	1,400	Factory Restoration/ Shelf Life Changeout
	$3,373/Yr/MSL	$1,673/Yr/MSL	

Figure 7. Before and After Cost Comparisons

But the real proof of the concept was that no HAWK unit has been placed in a non-operational status due to the lack of ready-to-fire HAWK missiles for the last 15 years.

PATRIOT, another dormant Air Defense System, has also maintained exceptionally high readiness rates validated by numerous annual services tests and the "combat" firings during Desert Storm. CNN broadcasts from Saudi Arabia and Israel verified time-after-time that dormancy was a winner for operational readiness by PATRIOT scoring successive successful intercepts of SCUD missiles.

CERTIFICATION PROCESS AND STOCKPILE RELIABILITY

Key to implementing the dormancy concept is the ability to validate the interceptor's reliability and availability through time. Reading the preceding section on dormancy one cannot help but ask the questions: How can we feel comfortable from the military point of view that the interceptor will launch and perform the intercept as required? Or more specifically, how do we know that the processors will function unless we test them?

The answer is the GBI's certified round program coupled with a responsive stockpile reliability program. Using this concept, the interceptor is validated through an up-front, comprehensive production testing and documentation process coupled with a stockpile reliability program that randomly tests and launches interceptors each year. This program is based on the precepts used in the Army's HAWK and PATRIOT certified round programs which have produced and maintained highly reliable missiles, year after year, validated by "nonprepped" test launches. Dormancy is not new to the missile field. In fact many of the Army's missiles are dormant under the wooden round concept: Pershing, Stinger, Multiple Launch Rocket, Chapparal, TOW, Hellfire, Dragon and others. They are produced to rigid production standards and flight tested for lot acceptance and also are tested throughout their life through a random sampling stock reliability program.

However, for complex missile systems, an added feature of certification is added to provide for detailed component testing on a routine basis throughout the system's life. This affects design, manufacturing and testing at all levels to forecast component degradation. In the case of GBI, plans are being formulated to have the GBI's certified round program require each piece part, component, and assembly to be manufactured, tested and documented to very rigid standards. The center-piece of this concept is that all major components are 100% functionally tested to assure they meet design specifications. Test data will be recorded for future reference in forecasting failure rates and determining life expectancy of the fielded missile.

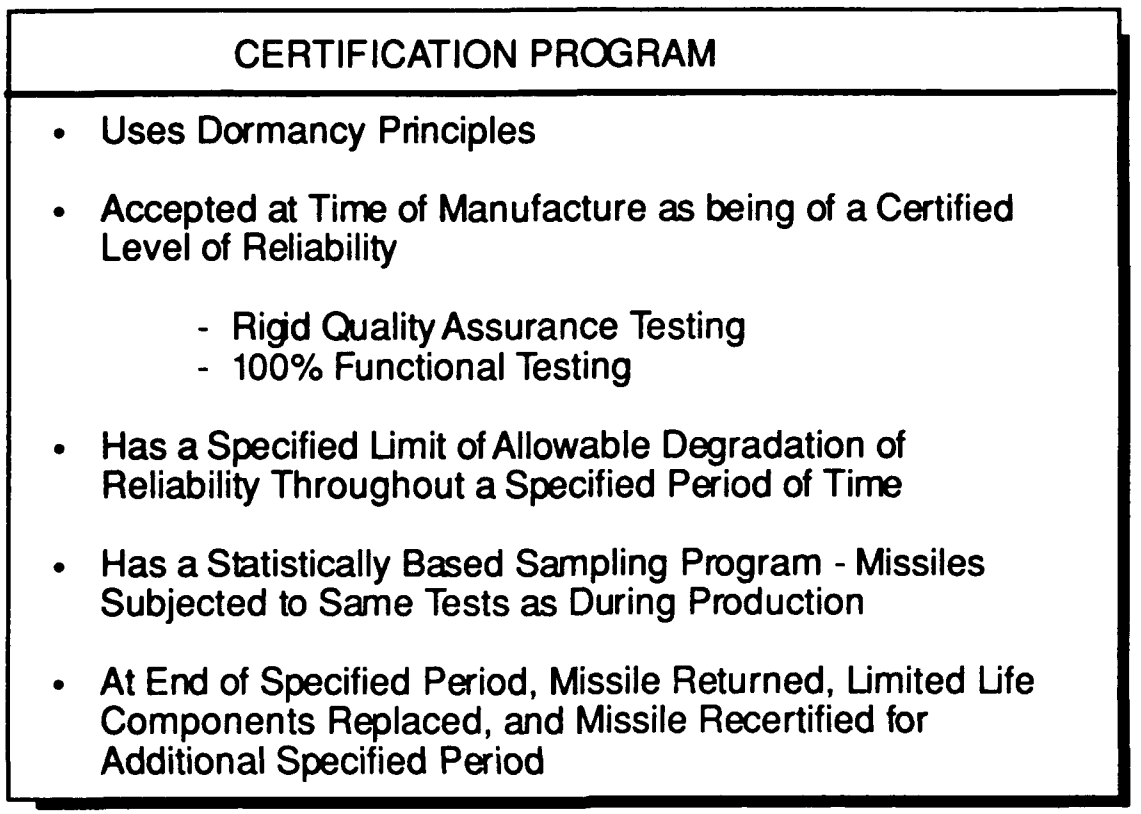


Figure 8. Certified Round Program

After a very rigid quality assurance program and deployment to the field, a random sample of a given number of interceptors will be recalled each year. Evacuated back to the GBI depot, they will be disassembled, and tested functionally similar to production acceptance tests and procedures. For example, the seeker will be disassembled from the kill vehicle assembly and placed in a test environment where a simulated scenario of the flight, engagement and intercept will be conducted. In the case of boosters, they will be physically inspected in detail and may be subsequently x-rayed to determine propellant abnormalities and degradation. Information received from these and similar tests will be compared with the original production acceptance test data and any other sampling test data to forecast the life of the material. After the functional and physical tests are completed, limited life components will be replaced as required, and the interceptor reassembled, recertified, and returned to the field for duty.

Stockpile reliability also plays a vital role in determining interceptor reliability. Under a random sampling selection process, a number of interceptors will be flight tested each year. Here, interceptors will be flown against threats in scenarios that emulate an actual engagement. Using added telemetry equipment, detailed flight and engagement data will be gained on the interceptors total flight process. Components will be functionally tested as they would in a true battle scenario.

LIFE CYCLE COST SAVINGS

GBI looks to realize significant savings by implementing the dormancy concept. Recognizing there are "no free lunches," GBI has been and is prepared to continue to make up-front investments to produce a mission capable dormant interceptor. Figure 9 illustrates the expected values to be realized across the board of a dormant and non-dormant system.

BROAD COST COMPARISON PARAMETERS
DORMANT VS NON-DORMANT

	COST COMP		COST VALUE
	DORMANT	NON-DORMANT	
MISSILE COST	↑	↓	M
UP FRONT TESTING	↑	↓	M
STOCKPILE RELIABILITY PROGRAM VS PERIODIC TESTING	↓	↑	H
TEST EQUIPMENT (1 VS 2 LEVEL)	↓	↑	H
REPAIR PARTS	↓	↑	M
LEVEL OF REPAIR (1 VS 2)	↓	↑	M
UPGRADED MANUFACTURING	↑	↓	M
FACILITIES	↓	↑	H
MIS DATA	→	→	M
DIAGNOSTICS-BIT	↓	↑	M
TRANSPORTATION	↓	↑	M
MANPOWER	↓	↑	H
FLOATS	↓	↑	L

COST COMP		COST VALUE
Same	= →	H = High
Increase	= ↑	M = Medium
Decrease	= ↓	L = Low

Figure 9. Dormancy Cost Expectations

KATECHIS, WALKER

SUMMARY

In conclusion, the dormancy concept has already validated significant life cycle cost savings for a GBI deployment. Technically, GBI has completed many functional validation requirements to being able to field a dormant interceptor. Cooling on the rise, the fiber optic gyro, long lasting calibration for the interceptor's seeker, and others have made it possible. GBI will be able to field without the traditionally large logistics force to maintain, supply and support the weapon system. Particularly in these austere times for military manpower and scarce O&S funding, GBI will be able to strategically defend the Nation against hostile RVs at a very affordable cost.

High-Power Impulse Generation Using a Photoconductively Switched Radial Transmission Line

*A. Kim (Dr.), R. Zeto (Dr.), R. Youmans (Mr.), M. Weiner (Dr.)
U.S. Army Electronics Technology and Devices Laboratory
LABCOM, Fort Monmouth, NJ 07703-5601
L. Jasper (Mr.)
U.S. Army Harry Diamond Laboratories
LABCOM, Adelphi, MD 20783-1197

I. Introduction

The free-space transmission of a short-duration pulse, namely, an impulse, produces a frequency spectrum that extends from several megahertz to several gigahertz. The use of this wide frequency spectrum offers new and unique system applications[1,2] in the areas of remote sensing, communication, electronic countermeasures, and directed energy weapons. Since the frequency spectrum is widely spread, the energy allotted per frequency interval is small. To overcome this limitation, the amplitude of the generated pulse becomes critical. Hence, the successful generation of a reliable and well-defined, high-peak-power impulse, with pulse repetition frequency capability, is a first step toward advancing this emerging ultrawideband technology.

Recently, significant progress has been made in the generation of high-peak-power pulses[3] with the use of an optically activated semiconductor switch. Pulses with subnanosecond risetime and multikilovolt amplitude have been successfully generated.[4-7] For the generation of fast-risetime (<1 ns), high-peak-power pulses, the major concern is the capability and performance of the switch, rather than the energy storage element. The subnanosecond switch turn-on time, in which the switch changes from a high-impedance state (switch-off state) to a conductive state (switch-on state), enables the stored electrostatic energy in the energy storage element to be delivered to the load impedance as a pulse with subnanosecond risetime. In the case where the recovery time is shorter than the pulse forming line (PFL), the switch recovery time determines the output pulsewidth. The output pulse amplitude depends on the switch hold-off voltage capability, the design of the

energy storage element, impedance matching conditions at the junction, and the effects of the stray inductance and stray capacitance.

Semi-insulating GaAs is characterized by a high-voltage hold-off capability as well as fast risetime, making GaAs switches a desirable candidate for fast, high-peak-power pulses. Normally, GaAs also exhibits a recovery in the nanosecond range, making it possible to produce narrow pulses. However, under high electric field strength, the GaAs switch exhibits a lock-on phenomenon.[7-9] This switching behavior results in a long switch conduction time with a long recovery tail, thus precluding the generation of a narrow pulse at high voltages.

One way to overcome the long recovery time in the GaAs switch is to employ a short section of PFL, so that the pulsewidth is determined by the two-way wave transit time in the PFL. The desired electrical pulsewidth can be generated from the energy in the PFL by use of miniaturized transmission lines or microstrip lines. In this mode of energy transfer, as soon as the switch is turned on, the storage element delivers all or most of the stored energy to the matched load so that a narrow pulse is obtained. However, the miniature physical dimensions may severely deteriorate the pulse output. In particular, the diminished size of the PFL emphasizes the stray capacitance and the stray inductance of these lines, which may result in a slower risetime or lower pulse amplitude, or both. Further, only one-half of the bias voltage, at most, is delivered to a matched load impedance.

Using a radial transmission line (shown in fig. 1) as the energy storage medium offers several advantages compared to the typical microstrip line or coaxial line structures. The geometrical configuration of the radial transmission line creates a large electrode area, so that a large electrostatic energy can be stored. In addition, the impedance transformation due to the impedance variation between the inner and outer radii of the radial transmission line produces an output voltage gain compared to the constant impedance transmission line. The concept of nanosecond pulse generation using a radial transmission line[10] has been successfully demonstrated.

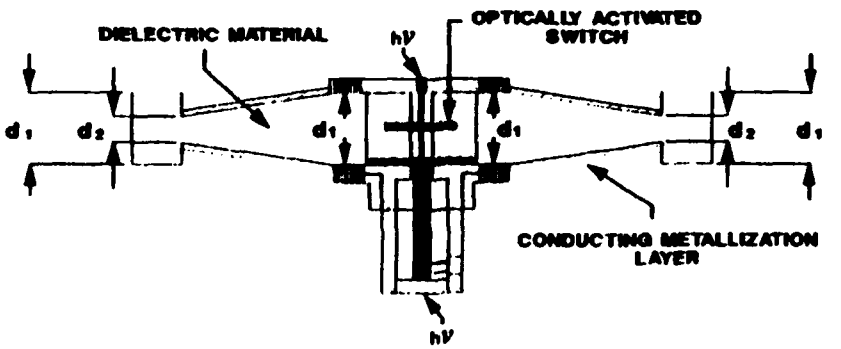


Figure 1. Diagram of photoconductive impulse-generating device using a tapered radial transmission line.

In this paper we investigate the two critical parameters affecting the pulse output, namely, the energy storage element and the photoconductive switch. The energy storage element, in the form of a tapered radial transmission line, produced reasonably well-defined pulses with nanosecond pulsewidths as well as output voltage enhancement. The voltage gain factor of the nonlinear radial transmission line was analyzed using the computer program SPICE. An equivalent circuit model, which considers parasitic parameters, was also developed and simulated using SPICE. Switching properties of the photoconductive GaAs switch with two opposite gridded electrodes were obtained with an electro-optic imaging system.[11] These properties were incorporated into the SPICE simulation. The output waveforms were measured by pulse biasing the device up to 50 kV and triggering the device with an optical pulse from a mode-locked Nd:YAG laser.

II. PHOTOCONDUCTIVE IMPULSE-GENERATING DEVICE USING A TAPERED RADIAL TRANSMISSION LINE

2.1 General Description

The diagram of the impulse-generating device using a tapered radial transmission line is given in figure 1. The tapered radial transmission line consists of two circular conducting plates, separated by a dielectric material. Unique to this tapered structure is an extremely large impedance variation. Because of the geometry of the radial line, the characteristic impedance is lowest at the outer edge of the radial line and increases toward the center. This large impedance transformation plays an important role in enhancing the output voltage.[12]

The tapered radial transmission line structure performs a dual role: it is a capacitor during the charging cycle and a nonuniform impedance transmission line during the discharging cycle. The capacitance is proportional to the square of the radius of the metallized electrode, but inversely proportional to the thickness of the dielectric medium in a given material. Hence, the capacitance can be substantially enhanced by enlarging the physical dimensions of the radial transmission line structure. However, a tradeoff exists because as the radius of the transmission line structure is increased, the pulsewidth increases due to the two-way wave transit time increase for the traveling wave.

2.2 Tapered Radial Transmission Line

The impedance of the nonuniform radial transmission line (fig.1) may be approximated if we use formulas derived for the linear microstrip line and correct for the differing geometries. In particular, for a given radius, we assume that the

corresponding circumference is equivalent to the width of the microstrip line. The characteristic impedance of the radial line[13] is

$$Z = (377 \times d)/[\sqrt{k} \times 2 \times \pi \times r] \ , \qquad (1)$$

where d is the thickness of the dielectric medium, k is the relative dielectric constant, and r is the distance from the center of the radial line to the electrode. As shown in equation (1), the characteristic impedance of the radial transmission line is inversely proportional to the radius of the electrode. Hence, the characteristic impedance at the outer edge is smaller than the characteristic impedance at the inner edge of the radial line. The microstrip impedance formula, equation (1), is most accurate when $d << 2\pi r$.

A closed-form simple solution for the output pulse from a radial transmission line[14] may be derived by treating the radial transmission line as a perfect step-up transformer. In this model the voltage gain factor, originating from the impedance variation between the inner and outer edges of the radial line, is linearly proportional to the square root of the ratio Z_{ir}/Z_{or}. Under certain conditions, where the impedance ratio is small (<2), this solution does allow one to make reasonably good estimates of the peak pulse output. However, this model ignores the substantial internal reflections due to the impedance variation.

The more accurate technique relies on the use of a SPICE simulation, in which the radial transmission line is simulated by considering sections of transmission lines with differing impedance levels, connected in series. This technique considers the internal reflections, and it is usually satisfactory provided the sections are chosen small enough to maintain accuracy.

The tapered radial transmission line may be regarded as a composite structure combining two types of transformation. The first transformation is linear, and is due to the linear variation of the thickness of the radial line. The second type of transformation follows the $1/r$ variation of the impedance, where r is the radius of the radial structure.

The transformer is divided into N sections of equal length, which connect the inner impedance, Z_{ir}, and the outer edge impedance, Z_{or}, which represents the lowest impedance. Each section is assumed to have a constant impedance.

The discrete impedance values for a 10-section tapered radial transmission line were obtained, with the innermost radius $r_{ir} = 2.25$ cm, the outermost radius $r_{or} = 13$ cm, inner height $d_{ir} = 3$ cm, outer height $d_{or} = 1$ cm, and $k = 2.07$. The discrete impedance values Z_n (where $n = 1, 2, ..., 10$) for the inner most impedance, Z_1, to the outermost impedance, Z_{in}, are 43.3, 29.2, 21.1, 16.0, 12.3, 9.71, 7.67, 6.07, 4.76, and 3.2 Ω, respectively.

2.3 *Photoconductive GaAs Switch with Two Opposite Gridded Electrodes*

In general, the high-power photoconductive switches can be categorized into two types: (1) surface switch and (2) bulk switch. The term "surface switch" applies to a switch in which the active carrier generation mechanism occurs at the surface of the switch. In contrast, the "bulk switch," with two opposite electrodes, generates the carriers inside the material. This paper is concerned with the bulk type of switch. In this switch geometry, the laser light is coupled into the switch active area from two opposite sides. The bias electric field strength is parallel to the incident triggering optical pulse. A unique feature of this switch geometry is nonuniform photon-generated carrier distribution.

Three illumination measurements were conducted: illumination of both sides simultaneously, as well as successive illumination of the high field side and the low field side (ground side). For the case of illumination from both sides, the laser light (200 μJ) from a mode-locked Nd:YAG laser was introduced into both sides of the switch. The one-sided illumination was obtained by blocking the appropriate beam; thus half of the energy was applied. The obtained transient electric field profiles, which show a sequential field collapse in time and space, are given in figure 2. The two-sided illumination produced the fastest electric field collapse, as expected.

2.4 *Equivalent Circuit Model And SPICE Simulation*

The photoconductive impulse-generating device (fig. 1) has a coaxial cable leading to the load impedance. The interconnections among the photoconductive switch, the tapered radial structure, and the outer coaxial cable leading to the load impedance all create parasitic parameters such as stray capacitance and stray inductance. In addition, there is the existence of an on-stage switch resistance in the semiconductor.

Parasitic parameters are categorized according to the location. The stray capacitance between the top electrode of the radial line and the switch, designated C_1, is parallel to the on-state switch resistance. The stray capacitance between the switch and outer coaxial cable, designated C_2, is parallel to the load impedance R_{on}. The stray inductance L is connected in series to the on-state switch resistance. The equivalent circuit model, if we consider the stray inductance, stray capacitances, and on-state switch resistance, is shown in figure 3.

First we simulated the output waveform with no parasitics, as a basis of comparison. A SPICE-simulated output pulse waveform with $R_{on} = C_1 = C_2 = L = 0$ is shown in figure 4. Next, figure 5 shows the simulated waveform with $L = 1$ nH, $C_1 = 2$ pF, $C_2 = 0$, and $R_{on} = 5\ \Omega$, in which the switch resistance dropped from 30 MΩ to 50 Ω in 100 ps, and from 50 to 5 Ω in 400 ps. Another simulated waveform is given in

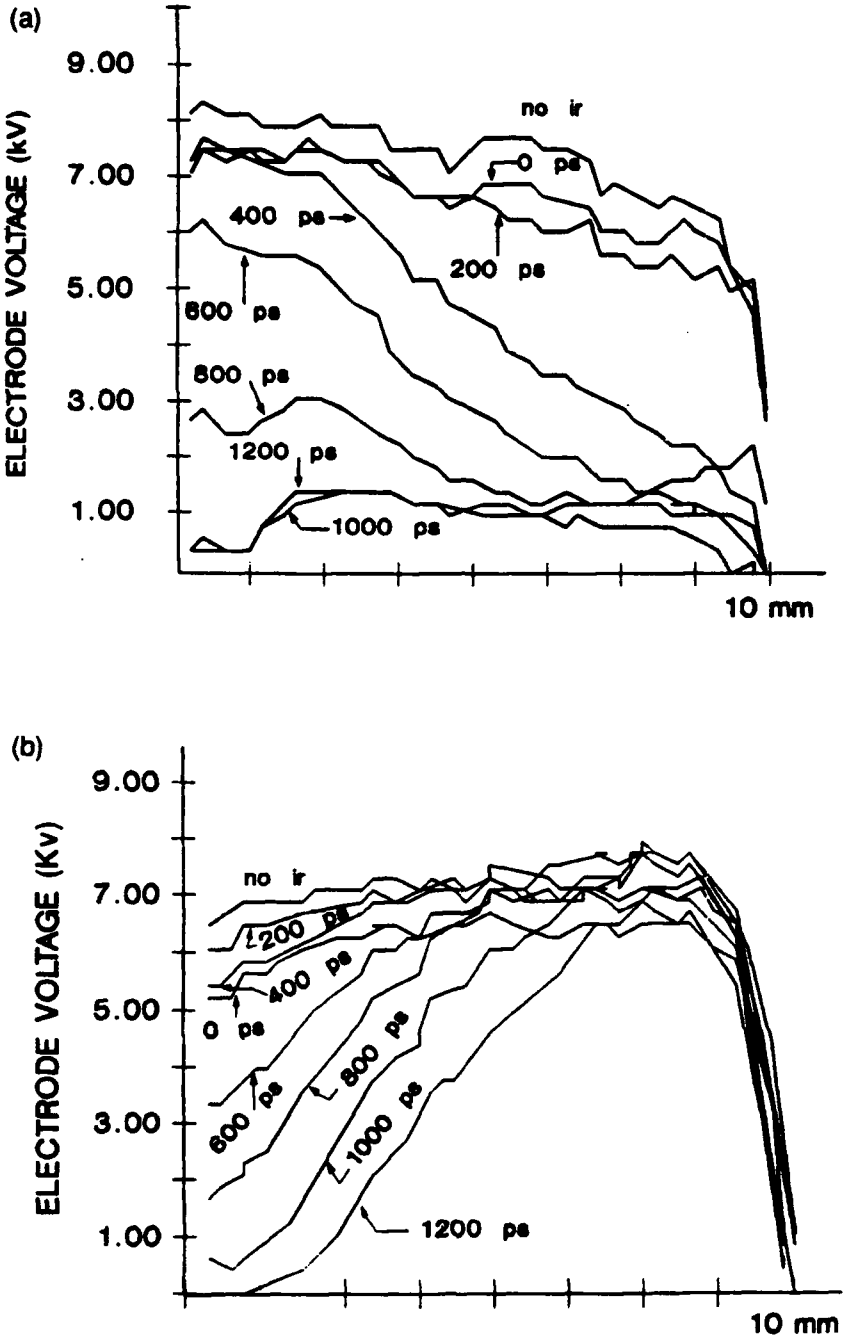


Figure 2. Transient electric field profiles obtained by an electro-optic imaging system: (a) illumination on both sides and (b) high-voltage-side illumination only.

figure 6, in which the switch resistance dropped from 30 MΩ to 50 Ω in 100 ps, and 50 to 5 Ω in 300 ps. When the voltage across the switch fell below a value of 10 kV, the switch resistance was allowed to recover to 50 Ω in 1 ns and then recover to 30 MΩ in 1 ns.

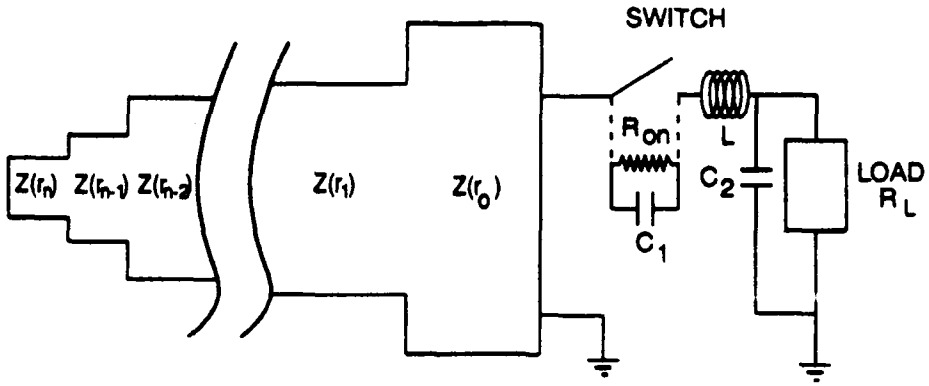


Figure 3. Equivalent circuit model for photoconductive impulse-generating device using a tapered radial line structure including stray reactance elements and switch resistance.

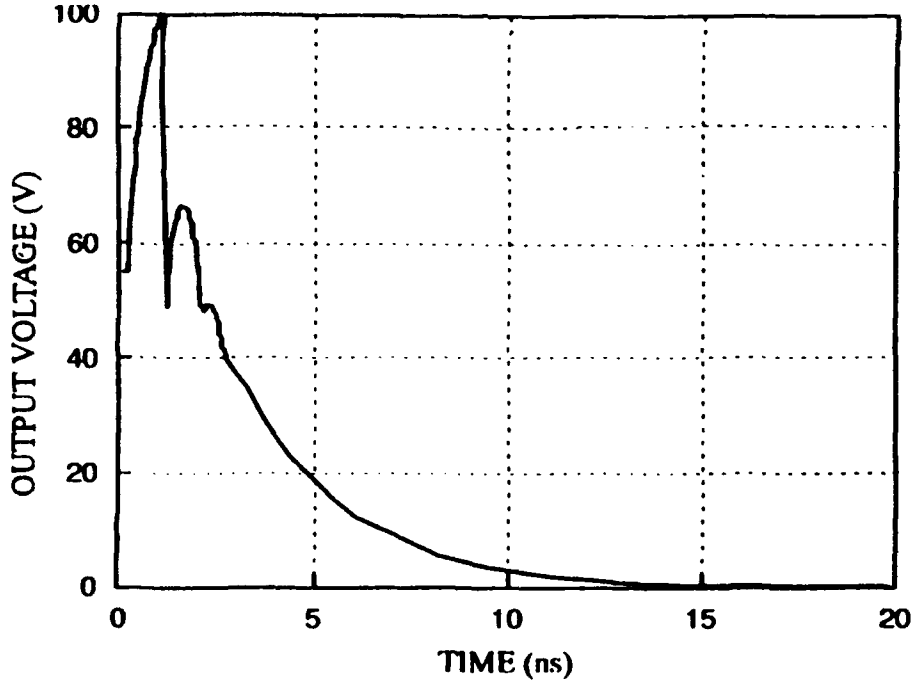


Figure 4. SPICE-simulated waveform obtained by assigning parasitic parameter values $R_{on} = 0$, $C_2 = C_1 = 0$, and $L = 0$ into a circuit model.

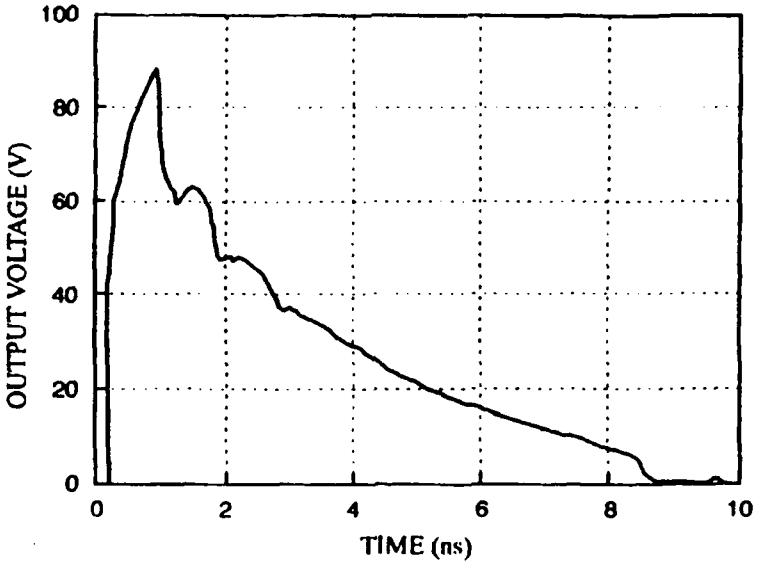


Figure 5. SPICE-simulated output waveform assigned by parasitic parameters R_{on} = 5 Ω, C_2 = 0, L = 1 nH, and C_1 = 2 pF (R_{on} changes from 30 MΩ to 50 Ω in 100 ps and from 50 to 5 Ω in 400 ps). When field across switch becomes less than 4.3 kV, switch resistance recovers to 50 Ω in 1 ns and to 30 MΩ in 2 ns.

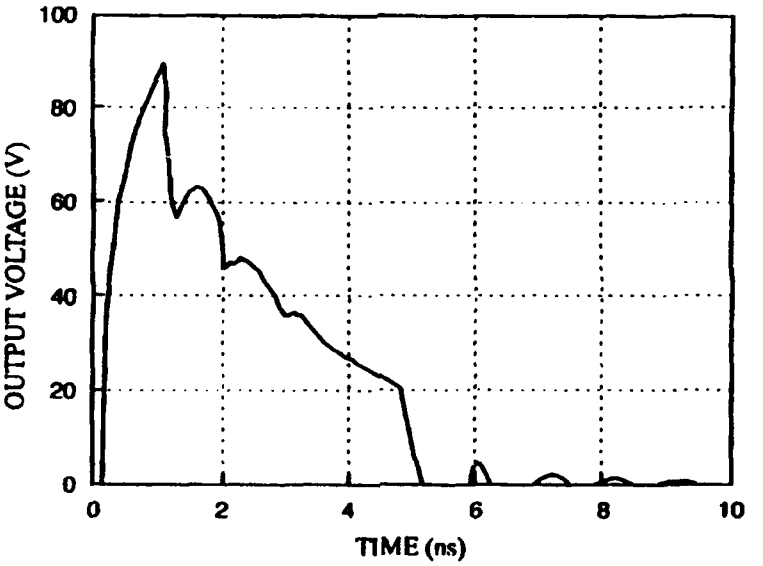


Figure 6 SPICE-simulated output waveform assigned by parasitic parameters R_{on} = 5 Ω, C_2 = 0, L = 1 nH, and C_1 = 2 pF (R_{on} changes from 30 MΩ to 50 Ω in 100 ps and from 50 to 5 Ω in 300 ps). When field across the switch becomes less than 10 kV, switch resistance recovers to 30 MΩ in 1 ns.

III. EXPERIMENTAL SETUP AND RESULTS

Surface flashover of the photoconductive switch under conditions of extremely high bias voltage made it necessary for the switch to be housed in an oil-filled cavity.

The tapered radial structure was fabricated by tapering down Teflon material from 3-cm thickness at a 2.25-cm radius to 1-cm thickness at 13-cm radius. Thus, the impedance variation between inner and outer edges was increased by a factor of 3, compared to the radial line with uniform thickness. The calculated output pulsewidth of the radial transmission line is approximately 1.3 ns. However, because of the relatively large transformation in the tapered radial transmission line with the above dimensions, the device is expected to have a wider pulsewidth than 1.3 ns, mainly due to internal reflections. If we substitute parameters such as $d_{ir} = 3$ cm, $d_{or} = 1$ cm, $k = 2.07$, $r_{ir} = 2.25$ cm, and $r_{or} = 13$ into equation (1), the characteristic impedances of the tapered radial transmission line at the inner and outer edges are 55.6 Ω and 3.2 Ω, respectively. With a load of $Z_L = 50$ Ω, the device was slightly negatively mismatched.

The test setup for the photoconductive impulse generator is shown in figure 7. A 1060-nm, mode-locked laser pulse, approximately 75 ps in pulsewidth, generated by a Quantel YG-501C Nd:YAG laser system, was delivered to the switch through the fiber-optic bundle. The output optical energy from each fiber-optic bundle was approximately 1.0 mJ.

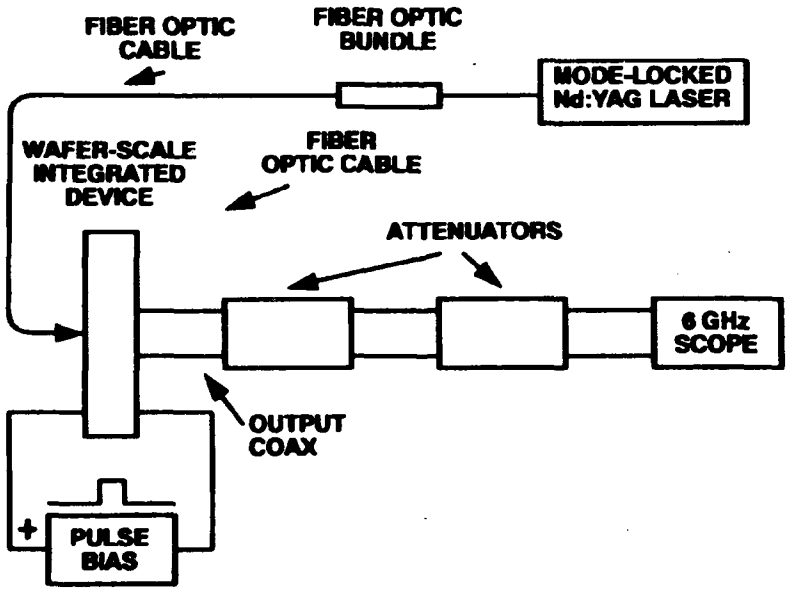


Figure 7. Experimental setup for photoconductive impulse-generating device.

The measurement sequence began by pulse charging of the radial line using an SCR-transformer circuit. After completion of the charging process, the laser was triggered, thereby initiating the switching action. The generated pulse was delivered to a specially constructed 50-Ω dummy load impedance. The dummy load had matched 50-Ω carbon resistors in series with a current viewing resistor (CVR), which had a 180-ps risetime. The signal from the CVR was conveyed via a delay line to a 6-GHz Textronix 7250 oscilloscope. The obtained output pulse, at 50-kV bias voltage to the tapered radial structure, is shown in figure 8. This pulse contains approximately 40 MW of peak power.

IV. DISCUSSION

The estimated capacitance based on the charged voltage waveform is about 80 pF. The impedance variation between the inner and outer radii of the tapered radial transmission line is about 17. With the tapered radial transmission line, a large capacitance and large impedance variation with miniaturized physical dimensions were realized.

Without parasitics and with almost instant switching action, large output amplitudes (with bias voltage set to 100 V), shown in figure 4, were obtained from SPICE simulation. With parasitic parameters (fig. 5 and 6), the output voltage obtained from using a bias voltage of 100 V is about 85 V. Therefore, the voltage gain factor amounts to about 1.7. Further, the SPICE-simulated waveform (fig. 6) indicates a significant role of lock-on threshold voltage during switch recovery. When the voltage across the switch reaches a certain threshold level, recovery of the switch resistance accelerates. The photoconductive impulse generator biased at 50 kV, and

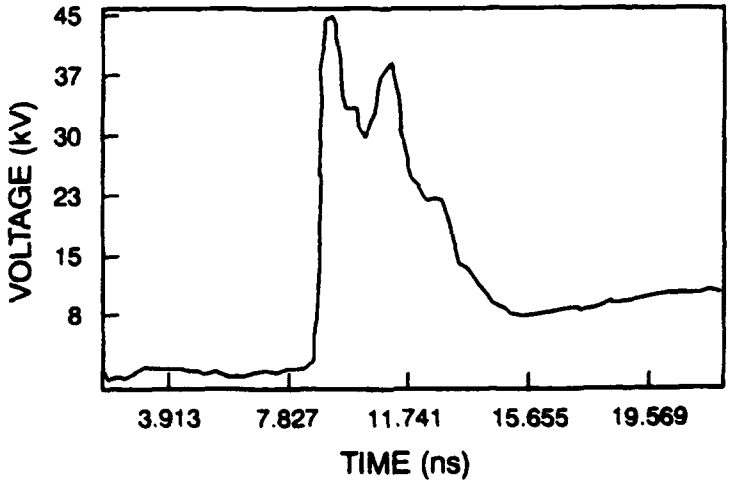


Figure 8. Obtained output pulse at 50-kV pulse bias voltage.

subsequently triggered by optical pulses from a mode-locked Nd:YAG laser, delivered a 44-kV output pulse to the 50-Ω load impedance. The experimental gain factor of the obtained waveform (fig. 8) was about 1.7. The SPICE simulation and experimental voltage gain are the same.

The measured risetime of 400 ps is actually larger than the true value. By taking into account the bandwidth of the various sensors, attenuators, etc., the actual risetime is estimated to be about 250 ps. The extremely fast risetime is believed to stem from the unique switch geometry, which may be operating in an avalanche switching mode. The output pulsewidth is about 4 ns, which is somewhat larger than the results of the SPICE simulation. Some of this discrepancy of the output pulsewidth may result from wave dispersion in the dielectric medium and imperfection of the model, as well as parasitics in the device. However, the main reason for pulse widening may be attributed to the mechanism of trapped energy in the PFL (due to the internal reflections), in combination with switch lock-on.

V. CONCLUSION

The unique properties of the tapered radial transmission line combined with a photoconductive GaAs switch with two opposite gridded electrodes enabled the generation of 40-MW pulses with about 250-ps risetime, 500-ps falltime, and pulsewidths of about 4 ns. A voltage gain of nearly two was obtained, in agreement with a SPICE simulation of the device. The tapered radial transmission line structure, which has a high energy storage capability, is uniquely suited for a compact, high-power, nanosecond impulse generator.

ACKNOWLEDGEMENT

The authors acknowledge J. Conrad and M. Wade for the switch fabrication.

REFERENCES

1. W. Scott, "UWB radar has potential to detect stealth aircraft," *Aviat. Week Space Technol.*, 38–41 (December 1989).
2. W. Scott, "Report critical of impulse radar triggers controversy," *Aviat. Week Space Technol.*, 18–20 (November 1990).
3. W. Nunnally and R. Hammond, "80-MW photoconductor power switch," *Appl. Phys. Lett.* **44** (10), 980–982 (1984).
4. G. Loubriel, F. Zutavern, W. Helgeson, and M. Buttram, "Light activated semiconductor switches for subnanosecond, high power impulse generation," *Proc. 5th Nat. Conf. High Power Microwave Technology* (West Point, NY) , June 1990.

5. M. D. Pocha and R. L. Druce, "35 kV GaAs subnanosecond photoconductive switches," *IEEE Trans. Electron Devices*, **37**, No. 12, 2473–2492 (December 1990).
6. G. Loubriel, W. Helgeson, D. McLaughlin, M. O'Malley, F. Zutavern, A. Rosen, and P. Stabile, "Triggering GaAs lock-on switches with laser diode arrays," *IEEE Trans. Electron Devices*, **38**, No. 4, 692–695 (April 1991).
7. A. Kim, M. Weiner, M. Dornath-Mohr, M. Wade, R. Youmans, R. Zeto, and G. Vezzoli, "Photonic switching and the lock-on phenomenon in single crystal semi-insulating GaAs," *Proc. Integrated Optical Circuit Engineering VI*, SPIE **993**, 137–140 (1988).
8. F. Zutavern, G. Loubriel, M. O'Malley, L. Schanwald, and D. McLaughlin, "Recovery of high-field GaAs photoconductive semiconductor switches," *IEEE Trans. Electron Devices*, **38**, No. 4, 696 (1991).
9. R. Brinkmann, K. Schoenbach, D. Stoudt, V. Lakdawala, G. Gerdin, and M. Kennedy, "The lock-on effect in electron-beam-controlled gallium arsenide switches," *IEEE Trans. Electron Devices*, **38**, No. 4, 701–705 (1991).
10. A. Kim, M. Weiner, R. Zeto, R. Youmans, L. Jasper and B. Lalevic, "Photoconductive nanosecond pulse generation utilizing radial transmission lines," *IEEE Trans. Electron Devices*, **37**, No. 12, 2506–2510 (December 1990).
11. W. Donaldson, L. Kingsley, M. Weiner, A. Kim, and R. Zeto, "Electro-optic imaging of the internal fields in a GaAs photoconductive switch," *J. Appl. Phys.* **68**(12) (December 1990).
12. R. A. Petr, W. C. Nunnally, C. V. Smith, Jr., and M. H. Clark, "Investigation of a radial transmission line transformer for high-gradient particle accelerators," *Rev. Sci. Instrum.*, **59**, 132–136 (January 1988).
13. F. Assadourian and E. Rimai, "Simplified theory of microwave transmission systems," *Proc. IRE*, **40**, 1651 (December 1952).
14. A. Kim, Ph.D. thesis, Rutgers–The State University of New Jersey, Ch. 2, January 1992.

Monolithic, Photoconductive Impulse-Generating Device

*A. Kim (Dr.), R. Zeto (Dr.), R. Youmans (Mr.), M. Weiner (Dr.)
U.S. Army Electronics Technology and Devices Laboratory
LABCOM, Fort Monmouth, NJ 07703-5601
L. Jasper (Mr.)
U.S. Army Harry Diamond Laboratories
LABCOM, Adelphi, MD 20783-1197

1. Introduction

Picosecond switching is of interest to the Army for generating narrow-band pulses with picosecond risetimes for synthetic aperture radars (SARs) and smart high-power microwave (HPM) jammers. The photoconductive, semi-insulating GaAs switch is a candidate for generating the desired pulsewidths because of its high-voltage hold-off capability, fast risetime, and large peak-power-handling capability. The low-jitter property of the switch, due to laser activation, gives the capability of phase arraying large numbers of switches for achieving gigawatt power levels at high repetition rates.

Recently a monolithic, photoconductive GaAs switch was described.[1] As shown in figure 1, the switch consisted of a circular GaAs wafer, typically 7.5 cm in diameter, with metallized disc electrodes on each side. On the high-voltage side, an aperture at the center was etched out so as to provide access to the semiconductor for the light pulse, which, by producing carriers in the central region, causes this region to transition from an insulating state to a conducting one. On the opposite, grounded side, an annular region close to the center region was etched out, providing contact to an output coaxial transmission line, which was pressure fitted to this side of the wafer. Once the light was introduced, a centrally located conducting channel between the disc electrodes was created, so that the energy stored between the electrodes was delivered to the output coaxial line in the form of a narrow pulse. One may view the device as a radial transmission line, which forms the energy-storing element. Once the light pulse produces the conducting channel, the radial line becomes connected to the output coaxial line, to which the pulse energy is

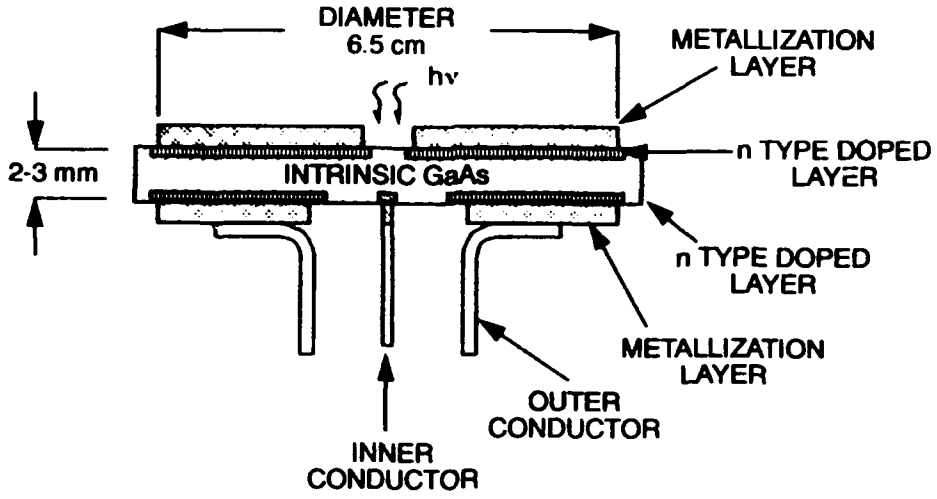


Figure 1. Schematic diagram of a monolithic, photoconductive GaAs impulse device.

delivered. The light source has a much narrower pulsewidth (75 ps) than the generated pulse, so as not to have any effect on the circuit response.

An important property of the radial structure has to do with the impedance transformation[2,3] that occurs as the wave proceeds from the outer to the inner radius of the structure. At the outer perimeter, the radial impedance is small, whereas the impedance increases near the center, varying as $1/r$, where r is the radius. If Z_o and Z_i are the characteristic impedance levels at the outer and inner radii, then for small transformation ratios ($Z_i/Z_o < 2$), a simple approximation[4] for the peak output voltage V_{out} applies. For a matched load, with no loss in the switch, V_{out} is

$$V_{out} = (V_o/2)(Z_i/Z_o)^{1/2}, \qquad (1)$$

where V_o is the charging voltage. For small transformation ratios, $(Z_i/Z_o)^{1/2}$ may be regarded as a gain factor (compared to the output of a uniform transmission line). For larger transformation ratios, however, more exact SPICE simulation shows that the gain factor saturates very quickly at a value equal to about two, so that the output itself saturates at the value $V_{out} = V_o$.

As the impedance variation between device impedance at the inner radius (Z_i) and the load impedance (R_L) increases, the output waveform more and more resembles that of an rf capacitor, with the change accompanied by a widening of the pulse but with an increase in the risetime. The resemblance to capacitor output becomes complete for a positive mismatch at the output, i.e., $R_L > Z_i$, where R_L is the output impedance. In this case, the pulse does indeed exhibit a fast risetime followed by an R_LC decay, where C is the capacitance of the radial structure. Figure 2

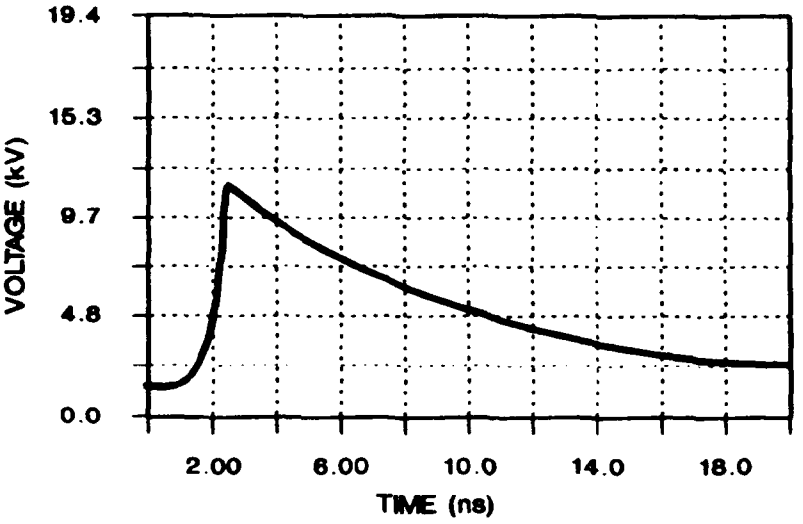


Figure 2. Output waveform obtained with a 3-mm-thick GaAs device. Switch is biased to 4 kV and triggered by a mode-locked Nd:YAG laser. Output is delivered to 50-Ω load impedance.

shows the experimental output for the radial device shown in figure 1, where Z_i is 6 Ω and R_L is 50 Ω, so that a positive mismatch prevails. The waveform in figure 2 closely resembles that of a discharging capacitor. Results of the SPICE simulation agree with the experimental waveform. For most experiments, the positive mismatch is typical for the high-power GaAs wafer device. This is because it is very difficult to obtain 50-Ω output device impedance levels with the GaAs wafer device, unless one resorts to extremely small inner radii (<1 mm). Such small radii became impractical because of voltage breakdown and problems associated with fabricating the transition to the output coaxial line.

Unfortunately, the relatively wide nanosecond pulsewidths obtained with the radial line limit the usefulness of the device for many applications where subnanosecond pulses are desirable. In this investigation, we observe a new mode of operation for the monolithic, photoconductive GaAs wafer device, which yields subnanosecond pulsewidths. With this new mode, the current is turned off in times much shorter than the R_LC decay time. The device, biased to 6 kV, produced pulses as narrow as 500 ps and risetimes as fast as 200 ps, with an output voltage of 5 kV. Narrower pulses at considerably higher voltages are foreseen.

2. Experimental Procedures and Results

The experimental procedures were similar to those reported previously,[1] for which a 50-Ω load was used. The laser light (either Nd:YAG laser or laser diode array) was conveyed to the GaAs wafer aperture by means of a fiber-optic bundle. The measurement sequence began with pulse charging of the radial line via an SCR-transformer circuit. Upon completion of the charging process, the laser was fired and the generated pulse was delivered to a specially constructed, high-voltage coaxial line. The pulse was attenuated using large-bandwidth Barth attenuators, and the final pulse delivered to a Tektronix SCD-5000 oscilloscope. For loads smaller than 50 Ω, a wideband current-viewing resistor (CVR) was used to monitor the pulse.

Figures 3 and 4 show the output of a 2-mm-thick device when biased to 4 kV and illuminated with a 75-ps-wide, 1-mJ pulse from an Nd:YAG laser. The inner and outer radii were 0.8 and 3.25 cm, corresponding to characteristic impedance values of 4 and 1 Ω, respectively. Figures 3 and 4 are for 50- and 10-Ω loads, respectively. Unlike that of the 3-mm-thick device, the 2-mm device output exhibits periodic peaks, separated by intervals of about 1 ns. The widths of the peaks themselves are less than 1 ns wide. In figure 3, which at 50 Ω has the greater mismatch, the dis-

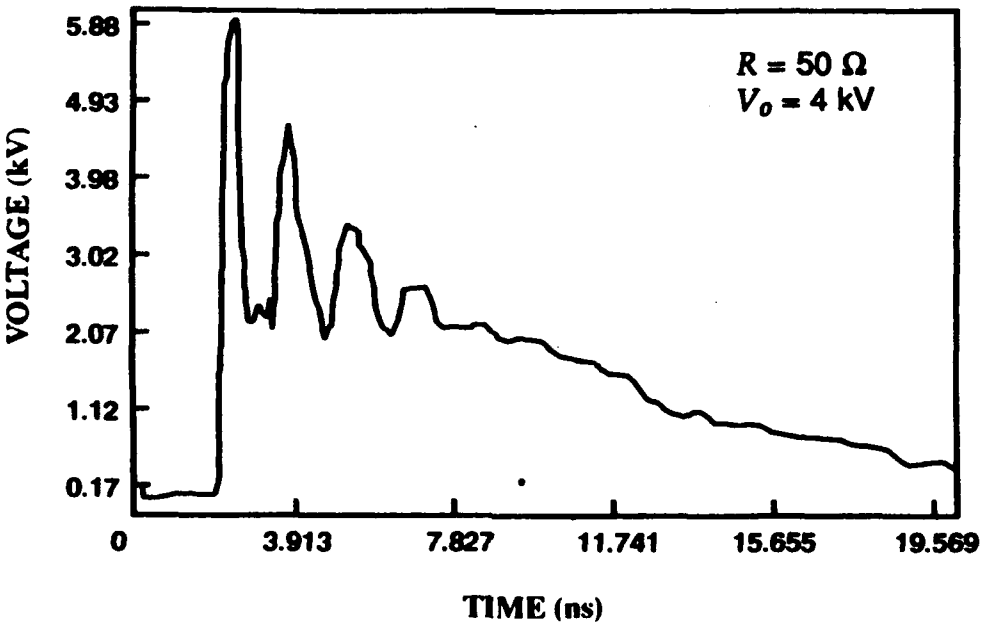


Figure 3. Output waveform obtained with a 2-mm-thick GaAs device. Switch is biased to 4 kV and triggered by a mode-locked Nd:YAG laser. Output is delivered to 50-Ω load impedance.

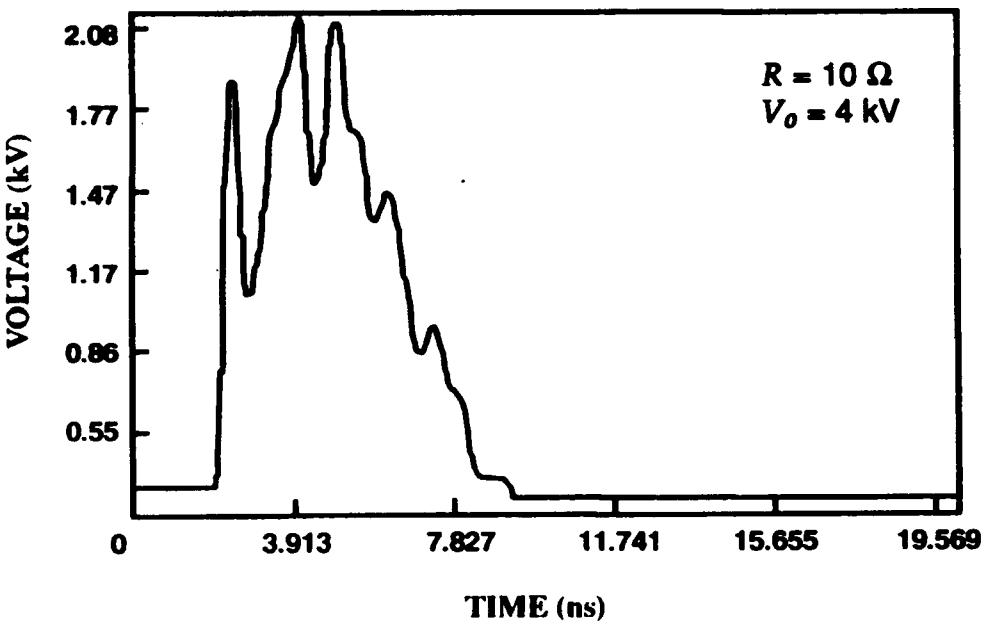


Figure 4. Output waveform obtained with a 2-mm-thick GaAs device. Switch is biased to 4 kV and triggered by a mode-locked Nd:YAG laser. Output is delivered to 10-Ω load impedance.

charge more closely follows an R_LC decay, modulated by the periodic peaks. With the 10-Ω load, the pulsewidth narrowed and the risetime increased. Figures 5 and 6 show the same device illuminated with a laser diode (wavelength of 904 nm), with a peak power of 100 W and a risetime of 3 ns. As before, the loads were 50 and 10 Ω. Unlike those with the Nd:YAG laser, the laser-diode-triggered waveforms did not exhibit any periodic peaks. As before, the waveform for the 50-Ω load exhibited a clear R_LC decay. The risetime, however, was somewhat slower compared to that of the Nd:YAG laser. For the 10-Ω load, the waveform was still reminiscent of an R_LC waveform, although the pulsewidth was somewhat narrower and the risetime slower. SPICE simulations bear out this trend when more closely matched loads are used.

We then obtained results for a device similar to the previous one, except that the inner diameter was 0.2 cm instead of 0.8 cm. As a result, the estimated inner impedance was 16 Ω instead of 4 Ω. As before, the computer simulations predicted an R_LC-type discharge for the 50-Ω load. However, the actual experimental waveforms obtained were quite different. Figure 7 gives the pulse for a 6-kV bias voltage and subsequent Nd:YAG laser illumination. The pulsewidth obtained was extremely narrow, about 500 ps, with a risetime of about 200 ps. It is as though the first peak,

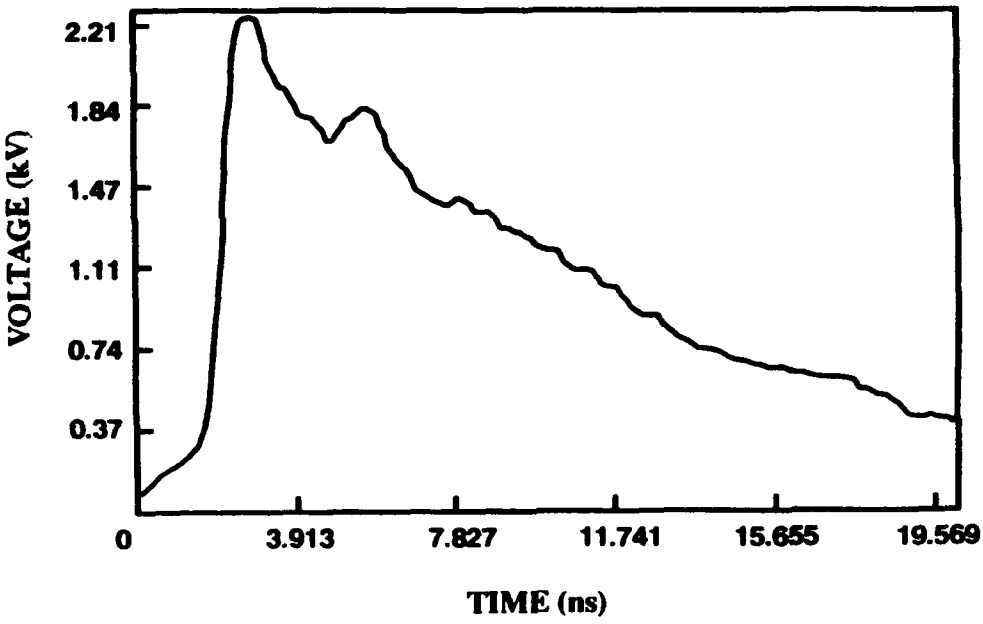


Figure 5. Output waveform obtained with a 2-mm-thick GaAs device. Switch is biased to 4 kV and triggered by a laser diode. Output is delivered to 50-Ω load impedance.

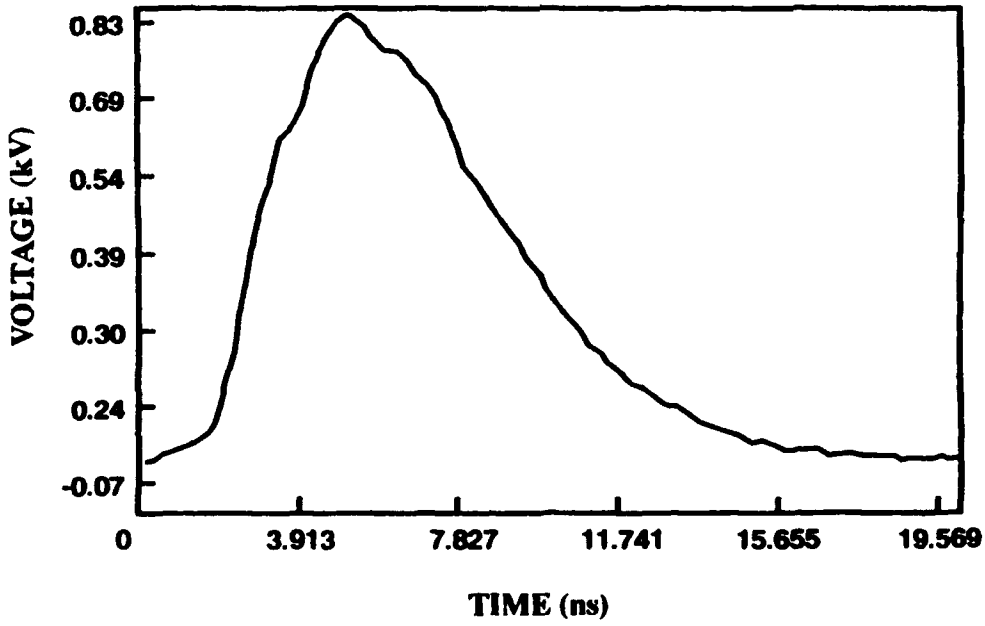


Figure 6. Output waveform obtained with a 2-mm-thick GaAs device. Switch is biased to 4 kV and triggered by a laser diode. Output is delivered to 10-Ω load impedance.

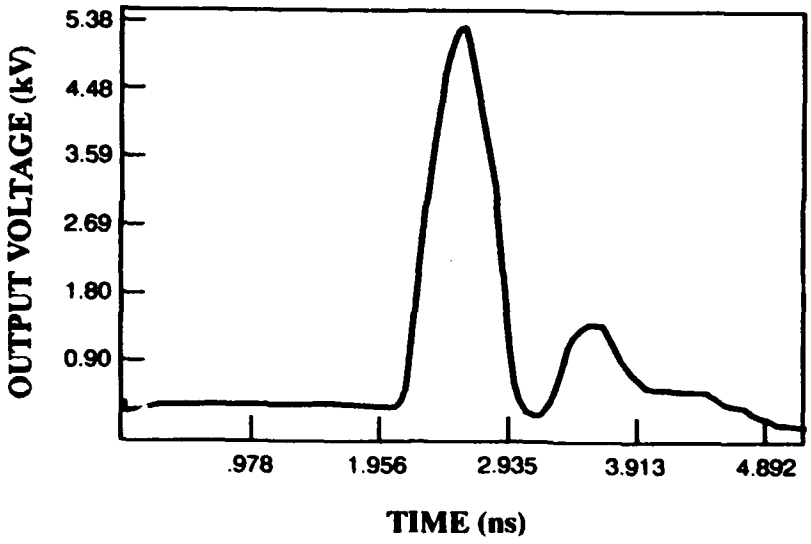


Figure 7. Output waveform obtained with a 2-mm-thick GaAs device. Switch is biased to 6 kV and triggered by a mode-locked Nd:YAG laser. Output is delivered to 50-Ω load impedance.

shown in figure 3, were allowed to pass through the device, while the remainder of the stored energy became trapped in the radial structure (the resistance in the input charging circuit eventually leaks off any trapped energy). As the bias voltage increases, the trend at higher bias voltages is to reduce the risetime and pulsewidth slightly, while increasing the relative output voltage. The waveform for laser diode excitation is shown in figure 8. In this case, a single pulse also is obtained, but it is wider compared to that generated by the Nd:YAG laser. Nevertheless, the pulsewidth produced by the laser diode was narrower (by at least a factor of 3×) compared to the expected $R_L C$ time constant.

3. Discussion and Proposed Model

As mentioned previously, waveforms were obtained that were different from those predicted either by simple theory or by SPICE simulations. In particular, the appearance of the output current spikes (fig. 3 and 4), or the single narrow pulses obtained in figures 5 and 6, cannot be explained by the circuit theory for the radial line. The most glaring discrepancy occurs when the switch current is turned off, leaving energy trapped in the radial line.

To account for the experimental facts, we suggest that the observed current turnoff, with the resultant pulse narrowing, is related to the creation of carriers in the annular region of the semiconductor. If sufficient carriers are created in this region, then the pulse output path will be shorted out, and the energy will be

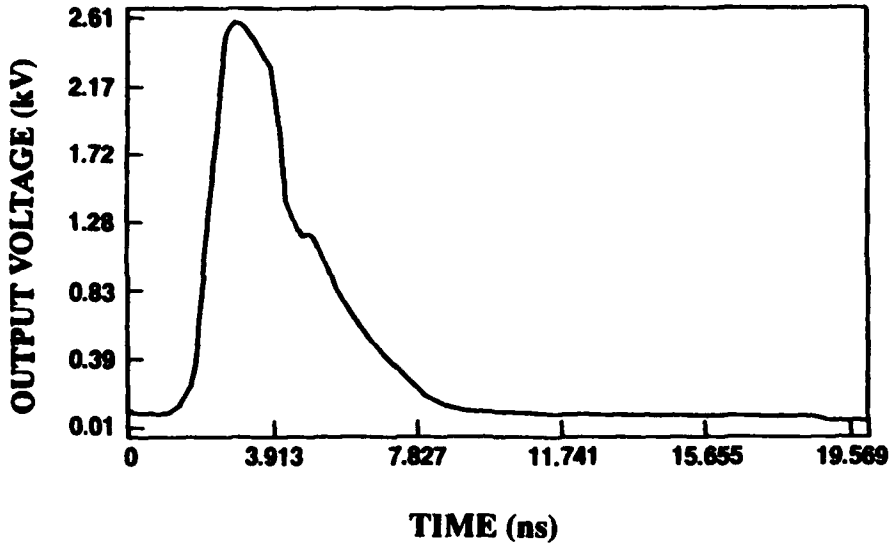


Figure 8. Output waveform obtained with a 2-mm-thick GaAs device. Switch is biased to 6 kV and triggered by a laser diode. Output is delivered to 50-Ω load impedance.

trapped in the radial structure. When the device is turned on by the laser light illumination, an electric field (field strength determined by the bias voltage and the length of annular region) is established in the annular region, between the center electrode and the inner radii, and this field helps to create a transverse conducting path. For the thicker device (3 mm) with the larger inner radius (0.8 cm), the Nd:YAG laser fails to interrupt the current pulse (fig. 2). For the 2-mm thickness, however, the current interruption begins to appear (fig. 3 and 4). Because of the large annular region and the multiple wave reflections, however, the interruption is not complete and oscillations of the peak current are observed. Mechanisms which might alternatively turn on and off the supply of carriers would include avalanche effects and rapid surface recombination. The initial carrier profile, determined by the laser pulse, may already partially bridge the gap between the inner and outer conductors, thus making it easy to short out the pulse for smaller inner gaps. One may speculate that if the closure of the annular region is caused by an avalanche, then the pulsewidth is determined by the delay time for the onset of the avalanche. For the same device (0.8 cm inner radius), it is also interesting to note that with the laser diode, there are no current spikes. This is attributed to the smaller light pulse energy, as well as the decreased light penetration of the 904-nm wavelength, which makes it more difficult to close the annular region. Under these conditions the R_LC decay is observed without any current spikes.

The situation is far different when the inner radius is reduced from 0.8 to 0.2 cm (maintaining the same center electrode radius of 1 mm). For the Nd:YAG laser, a single pulse is observed, and we suggest that the annular region remains completely closed off because of the smaller spacing. This smaller spacing will result in a large field strength, as well as a shorter plasma spreading path. The combination of these two factors may initiate an early onset of an avalanche and contribute to the generation of subnanosecond impulses. With the laser diode excitation, a single pulse is again achieved, but with a somewhat larger pulsewidth (2 ns). The larger pulsewidth may be connected with the larger delay time for the onset of an avalanche. Such a long delay may be expected, since the laser diode produces fewer carriers to seed the avalanche initially.

Other possible mechanisms may account for the closure of the device, including breakdown in the output coaxial line. The observed current turnoff, however, occurs at electric fields as low as 5 to 10 kV/cm in the coaxial line. The relatively small field, combined with the speed of the event, makes it unlikely that the source of the closure is the breakdown in the line.

4. Conclusions

Narrow-pulse generation has been achieved in monolithic, GaAs switches fabricated in the form of a radial transmission line. Pulses as narrow as 500 ps with 5-kV output into a 50-Ω load have been achieved. The large voltage enhancement (a 5-kV output voltage out of a 6-kV bias voltage) is attributed to the impedance transformation of the radial transmission line. The narrow-pulse operation is attributed to the plasma spreading into the annular region of the GaAs region. The presence of these carriers probably leads to avalanching, which disconnects the radial line from the output after a fraction of a nanosecond. Using this mode of operation, considerable improvements in peak voltage output, as well as further reduction in pulsewidth, are envisaged. In addition, extremely fast repetition rates may be possible because of the trapped energy in the transmission line, which would reduce subsequent charging times.

Acknowledgement

The authors acknowledge J. Conrad and M. Wade for the switch fabrication.

Appendix. SPICE Simulation of Monolithic Photoconductive GaAs Pulser

A SPICE simulation may be used to obtain the response of the radial structure of the pulser. This appendix describes general technique, which may then be particularized to the radial line. The technique involves first breaking up the transformer line into a number of segments of equal length, each of which has a constant impedance. The impedance value of each section follows the functional dependence of the transformer line with length, connecting the initial impedance to the final value. Once this is done, standard software, such as SPICE, may be used to find the transformer line response. In figure A-1, which demonstrates the technique, the impedance is a function of length $Z(x)$. Accordingly, the transformer line is divided into N sections (N = a positive integer), defining a set of lengths x_n given by

$$x_n = x_i + n(x_f - x_i)/N\,, \tag{A-1}$$

where x_i is the initial value of the device length, x_f is the final value, and n is an integer: 0, 1, 2, ..., N. The length of each section is $(x_f - x_i)/N$, and the average impedance of each section is $Z_m = Z(x_m')$, where

$$x_m' = (x_m + x_{m-1})/2 \tag{A-2}$$

and $m = 1, 2, ..., N$.

The above technique may be applied to the radial line, where the radial characteristic impedance of the wave may be approximated by

$$Z(x) = (377d)/(2pe^{1/2}x)\,, \tag{A-2}$$

where x is the radius, d is the separation between the circular plates, and e is the relative dielectric constant. Equations (A-1) and (A-2) are then used to set up the SPICE simulation.

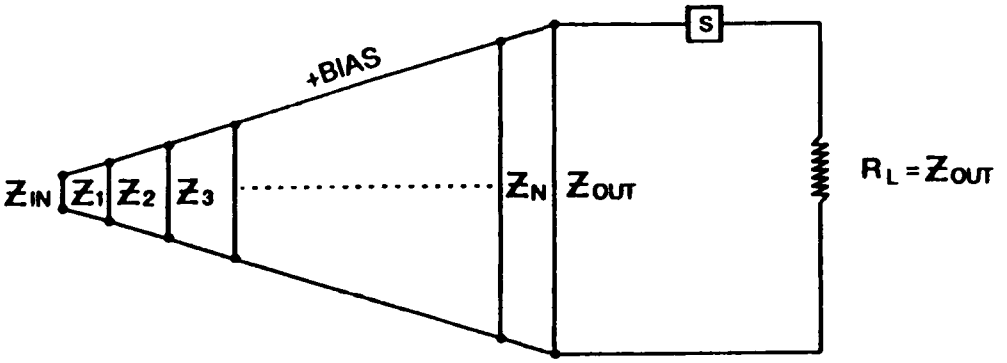


Figure A-1. Division of transformer into a number of sections. Each section is approximated by a constant impedance, thus allowing for a simple computer simulation (such as SPICE).

References

1. A. Kim, R. Zeto, R. Youmans, J. Conrad, C. Kondek, and M. Weiner, H. Brisker, J. Agee, L. Jasper, and B. Lalevic, *Appl. Phys. Lett.* **58**, 2812 (1991).
2. R. A. Petr, W. C. Nunnally, C. V. Smith, and M. H. Clark, *Rev. Sci. Instrum.* **59**, 132 (1988).
3. C. Bamber, W. Donaldson, T. Juhasz, L. Kingsley, and A. C. Melissinos, *Particle Accelerators* **23**, 255 (1988).
4. A. Kim, Ph.D. Thesis, Rutgers University, 47 (1992).

LANE

Establishing Integrated Two-Man Crew Station (ITCS) Requirements

Mr. Gerald R. Lane
U.S. Army Tank-Automotive Command
Research, Development & Engineering Center
ATTN: AMSTA-ZR
Detroit Arsenal, MI 48397-5000

PROJECT OBJECTIVE:

Technology advances in automation and robotics for vehicular applications have reached a threshold in demonstrated performance. Now, two-man combat vehicle operation can be a battlefield reality. Driving automation developed under the Unmanned Ground Vehicle Program can contribute significantly to driver workload reduction/part-time elimination. The U.S. Army Tank-Automotive Command is drawing plans for an Automated Two-Man Testbed (AUTOMATE) project to investigate critical crew performance interaction and task performance when the crew is augmented with driver aids and advanced integrated control and displays.

BACKGROUND:

AUTOMATE is one of several ongoing reduced crew programs throughout AMC developing critical combat vehicle technologies for two-man crew operation. The TACOM Integrated Two-Man Crew Station (ITCS) Project encompasses all aspects of combat vehicle operation including comprehensive soldier-machine interface studies by the Human Engineering Laboratory (HEL) and the automation of the fire control by the Armament Research Development and Engineering Center (ARDEC). The AUTOMATE project will provide the ITCS program with critical driver automation field validated performance data. This data will be used in crew task performance simulation to optimize crew station/vehicular automation subsystem design alternatives.

The AUTOMATE concept was initially suggested by MG Rick Brown in 1986, when he was Commander of the Armor Center. In a meeting with Defense Advanced Research Projects Agency (DARPA) and TACOM Research Development and Engineering Center (TARDEC) representatives, MG Brown stated the autoloader could reduce the crew to three. With the addition of driver's aids (e.g., Computer- Aided Remote Driving (CARD) and an Automated Route Planner), another crew member could be eliminated. The concept was for the driver to drive in noncombat scenarios and during combat with automated driving-- the driver could perform gunner and/or commander functions.

Crew Reduction

Program Application	Crew Size	Technology
M-1 Tank	4 Men	Manual Operation
ASM	3 Men	+ Autoloader
ALCV	2 Men	+ Automated Driving
		+ Int. Display & Ctrl
		+ Auto Tracker
		+ ATR
UGV	0 Men	+ Remote Ctrl/Teleop

Since 1986, significant advances have been made in autonomous vehicle driving under the Office of Secretary of Defense (OSD) Unmanned Ground Vehicle (UGV) Program. AUTOMATE builds on MG Brown's original two man crew concept. It leverages and applies ongoing OSD sponsored UGV research to a manned combat vehicle need.

PROJECT PLAN:

The AUTOMATE project will be designed, developed and tested by the TACOM TARDEC. It will integrate technologies primarily from the OSD UGV program together with a M113 automation project developed for Operation Desert Storm.

UGV technologies planned for integration in AUTOMATE are (1) Automated Route and Mission Planner, (2) CARD, (3) Autonomous Road Following, (4) Retrotraverse, (5) Autonomous Cross Country, (6) Peripheral and Stereo Video displays, (7) Electronic Map (8) Menu Driven Touch Panels and (9) an Intelligent Message Router.

AUTOMATE will use a M113 standard armored personnel carrier for the

testbed. During Operation Desert Storm, TACOM developed a M113 remote control system to be used as a lead/point vehicle for mechanized unit or convoy operation. The robotic M113 would trip booby traps, mines or draw fire. AUTOMATE will use the remote control automotive actuator system to give both crew members a drive-by-wire capability.

The AUTOMATE concept is shown in Figure 1. Note the mid-vehicle, side-by-side crew station location. Both crew members will have completely redundant control stations. Each crew member can drive or act as the gunner or commander. The crew station will be built similar to the driver station developed for the UGV program Robotic Command Center, shown in Figure 2 and Figure 3.

Another subsystem imported from the UGV program is a stabilized sensor platform. For the UGV program, the sensor platform was used for Autonomous Driving, for AUTOMATE, it will serve as the Target Acquisition with a second stabilized sight used as a surrogate for the tank gunner's primary sight. The two sights will allow the crew task function evaluation to assess the ability of the crew to perform target acquisition and gunnery while the driving is under full automation.

OPERATIONAL CONCEPT:

The AUTOMATE project objective is to demonstrate the concept feasibly and develop a validated field test data base for two-man crew operation with driver automation. The initial operational concept of AUTOMATE is to accomplish the full capabilities of a combat vehicle operation/tank with two crew members. It assumes that automatic ammunition loader and automated fire control techniques will be available.

Operational Concept

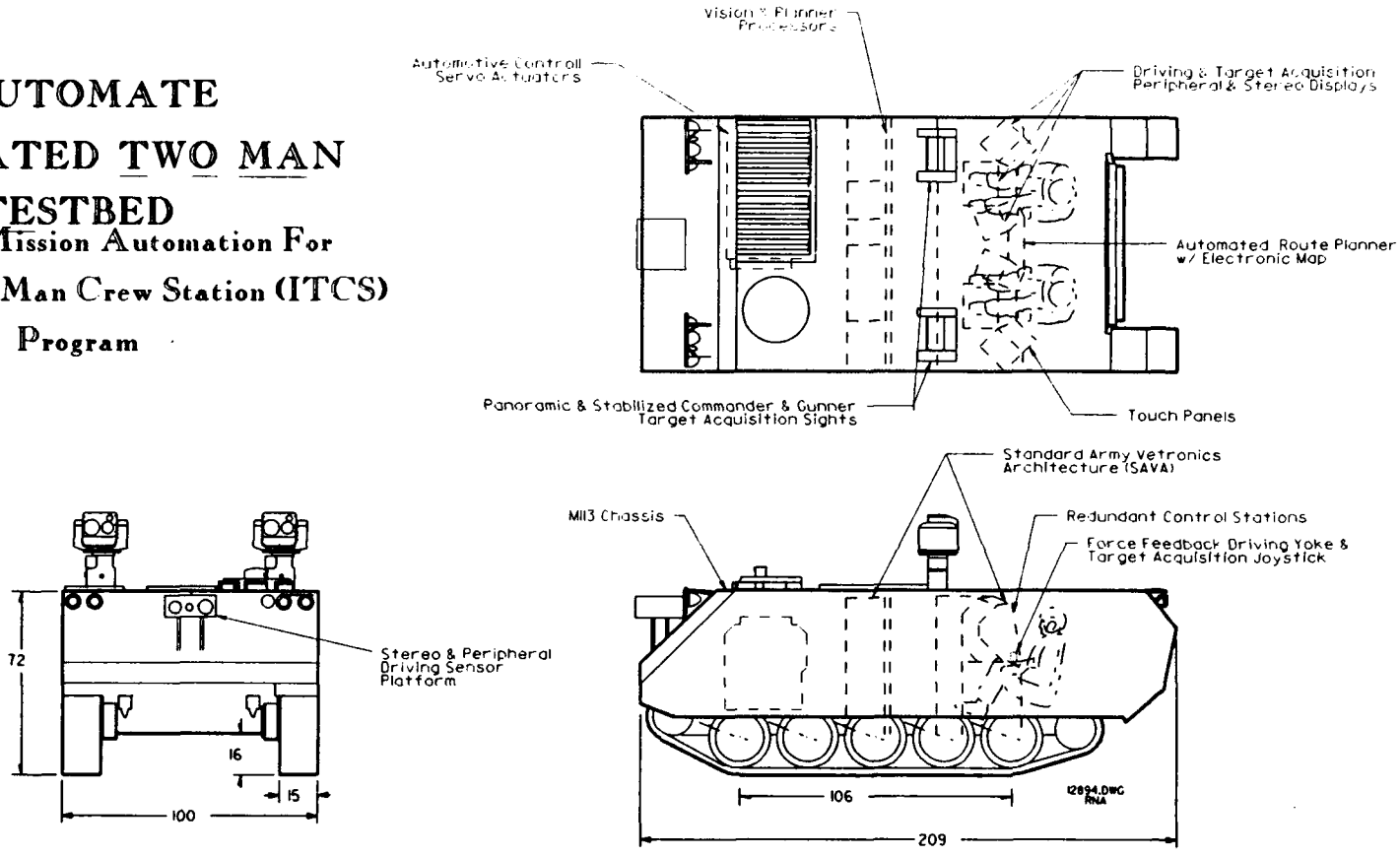
	Normal	w/Driver Aid
Driver	Manual	Automated
Gunner	——	Manual
Commander	Manual	Manual

NONCOMBAT. During noncombat operations, AUTOMATE'S two-man crew can operate in two modes-- manual or automated driver. In the normal mode; manual, the driver would perform typical functions. In the automated mode, the driver could tend to administrative functions or use an acquisition sight for surveillance. The commander/gunner would perform normal command functions. In vehicle road marches, the automated mode may be particularly helpful as a safety aid to prevent accidents during extended

AUTOMATE

AUTOMATED TWO MAN TESTBED

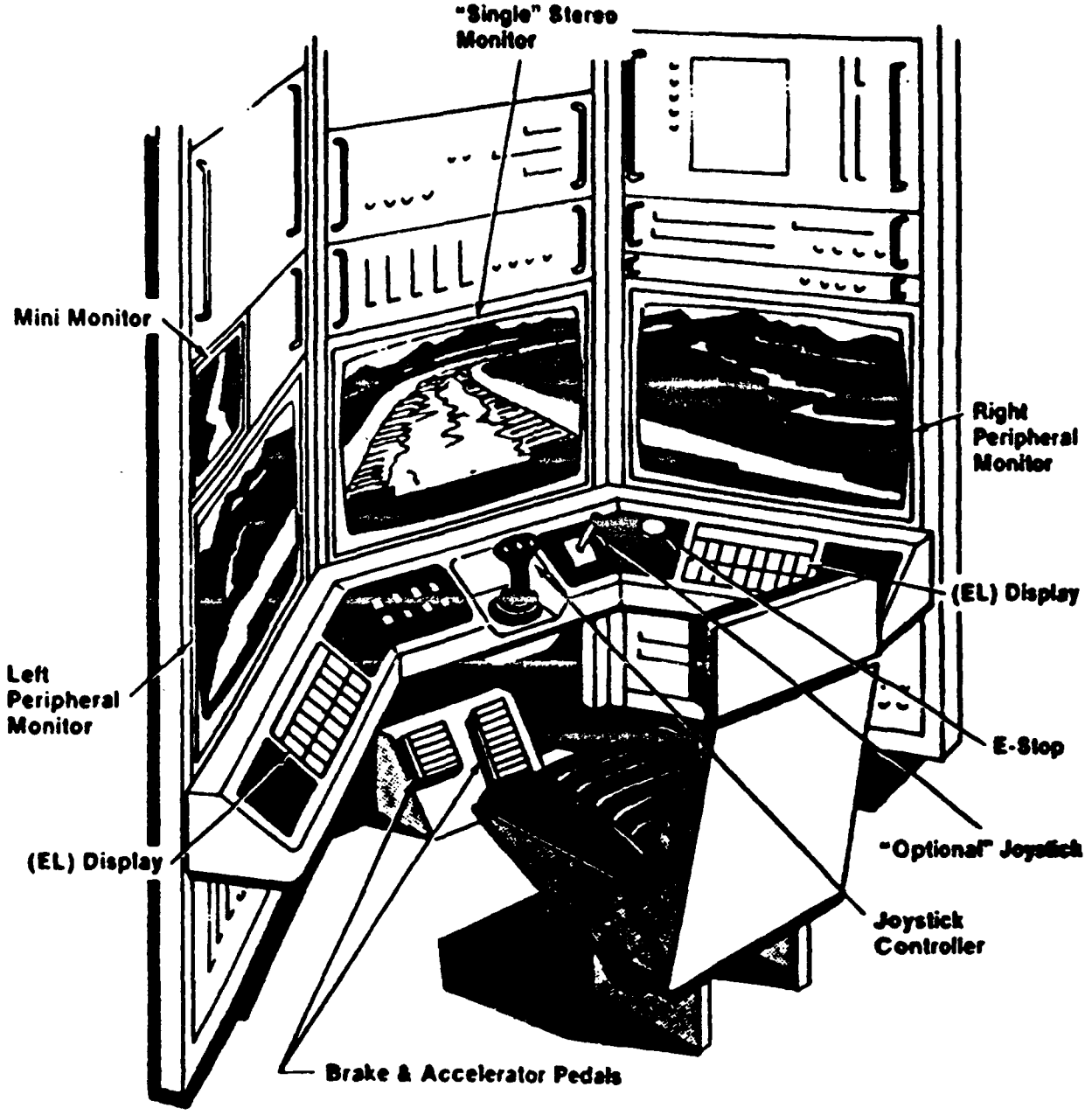
Driving & Mission Automation For Integrated Two Man Crew Station (ITCS) Program



ASCP

TACOM

FIGURE 1



RCC DRIVER STATION

FIGURE 2

TAC-2061-88

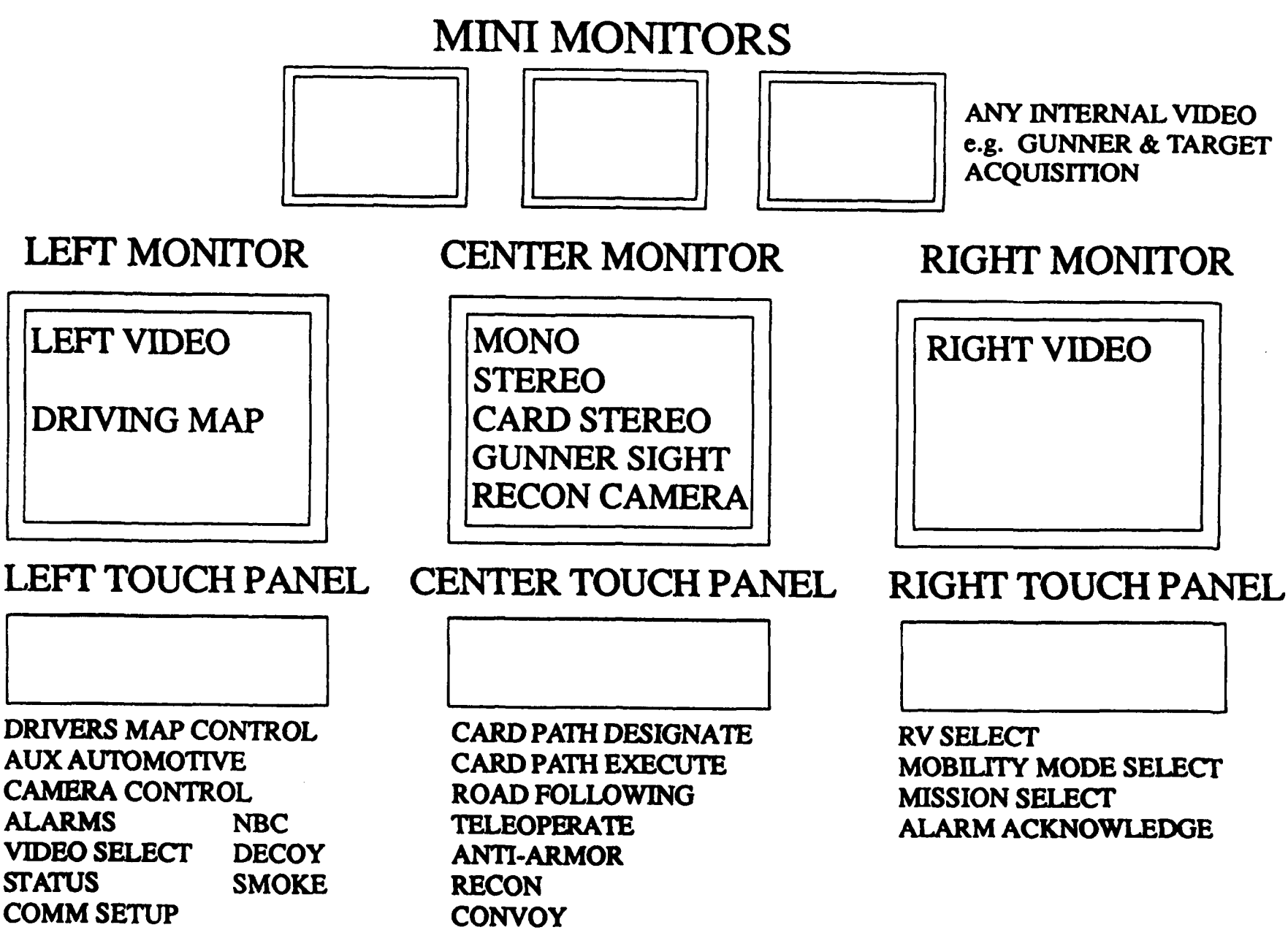


FIGURE 3 RCC DRIVER VIDEO & TOUCH PANEL

combat operations from driver fatigue/falling asleep. This mode enables supervised driving, not reducing full-time continuous manual control. Autonomous road following would be activated to perform this noncombat driver automation.

A second noncombat capability is the use of the route planner and electronic map. The route planner, coupled with the electronic map can assist the crew in normal travel to and from the engagement area.

Another automated noncombat mode relieves the crew almost totally. For use in road convoys, the robotic convoy mode could allow the crew to rest, possibly sleep, while following a lead vehicle. Today the robotic convoy mode is being designed to operate on well-defined roads through the implementation of an automated neural network driver.

COMBAT. AUTOMATE's primary objective is operation in a combat mode. The basic principle is to automate the driving function, thus freeing the driver to perform gunnery or target acquisition. The combat operation to initiate automated driving is broken down as follows:

1. Crew receives mission orders.
2. Driver or commander plans a global route on the electronic map digital terrain display inputting; start, objective, way points, mission type, vehicle type, threat location (probably automated input from higher-level systems), season and environmental conditions.
3. The route is planned and displayed on the electronic map. The crew would then iterate/adapt the route, if desired.
4. The driver selects the desired driver aid-- Autonomous Road Following, if on a road; CARD, if bounding from concealment position to concealment position; Autonomous Cross-Country, if the route planned and displayed on the video is acceptable with no modifications.
5. For the CARD mode, the driver would designate a path between points of concealment. CARD can also be used for multiple steps but requires a replanning mode, which may be difficult on the move. CARD could be used during the autonomous cross country travel when the driver selects a major deviation to the route plan displayed.
6. The driver initiates the automated driving mode. The driver confirms vehicle movement on the real-time electronic map.
7. The driver begins gunnery or target acquisition functions.
8. Either crew member can monitor driving video or position, heading and field-of-view coverage on the electronic map.
9. The crew can select the retrotraverse mode at any time. This will command the vehicle to retrace the last segment of the route travelled.

Retrotraverse is particularly useful to automatically move the vehicle to secondary firing position.

10. Either crew member can replan a path segment as mission changes or new terrain or threat information is obtained.

11. Critical mission threat, and vehicular position, heading and field-of-view information is automatically passed between vehicles in a unit/platoon to facilitate inter-vehicular cooperation.

TECHNOLOGY:

AUTOMATE technology is primarily being developed for Unmanned Ground Vehicles (UGV). UGV's have similar driving needs to the AUTOMATE driving concept; driving is performed by the computer. AUTOMATE's driving performance, however, can be improved because the driver is on board and can sense the full vehicle feedback; all vibration, shock, noise, vision, attitude smoke, and smell. Implementation of intelligent driving capability on manned tank-automotive materiel is a logical first step.

Driving Aid Hierarchy

1. Electronic Real Time Map
2. Automated route Planner
3. Retrotraverse
4. Computer Aided Remote Driving
5. Autonomous Road Following
6. Robotic Convoy
7. Autonomous Cross-Country

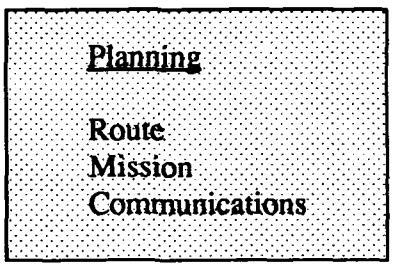
Driving Automation Elements

Planning
Perception
Control
Display

The principle driving automation elements are planning, perception, control and display. The UGV program has research in all these elements. Although not a final design in controls and displays, the driver's station in the Robotic Command Center offers a proven capability at a cost and schedule efficient price.

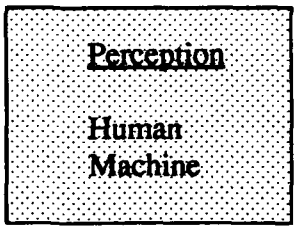
The critical driving aids which are the cornerstone capability of AUTOMATE are listed in a hierarchical order of development and driving automation performance. Each of the driving aids will be discussed separately in each driving automation element.

PLANNING. The planning is broken down into three subelements: Route, mission and communication. Planning efforts underway emphasize multiple-vehicle cooperation and reactive planning. Hughes Research Laboratories is developing a planning system with direct connection to the perception system, Advanced Decision Systems Inc. is developing a mission planner based on variable behavior control, and the University of Michigan is developing a planner based on the coordination of multiple vehicles. All three aspects will be fused in a higher-level planner system in the UGV Demonstration II scheduled for 3rd qtr FY95.



All these planner systems are built upon the initial multi-vehicle route planner developed for the Robotic Command Center. This planner embodied expert system knowledge to plan the mission; tactical reconnaissance, movement to contact, decoy, convoy, evacuation and recovery. The route planner displays in real-time; the position, heading and field of view of its unit vehicles on a tactical map display based on a 30-meter digital terrain database.

PERCEPTION. *There are two* AUTOMATE Perception modalities; human and machine. Each perception mode can interact with each driving automation element. AUTOMATE's operators selectively move in and out of both perception modes as the mission progresses. The current perception mode performance capabilities are primarily driven by UGV research ongoing for UGV Demonstration II in 1995.



The first step in execution of a mission order is the overall perception of the local scene/environment or battlefield. Today and for the foreseeable future, this is best performed by the operator. Through the use of many vision modes; peripheral, stereo, thermal, etc., the operator perceives the local scene/environment, and with the use of map data, determines the optimum driving-aid methodology. In AUTOMATE, the driver can select from the driving aid Hierarchy; *(for planning the driver can use)* Electronic Map or Automated Route Mission & Mission Planner, *(or for driving select from)* Retrotraverse, Computer Aided Remote Driving (CARD), Autonomous Road Following (ARF), Robotic Convoy or Autonomous Cross-Country.

The driving aid technologies are primarily developed under the OSD

UGV program. *Retrotraverse* developed by the National Institute of Standard and Technology (NIST) under the Army Research Laboratory sponsorship has demonstrated retrace of a vehicles path automatically with path deviations less than one meter in 100/-200 meter retraces. *Computer-Aided Remote Driving (CARD),* developed by the Jet Propulsion Laboratory under TACOM sponsorship, allows the driver to predrive/designate the vehicles path in a video scene. The driver moves the vehicle icon in a 3-D/stereo video scene, as the vehicle moves away (into the scene), a path (full vehicle width) is displayed. At the end of the path designation the driver executes CARD. The vehicle now drives the path under full computer control. A passive stereo vision based obstacle detection/avoidance system is added to correct for navigation or path designation errors. The vehicle will automatically stop if it detects an obstacle in its path. In the future, when the earlier discussed reactive planning algorithms are added, the vehicle will avoid/go-around the obstacle.

Autonomous Road Following (ARF) performance has increased dramatically under the DARPA-sponsored TACOM contract with Carnegie Mellon University (CMU). CMU using a technique called Autonomous Land Vehicle in a Neural Network (ALVINN) has demonstrated ARF at speeds up to 104KPH and distances over 20 miles per segment. ALVINN follows defined road edges using a neural network approach which is trained by the operator. Because its trained by the driver, it reacts very similar to him, copying turning and lane position styles. Other ARF techniques available when road edges are not adequately defined include: color segmentation, following the distinct color of the road; texture matching, comparing the road texture to the background; or shape/profile tracking, use range reflectance to recognize the unique road profile/crown.

Robotic Convoy research underway by TACOM is focused today on tactical vehicle applications. For AUTOMATE, it offers the crew additional safety and administrative time. The TACOM research contractor, Redzone Inc., plans to use the CMU ALVINN coupled with a hybrid Kalman filter-based navigation (GPS and inertial navigation) system to implement a robotic convoy. The lead-vehicle ALVINN processor will learn the path from its driver. It will transmit the path and navigation data to the following vehicle, and so on back to all convoy vehicles. This TACOM research is scheduled for demonstration in the 2nd Qtr FY94. At that time, it will be available for AUTOMATE integration.

The robotic convoy techniques must become more robust for the battlefield and traffic. TACOM plans to supplement it with ongoing

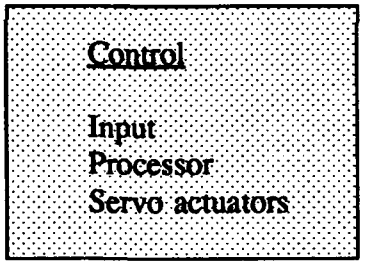
Department of Transportation (DOT) research into blind spot/side, rear and front obstacle detection. DOT is also researching lane-weave correction techniques which could supplement ALVINN.

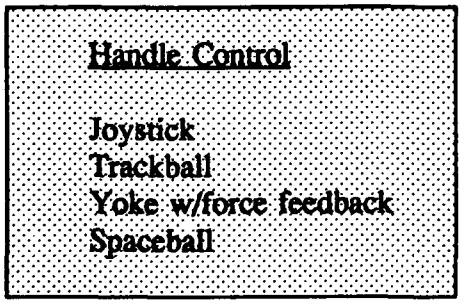
Autonomous Cross-Country provides the greatest technical challenge, but it is absolutely essential for AUTOMATE to demonstrate battlefield potential. CMU research is now focusing on cross-country movement. TACOM provided a GFE HMMWV for the CMU research. Entitled NAVLAB II, the robotic HMMWV has began traversing mild cross-country terrain. The CMU work is supplemented with the Hughes Research Lab (discussed earlier) reactive planning work. The reactive planning will facilitate AUTOMATE maneuvering to avoid obstacles and remain on its route plan, if possible.

Other complementary research includes the Joint TACOM/DARPA sponsored JPL Passive Stereo Vision research in obstacle detection and general scene understanding.

These projects are complimented by the TACOM contract with the University of Massachusetts (UMASS). Similar to CMU, UMASS will configure a robotic HMMWV to perform perception research. The UMASS Perception Lab will focus research in natural landmark recognition. AUTOMATE will use the local landmarks as waypoints in execution of its route/mission plan.

CONTROL. The Control element in AUTOMATE include both human and machine modes. The control element has three subelements: Input devices; vision and planning processors; and the servo actuator automotive controls. The human control mode is primarily divided into two categories; hand controls and data entry.



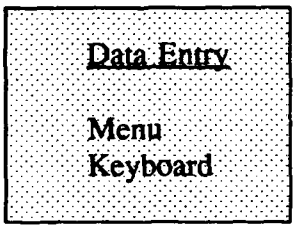


AUTOMATE will provide the operator with multiple *hand controls* specifically selected for optimized performance in different driving and mission tasks. The *Joystick* has demonstrated superior target acquisition and gunnery performance; the *Trackball* provides the best means to move the cursor on an electronic map or menu display; the *yoke* with force feedback provide the operator with the best feel of the road; and the *spaceball* enables CARD to move the cursor in three dimensional space, required for

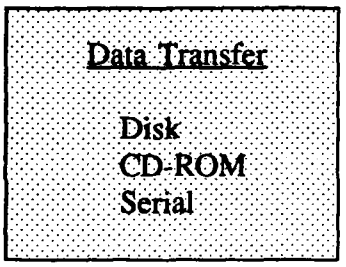
accurately designating the range to vehicle turning point.

Whenever possible, each hand control will be reprogrammable to evaluate the side-by-side performance. Quantification of performance for each automation mode controlled by various hand control approaches will be obtained.

Two methods of *data entry* will be available; menu and keyboard. *Menu* will be available on the touch-sensitive videos, high resolution electronics maps and touch panels. Some data, however, is more efficiently entered via keyboard. MENUs will be designed to minimize AUTOMATE keyboard entry. The Keyboard, however will provide a diagnostic and inter-vehicle communication capabilities.



The control *processor* will be a hybrid of vision and general-purpose processors including Improved Warp (I-WARP), datacubes, silicon graphics, LISP, etc. The DARPA Image Understanding Architecture (IUA) may be an eventual AUTOMATE upgrade, as processing requirements grow based on planning, perception control and display capabilities expand.

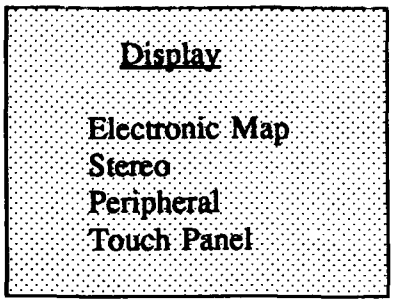


Data Transfer for AUTOMATE is also a requirement with three alternatives: simple floppy disks, CD ROM or serial. Both techniques will facilitate ease of data input from a variety of supporting research projects/organizations. AUTOMATE's Standard Army Vetronics Architecture (SAVA) will facilitate data transfer via a serial link. The serial link will be used primarily for diagnostics or up/down loading of programs. Whereas the floppy disk/CD ROM will be used for data transfer, e.g., digital terrain data base, threat information, etc.

The *servo-actuator* control system is an adaptation of the M113 TACOM developed for Operation Desert Storm. It controls the basic automotive functions; start engine, acceleration, braking, steering and transmission gear select. The actuators are both electrically and hydraulically actuated. The steering, braking and throttle are hydraulically actuated with pressure from the crew ramp hydraulic system.

DISPLAY. The crew displays specification selected for AUTOMATE are from the Robotic Command Center; high resolution for Electronic Map, Silicon Graphics for 3-D Stereo display, three-screen wrap around for 180° peripheral vision and

Touch Panels for menu-driven feature selection. Graphic video overlays and some touch-screen video (left and right monitors only) will be integrated. Additional video monitors will be provided for instantaneous monitoring of mission sensors or driving sensors if in the automated driving mode. AUTOMATE will also employ a video overlay (multiple video on one screen) capability to improve crew performance. This will allow the crew to concentrate on one versus multiple screens.



OTHER TECHNOLOGIES

The AUTOMATE system will be implemented on the *Standard Army Vetronics Architecture (SAVA).* This will require recoding much of the research software. However, it is minimized because most research organizations use the VME architecture for development. SAVA is an expanded VME system with built-in diagnostic capability.

Two stabilized sensor platforms will be provided offering daylight and thermal vision capabilities to perform target acquisition and gunnery tasks. *Both will have full 360° panoramic vision.*

The *driving sensor* platform will contain peripheral and stereo camera arrays on a stabilized mount. The driving sensor will be limited for daylight vision at this time, because most UGV research has selected this spectrum as a first step. A supplemental laser scanner will be mounted to provide range measurements for ARF and obstacle detection.

SIMULATION

ITCS and AUTOMATE will be guided in all three phases through both analytical and physical simulation. Initial crew operational requirements will be modeled on CRAVES (Army Crewstation Design Tool) developed by the Human Engineering Laboratory. The CRAVES model will identify the impact of alternative crew station design, assess impact of crew-reduction on mission effectiveness and assess overall system performance with various crew workload distribution. The CRAVES model will help assess crew reduction potential of driving automation demonstrated in AUTOMATE.

The Vetronics Crew Display Demonstrator (VCDD) will feed the AUTOMATE crew stations design and be adapted to simulate driver

automation.

Both CRAVES and VCDD simulations will help specify initial crew station design specification, support development and will be validated with field trials from AUTOMATE, SMI from HEL, and fire control capabilities from ARDEC.

SUMMARY

The technology is available to Establish Requirements for an Integrated Two-Man Crew Station (ITCS). The technology will come from several AMC sources with appropriate subsystem expertise; Soldier-Machine Interface from Human Engineering Laboratory, Fire Control Automation from Armament Research Development and Engineering Center and Tank-Automotive Command implementing overall system integration and driving automation.

AUTOMATE will serve as the driver automation testbed to develop the required driver's aids for two-man crew operation, in parallel to other ITCS activities.

AUTOMATE reduces the ITCS program risk by independently specializing on the driving automation topic through the implementation of the OSD UGV research initiatives.

Power Spectral Characterization of Fracture Surfaces of TNT, Composition B, and Octol

*M. Yvonne D. Lanzerotti+, Dr.
James J. Pinto+, Mr.
Allan Wolfe++, Prof.
David J. Thomson+++, Dr.
+U. S. Army Armament Research, Development and Engineering Center
Picatinny Arsenal, NJ 07806-5000
++New York City Technical College, Brooklyn, NY 11201-0000
+++AT&T Bell Laboratories, Murray Hill, NJ 07974

BACKGROUND

Knowledge of the fracture phenomena of energetic materials subjected to high acceleration (high *g*) is of paramount importance to the Army in such fields as ordnance and propulsion applications where energetic materials are often subjected to high, fluctuating, and/or sustained accelerations. For example, explosives in projectiles are subjected to high setback forces (up to 50,000 *g*) and centrifugal forces due to rifling during the gun launch process. These high setback forces can cause fracture and premature ignition of explosives. Interaction of the stresses produced in projectile explosives by the setback forces and charge motion can lead to rapid temperature rise at the location of voids, cracks, and other defects. These sites of localized temperature rise are potential ignition sites, that can result in premature ignition. Grain fracture of solid propellants during acceleration can result in abnormal burning and performance. Therefore, achievement of a more fundamental understanding of the behavior of energetic materials subjected to high acceleration is a key to better practical ordnance designs that solve the problems of abnormal propellant burning and premature ignition of explosives during gun launch.

High explosive materials for military applications must have a uniform density and composition. Early detection of defects in hign explosive pressings can reduce the cost and improve the safety of the high explosive production cycle. Better high explosive characterization can lead to improvements in production processes and explosive performance. Laboratory and statistical techniques described herein have been pioneered to characterize the grain sizes in energetic material fracture surfaces during high acceleration. These techniques have also contributed to understanding the fundamental physics of the fracture process during high acceleration.

INTRODUCTION

We have created several new fields of research to study the mechanical behavior of energetic materials during high acceleration by using an ultracentrifuge (1-6), and to characterize the mechanical failure surfaces of energetic materials by power spectral statistical techniques (7-12). This is the first time that ultracentrifuge techniques and the new science of fractals have been used to study the fracture physics of energetic materials during high acceleration. The objectives are to define the nature of the mechanical failure during high acceleration and to determine the role of fracture in explosive sensitivity. Extensive studies have yielded highly significant results that are of critical importance in the understanding of major munitions problems. This work is particularly relevant to the future development of insensitive energetic materials used in weapons with higher acceleration and longer range, such as rocket assisted projectiles and liquid propellant weapon systems.

The mechanical behavior of energetic materials subjected to acceleration greater than $10^4 g$ has been studied in an ultracentrifuge (1-6). The energetic material experiences a pressure gradient both during acceleration in the gun and under *g*-loading in the ultracentrifuge. The pressure gradient experienced by the explosive during acceleration in the gun and under *g*-loading in the ultracentrifuge is unique and leads to different kinds of behavior and failure than under other material test conditions. Therefore, the pressure gradient experienced by the explosive under *g*-loading in the ultracentrifuge is more relevant to the gun launch process than are the behaviors and failures of materials studied under other test conditions.

Under *g*-loading, TNT, Composition B, and Octol have been discovered to behave as brittle solids. Millimeter size particles have been discovered to break loose from the as-cast surface when the tensile or shear strength is exceeded. The fracture threshold of TNT, Composition B, and Octol has been found to vary between 10,000 *g* and 60,000 *g* during high acceleration. A hemispherical fracture surface is found to form. The millimeter size fractured particles are found to break up into 10-20-μm average size particles of explosive during their impacts on the opposite side of the void (1-6).

A new mode of void collapse, not previously known or considered in ordnance technology, has been discovered (1). This finding is a highly significant result for the Army and the scientific community in that previously it had been thought that adiabatic compression due to viscoplastic behavior of the explosive could cause temperature rise and ignition (13). Our results reveal, however, that fragmentation may have a significant role in premature ignition. Specifically, the particles of explosive in a projectile that break loose under acceleration are confined to a void. These particles impact the opposite side of the void and break up into micron size pieces. These fine particles of explosive can be more easily ignited (i. e., premature ignition) than the bulk explosive. This is a very important consideration for premature ignition since voids can fill with fragments of explosive during high acceleration.

Projectiles were fired with embedded TNT samples and recovered to determine the details of fracture during the actual gun launch process (5). About 10% of the TNT sample

was found to fracture into 80-μm average size particles. This experiment has thus shown that a significant number of particles can fracture during the millisecond time frame of the gun launch if the fracture threshold is exceeded.

An explosive crack problem was experienced with the Octol in the STAFF (Smart Target Activated Fire And Forget) warhead. During *g*-loading in an ultracentrifuge, the fracture acceleration of Octol STAFF warhead samples varied from 23,000 *g* to 60,000 *g* as a function of location in the warhead. Corresponding values of effective tensile stress at failure were calculated from these values of fracture acceleration and the fractured mass. Knowledge of the tensile strength of the STAFF Octol was an essential contribution to the solution of the explosive crack problem in the STAFF warhead (6).

We have underway a program for studying in detail the fracture behavior of TNT, Composition B (59% cyclotrimethylene-trinitramine (RDX), 40% TNT, and 1% wax), and two types of Octol (70% cyclotetramethylene-tetranitramine (HMX) and 30% TNT; ≈ 83% HMX and ≈ 17% TNT) using an ultracentrifuge (7-12). This program includes especially the study of the characteristics of the fracture surfaces of the materials. In this paper, the results and conclusions obtained from statistical studies of the fracture surface topography along several parallel scans across the fracture surfaces of TNT, Composition B, and Octol are described.

TECHNIQUE

The fracture surfaces of the energetic materials under study are obtained by accelerating prepared samples in a Beckman preparative model L8-80 ultracentrifuge. The sample under study can be rotated up to 60,000 rpm (500,000 *g*). When the tensile or shear strength of the object material is exceeded, a fracture surface is obtained. A diamond stylus profilometer has been used to measure the topography of the surface after removal of the sample. Profiles have been obtained at several locations spaced approximately 1.0 mm apart across the surface. Spatial wavelengths along each trace of 1.0 μm to approximately 1.0 cm have been statistically characterized. Spatial power spectra have been calculated from the data of the individual traces using a prolate spheroidal data window (14-15) applied in the horizontal space domain prior to using a fast Fourier transform algorithm (16). The spatial frequency bands for which power spectral amplitudes are reported are designated A and C in Table 1.

RESULTS

TNT. The sample of melt-cast TNT reported here was found to fracture at an acceleration of 50,000 *g* at 25°C. The spatial power spectra of the fracture surface profiles across traces g and i (see inset) are shown in Figure 1. The distribution of peaks in the spatial power spectra show the nonuniform TNT grain size. The TNT grain size varies discontinuously from ≈ 0.1 mm to ≈ 1.0 mm. Figure 1 also shows how the nonuniform TNT grain size varies with location across the surface. As has been shown previously (8-12), the spatial power spectra for TNT falls off with increasing spatial frequency. The solid and starred lines superimposed on the spectra traces are linear fits to the sections of the spectra corresponding to the line lengths.

Table 1. Spatial frequency bands for which TNT power amplitudes are plotted in Figure 3

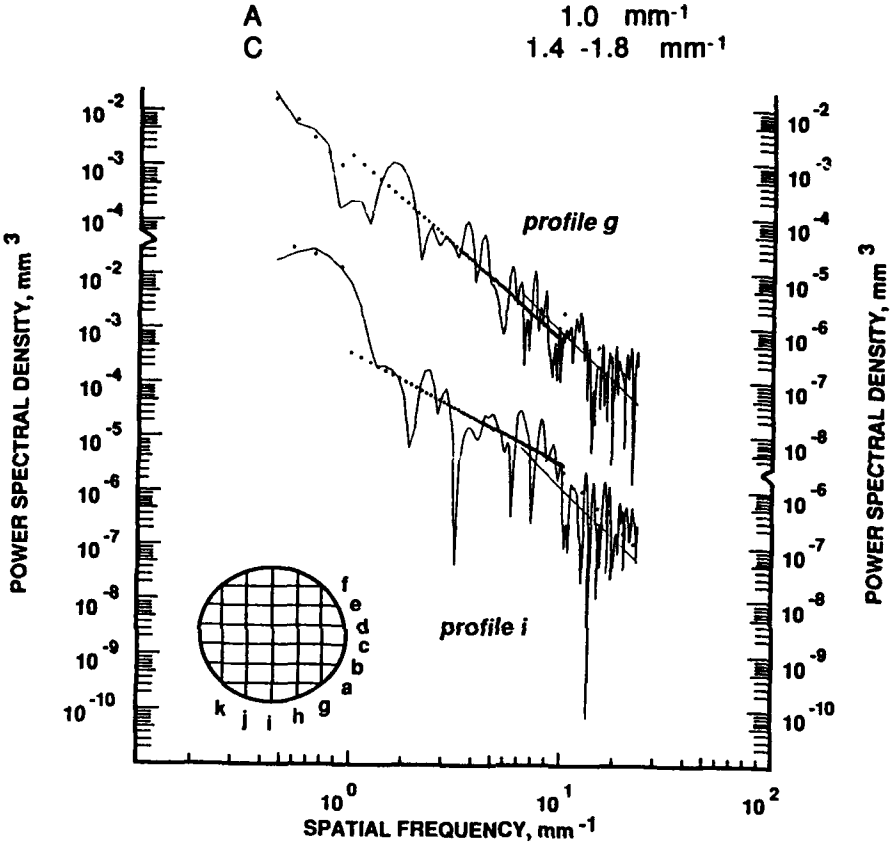


Figure 1. Spatial power spectra of the fracture surface profiles for sample traces g and i of TNT. The locations of the profile traces across the sample surface are shown in the inset. Diamond stylus horizontal spacing is 0.004 mm.

The variation of the spectral slope across the surface is of considerable interest. In the frequency interval 10^0 mm^{-1} to 10^1 mm^{-1} the spectral slope of trace g is -3.5. We have found values similar to this for the spectral slopes in our previous studies of TNT (8-12). The spectral slope of trace i is -2.1 in this frequency interval. The absolute value of the spectral slopes for profiles g through k of Figure 1 are plotted in Figure 2 as a function of the sample trace. Similar results are obtained for sample traces a through f.

For spectral slopes ≤ -3, the fractal dimension of the respective wavelength interval equals one (17-18). Such a slope suggests that the failure process is deterministic. As we have noted before, this interpretation is consistent with the occurrence in the spectrum of quasi-harmonic peaks. That is, the occurrence of the peaks, which correspond to TNT grain sizes, as well as the magnitude of the spectral slopes, indicate that, in general, the fracture is occurring at grain boundaries: intergranular failure.

For spectral slopes, s, that lie in the range $-3 < s \leq -2$, the fractal dimension is given (17-20) by

$$s = - (5-2D), \qquad (1)$$

where D is the fractal dimension. The fractal dimensions for profiles i and d are 1.5 and 1.8, respectively. This indicates that the failure is non-deterministic chaos. That is, transgranular failure occurs through the TNT grain. The region of non-deterministic chaos suggests that the origin of the failure is in the center of the sample.

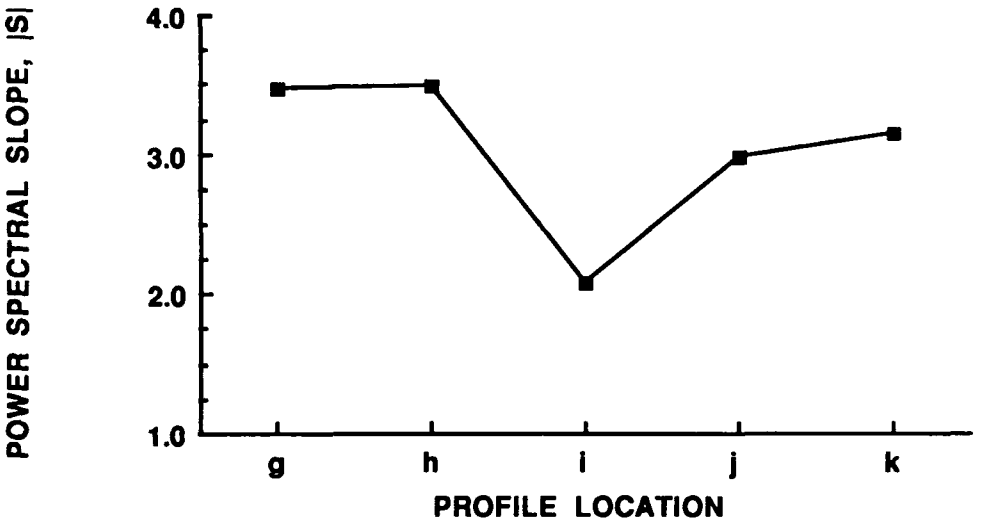


Figure 2. Spectral slopes versus profile location g through k of TNT for spatial frequency interval 10^0 mm^{-1} to 10^1 mm^{-1}.

Also of considerable interest is the nature of the spectral variations across the fracture surface. This has been investigated by determining the frequencies and amplitudes of the spectral peaks in each of the spectra computed from profile data taken for the sample traces a through k. The most prominent spectral peak is listed in Table 1 and is designated C. Spectral value A corresponds to a low frequency spectral value selected outside a distinct peak.

The power amplitudes for each of the spectral values of Table 1 are plotted in Figure 3 as a function of the sample trace. There is a distinct tendency for the high peak power levels to occur for those traces taken closer to the sample edges (i.e., the spectra from profiles g and k) for the spatial frequency value, C, corresponding to spectral peaks. The lowest spectral frequency has a peak amplitude in the center of the sample (profile trace i). Similar results are obtained for traces a through f. The high peak power levels occur for edge traces for the spatial frequency values, 2.1-2.6 mm^{-1}, corresponding to spectral peaks. The lowest spectral frequency, A, also has a peak amplitude in the center of the sample (profile trace d).

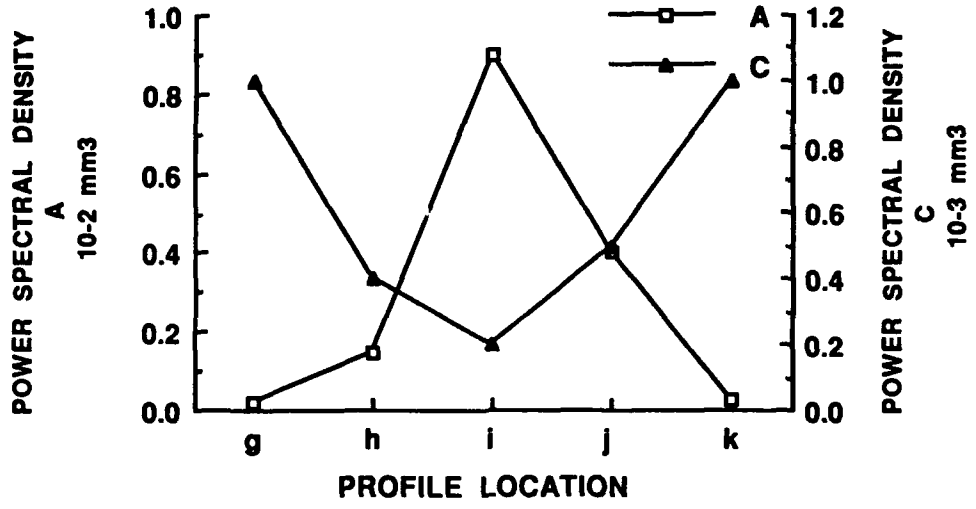


Figure 3. Power amplitudes for the spatial frequencies A and C in Table 1 plotted as a function of profile location g through k of TNT in Figure 1.

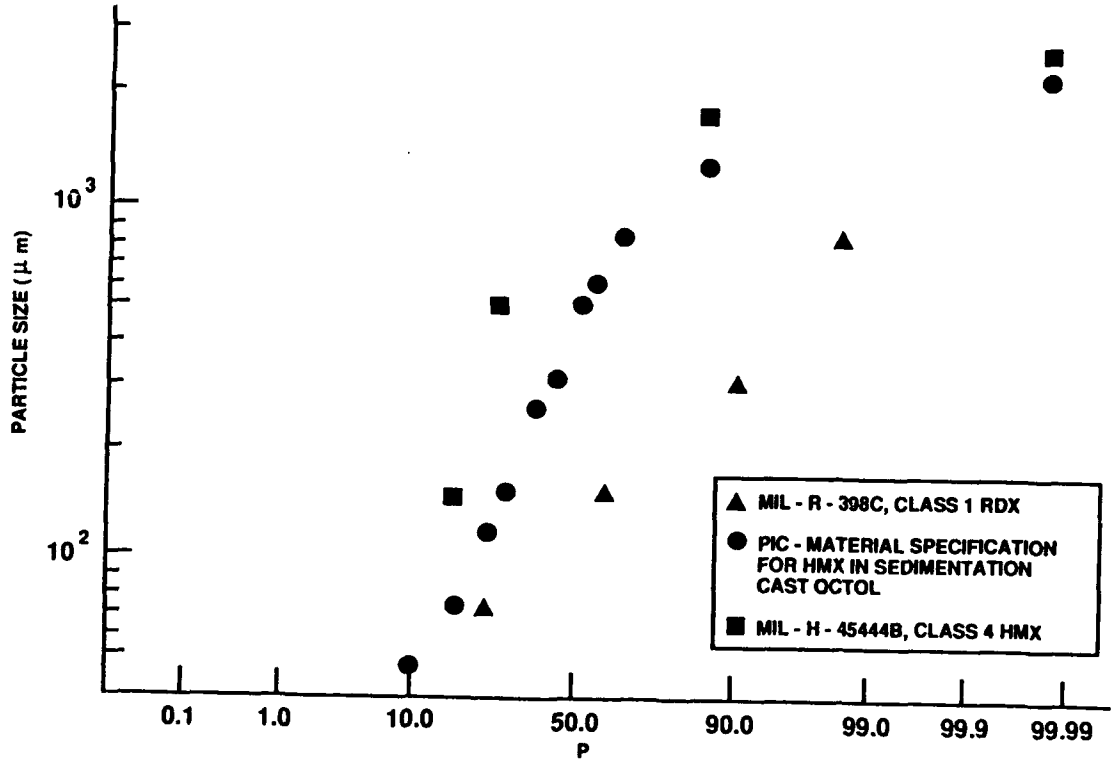


Figure 4. RDX and HMX Particle Size Distribution

Composition B. The sample of melt-cast Composition B reported here was found to fracture at an acceleration of 41,000 *g* at 25°C. The logarithmic normal crystal size distribution of RDX as specified by military specification Class 1 (21) is shown in Figure 4. Figure 5 shows the spatial power spectra of the fracture surface profiles across traces i and j (see inset). The distribution of peaks in the spatial power spectra shows the inhomogeneous grain size of the RDX in the Composition B. The RDX grain size varies discontinuously from ≈ 0.1 mm to ≈ 1.0 mm in contrast to Class 1 RDX particle sizes in Figure 4. Figure 5 also shows how the inhomogeneous RDX particle size varies with location across the surface.

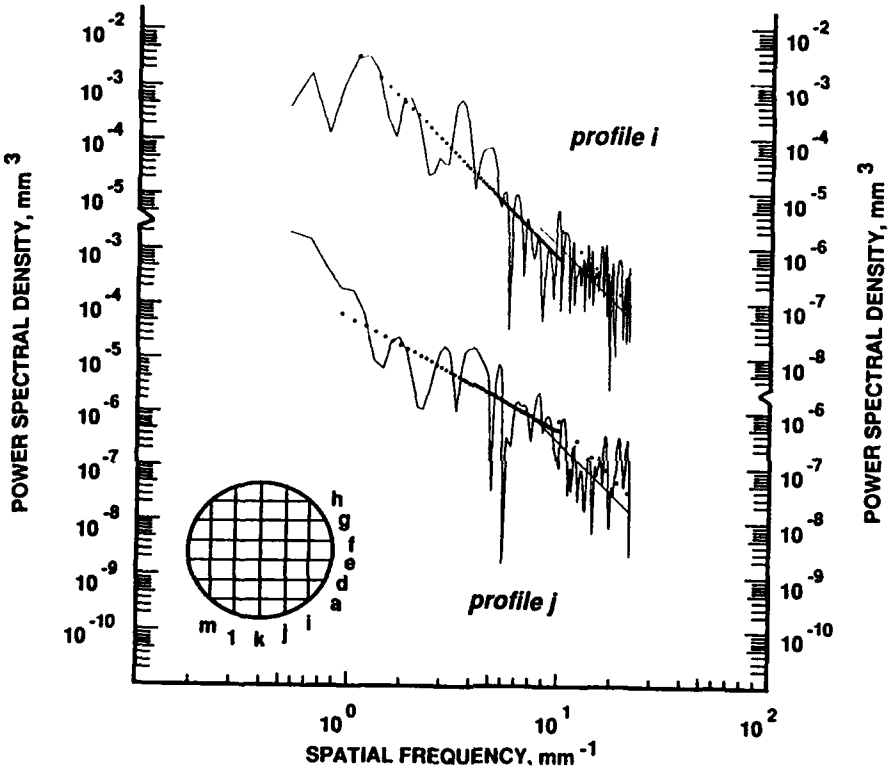


Figure 5. Spatial power spectra of the fracture surface profiles for sample traces i and j of Composition B. The locations of the profile traces across the sample surface are shown in the inset. Diamond stylus horizontal spacing is 0.003 mm.

In the frequency interval 10^0 mm^{-1} to 10^1 mm^{-1} the spectral slope of trace i is -4.0. We have found values similar to this for the spectral slopes in our previous studies of Composition B (8-9). The spectral slope of trace j is -2.1 in this frequency interval. The absolute value of the spectral slopes for profiles i through m are plotted in Figure 6 as a function of the sample trace. Similar results are obtained for sample traces a through h.

As noted above, for spectral slopes ≤ -3, the fractal dimension = 1 (17-18). As we have noted before, this interpretation of a deterministic failure process is consistent with the occurrence in the spectrum of quasi-harmonic peaks. That is, the occurrence of the

peaks, that correspond to RDX crystal sizes, as well as the magnitude of the spectral slopes, indicate that, in general, the fracture is occurring between TNT and RDX grains: intergranular failure.

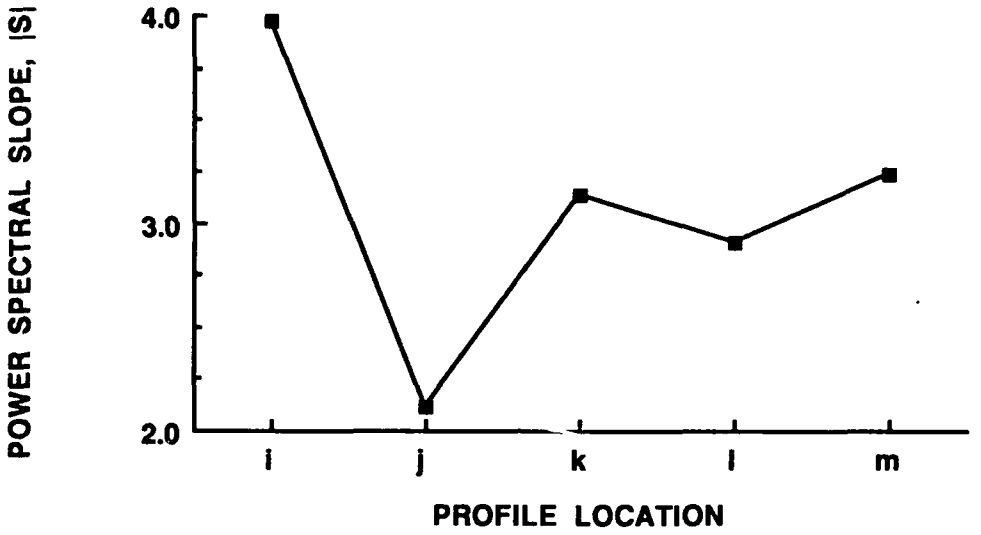


Figure 6. Spectral slopes versus profile location i through m of Composition B for spatial frequency interval 10^0 mm^{-1} to 10^1 mm^{-1}.

Equation (1) indicates that the fractal dimension for profiles j and e are 1.4 and 1.3, respectively. This indicates that the failure is chaotic and transgranular failure occurs through the grain. The chaotic region suggests that the origin of the failure, as for TNT, is in the center of the sample.

Octol. The sample of melt-cast 70/30 Octol reported here was found to fracture at an acceleration of 41,000 *g* at 25°C. The logarithmic normal crystal size distribution of HMX as specified by military specification Class 4 (22) is shown in Figure 4. The spatial power spectra of the fracture surface profiles across traces g and h (see inset) of 70/30 Octol are shown in Figure 7. The distribution of peaks in the spatial power spectra show the inhomogeneous grain size of the HMX in 70/30 Octol. The HMX grain sizes vary discontinuously from ≈ 0.1 mm to ≈ 1.0 mm in contrast to Class 4 HMX particle size in Figure 4. Figure 7 also shows how the inhomogeneous HMX grain size distribution changes with location across the surface.

In the frequency interval 10^0 mm^{-1} to 10^1 mm^{-1} the spectral slope of trace g is -3.9. The spectral slope of trace h is -2.7 in this frequency interval. The absolute value of the spectral slopes for profiles c through h are plotted in Figure 8 as a function of the sample trace. The spectral slopes for profiles i through m and c through g are ≤ -3. The fractal dimension equals one for these profiles. Peaks, that correspond to HMX grain sizes, as well as the fractal dimensions of one indicate that, in general, intergranular failure is occurring at grain boundaries.

Equation (1) indicates that the fractal dimension for profile h is 1.1. This chaotic failure indicates that transgranular failure occurs through the grain. The chaotic region suggests that the origin of the failure is at the edge of the sample.

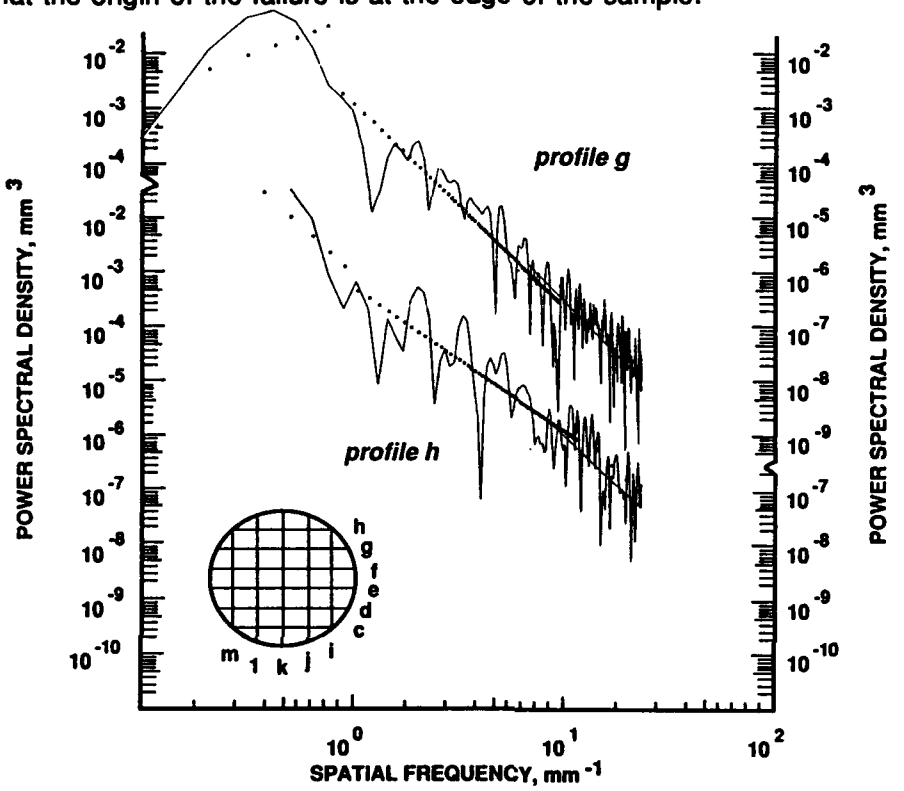


Figure 7. Spatial power spectra of the fracture surface profiles for sample traces g and h of 70/30 Octol. The locations of the profile traces across the sample surface are shown in the inset. Diamond stylus horizontal spacing is 0.004 mm.

The sample of sedimentation-cast 83/17 Octol reported here was found to fracture at an acceleration of 60,000 *g* at 25°C. The logarithmic normal crystal size distribution of HMX as specified by PIC Material Specification for Sedimentation Cast Octol (23) is shown in Figure 4. The spatial power spectra of the fracture surface profiles across traces a and f (see inset) of 83/17 Octol are shown in Figure 9. The distribution of peaks in the spatial power spectra shows the inhomogeneous grain size of the HMX in 83/17 Octol. The HMX grain size varies discontinuously from ≈ 0.1 mm to ≈ 1.0 mm in contrast to to the HMX particle size distribution used in sedimentation cast Octol as shown in Figure 4. Figure 9 also shows how the inhomogeneous HMX grain size distribution changes with location across the surface.

In the frequency interval 10^0 mm^{-1} to 10^1 mm^{-1} the spectral slope of trace f is -4.1. The spectral slope of trace a is -2.4 in this frequency interval. The absolute value of the spectral slopes for profiles a through f of 83/17 Octol are plotted in Figure 10 as a

function of the sample trace. The spectral slopes of sample traces g through l and c through f are ≤ -3. The fractal dimension equals one for these sample traces. Peaks that correspond to HMX grain sizes, as well as fractal dimensions of one, indicate that, in general, intergranular failure is occurring at grain boundaries.

The fractal dimension for profiles a and b are 1.3 and 1.2, respectively. This chaotic failure indicates that failure occurs through the grain: transgranular failure. The chaotic region suggests that the origin of the failure is at the edge of the sample.

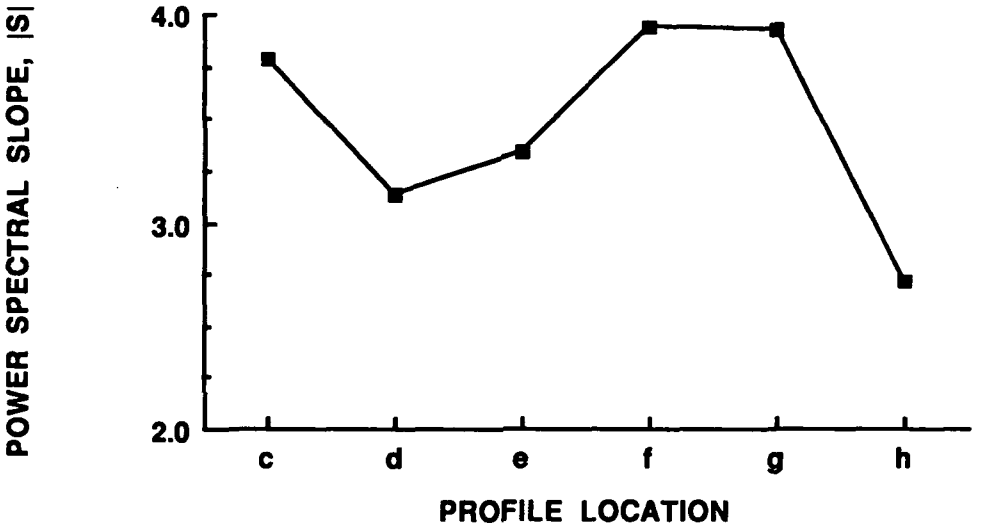


Figure 8. Spectral slopes versus profile location c through h of 70/30 Octol for spatial frequency interval 10^0 mm^{-1} to 10^1 mm^{-1}.

DISCUSSION

The results of Figures 1-3 as well as our previous work (8-12) show clearly that the fracture process under high acceleration of our TNT sample tends to occur at grain boundaries. The results of Figure 2 show that the grain-boundary determination of the fracture is strongest at the edges of the sample. That is, the spectral peaks have their largest values for the profiles near the edges of the samples. A fractal dimension of 1 for these profiles is consistent with this interpretation.

Further, the largest power level at the lowest frequency occurs for two central traces. This lowest spatial frequency is of the order of the sample size. At the center of the sample, therefore, the fracture occurs in the bulk: transgranular failure. Such a failure mode is non-deterministic, i. e. chaotic. Fractal dimensions found for two traces through the center are consistent with this interpretation.

The results of Figures 5-6 as well as our previous work (10-11) show clearly that the fracture process under high acceleration of our Composition B sample tends to occur

between the TNT and RDX grains. The results of Figure 6 show that the grain-boundary determination of the fracture is strongest at the edges of the sample. A fractal dimension of 1 for these profiles is consistent with this interpretation. Figure 6 shows that transgranular failure occurs in the center of the Composition B sample. Fractal dimensions found for two traces through the center are consistent with this interpretation.

The results of Figures 7-8 show clearly that the fracture process under high acceleration of our 70/30 Octol sample tends to occur between the TNT and HMX grains. The results of Figure 8 show that the grain-boundary determination of the fracture is strongest at the center and at three of the four edges of the sample. A fractal dimension of 1 for these profiles is consistent with this interpretation. Figure 8 shows that transgranular failure occurs at one edge of the 70/30 Octol sample. The fractal dimension found for the other edge trace is consistent with this interpretation.

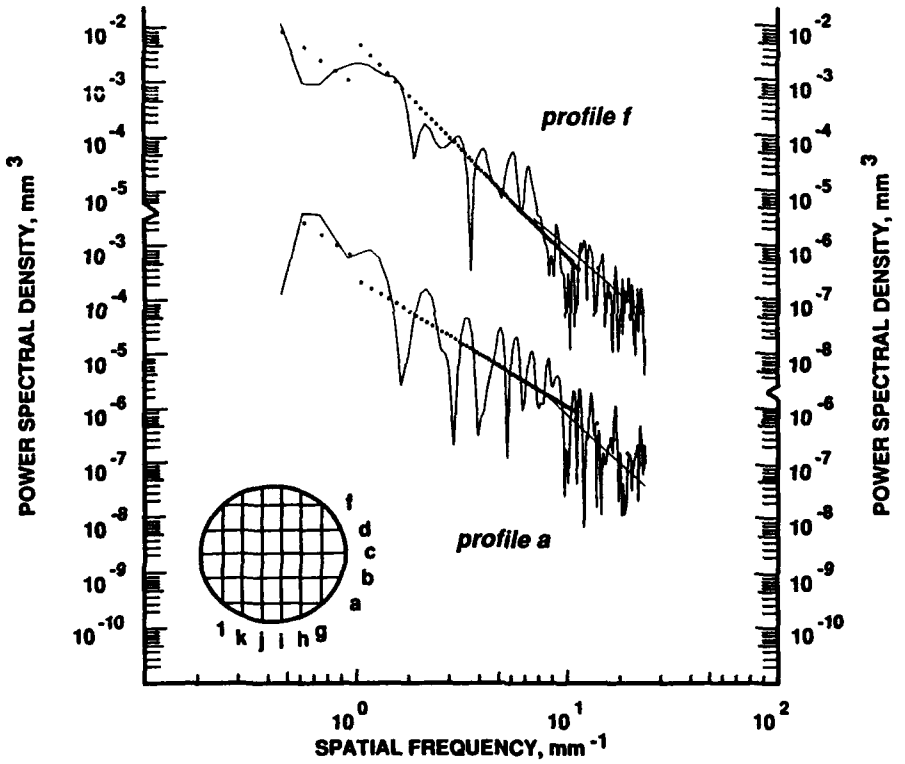


Figure 9. Spatial power spectra of the fracture surface profiles for sample traces a and f of 83/17 Octol. The locations of the profile traces across the sample surface are shown in the inset. Diamond stylus horizontal spacing is 0.004 mm.

The results of Figures 9-10 show clearly that the fracture process under high acceleration of our 83/17 Octol sample tends to occur between the TNT and HMX grains. The results of Figure 9 show that the grain-boundary determination of the fracture is strongest

at the center and at three of the four edges of the 83/17 Octol sample. A fractal dimension of 1 for these profiles is consistent with this interpretation. Figure 10 shows that transgranular failure occurs at the other edge of the sample. The fractal dimensions found for two traces at the same edge are consistent with this interpretation.

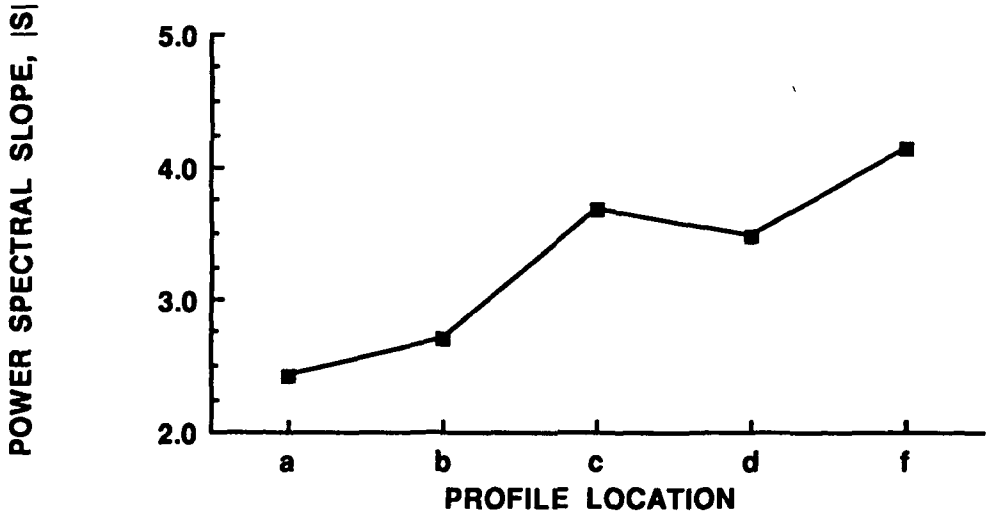


Figure 10. Spectral slopes versus profile location a through f of 83/17 Octol for spatial frequency interval 10^0 mm^{-1} to 10^1 mm^{-1}

CONCLUSIONS

Understanding of the fracture and rupture of energetic materials subjected to high acceleration is a key to better practical designs in several fields, including ordnance, the extraction industry, and space propulsion. In such applications the materials can often be subjected to high, fluctuating, and/or sustained accelerations. We have introduced fractal concepts and techniques to characterize the fracture surface topography of energetic materials obtained at high acceleration. These concepts have provided new statistical ways to characterize grain sizes of explosives. Of particular importance, these experimental and analysis techniques now allow quantitative comparisons to be made between different energetic materials so that improvements and changes in cast and composition of energetic materials can be verified. These concepts have also provided unique insights into the fundamental physics of the fracture process of explosives. For example, we have demonstrated that in general intergranular fracture occurs at grain boundaries or between grains in a mixture. We have also observed chaotic transgranular fracture through the grain that may indicate the origin of the failure in the sample. We have shown that the chaotic failure through the sample center appears to be similar for Composition B and TNT. We have also shown that the chaotic failure at the sample edge appears to be similar for 70/30 Octol and 83/17 Octol.

This work, both in the laboratory and in the field verification, is a major breakthrough in improving the understanding of the behavior of energetic materials during

the gun launch process. This work also confirms that the value of relatively inexpensive laboratory experiments conducted in an ultracentrifuge, combined with sophisticated analysis and statistical techniques, can provide important new discoveries that are relevant and important to the gun launch process. We believe the experimental and analysis techniques used here have wide applicability in future studies of energetic materials.

REFERENCES

1. M. Y. D. Lanzerotti and J. Sharma, App. Phys. Lett. 39 (1981) 455.

2. M. Y. D. Lanzerotti, "Fracture Phenomena of Energetic Materials During High Acceleration", in Proc. Symposium on Thermomechanical Properties of Energetic Materials and Their Effects on Munitions Survivability, Vol. II, (Naval Weapons Center, China Lake, CA, March 1985), pp. 1-43.

3. M. Y. D. Lanzerotti, "Mechanical Behavior of Gun Propellant During High Acceleration", in Proc. 21st JANNAF Combustion Meeting, Vol. 1, (Johns Hopkins University, Laurel MD, 1984), pp. 275-280.

4. M. Y. D. Lanzerotti and J. Pinto, "Chemical Reaction of Energetic Materials During High Acceleration", in Shock Waves in Condensed Matter, Y. M. Gupta, Ed., (Plenum Publishing Corp., New York, 1986), pp. 909-916.

5. M. Y. D. Lanzerotti, "Dynamic Response and Defect Study by Ultracentrifuge Simulation of the Gun Launch Process." FY82 Final ILIR Report, August 1982.

6. M. Y. D. Lanzerotti, et al, "Explosive Properties of Octol Castings (STAFF)", May 1988.

7. M. Y. D. Lanzerotti and J. Pinto, "Fractal Dimension of Fracture Surfaces of Energetic Materials", in Fractal Aspects of Materials II, D. W. Schaefer, R. B. Laibowitz, B. B. Mandelbrot and S. H. Liu, Eds., (Materials Research Society, Pittsburgh PA, 1986) pp. 133-4.

8. M. Y. D. Lanzerotti, J. Pinto, A. Wolfe, and S. Morrow, "Fractal Properties of Fracture Surfaces of Energetic Materials", in Proceedings of the 1988 Army Science Conference, Ft. Monroe, VA, 25-28 October 1988, Volume II, pp. 113-125.

9. M. Y. D. Lanzerotti, J. Pinto and A. Wolfe, "Broad Bandwidth Study of the Topography of the Fracture Surfaces of Energetic Materials", in New Materials Approaches to Tribology: Theory and Applications, L. W. Pope, L. L. Fehrenbacher, W. O. Winer, Eds., (Materials Research Society, Pittsburgh PA, 1989), pp. 411-416. (Volume 140, Materials Research Society Symposium Proceedings)

10. M. Y. D. Lanzerotti, J. Pinto and A. Wolfe, "Broad Bandwidth Study of the Topography of the Fracture Surfaces of Explosives", in The Ninth Symposium (International) on Detonation, Volume 1, pp. 355-361, 28 August - 1 September 1989, Portland OR.

11. M. Y. D. Lanzerotti, J. Pinto and A. Wolfe, Characterization of the Mechanical Failure Surfaces Of Energetic Materials By Power Spectral Techniques, in Fractal Aspects of Materials - 1989, eds. J. K. Kaufman, J. E. Martin, and P. W. Schmidt (Mater. Res. Soc. EA-20, Pittsburgh PA, 1989) pp. 225-7.

12. M. Y. D. Lanzerotti, J. Pinto, and A. Wolfe, "Fracture Surface Topography of Cast TNT", in Proceedings of the Seventh APS Topical Conference on Shock Waves in Condensed Matter, Williamsburg VA, 16-60 June 1991, in press.

13. F. P. Bowden and A. D. Joffe, Initiation and Growth of Explosion in Liquids and Solids (Cambridge University Press, 1952).

14. D. J. Thomson, Spectral Analysis of Short Series, Ph. D. Dissertation, Department of Electrical Engineering, Polytechnic Institute of Brooklyn, Brooklyn NY, 1971.

15. D. J. Thomson, Proc. IEEE, 70, 1055 (1982).

16. D. J. Thomson, M. F. Robbins, C. F. Maclennan, and L. J. Lanzerotti, Physics of the Earth and Planetary Interiors, 12, 217 (1976).

17. S. R. Brown and C. H. Scholz, J. Geophys. Rev. 90, 12,575 (1985).

18. S. E. Hough, Geophys. Res. Lett. 16, 673 (1989).

19. B. B. Mandelbrot, The Fractal Geometry of Nature (Freeman, NY, 1983).

20. B. B. Mandelbrot, Physica Scripta, 32, 257 (1985).

21. RDX, MIL-R-398C, 22 August 1962.

22. HMX, MIL-H-4544B (PA), 27 February 1974.

23. PIC Material Specification for Sedimentation Cast Octol, 1986.

Thermodynamic Analyses of Transmembrane Channel Formation by Botulinum Toxin

FRANK J. LEBEDA*, Ph.D.
Neurotoxicology Branch, US Army Medical Research Institute of Chemical Defense, Aberdeen Proving Ground, MD 21010-5425

MARY K. GENTRY, B.S.
Biological Chemistry Branch, Walter Reed Army Institute of Research, Washington, D.C. 20307-5100

DALLAS C. HACK, MAJ, M.D.
Virology Division, US Army Medical Research Institute of Infectious Diseases, Frederick, MD, 21702-5011

INTRODUCTION

The peptide neurotoxins from *Clostridium botulinum* (botulinum toxin; BoTX) are the most lethal substances known. At the present time, seven immunologically distinct serotypes (A-G) have been identified and are the subjects of numerous and extensive reviews[1-4]. The BoTX-A serotype is the most toxic species and can now be massively cloned with low-cost genetic engineering techniques[5]. Although a pentavalent toxoid is available (for the A, B, C1, D and E serotypes), clinically relevant prophylaxis takes months to develop. Immunizing personnel directly by injecting antibodies for the different serotypes is currently being developed, but despite this progress there are no effective antidotes that can be administered to exposed individuals who are either unvaccinated or who have insufficient titers of the circulating, neutralizing antibodies. Thus, because of their extreme potency, availability, and the lack of an adequate set of treatment compounds, development of medical countermeasures to BoTX is considered by the Medical Biological Defense Program as high-priority in accordance with the seriousness of this agent's threat to U.S. military personnel.

These concerns raise additional issues. For example, other, but as yet uncharacterized, naturally occurring toxin serotypes may exist and that the present approach to counteract these new species with specific antibodies would require further developmental efforts and attendant high costs. Furthermore, site-directed mutagenesis of genetic material could conceivably produce an infinite number of novel toxins that could conceivably circumvent the presently available set of neutralizing antibodies. In light of these problems, an alternative strategy to discover and develop new countermeasures forms the basis of the present research. This study's ultimate goal is to identify and develop single pharmaceuticals or other medical products for clinical care applications that can counteract various BoTX serotypes (rather than one for each serotype) regardless of their biological origin or source. To achieve this goal, a detailed knowledge of the toxin mechanism of action is required.

The primary toxic effect of all the BoTX serotypes (excluding the C2 and C3 serotypes[6]), is flaccid paralysis of skeletal muscles. The toxin prevents the release of acetylcholine in the periphery, blocks neurally induced muscle contraction[7,8] which can eventually cause respiratory failure and death. The toxin also inhibits the release or exocytosis of other neurally active substances in cholinergic and noncholinergic systems, but at concentrations that are much higher than those required to exert effects at the cholinergic peripheral synapse. This discriminating action is presumably due to the selective binding of the toxin to receptors that are predominantly localized at these presynaptic endings[3].

Like diphtheria toxin (DiTX) which inhibits protein synthesis[9], the *Clostridial* neurotoxins (including tetanus; TeTX) block neurochemical transmission by a complex series of reactions that is related to their tripartite structure. The C-terminal half of the heavy (H) chain binds to the target site (peripheral presynaptic nerve terminal), the N-terminal half is involved in the endocytotic internalization process, while the light (L) chain exerts the toxic effect. Once internalized, the release (translocation) of the toxic L-chains from the endocytotic vesicles involves a *p*H-dependent process that may include the formation of ionic channels. Although the heavy chains of TeTX and DiTX (and their N-terminal halves) are capable of forming functional ionic channels in artificial membranes[10-16], studies with DiTX indicate that hydrophobic segments of the light chain can also interact with the membrane[17] and take part in translocation.

If a process in this multi-reaction scheme is common to all of the serotypes, then it is possible that a single treatment drug could be found and developed for clinical use. Evidence from biochemical studies suggests that the first step in the scheme outlined above differs for the various serotypes. At least two different serotype binding sites exist[2]. From electrophysiological data, the pattern of the final, toxic reaction also varies somewhat among the serotypes[18]. In contrast, the intermediate process, the toxin translocation step, is not as well characterized and may yield information on a reaction that the neurotoxin serotypes could have in common.

Because ionic channel formation by the toxins is believed to play a role in their internalization we focused on identifying regions within the heavy chains of BoTX, TeTX and DiTX that could be important in forming ionic channels. Our approach in localizing channel-forming regions was to use several different computer-assisted calculations to determine free energies of peptide insertion into biological membranes. The main assumption in this study was that channels are comprised of sequences containing both hydrophobic and hydrophilic residues[19-22]. Our aim was to determine where sufficiently long hydrophobic regions occurred and whether they included amphipathic segments. In conjunction with the presently known electrophysiological properties of these channels, our calculations support the conclusion that short segments located in the N-terminal half of the toxin heavy chains could participate in channel formation.

METHODS

Amino acid sequences of the BoTX-A, C1, and D serotypes and TeTX were obtained from the National Biomedical Research Foundation (NBRF) database at the Advanced Scientific Computing Laboratory (National Cancer Institute, Frederick, MD). Peptide search, comparison and analysis software included FASTA[23], the Sequence Analysis Software Package (Genetics Computer Group, Inc.[24]), the IDEAS package[25] and the programs supported by the Protein Identification Resource (NBRF, Georgetown University Medical Center). Programs on personal microcomputers included the sequence alignment program MACAW (Center for Biotechnology Information[26]) and the MacVector analysis package (IBI, New Haven, CT).

The Kyte-Doolittle hydropathic index[27] was calculated using the authors'

SOAP program and amino acid hydropathy scale (*i.e.*, the free energy of transfer of an amino acid from a nonpolar medium to water). This index was calculated as a sum of hydrophobicity values of short peptide segments contained within a consecutively moving window. A second program (GES) was written by us to calculate this free energy of transfer using the amino acid polarity scale, window size and method described by Engleman *et al.*[28].

Analysis of peptide amphipathicity used the MOMENT program which uses the Fourier transform to calculate the sequence hydrophobic moment:

$$<\mu_H> = \{[\sum_{n=1}^{N} H_n \sin(\delta n)]^2 + [\sum_{n=1}^{N} H_n \cos(\delta n)]^2\}^{1/2} \qquad \text{(eq. 1)}$$

where N = window size, H_n = hydrophobicity of the n^{th} residue, δ = an angle from 0-180° (in radians) and a consensus scale for the free energy of transfer of an amino acid residue from a nonpolar to a polar medium[19,20,29,30]. Multimeric TM segments were defined on the basis of values for the moment and the hydrophobicity of the sequence of amino acids within a window. Two different amino acid scales were used with MOMENT. The consensus hydrophobicity scale of Eisenberg was used initially. For comparison, we also developed a novel amino acid scale that is based on the frequency of an amino acid's occurrence in the (nonhelical) nontransmembrane and helical (transmembrane) segments of the photosynthetic reaction center (see Table II in ref. 31). The values for N for each of these algorithms were selected on the basis of test runs using the amino acid sequence of the M-chain of the photosynthetic reaction center of *R. viridis*[31]) when all five of the known membrane-spanning regions of this peptide were detected.

RESULTS

Sequence alignments

Blocks of similar sequences within the toxins examined in this study were determined with the MACAW program[26] using a modified PAM-120 matrix[32]. Regions of amino acid similarity are represented as the wider blocks in Fig. 1. Gaps were automatically inserted by MACAW to optimize the alignment. The GCG GAP program[33] was used to independently quantify the amount of similarity in the aligned peptide pairs (Table 1). Similarity in the single chain sequences of

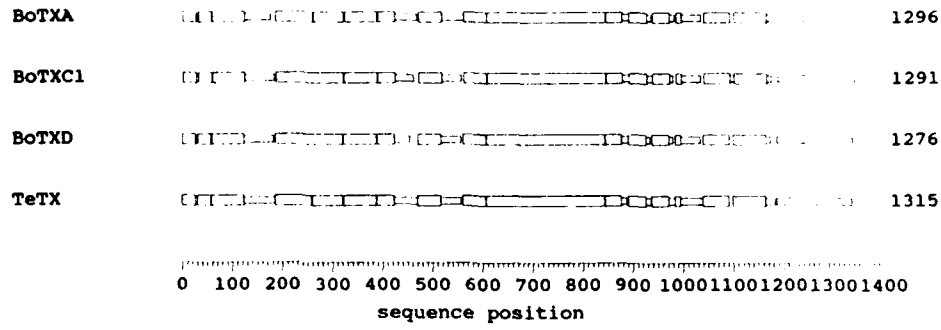


Fig.1. Schematic alignment of the primary structures of BoTX-A, C1 and D and TeTX. The larger rectangles represent regions of similar residues. The N-terminal of the L-chain is on the left and the C-terminal of the H-chain is on the right. The H-chain of BoTX-A starts at residue 448. Gaps were optimally inserted by MACAW.

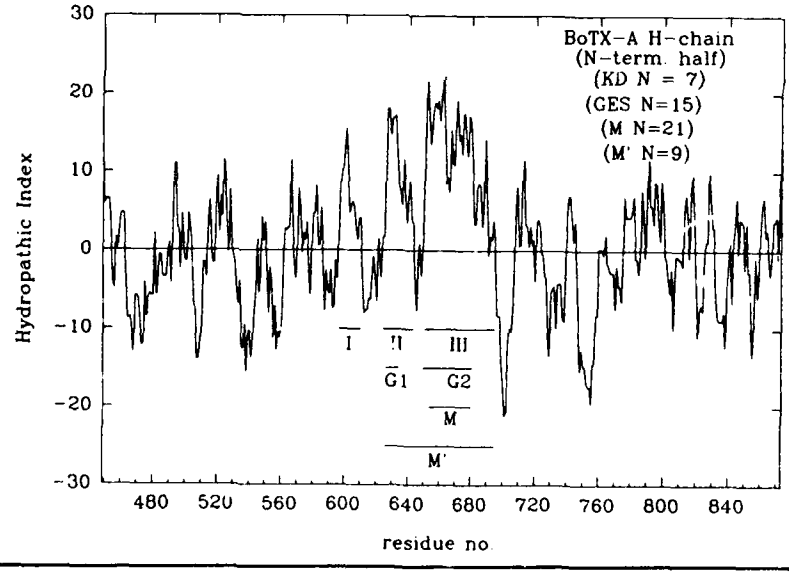


Fig. 2. The Kyte-Doolittle hydropathic analysis reveals potential membrane-spanning sections on the N-terminal half of the H-chain in BoTX-A. Three relatively long hydrophobic regions are indicated by the horizontal lines labelled I, II and III. Window size (N) = 7. Regions calculated to be transmembrane by GES (G1 and G2) and MOMENT (M and M') are also shown with their respective window sizes.

BoTX (A, C1, D) and TeTX ranged from 55.2 to 68.9%. The highest similarity scores (73, 69, and 61%) were obtained between BoTX-C1 and BoTX-D. The interchain comparison between the L- and the H-chains of TeTX and BoTX-A, respectively, showed about a 40% similarity (not shown). Because comparable percentages resulted when randomized sequences were used, data from the L- and H-chain comparisons were not considered to be significant. These results contrast with the conclusion that the two chains of TeTX are similar[34]. In contrast to the similarities in the primary structures between the BoTX serotypes and TeTX, only a short segment in the B-chain of DiTX was similar to aregion in the H-chains of the *Clostridial* toxins. The importance of this region was revealed during the amphipathic analysis (see below).

Hydropathic analysis

The KD algorithm and amino acid hydrophobicity scale developed by Kyte and Doolittle[27] were used in this portion of the study to obtain the first estimate of the number of possible TM segments in the *Clostridial* toxin H-chains. The horizontal bars in Fig. 2 show the extent of the three calculated hydrophobic regions in the N-terminal half of the BoTX type A H-chain. From the MACAW analysis, these regions, in general, were similar in the other toxins examined. For convenience, these three sequences will be referred to in this paper as toxin hydrophobic regions I (residues 596-610), II (624-643) and III (650-695).

The computational approach developed by Engleman, Steitz and Goldman[28] was the second method used to estimate the number of TM segments. It is similar to the KD algorithm except that it uses a different amino acid scale and required a larger window size. The results from the GES program shown in Fig. 3 indicate that there are two potential membrane insertion regions within the BoTX-A H-chain. These regions have residues in common with hydrophobic region II (625-630; labelled G1 in Fig. 2) and the N-terminal half of hydrophobic region III (IIIa; 647-676; labelled G2 in Fig. 2). Although the first region has too few residues to span a biological membrane, it was still considered to be an important area (see below). As with the KD algorithm, similar results were obtained with the other toxins examined which again emphasizes the similarities of the different BoTX serotypes when considering these TM, potentially channel-forming segments.

Table 1

Comparison of whole toxin (W), light (L) and heavy (H) chains of selected botulinum toxin serotypes and tetanus toxin

	BoTX-A(W)	BoTX-C1(W)	BoTX-D(W)	TeTX(W)
BoTX-A	1/1[¶]	0.552/0.339	0.552/0.343	0.552/0.354
BoTX-C1	-	1/1	0.689/0.536	0.562/0.351
BoTX-D	-	-	1/1	0.553/0.341
TeTX	-	-	-	1/1
	BoTX-A(L)	**BoTX-C1(L)**	**BoTX-D(L)**	**TeTX(L)**
	(1-448)[$]	(1-449)	(1-445)	(1-457)
BoTX-A(L)	1/1	0.526/0.345	0.544/0.341	0.528/0.321
BoTX-C1(L)	-	1/1	0.611/0.467	0.555/0.348
BoTX-D(L)	-	-	1/1	0.549/0.345
TeTX(L)	-	-	-	1/1
	BoTX-A(H)	**BoTX-C1(H)**	**BoTX-D(H)**	**TeTX(H)**
	(449-1296)	(450-1291)	(446-1276)	(458-1315)
BoTX-A(H)	1/1	0.565/0.337	0.558/0.343	0.572/0.379
BoTX-C1(H)	-	1/1	0.730/0.572	0.566/0.354
BoTX-D(H)	-	-	1/1	0.561/0.345
TeTX(H)	-	-	-	1/1

[¶] Values represent the fraction of similar/the fraction of identical amino acid residues for a given pair of aligned sequences.
[$] Values represent the range of amino acid positions.

Hydrophobic moment analysis

Transmembrane amphipathic regions were identified based on values of their sequence hydrophobic moment (eq. 1) and hydrophobicity[22]. Both the consensus amino acid free energy scale derived by Eisenberg *et al.*[29] and the scale that was developed by us (see Methods) were used with MOMENT. Using the Eisenberg

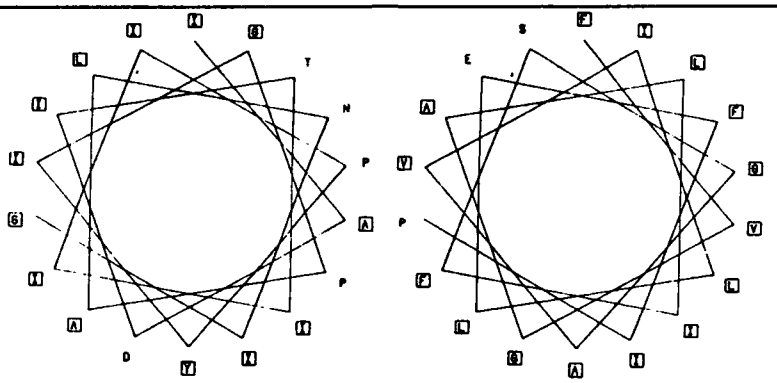


Fig. 3. Helical wheel plots of BoTX-A show the relative orientation of the hydrophobic (boxed) and hydrophilic (unboxed) residues in region II (residues 627-644; left panel) and the N-terminal half of region III (652-669; right panel).

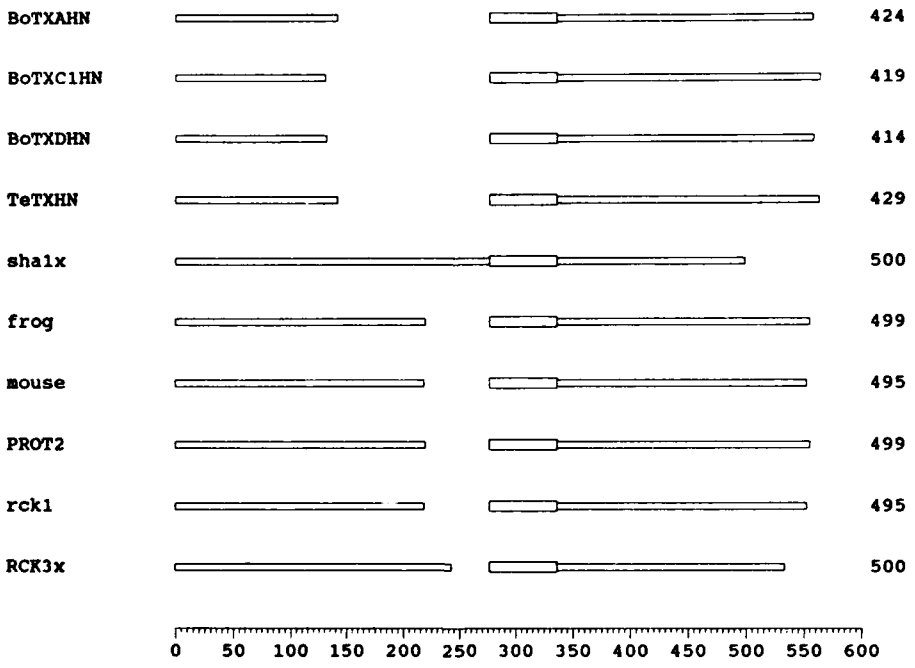


Fig. 4. Sequences of amino acid residues in the hydrophobic regions of BoTX-A, C1 & D and TeTX are similar to those of transmembrane segments S2 and S3 of peptides forming voltage-gated potassium channels from a wide variety of species. Only the N-terminal halves of the H-chains (HN) are illustrated. The residue positions are relative to their location on the H-chain.

consensus scale, a single region (BoTX-A residues 625-695) was predicted by the MOMENT program to be amphipathic. This region corresponded to hydrophobic region IIIa (labelled M in Fig. 2). In additional calculations, the region found by MOMENT using our novel amino scale also included hydrophobic region II (labelled M' in Fig. 2).

Although the low values of $<\mu_H>$ for BoTX could not be categorized as strongly reflecting either an amphipathic helix or ß-strand in these single regions, a helical arrangement for this segment was considered. One possible configuration of this amphipathic region was visualized by the helical wheel display[35]. Plots shown in Fig. 3 illustrate the relative orientation of hydrophobic (boxed) and hydrophilic (unboxed) residues in regions II and IIIa (the N-terminal half of region III) of BoTX-A. These hydrophobic residues tend to be clustered within one quarter of a possible helical conformation, an arrangement that will be discussed below.

Additional sequence alignments

Interestingly, one of the two amphipathic regions in the DiTX heavy (B) chain (not shown) was found by MACAW to be similar to the single amphipathic sequence (region III) in both BoTX-A and TeTX. This alignment suggests that the channel-forming regions of these three different toxins may have some common structural features.

A final set of sequence comparisons was performed for which one result is illustrated in Fig. 4. Amino acid residues in hydrophobic regions I and II of BoTX-A and TeTX are similar to those in the putative TM S2 and S3 segments of voltage-gated K^+ channels in vertebrate and invertebrate neurons[36]. In addition, there were significant matches between these toxin regions and the M2/M3 segments within the ß- and γ-subunits of peripheral nicotinic receptors[37] (not shown). The toxin hydrophobic regions, however, did not resemble the sequences that are considered to line the ionic conduction pathways and function as selectivity filters. Similarity between peptide toxins and membrane channels, nevertheless, lends plausible evidence in support of the hypothesis that amphipathic regions in these toxins are involved in the formation of ionic channels.

DISCUSSION

Based on three different sets of calculations, a section containing the necessary hydrophobic and hydrophilic residues for channel construction was localized in the N-terminal portions of the H-chains of all the toxins examined. Two regions within this amphipathic section could span biological membranes. It is not clear from these calculations, however, if these two regions are helices or ß-strands. Experiments with TeTX H-chain fragments[38] showed that the N-terminal portion of the H-chain is involved in a *p*H-dependent translocation process and according to Binz *et al.*[5], this region includes regions II and III as defined by the KD analysis.

Possible TM regions in the H-chains that were identified in this study lack the series of charged residues that exist in voltage-gated K^+ channels[36]. It is possible that the toxin segments serve only as structural components[39]. On the other hand, the importance of uncharged hydrophilic residues lining ionic channels has been emphasized by Eisenman and coworkers[40]. Because of the similarity of some residues with those in the TM regions of biological channels, it is likely that these regions adopt similar secondary structures. Despite apparent similarities in peptide structure between toxins and ionic channels, the cautionary remarks of Doolittle[41] on spurious matches and inferences about ancestral relationships must be also considered.

The schematic diagrams in Fig. 5 may be used as a starting point in portraying the formation of ionic channels by toxin H-chains. If only a relatively short portion of the H-chain traverses the membrane to form two antiparallel segments, it would be expected that both the N- and C-terminals would remain on the *cis* side of the membrane (Fig. 5A-C). The single amphipathic region that was predicted was divided into two hydrophobic regions (II, IIIa). The steps for binding (5A), insertion (5B) and alignment (5C) are shown from a view lateral to the membrane (parallel lines). Further details in the alignment process are illustrated in the view normal to the membrane (Fig. 5D). Shaded areas that correspond to the hydrophilic residues depicted in the helical wheel plots (Fig. 3) are aligned toward the hydrophilic residues located within the opposite region.

From the electrophysiological evidence obtained with artificial membranes, the concentration-effect relation for DiTX and BoTX is a power function[11,14], a

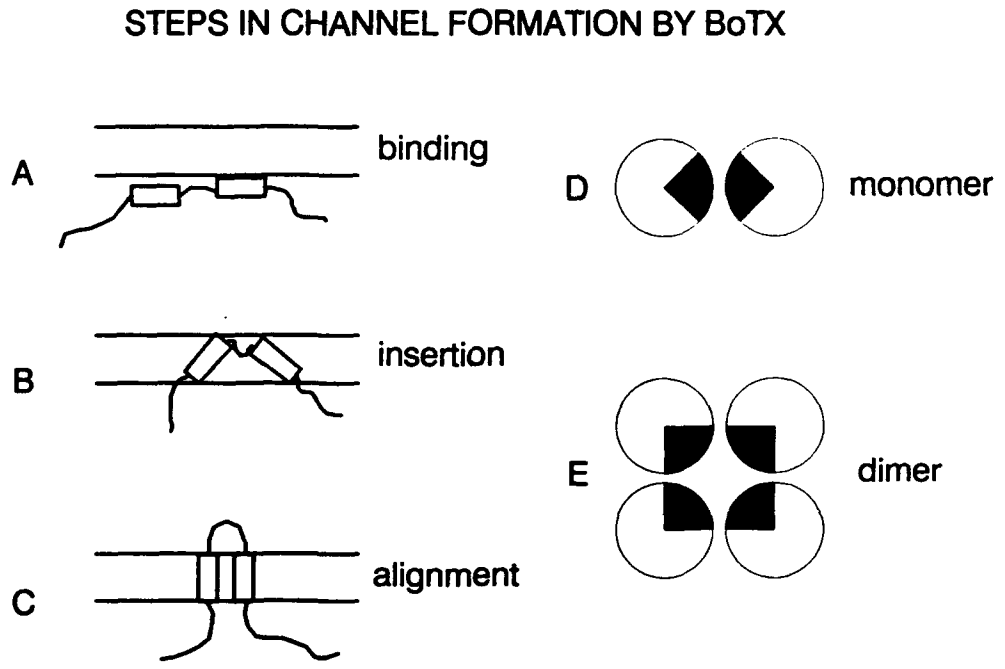


Fig. 5. Two hydrophobic regions (rectangles) within the single amphipathic region are shown to bind (A), insert (B) and align (C) themselves starting on the cis side of the membrane. The resulting transmembrane topology predicts that both the N- and the C-termini of the toxin H-chain are located on the cis side. The top view (D) shows how the hydrophilic residues (shaded areas) could be aligned. E: A dimer forms a membrane-embedded hydrophobic structure whose core contains hydrophilic residues.

result which suggests that there must be at least two toxin molecules involved in a conducting channel. The illustration of a dimeric arrangement in Fig. 5E is based on the postulated topology of peptides that form voltage-[42,43] and agonist-gated channels[44,45]. These segments are depicted to have residues with hydrophilic side chains that line a core, interleaved between residues with hydrophobic side chains that face toward the lipid environment[46].

Toxin insertion was viewed in our study as a spontaneous, thermodynamically favorable reaction, an approach that was prompted by the signal peptide hypothesis[47]. We have intentionally focused only on the toxin H-chains and have not incorporated other data that are available on the translocation process. For example, both chains of DiTX have been demonstrated to interact with lipids[17,48] or undergo structural alterations that occur with changes in *p*H[48,49]. Further, hydrophobic and other noncovalent interactions between the chains may also play a vital role. Misler[12] has suggested that the DiTX B-chain envelopes the A-chain either during synthesis or membrane insertion. Once this complex is inserted, the A-chain migrates to the *trans* side, leaving the B-chain within the membrane to serve as a residual structure that has properties of an ionic channel.

The main goal of this study was to localize regions involved in the ionic channels formed by the H-chains of BoTX and TeTX. Because the amino acid sequences in the H-chains of these toxins are more than 50% similar (Table 1) and that all are predicted to have hydrophobic and amphipathic TM segments in the same region (Fig. 2), it is possible that the ionic channels formed will be structurally and functionally similar. Further, we hypothesize that the toxin-induced ionic channels represent targets common to all of the diverse serotypes and that these structures serve as focal points for the development of therapeutic drugs. Rather than develop drugs that are unique for each serotype (because of their different antigenic, binding and toxic properties) it is proposed that a single drug could be developed that would be an effective countermeasure against all of the serotypes that form ionic channels.

ACKNOWLEDGEMENTS

Staff support and allocation of computer time was generously provided by the National Cancer Institute's Advanced Scientific Computing Laboratory of the Frederick Cancer Research Facility.

REFERENCES

1. Simpson, L.L. *Pharmacol. Rev.* **33**:155-188 (1981).
2. Habermann, E. & Dreyer, F. *Curr. Topics Micro. Immunol.* **29**:93-179 (1986).
3. Middlebrook, J.L. in *Botulinum neurotoxin and tetanus toxin*, Academic Press, New York, 1989), pp. 95-119.
4. Niemann, H. in *Sourcebook of bacterial protein toxins*, J.E. Alouf & J.H. Freer, Eds., (Academic Press, London, 1991), pp. 303-348.
5. Binz, T., Kurazono, H., Wille, M., Frevert, J., Wernars, K. & Niemann, H. *J. Biol. Chem.* **265**:9153-9158 (1990).
6. Aktories, K. & Hall, A. *Trends Pharmacol. Sci.* **10**:415-418 (1989).
7. Dickson, E.C. & Shevky, E. *J. Exp. Med.* **38**:327-346 (1923).
8. Guyton, A.C. & MacDonald, M.A. *Arch. Neurol. Psychiat.* **57**:578-592 (1947).
9. Collier, R.J. (1982) in *ADP-ribosylation reactions*, D. Hayaishi, D. & K. Ueda, K., Eds. (Academic Press, New York, 1982), pp. 575-592.
10. Donovan, J.J., Simon, M.I., Draper, R.K. & Montal, M. *Proc. Natl. Acad. Sci. USA* **78**:172-176 (1981).
11. Kagen, B.L. Finkelstein, A. & Colombini, M. *Proc. Natl. Acad. Sci. USA* **78**:4950-4954 (1981).
12. Misler, S. *Biophys. J.* **45**:107-109 (1984).
13. Hoch, D.H., Romero-Mira, M., Ehrlich, B.E., Finkelstein, A. DasGupta, B.R. & Simpson, L.L. *Proc. Natl. Acad. Sci. USA* **82**:1692-1696 (1985).
14. Donovan, J.J. & Middlebrook, J.L. *Biochem.* **25**:2872-2876 (1986).
15. Blaustein, R.O., Germann, W.J., Finkelstein, A. & DasGupta, B.R. *FEBS Let.* **226**:115-120 (1987).
16. Shone, C.C., Hambleton, P. & Melling, J. *Eur. J. Biochem.* **167**:175-180 (1987).
17. Montecucco, C. Sciavo, G. & Tomasi, M. *Biochem. J.* **231**:123-128 (1985).
18. Molgo, J., Comella, J.X., Anggaut-Petit, D., Pecot-Dechavassine, M., Tabti, N., Faille, L., Mallart, A. & Thesleff, S. *J. Physiol. (Paris)* **84**:152-166 (1990).
19. Eisenberg, D., Schwarz, E., Komaromy, M. & Wall, R. *J. Mol. Biol.* **179**:125-142 (1984a).
20. Eisenberg, D., Weiss, R. & Terwilliger, T.C. *Proc. Natl. Acad. Sci. USA* **81**:140-144 (1984b).
21. Finer-Moore, J. & Stroud, R. *Proc. Natl. Acad. Sci. USA* **81**:155-159 (1984).
22. Jähnig, F. in *Prediction of protein structure and the principles of protein conformation*, G.D. Fasman, Ed. (Plenum Press, New York, 1989), pp. 707-717.
23. Lipman, D.J. & Pearson, W.R. *Science* **227**:1435-1441 (1985).

24. Devereux, J., Haeberli, P. & Smithies, O. *Nucl. Acids Res.* 12:387-395 (1984).
25. Kanehisa, M., Fickett, J.W. & Goad, W.B. *Nucl. Acids Res.* 12:149-158 (1984).
26. Schuler, G.D., Altschul, S.F. & Lipman, D.J. *Proteins Struct. Funct. Gen.* 8:180-190 (1990).
27. Kyte, J. & Doolittle, R.F. *J. Mol. Biol.* 157:105-132 (1982).
28. Engleman, D.M., Steitz, T.A. & Goldman, D. *Ann. Rev. Biophys. Biophys. Chem.* 15:321-353 (1986).
29. Eisenberg, D., Weiss, R. M. Terwilliger, T.C. & Wilcox, W. *Faraday Symp Chem. Soc.* 17:109-120 (1982).
30. Eisenberg, D., Wesson, M. & Wilcox, W. in *Prediction of protein structure and the principles of protein conformation*, G.D. Fasman, Ed. (Plenum Press, New York, 1989), pp. 719-759.
31. Michel, H., Weyer, K.A., Gruenberg, H., Dunger, I., Oesterhelt, D. & Lottspeich, F. *EMBO J.* 5:1149-1158 (1986).
32. Dayhoff, M.O., Schwartz, R.M. & Orcutt, B.C. in *Atlas of protein sequence and structure*, M.O. Dayhoff, Ed. (Natl. Biomed. Res. Found., Washington, D.C., 1978), Vol. 5, Suppl. 3, pp. 345-352.
33. Needleman, S.B. & Wunsch, C.D. *J. Molec. Biol.* 48:443-453 (1970).
34. Taylor, C.F., Britton, P. & Van Heyningen, S. *Biochem. J.* 209:897 (1983).
35. Schiffer, M. & Edmundson, A.B. *Biophys. J.* 7:121-135 (1967).
36. Temple, B.L., Papzian, D.M., Schwarz, T.L., Jan, Y.N. & Jan, L.Y. *Science* 237:770-775 (1987).
37. Imoto, K., Busch, C., Sakmann, B., Mishina, M., Konno, T., Nakai, J., Bujo, H., Mori, Y., Fukuda, K. & Numa, S. *Nature* 335:645-648 (1988).
38. Rao, J.K. & Boquet, P. *J. Biol. Chem.* 260:6827-35 (1965).
39. Guy, H.R. & Conti, F. *Trends Neurosci.* 13:201-206 (1990).
40. Eisenman, G., Oberhauser, A. & Bezanilla, F. in *21th Jerusalem symposium on quantum chemistry*, A. Pullman, J. Jortner & B. Pullman, Eds. (Kluwer Academic, Dordrecht, Netherlands, 1989) pp. 27-50.
41. Doolittle, R.F. in *Prediction of protein structure and the principles of protein conformation*, G.D. Fasman, Ed. (Plenum Press, New York, 1989), pp. 193-301.
42. Oiki, S., Danho, W. & Montal, M. *Proc. Natl. Acad. Sci. USA* 85:2393-2397 (1988).
43. Hartmann, H.A., Kirsch, G.E., Drewe, J.A., Taglialatela, M., Joho, R.H. & Brown, A.M. *Science* 251:942-944 (1991).
44. Montal, M., Montal, M.S. & Tomich, J.M. *Proc. Natl. Acad. Sci. USA* 87:6929-

6933 (1990).
45. Galzi, J.-C., Revah, F., Bessis, A. & Changeux, J.P. *Annu. Rev. Pharmacol.* **31**:37-72 (1991).
46. Sansom, M.S.P. *Prog. Biophys. Molec. Biol.* 55:139-235 (1991).
47. Von Heijne, G. *Curr. Top. Mem. Transp.* **24**:151-179 (1985).
48. Jiang, J.X., Abrams, F.S. & London, E. *Biochem.* **30**:3857-3864 (1991).
49. Cabiaux, V., Brasseur, R., Wattiez, R., Falmagne, P., Ruysschaert, J.-M. & Goormaghtigh, E. *J. Biol. Chem.* **264**:4928-4938 (1989).

Ohmic Contacts to Heavily Carbon-Doped p^+-GaAs Using Ti/Si/Pd

*H.S. Lee,Dr., M.W. Cole,Ms., R.T. Lareau,Dr., L. Casas, Mr. R.J. Thompson,Mr., A. DeAnni,Mr., and K.A. Jones,Dr.
U.S. Army Electronics Technology and Devices Laboratory, Fort Monmouth, NJ 07703-5601
W.Y. Han,Mr., Y. Lu,Dr.
Dept. of Electrical and Computer Engineering, Rutgers University, New Brunswick, N.J. 08855
L.W. Yang,Dr.
Ford Microelctronics, Colorado Springs, CO 80921; Current Address: General Electric Co., Electronics Laboratory, Syracuse, N.Y. 13221-4840.

INTRODUCTION:

The fabrication of high speed heterojunction bipolar transistors (HBT) requires a heavily doped p-type base region. [1,2] The use of carbon as an acceptor dopant is of interest in HBTs due to the ability to incorporate high dopant levels ($> 10^{19}$ cm^{-3}) and the low diffusion coefficient ($\sim 6 \times 10^{-15}$ cm^2/sec at 900 C) [3] of carbon as compared to the more commonly used Zn and Be dopants. This low diffusion coefficient for carbon allows implant activation to be done at high temperatures that would otherwise lead to device degradation via Zn or Be diffusion. Accordingly, stable ohmic contacts to the p^+-GaAs are desirable. Silicides have been shown to be stable to Si [4] as well as GaAs and offer low resistivity contacts and Schottky contacts [5] that are stable at temperatures above 400 C. [6-8] Alternative contact metallizations to the Au-based alloyed systems for p-GaAs have included Pt/Ti [9,10], Pd/Sb[11] and WSi_X.[8] In this study ohmic contacts were formed to p^+ GaAs using the Pd/Si/Ti/GaAs system. The use of

LEE, COLE, LAREAU, CASAS, THOMPSON, DeANNI, JONES, HAN, LU, and YANG

Pd allows for the formation of PdSi or Pd_2Si at low temperatures and produces a low resistance silicide that is highly resistant to oxidation. The usefulness of a silicide based ohmic contact system to p^+ carbon-doped GaAs will be explored.

EXPERIMENTAL:

Epitaxial layers of GaAs were grown on (100) semi-insulating (SI) GaAs substrates by low pressure organometallic vapor phase epitaxy (LPOMVPE). Trimethylgallium and arsine are the group III and V sources respectively, with CCl_4 as a carbon doping source. A 3500 Å undoped buffer layer was grown first followed by the growth of a 0.8 μm thick carbon-doped (5×10^{19} cm^{-3}) GaAs layer. The samples were then degreased, etched in $NH_4OH:H_2O$ (1:10), rinsed in de-ionized water, blown dry and loaded into an electron-beam evaporator. The sample structure evaporated was Pd(450 Å)/Si (1250 Å)/ Ti (250 Å)/ GaAs substrate. The thicknesses were chosen to insure that a sufficient supply of Si was available for the formation of both PdSi or Pd_2Si and $TiSi_2$ layers. The samples were heated using rapid thermal annealing (RTA) in forming gas (12% H_2: 88% N_2 in the temperature range 400-700 C for 10 sec. The Transmission Line Model (TLM)[12] was used to measure contact resistance parameters. Interfacial reactions, microstructure, and elemental diffusion in the contact were studied using transmission electron microscopy (TEM) and Auger electron spectrometry with depth profiling.

RESULTS AND DISCUSSION:

Figures 1(a), (b), and (c) summarize the electrical measurements made using the TLM method. The contacts exhibited ohmic behavior in the as-deposited state. Fig. 1(a) shows the variation in sheet resistance with RTA temperature. The sheet resistance decreases from 24.9 Ω/square to 11.7 Ω/square after a 700 C anneal. The contact resistivity shows a decrease in Fig. 1(b) from 0.45 Ω-mm for the as-deposited contact to 0.061 Ω-mm after the 700 C anneal. Similarly, the specific contact resistivity shown in Fig. 1(c) also exhibits a decreasing

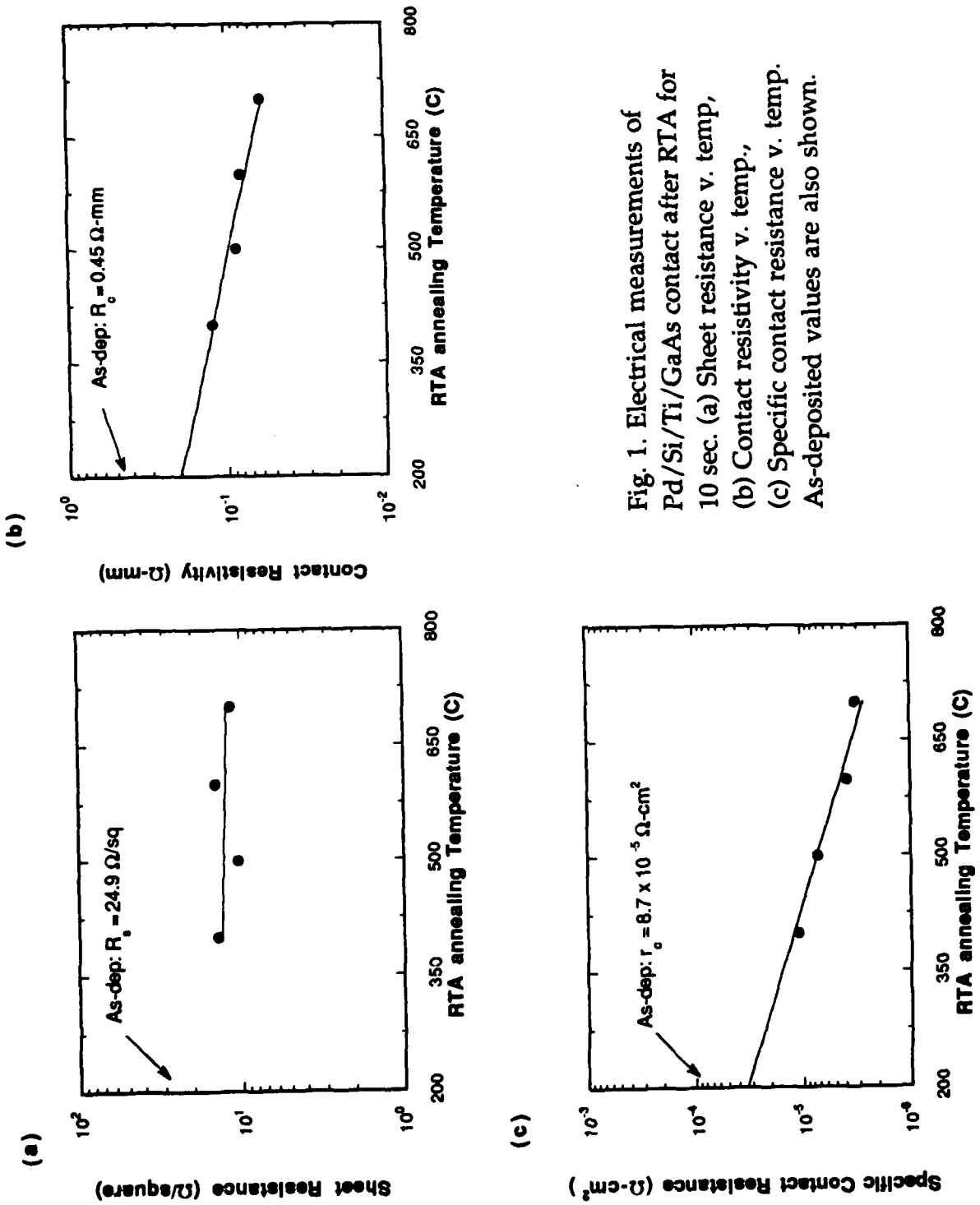


Fig. 1. Electrical measurements of Pd/Si/Ti/GaAs contact after RTA for 10 sec. (a) Sheet resistance v. temp, (b) Contact resistivity v. temp., (c) Specific contact resistance v. temp. As-deposited values are also shown.

dependence on RTA temperature. The specific contact resistance falls from the as-deposited value of 8.7×10^{-5} Ω-cm^2 to a value of 3.2×10^{-6} Ω-cm^2. These contact results compare very favorably with previous work on Pt/Ti on p^+ GaAs (25 Ω/square, 0.08 Ω-mm, 4×10^{-6} Ω-cm^2) [9] and are only slightly higher than typical Au-based p-type contacts (~ 1×10^{-6} Ω-cm^2).[13]

Auger depth profiles of the Pd/Si/Ti/GaAs contact structure are shown in Figures 2 through 6. The as-deposited sample is shown in Fig. 2 and is composed of distinct Pd/Si/Ti layers on GaAs. The Auger depth profile for the contact after annealing at 400 C for 10 sec is shown in Figure 3. The Pd peak height near 2 mins has decreased in relation to the Si signal height near 7 mins. In combination with the formation of a plateau in the Si signal near 2 mins the Auger results show the formation of Pd silicide, most likely Pd_2Si. The Si in contact with the Ti layer is unreacted and the GaAs shows no reaction with the Ti.

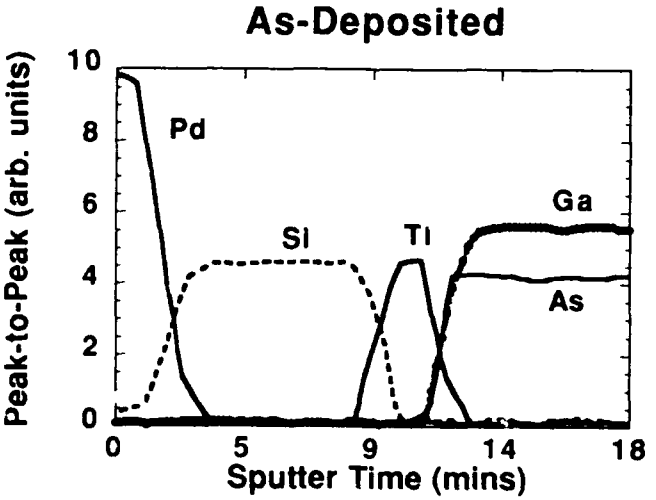


Fig. 2. Auger depth profiles of as-deposited Pd/Si/Ti contact.

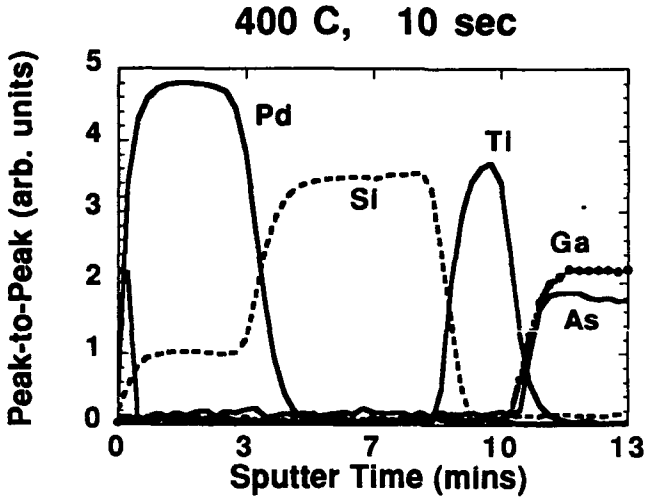


Fig. 3. Auger depth profiles of contacts annealed at 400 C.

Fig. 4 shows the depth profile of the sample after a 500 C RTA heat treatment. The reaction between Pd and Si near the surface is apparent and results in the continued formation of Pd_2Si. The second Pd peak near 10 min suggests the development of a thin Pd silicide layer below the Si layer. Based on the relative heights of the Pd and Si signals in the region near 10 mins this is probably a thin PdSi layer. The Ti signal has broadened slightly in the same region as a Ga peak near 12.5 mins. This is likely due to Ti-Ga interaction across the interface. In addition the rise of the As signal near 14 mins also suggests As out-diffusion at the Ti/GaAs interface.

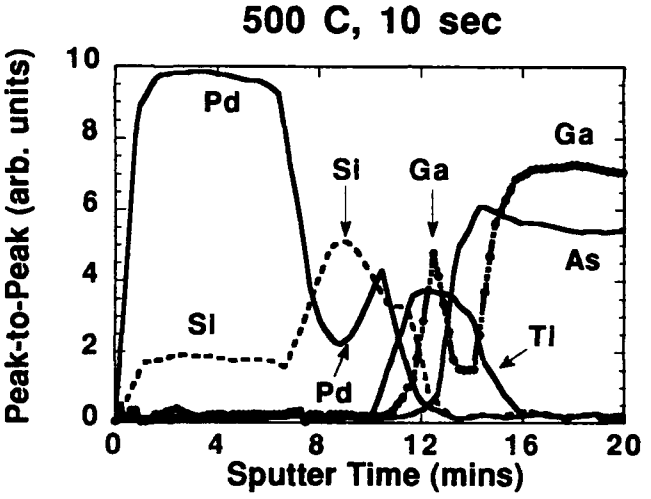


Fig.4. Auger depth profiles of contacts annealed at 500 C.

Fig. 5 shows the Auger depth profile of the sample after a 600 C anneal. The Pd_2Si layer appears stable while the Ga and As peaks near 3.5 mins and 4.0 mins respectively are still indicative of Ti-Ga-As reactions at the Ti/GaAs interface. Finally, Fig. 6 shows the Auger depth profile after a 700 C anneal where the Pd_2Si remains stable at the surface with increased Si-Pd interdiffusion below the Si layer near 3 mins. We attribute this increased interdiffusion to PdSi layer growth. Increased Ti-Ga-As interdiffusion can also be seen at the Ti/GaAs interface near 4 mins, indicative of increased Ga and As outdiffusion and subsequent reaction with Ti.

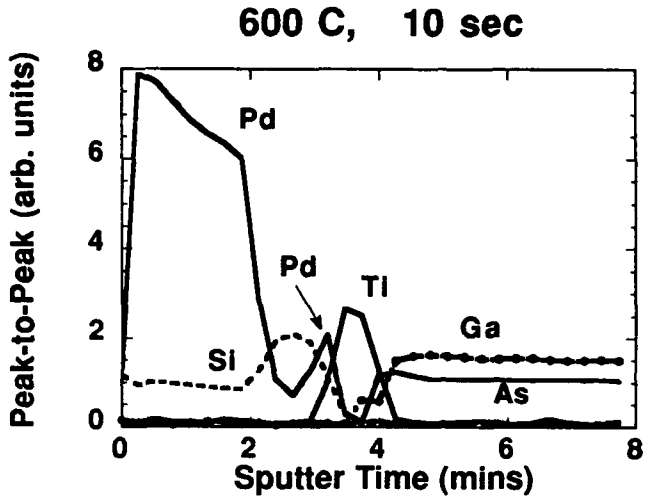


Fig. 5. Auger depth profiles of contacts annealed at 600 C.

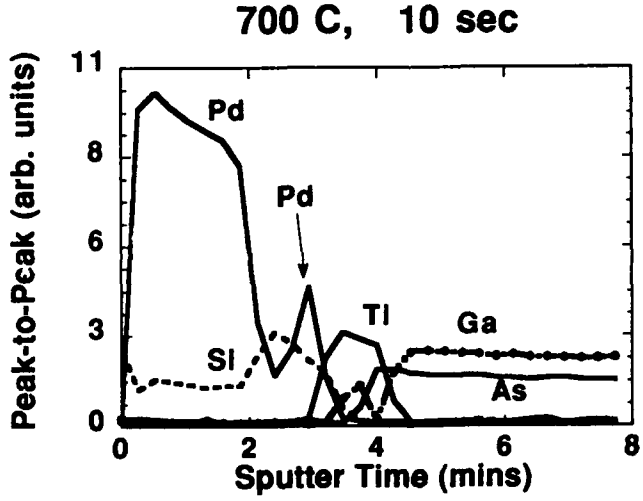


Fig. 6. Auger depth profiles of contacts annealed at 700 C.

A JEOL 2010 operating at 200 eV was used for TEM analysis. Figure 7 shows a TEM cross-section of the contact after a 500 C RTA heat treatment. Pd can be seen as reacting non-uniformly with Si layer as evidenced by the protrusions from the surface down to the the underlying Si. The Si layer appears to have recrystallized as a result of the annealing. The Ti layer above the GaAs substrate appears uniform and there is no evidence of spiking at the the Ti/GaAs interface. Energy dispersive analysis by x-ray (EDAX) analysis of the sample for qualitative determination of composition was made at selected areas denoted by the numbers 1-4. Area 1 was predominately Pd_2Si which was also suggested by the Auger analysis. Area 2 was determined to be predominately Si as expected. Primarily unreacted Ti on GaAs was detected in area 3 and area 4 analysis revealed GaAs only, with no evidence of Ti or Si diffusion into the substrate. Determination of PdSi and Ti-Ga-As reactions were difficult to resolve with EDS due to beam spreading of the probe. Scanning electron microscopy of the contact revealed no surface degradation after the 700 C anneal.

Fig. 7. TEM cross section of contact sample annealed at 500 C for 10 sec. EDAX analysis areas are denoted by numbers 1 through 4.

The Pd/Si/Ti contacts to p^+ GaAs exhibited ohmic behavior in the as-deposited state and showed a general decrease in specific contact resisitivity with RTA temperature. At this doping level we attribute the ohmic behavior of the as-deposited contact to both field emission and thermionic field emission.[14] The exact mechanism for ohmic contact behavior cannot be accurately determined from the results of this study but an estimate of the dominant effects can be made by evaluating the parameter E_{00} which describes quantum mechanical tunneling or field emission through the Schottky barrier [14]:

$$E_{00} = 18.5 \times 10^{-15} (N/m_r e_r)^{1/2},$$

where N is the carrier concentration, m_r is the effective mass ratio, and e_r is the permittivity ratio. Taking $m_r = 0.082$ and $e_r \sim 13.1$[15] we obtain $E_{00} \sim 0.126$ eV which is larger than the thermal voltage $kT/q \sim 0.026$ eV by a factor of about 5. At this doping level we find that tunneling and thermionic emission may contribute to the transport mechanism. In addition, it was found that the specific contact resistivities of WSi_x/p^+GaAs could be modeled by tunneling transport of light holes.[8] Increasing the doping level will result in lower contact resistivities but performance degradation may not be as severe as in Zn-doped GaAs contacts [16] due to the much lower carbon diffusion coefficient in GaAs.

The interfacial reactions that occurred over the temperature range 500-700 C included Pd-Si and Ti-GaAs reactions. Pd_2Si and PdSi formation was observed in the Auger depth profiles after 500 C RTA treatments. Pd_2Si formed near the contact surface while a thin PdSi layer below the Si layer. Pd reactions with Si are known to occur at temperatures < 400 C which are lower than the temperatures required for $TiSi_2$ formation.[4] Consequently, Pd-Si reactions are thermodynamically favored. The decrease in contact resistivities with RTA coincide with the Pd silicide formation which suggests the importance of this silicide formation for lower contact resistivities. Microanalysis revealed that the Pd-Si reactions were non-uniform in nature but that this silicide spiking was limited to the Si layer exclusively. This is important since silicide formation can

occur resulting in lower contact resistivities and no spiking into the GaAs substrate.

Ti reactions with GaAs can occur at 350 C with the reaction products TiAs and Ti_xGa_{1-x}.[17] Examination of the Auger depth profiles in Fig. 4 reveal the outdiffusion of Ga after the 500 C anneal resulting in a thin Ti_xGa_{1-x} phase. The stability of such a Ti-TiAs-Ti_xGa_{1-x} system is also supported by the Ti-Ga-As ternary phase diagram.[18] The presence of the Ga within the Ti region in Figs. 4-6 indicates the Ti-Ga reactions as was documented by others in the Pt/Ti/GaAs system.[10] However, in contrast to that system, heating the Pd/Si/Ti contacts at temperatures above 500 C did not induce further Ga out-diffusion through the TiAs layer. Further reductions in the contact resistivities may be due in part to the increased Ti-As and Ti-Ga interaction. The interface between the Ti-based layer and the GaAs remained smooth during the RTA treatments.

CONCLUSION AND SUMMARY:

Low specific resistance ohmic contacts using Pd/Si/Ti to heavily carbon-doped p^+ GaAs were studied. Contact resistivities as low as 0.061 Ω-mm and specific contact resistances as low as 3.2×10^{-6} Ω-cm^2 were achieved. Pd_2Si and PdSi formation was observed and may be important in achieving low contact resistivities. Non-uniform Pd-Si reactions occur but no spiking into the GaAs substrate was observed. Increased Ti-Ga and Ti-As reactions may also play a role in contact resistance reduction. The usefulness of the Pd-based silicide contact to p^+ GaAs is manifested in lower contact resistances and lower temperature of formation. However, this lower temperature characteristic also implies potential thermal stability problems if the contact is subjected to temperatures above 700 C for extended periods of time.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge D. W. Eckart for invaluable assistance.

REFERENCES

1. C.R. Abernathy, S.J. Pearton, R.Caruso, F.Ren, and J. Kovalchik, *Appl. Phys. Lett.* **55**, 1750 (1989).
2. K. Morizuka, R. Katoh, K. Tsuda, M. Asaka, N. Iizuka, and M. Obara, *Electron Device Lett.* **EDL-9**, 570 (1988).
3. K. Saito, E. Tokumitsu, T. Akatsuka, M. Miyauchi, T. Yamada, M. Konagai, and K. Takahashi, *J. Appl. Phys.***64**, 3975 (1988).
4. S.P. Murarka, in *Silicides For VLSI Applications*, (J. Wiley, New York, 1983).
5. C.P. Lee, T.H. Liu, T.F. Lei, and S.C. Wu, *J. Appl. Phys.* **65**, 4062 (1989).
6. E.D. Marshall, C.S. Wu, D.M. Scott,S.S. Lau, and T.F. Kuech, in *Thin Films and Interfaces II*, edited by J.E. Baglin, D.R. Cambell, and W.K. Chu (Mater. Res. Soc. Proc. **25**, Pittsburgh, PA 1984) pp. 283.
7. S.S. Lau, W.X. Chen, E.D. Marshall, C.S. Pai, W.F. Tseng, and T.F. Kuech, *Appl. Phys. Lett.* **47**, 1298 (1985).
8. T. Usagawa, M. Kobayashi, T. Mishima, P.D. Rabinzohn, A. Ihara, M. Kawata, T. Yamada, E. Tokumitsu, M. Konagai, and K. Takahashi, *J. Appl. Phys.***69**, 8227 (1991).
9. A. Katz, C.R. Abernathy, and S.J. Pearton, *Appl. Phys. Lett.* **56**, 1028 (1990).
10. A. Katz, S. Nakahara, W. Savin, and B.E. Weir, *J. Appl. Phys.***68**, 4133 (1990).
11. C.C. Han, X.Z. Wang,S.S. Lau, R.M. Potemski, M.A. Tischler and T.F. Kuech, *Appl. Phys. Lett.* **58**, 1298 (1991).
12. H.H. Berger, *J. Electrochem. Soc.* **119**, 507 (1972),
13. A. Katz, P.M. Thomas, S.N.G. Chu, J.W. Lee, and W.C. Dautremont-Smith, *J. Appl. Phys.***66**, 2056 (1989).

14. E.H. Rhoderick and R.H. Williams, in *Metal-Semiconductor Contacts*, (Clarendon Press, Oxford, 1988), p. 111.
15. S.M. Sze, in *Physics of Semiconductor Devices*, (J. Wiley, New York, 1981) p. 850.
16. F. Ren, S.J. Pearton, W.S. Hobson, T.R. Fullowan, A.B. Emerson, and D.M. Schleich, *Appl. Phys. Lett.* **58**, 1158 (1991).
17. K.B. Kim, M. Kniffin, R. Sinclair, and C.R. Helms,*J. Vac. Sci. Technol.***A6**, 1473 (1988).
18. R. Schmid-Fetzer, *J. Electron. Mater.* **17**, 193 (1988).

A Small Angle Neutron and X- Ray Scattering Study of the Onset and Nature of Fracture of Uniaxially Compressed Gun Propellants

Dr. Robert J. LIEB*
U. S. Army Ballistic Research Laboratory
Aberdeen Proving Ground MD 21005-5066

Dr. Samuel F. TREVINO
U.S. Army Research Development and Engineering Center
Picatinny Arsenal, NJ 20895-5000

1. INTRODUCTION

Pressure generation within guns is known to depend on the amount of surface area of the propelling charge. Generation of fracture surface area can occur when mechanical loads are applied to the propellant grains during ignition and combustion. Many studies[1,2,3,4,5,6] have been undertaken to establish the link between the mechanical response to deformation and the surface area generation that results from mechanical damage. Results of these studies have lead to improvements in performance and propellant vulnerability responses, but have not revealed the nature of fracture onset and growth within the propellant.

The measurement of surface area created by mechanical damage has been performed in various ways. Picnometery and the closed bomb burning of damaged specimens have been used and they offer insight into the extent of fracture damage of propellant specimens. However, in order for the results to be optimally used, the experimental results must be related to the influence that the damage has on propellant combustion. Problems occur when the results are interpreted. For example, the surface area measured by the picnometer may not be accessible to the flame, or, conversely, the flame may avail itself to much more area than the picnometer can measure. Closed bomb analysis seems to provide the most direct measurement, but, even there, dynamic effects, such as the ease of the ignitibiliy of freshly fractured surfaces or the nature of the generation of fracture surface area under the influence of the dynamic environment, raise questions about applicability of the experimental results to the operational environment situation. It is imperative, therefore, that as much as possible be known about the fracture process, including its onset and early development so that the applicability of test methods and results can be properly made. One method that can be used to investigate the boundary surfaces of a system is by measuring the effects of the scattering of radiation by that system.

Neutron and x-ray scattering by a specimen is dependent on the scattering cross-section of the materials in the specimen and on the contrast between materials at boundaries. In a homogeneous material, fracture will increase the scattering cross section by the introduction of the fracture surface. In a composite composed of materials of different contrast, scattering will occur at interface boundaries. This scattering will be increased by fracture occurring either by delamination of the component materials or by bulk fracture.

These scattering features offer a method by which the onset and nature of the fracture process can be investigated. Since the greatest change in scattering due to fracture would be expected to occur in composite materials, composite propellants, M30 and M43, were selected for this investigation. Specimens were prepared and uniaxially compressed at a high rate of deformation to various strain levels that ranged from 0 to 20 percent. The changes in the structure of these materials caused by this treatment were investigated by small angle scattering of both neutrons and x-rays. These measurements revealed the onset and nature of growth of the damage. This information will offer significant insight into the fracture process and will help direct future research efforts aimed at uncovering important processes that occur during propellant combustion.

2. NEUTRON AND X-RAY SMALL ANGLE SCATTERING

The theory and practice of scattering of neutrons and x-rays is well established[7]. Improvements in the instrumentation and adaptations of the theory tounderstand the physics in new materials continues. Here, we will focus on those aspects which address the present work.

Both types of radiation are able to scatter coherently from centers of which the material is composed. For neutrons these scattering centers are the nuclei and for x-rays they are the electrons of the atoms in the material. Since the scattering is coherent, it carries information about the geometric properties of the conglomerate of scattering centers. An essential quantity which describes the radiation is the wave vector **Q** whose magnitude is given by

$$Q = \frac{4\pi}{\lambda} \sin \theta, \qquad (1)$$

where λ is the wavelength of the radiation and 2θ is the scattering angle. The scale, r_s, of the geometry to which the radiation is sensitive is of order

$$r_s = Q^{-1}. \qquad (2)$$

Typical values of the radiation wavelength range from 0.1 to 1 nm (1 to 10 Å) so that the study of structures with geometric scales of 100 nm requires scattering angles of about 1° or less (thus small angle scattering). The structures of this scale size will reveal themselves if they exist in a matrix which scatters the radiation, i.e. has contrast. This matrix may consist of different ma-

terials or a material and voids. The only requirement is that the contrast, $\Delta\rho_{ij}$, between the two materials i and j be non-zero. The contrast ρ_i for material i can be calculated from measured properties of the scattering centers of which it is composed, and is given by

$$\rho_i = \frac{N_A D_i \sum_k n_k b_k}{\sum_k W_k} \qquad (3)$$

where N_A is the Avogadro constant, D_i is the density of material i, the sums are over the k atom types in material i, n_k is the number of atoms of type k, W_k is the atomic weight of atom type k and b_k is the scattering length of atom type k for the radiation of wavelength λ. Note that the dimension of ρ is L^{-2}. For neutrons, the values of b_k for the various materials are available as measured quantities, whereas for x-rays $b_k = Z_k b_e$, where Z_k is the atomic number of atom k and b_e is the scattering length of a single electron for x-rays, viz. the classical electron radius. These quantities are usually available for the material of interest so that the contrast is calculable. Note that the contrast of material i is different for neutrons and x-rays, but of more importance is that the relative contrast $\Delta\rho_{ij}$ of materials i and j will be different (unless one is singularly unfortunate) for neutrons and x-rays. This last observation is particularly useful when the system under study consists of three components (as is the case here).

The dependence of the scattering intensity I(Q) on Q is different for different magnitudes of Q even for the same material. This can be understood from Equation 2. The region of interest in the present work is known as Porod scattering and is valid for the range of Q sensitive to the surface area of the geometric structure. In this region the intensity is given as

$$I(Q) = 2\pi \frac{\sum_{ij} \Delta\rho_{ij}^2 S_{ij}}{Q^4} , \qquad (4)$$

where the sum is over pairs of materials ij between which there exists a boundary of specific surface S_{ij} (i.e. surface area per unit volume of sample). In the present work the propellants consist of three materials, the filler, F, the binder, B, and voids, V. There are therefore three specific surface areas to be determined S_{FB}, S_{FV} and S_{BV}. There are only two measurements which are available, viz. I_n and I_x, the neutron and x-ray intensities. As indicated above, voids can exist in the binder, filler, or between binder and filler. If these voids had been accessible to the outside, they could have been filled with a contrasting liquid making available independent data. This was tried without success. A simple model, in which each void presents an equal surface to each of the filler and binder materials, is used here. During manufacture the filler is wet by the binder and extruded. Voids *within* the binder or filler alone comprise a very small amount of the interface

surface. It is thought that any new fracture surface would occur *between* the binder and filler, as is usually seen in these fractured materials. These observations help justify considering this model, in which $S_{FV} = S_{BV} = S_1$. Now, the two measurements of the neutron and x-ray intensities yield the equation

$$I_m = 2\pi \frac{\left[(\rho_{F_m}^2 + \rho_{B_m}^2) S_1 + (\rho_{F_m} - \rho_{B_m})^2 S_{FB} \right]}{Q^4}, \qquad (5)$$

where m represents either neutrons or x-ray radiation. Inversion of these equations yields a direct measure of the two specific surfaces.

3. EXPERIMENTAL PROCEDURE

A. Preparation of Propellant Specimens

The specimens were prepared by cutting extruded solid sticks of propellant into right circular cylinders (diameter = 12.7 mm) with a length of 10.0 mm using a diamond saw. Two sets of five specimens were prepared for each propellant and each set was uniaxially strained to 0 (undamaged), 2, 5, 10 and 20 percent. The compression was conducted in a servohydraulic test fixture, illustrated in Figure 1. The unique feature about this machine is that compression can be arrested at a selected strain by adjusting the height of the anvil. As the actuator moves down the force gage comes into contact with the specimen and compression begins. It is halted when the inside surface of the impact bell contacts the mated surface of the impact cone. The nitrogen cylinder then absorbs the residual system energy by moving the piston. Therefore, the level of strain can be selected by adjusting the anvil height. The compressive strain rate was 100 s^{-1}.

The nominal formulation for each propellant is given in Table 1. The filler in M30 is nitroguanidine (NQ), which is in the form of long needle-like crystals of diameter from 5 to 10 μm. The nitrocellulose (NC) and nitroglycerin (NG) form an intimate mixture and act as a binder for the system. The filler in M43 is RDX and is of significantly different form than the filler in the M30 system. The RDX is ground into small (1-20 μm) ellipsoidal crystals which are held together by the intimately mixed CAB/NC/plasticizer binder. The adhesive binding strength in this system is much weaker than that of the M30 system. Figure 2 shows scanning electron micrographs of split propellant specimens which reveal the undamaged, intrinsic, propellant morphology.

After compression, the propellant cylinders were cut into 1-mm thick scattering specimens, as illustrated in Figure 3. Two types of specimens were made. Type A was cut perpendicular to the cylinder axis and Type B was cut parallel to the axis and through the center of the cylinder. The two types were needed to detect in greater detail the way in which damage

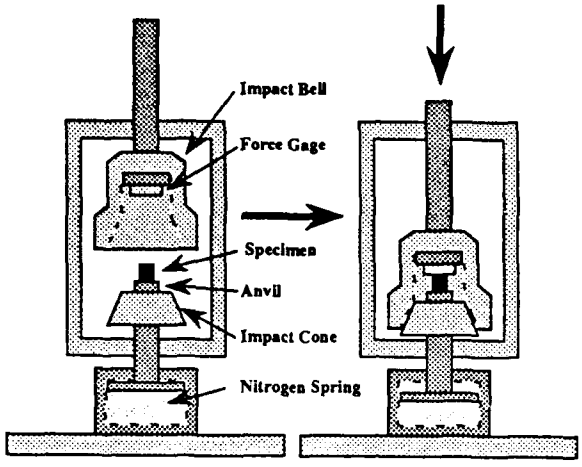


Figure 1. Servohydraulic Tester

Table 1. Nominal Percent Composition of Propellants

	M30	M43
Nitrocellulose (NC)	28	4
NC Nitration Level	12.6	12.6
Nitroglycerin (NG)	22	
Nitroguanidine (NQ)	48	
Ethyl Centralite (EC)	2	
RDX (Ground)		76
Cellulose Acetate Butyrate		12
Plasticizer		8

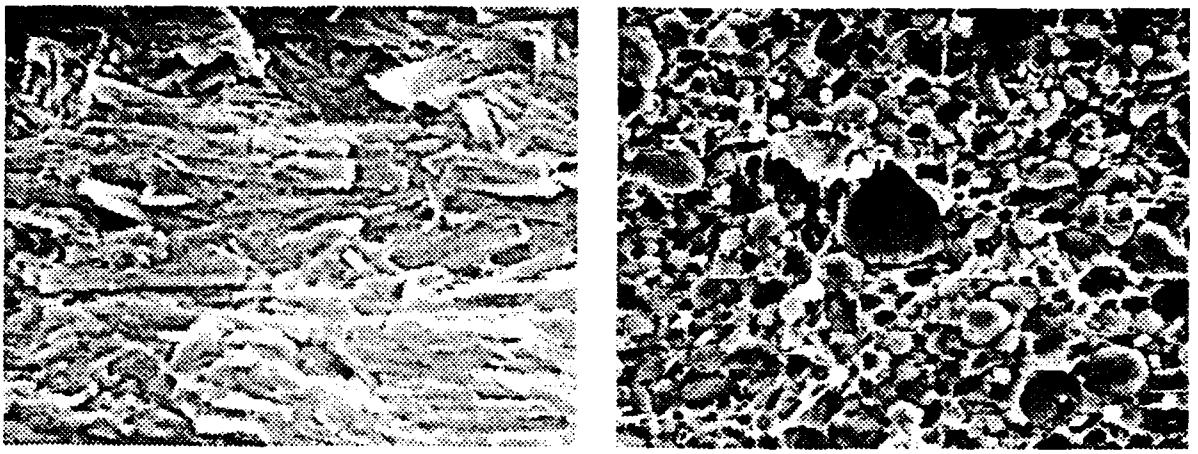
occurs. Note that the alignment of the NQ, which occurs during the extrusion process, is along the axis of the grain. This orientation is presented differently to the beams in the Type A and Type B specimens (hereafter referred to as simply A or B), as shown in Figure 4.

B. Scattering Measurements

Both the neutron and x-ray spectrometers are located at NIST. Figure 4 is a schematic appropriate for both instruments. The method of obtaining the monochromatic beam of radiation differs for each technique but the subsequent characteristics are essentially the same. The neutron beam wavelength was 8 Å and the x-ray 1.5 Å. The direction of the neutron beam is defined by a series of circular slits separated by distances of the order of 1 m. The distance between the sample and detector is about 8 m. These two geometries determine the angular resolution of the instrument. The detector is a 2-dimensional, position sensitive device which collects data simultaneously in a plane perpendicular to the incident beam. The data is stored in digital form for subsequent analysis. This includes the ability to obtain averages for constant scattering angle as functions of the azimuthal angle. If there exists asymmetry in the scattering pattern, sector averages for azimuthal angles are readily performed. The data collection time for the present set of measurements totalled approximately six hours for each scattering run.

4. RESULTS

Both the neutron and x-ray scattering patterns were qualitatively similar. Different amounts of spherical asymmetry were obtained for the different degrees of strain. The B specimen of unstrained M30 produced the greatest degree of asymmetry. A plot of the raw 2-dimensional data is presented in Figure 5. The asymmetry is due to the orientation of nitroguanidine needle-like crystal along the specimen axis. For all the B specimen data, sector



a. M30 b. M43

Figure 2. SEM Micrographs of the Propellant Specimens

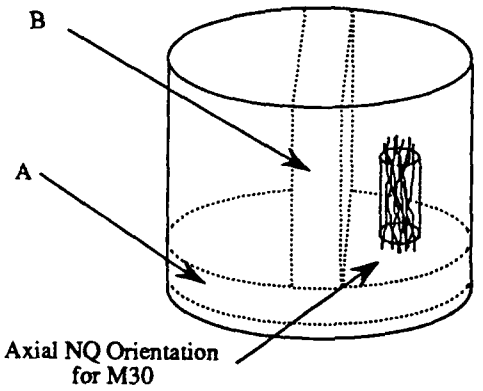


Figure 3. Illustration of Type A and B Specimens

averages, over 30° centered at ϕ, with ϕ equal to 0, 30, 60, and 90°were obtained. The data was corrected for detector efficiency, for background and was normalized to standard samples from which absolute cross-sections are determined. Sector averages were not performed for the A specimens because they showed symmetric scattering. The coefficient of Q^{-4}, which is associated with the contrast and specific surface factor in Equation 5, is obtained by plotting IQ^4 vs Q^4, fitting a straight line, and determining the intercept. This was done for both neutron and x-ray data. Equation 5 was inverted to obtain S_1 and S_{BF}. Figures 6 and 7 present the values of these specific surface areas for both propellants as functions of strain for the A and B specimens, and for sector averages for B specimens. For the B specimens, ϕ equal to 0° corresponds to the specimen axial direction.

5. ANALYSIS

A. M30

Figure 6 shows the specific area S_1 and S_{FB} plotted against strain and axial orientation. The specific area is the interface area per unit volume of material and is given in units of cm^{-1}. S_1 for the A specimen indicates that there were some voids present in the undamaged material.

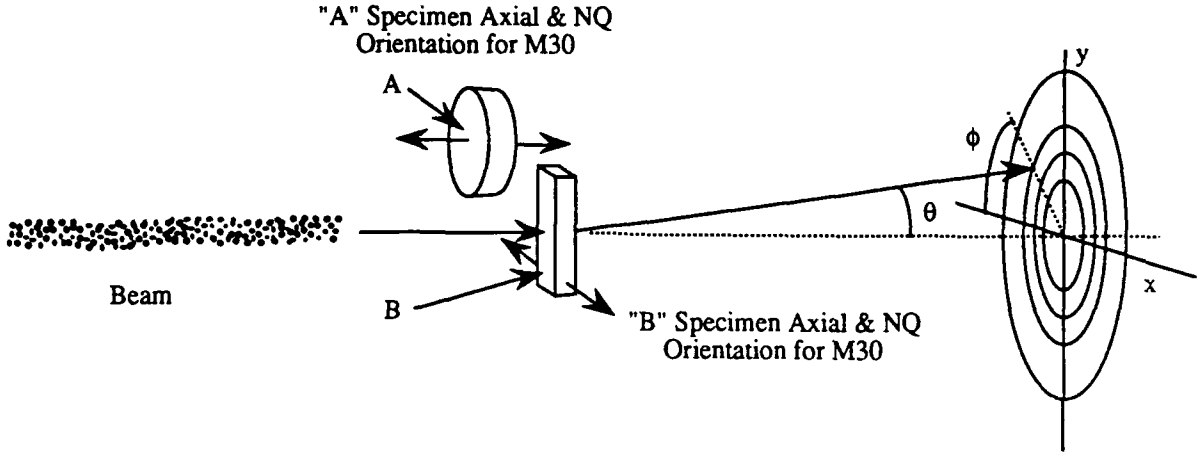


Figure 4. Schematic of the Scattering Experiment

Since previous M30 propellant micrographs indicated that there was no voids in the binder, micrographs of crystalline NQ from another investigation were studied. Figure 8 shows voids that are present on the surface of the NQ crystals. This micrograph and others from this series indicate that hollow NQ crystals could result from these elongated voids being internalized. The mismatch in the S_1 values for the A specimen and the B specimen at ϕ equal to 90°could be explained from such internal voids. Now, if spherical voids were present, or if the detector had infinite resolution, these values would be the same. However, with elongated voids, as indicated here, larger surfaces are measured when averaging S_1 along the axial instead of the direction perpendicular to it and the area vector.

As strain increased, the A specimen showed an increase in S_1 which levels off and then increased again at higher strain. This is consistent with delamination of binder and filler with the internal voids of the NQ staying unchanged. For the B specimen at ϕ equal to 0°, not much change was observed, possibly because of the insensitivity to delamination in this direction. However, in the other three directions for ϕ there was a slight decrease in S_1, with a significant rise then a decrease or leveling out, and then a final increase at high strain. This can be thought of as an early collapsing of voids or a closing of previously delaminated surfaces, then, after 2-percent strain, a signifi-

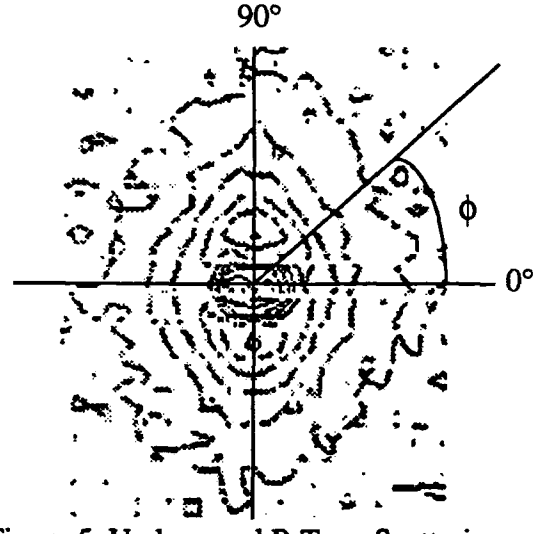
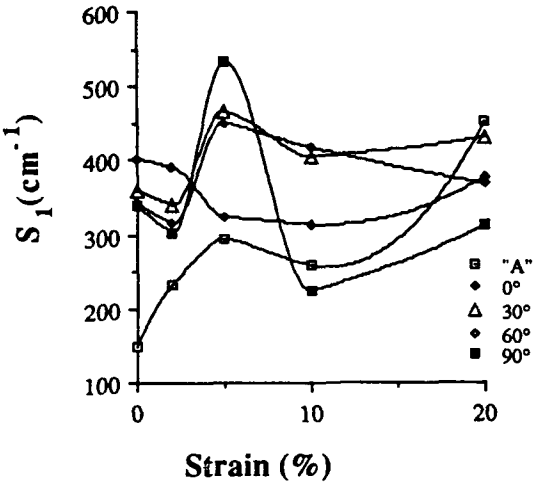
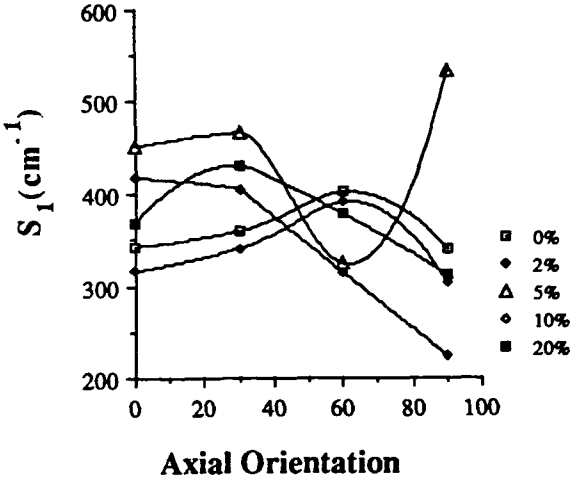


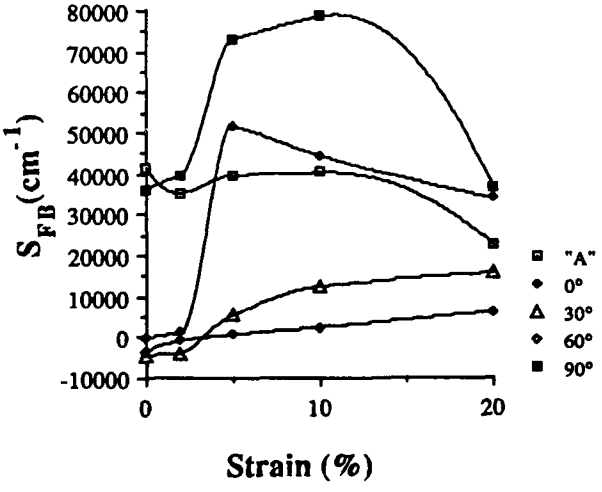
Figure 5. Undamaged B Type Scattering Results for M30 Showing NQ Alignment



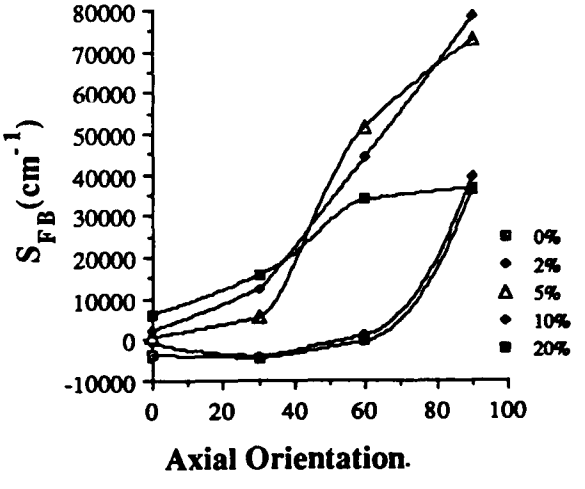
a. S_1 vs Stain
for the A and B Specimen



b. S_1 vs Axial Orientation
for the B Specimen

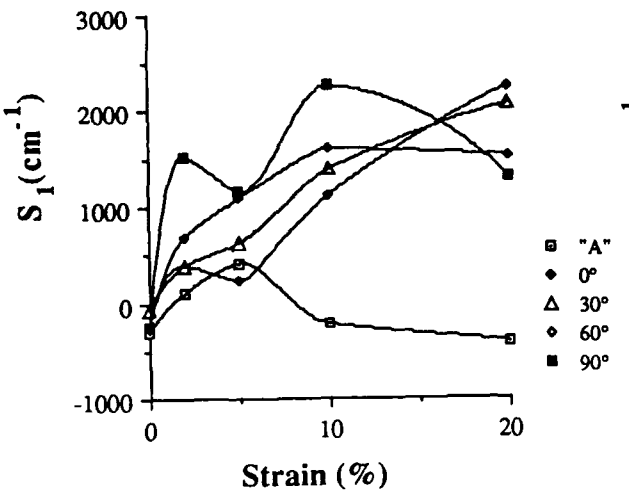


c. S_{FB} vs Stain
for the A and B Specimen

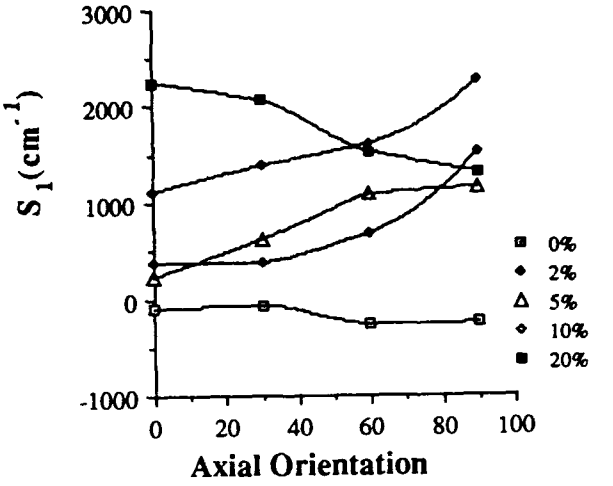


d. S_{FB} vs Axial Orientation
for the B Specimen

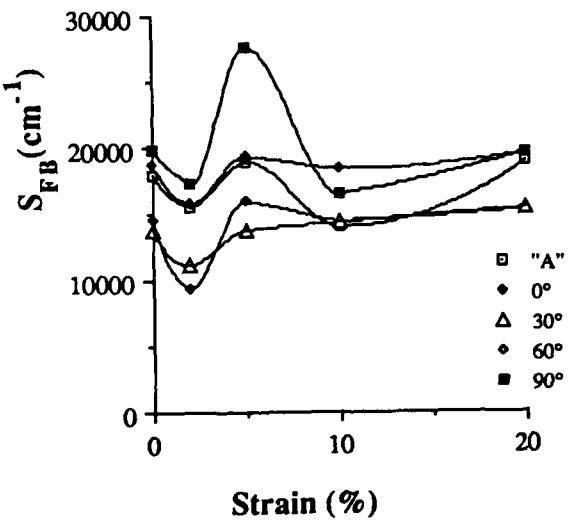
Figure 6. Specific Surface Area for M30 Propellant



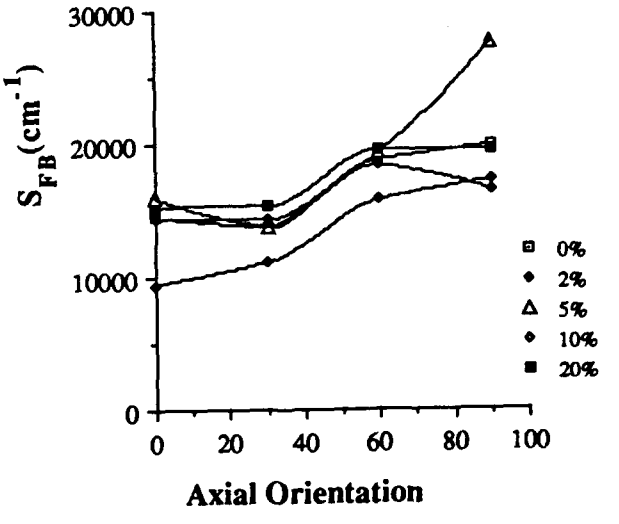
a. S_1 vs Stain
for the A and B Specimen



b. S_1 vs Axial Orientation
for the B Specimen



c. S_{FB} vs Stain
for the A and B Specimen



d. S_{FB} vs Axial Orientation
for the B Specimen

Figure 7. Specific Surface Area for M43 Propellant

cant delamination and reorientation of the NQ crystals occurred within the solid. The most difficult feature to explain is the rapid drop in S_1 at ϕ equal to 90° at 5-percent strain. There is no apparent reason as to why the value is reduced so much.

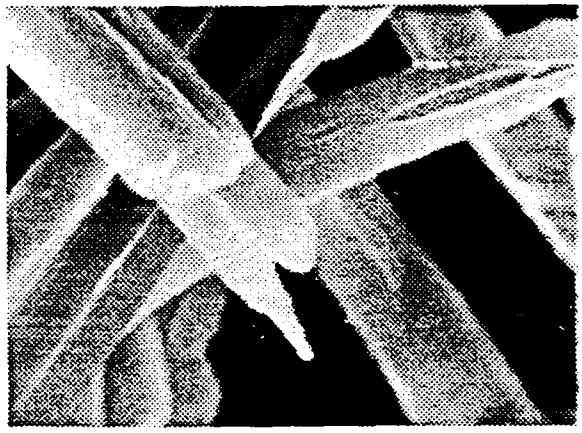


Figure 8. Micrograph of NQ Crystals Showing Surface Voids

In Figure 6b the S_1 values for the B specimen are plotted against axial orientation with strain level of each curve being held constant. The general trend indicated that initially there was a slight reduction in the S_1 due to void collapse or recombination of surfaces. At 5-percent strain significant delamination occurred which recombined as strain was increased. At 20 percent, S_1 increased again and possible reorientation of the NQ had begun.

In Figure 6c, the S_{FB} values at zero strain show a good match between the A specimen and B specimen at ϕ equal to 90°. Also the values were very low along 0° and 30°, as they should be with little surface between binder and filler in this direction. The rise in values with increasing strain may be attributed to the reorientation of NQ crystals. The perplexing feature of this plot is the enormous increase in S_{FB} at 60° and 90° at 5-percent strain. A mechanism accounting for this increase could be the precipitation of NG, which is a liquid, from the binder mixture. This would create new binder-filler surface and still allow delamination surfaces to be created. The decrease of S_{FB} at higher strains is in line with continued delamination, and the magnitudes need not be reflected in corresponding increases of S_1.

The change in values of S_{FB} with axial orientation follow what was expected at each strain level. All values were low in the ϕ equals 0° direction (axial direction) and increase as ϕ increases. This was expected and reflected the alignment of the NQ crystals. However, the unexpected increase in S_{FB} with strain is again clearly displayed.

B. M43

The specific area values for M43 are displayed in Figure 7. This propellant showed no voids in the undamaged specimens. In the B specimen, S_1 showed increasing values at increasing strains, and generally higher values in damaged specimens at increasing ϕ values. This indicated delamination of the binder-filler interface, since the binder deforms plasticly at 20°C and no fractured RDX has been observed under these conditions in micrographs of damaged grains. Two features in the S_1 curves need to be commented upon. The first concerns the divergence of the S_1 values for the A specimen and the B specimen for ϕ equals 90° at increasing strain. These

values should track with each other because the surface area created here is due to binder-filler delamination and axial alignment is much reduced from that found in the M30 propellant. These data suggested a significant difference in the degree of damage in the two scattering specimens, which was subsequently shown to be the case. The mechanism of failure is by shear failure along 45° cones, as illustrated in Figure 9. Since the A specimen was taken from the end of the compression specimen, the center region of A had few areas of failure. The B specimen, on the other hand, contains the central portion of the compression specimen, at which the 45° failure cone is directed. So the B specimens would be expected to contain greater amounts of damaged material. The second feature concerns the decreasing S_1 value with axial orientation at 20-percent strain (see Figure 7b). It is thought here that RDX crystals were becoming so loose that they were able to rotate. This would cause what was observed, a more rapid increase in S_1 at lower ϕ values and reduction in S_1 at higher ϕ values, due to the ellipsoidal shape of the RDX.

The S_{FB} curves plotted against strain show similar changes at increasing strain levels, and generally show higher values as ϕ increases. Note that in this case the A specimen and the B specimen at ϕ equal to 90° matched well, as expected. The curves show an initial decrease, indicating delamination, as did S_1. However, at 5-percent strain a dramatic increase in S_{FB} was found which was followed by a decrease and then leveling. A possible mechanism for this observation is illustrated in Figure 10. From SEM micrographs it is known that many small RDX crystals are present in the matrix. If tightly packed agglomerations of small particles lodge between larger crystals, they could burst during the plastic shear failure and create new S_{FB} surface, as shown in Figure 10. The delamination process continues, as indicated in the S_1 values, but this process is in competition with the shear process. The changes in S_{FB} with axial orientation showed these same features and made another observation clear, as well. The ellipsoidal shapes of the RDX crystals appeared to align with the axial direction of the specimen. This was indicated by the increasing values of S_{FB} at increasing ϕ for each strain level (see Figure 7d). These observations provided a consistent picture and added significant insight into the failure process for M43 propellant.

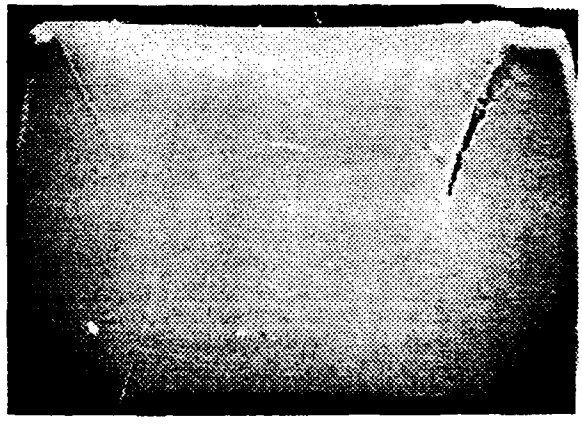
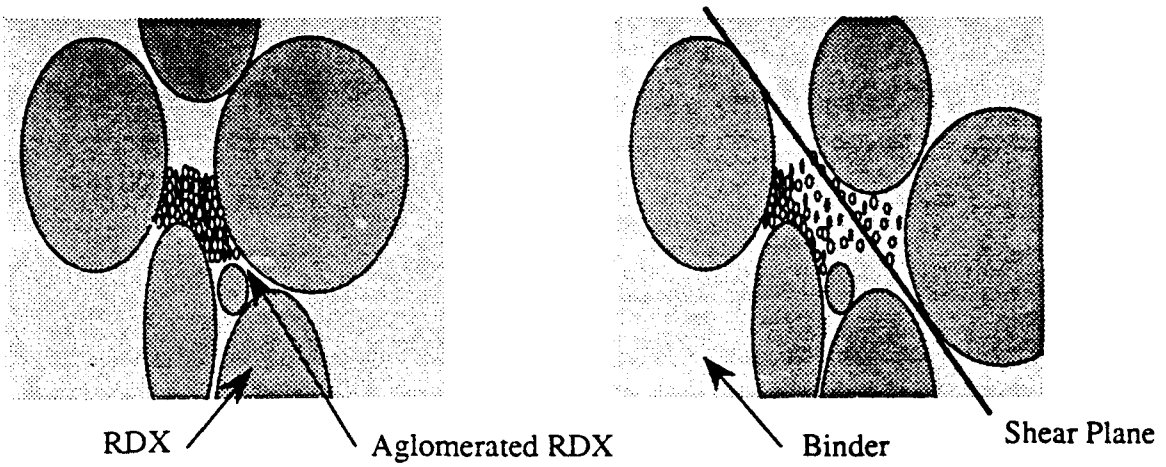


Figure 9. Type B Specimen at 20 % Strain Showing 45° Failure Surfaces

To find what this process requires, assume that 1 percent of the RDX is agglomerated and the eccentricity of the particles making up the agglomerate is 0.5. This is reasonable since MicroTrac® data[8] indicates that 3.1 percent of the crystals are smaller than 0.9 μm for this type of RDX. This is the lowest size distribution category reported. Now, the aver-



a. Undamaged M43 b. M43 after Shear Failure

Figure 10. Mechanism that Would Increase S_{FB} with Increasing Strain

age increase in surface area per cm^3 of propellant from 2 to 5-percent strain is 5.7×10^3 cm^{-1}. A quick calculation of the particle size in these agglomerations, required to produce the area increase indicated in these calculations, gives the following. Using the surface area and volume formulas for the ellipse, the major axis would be need to be about 0.05 μm in order to produce enough area. This is quite a small size. However, if precipitation of RDX occurred when the solvent from the extrusion process was evaporated, particles of this size could be deposited.

One additional piece of information that adds credence to the interpretation of these measurements is found in a recent determination of the surface area of the RDX crystals[9], which as made before mixing the crystals and binder. The value of the surface area density, S_d, was found to be 1.86×10^4 cm^2/g. This can be converted to specific density for M43 by using

$$S_{FB} = S_d \, \delta_F \, v_F, \qquad (6)$$

where δ_F and v_F are the mass density and volume fraction of the filler in the M43, respectively. The density of RDX is 1.82 g/cm^3 and its volume fraction in M43 propellant is 0.675. When these values are used, the specific surface value received is 2.3×10^4 cm^{-1}. This value is in excellent agreement with the values measured here for the undamaged specimens (see Figure 7c). The fact that this value is larger than observed in the scattering experiments lends support for the agglomeration model.

6. CONCLUSIONS

A new method has been found for investigating the changes in surface morphology of filled energetic materials. Neutron and x-ray scattering were used to measure the binder-filler

and void surfaces in M30 and M43 gun propellant as a function of strain and axial asymmetry of the specimen.

For M30 propellant, scattering measurements showed many features of the propellant system, including: previously noted axial alignment of the NQ crystals; the large binder-filler surface area orthognal to the crystal axis; the strain level at which fracture begins; and the nature of the failure mechanism, i.e. binder-filler delamination. Also indicated in the data were internal voids within the NQ crystals, closing of delaminated portions of damaged propellant after a certain strain level, and the reorientation of NQ crystals at high strain. One result that is not well understood was the indication of dramatic increases in binder-filler surface area. This was unexpected and would require an unusual occurrence, such as the precipitation of nitroglycerin (a liquid) from the binder mixture. Such an occurrence would help resolve the apparent inconsistencies in the data. Efforts to understand these results are continuing.

It seems that for M30, the collapse of preexisting voids, or the mating of preexisting binder-filler separations began immediately upon deformation. Between 2 and 5-percent strain significant delamination damage began, and at the same time some mechanism (such as the precipitation of NG) created more binder-filler surface. As strain continued beyond 10 percent, delamination continued and the trend was for greater void area and decreased binder-filler surface.

The most significant morphological features observed for M43 propellant were the following. First, no voids were indicated before compression. Second, axial alignment of the major axis of the ellipsoidal RDX particles along the specimen axis was indicated. This alignment was much less than that found for M30, but still observable. Third, damage in the A and B specimens was markedly different due to the shear failure occurring in the specimen and the method of specimen preparation.

Results suggested that as deformation occurred in M43, delamination began and generally continued throughout the deformation. Some new binder filler surface seems to be generated at about 5-percent strain. This could have arisen from either a recombination of a separated interface, precipitation of plasticizer on a much smaller scale than for M30, or the breakup of agglomerated RDX, as discussed above. As compaction continued the delamination grew, and there was an indication that, somewhere between 10 and 20-percent strain, RDX crystals became free to rotate in the matrix.

This information was largely unknown before these experiments were performed. These new insights will help guide the development of propellants and aid in the development of methods for increased performance and reduced vulnerability response.

The scattering techniques that were developed will permit more extensive progress in future scattering efforts. This work should include a study of low temperature deformation to investigate the agglomerated RDX conjecture. If the suggestion is correct, the increase in binder-filler surface noted in this study should not appear in the specimens deformed at lower temperature. The lower temperature should not permit plastic flow of the binder and should thereby eliminate the creation of new surface. Other experiments on single and double-based propellants will offer similar insight into the fracture mechanisms of those systems, and provide methods for improvement.

7. REFERENCES

1. G. A. Gazonas, A. Juhasz, and J. C. Ford, "Strain Rate Insensitivity of Damaged-Induced Surface Area in M30 and JA2 Gun Propellants," Technical Report BRL-TR-3251, USA Laboratory Command, Ballistic Research Laboratory, Aberdeen Proving Ground, Maryland, August 1991.

2. P. Lu, B. Strauss, S. Moy, and R. Lieb. "Shaped Charge Jet Impact on Gun Propellants Study I - Temperature and Mechanical Properties Effects," 1991 Propulsion System Hazards Subcommittee Meeting, March 1991.

3. R. J. Lieb, and M. G. Leadore, Mechanical Failure Parameters in Gun Propellants," Technical Report BRL-TR-3296, USA Laboratory Command, Ballistic Research Laboratory, Aberdeen Proving Ground, Maryland, November 1991.

4. G.A.Gazonas, "The Mechanical Response of M30, XM39, and JA2 Propellants at Strain Rates fr m 10-2 to 250 sec-1." BRL-TR-3181, USA Ballistic Research Laboratory, Aberdeen Proving Ground, Maryland, January 1991.

5. R. J. Lieb, "High Rate Intrinsic Bed R_sponse of Gun Propellant," 1987 JANNAF Structures & Mechanical Behavior Subcommittee Meeting, **1**, CPIA Publication 463, pp 51-62, March 1987.

6. A.W. Horst, "The role of Propellant Mechanical Properties in Propelling Charge Phenomenology," 1981 JANNAF Structures and Mechanical Behavior Subcommittee Meeting, **1**, CPIA Publication 351, December 1981.

7. O. Glatter and O. Kratky, Ed., *Small Angle X-ray Scattering*, Academic Press, NY 1982.

8. R. Simmons, Naval Surface Warfare Center - Indian Head, Private Communication.

9. S. Caulder, Naval Surface Warfare Center - Indian Head, Private Communication.

LISTON, ROBERSON

Exoatmospheric Intercept: GBI Program and the ERIS FTV Flight Tests

Mr. John E. Liston*
Chief, GBI System Engineering and Analysis Division
U. S. Army Strategic Defense Command
Huntsville, AL 35807-3801

Mr. William E. Roberson
Chief, GBI Interceptor Design and Development Division
U. S. Army Strategic Defense Command
Huntsville, AL 35807-3801

Overview

On January 28, 1991 the Strategic Defense Initiative (SDI) non-nuclear Exoatmospheric Reentry Vehicle Interceptor Subsystem (ERIS), managed by the U.S. Army Strategic Defense Command (USASDC), acquired, designated (in the presence of decoys), tracked, and kinetically destroyed a simulated threat reentry vehicle over the middle of the Pacific Ocean. This exoatmospheric non-nuclear intercept proved the feasibility and lethality of a cost effective exoatmospheric interceptor subsystem operational concept.

Design of the ERIS Functional Technology Validation (FTV) flight test vehicle was based on functionally emulating an operational ballistic missile defense system including: targets, early warning, precommit sensors, booster flyout, inflight updates, acquisition, track, and RV engagement. This paper will describe how ERIS FTV fits within the Ground Based Interceptor (GBI) Program and will review in detail the ERIS mission objectives, the FTV flight test sequence, and the major accomplishments and significant results that were achieved from this flight test.

The GBI Program

The Ground Based Interceptor (Figure 1) is a low cost, non-nuclear, hit to kill interceptor under development by the USASDC for the Strategic Defense Initiative Organization (SDIO). The GBI's battlespace is in the region above the Earth's atmosphere. The GBI is

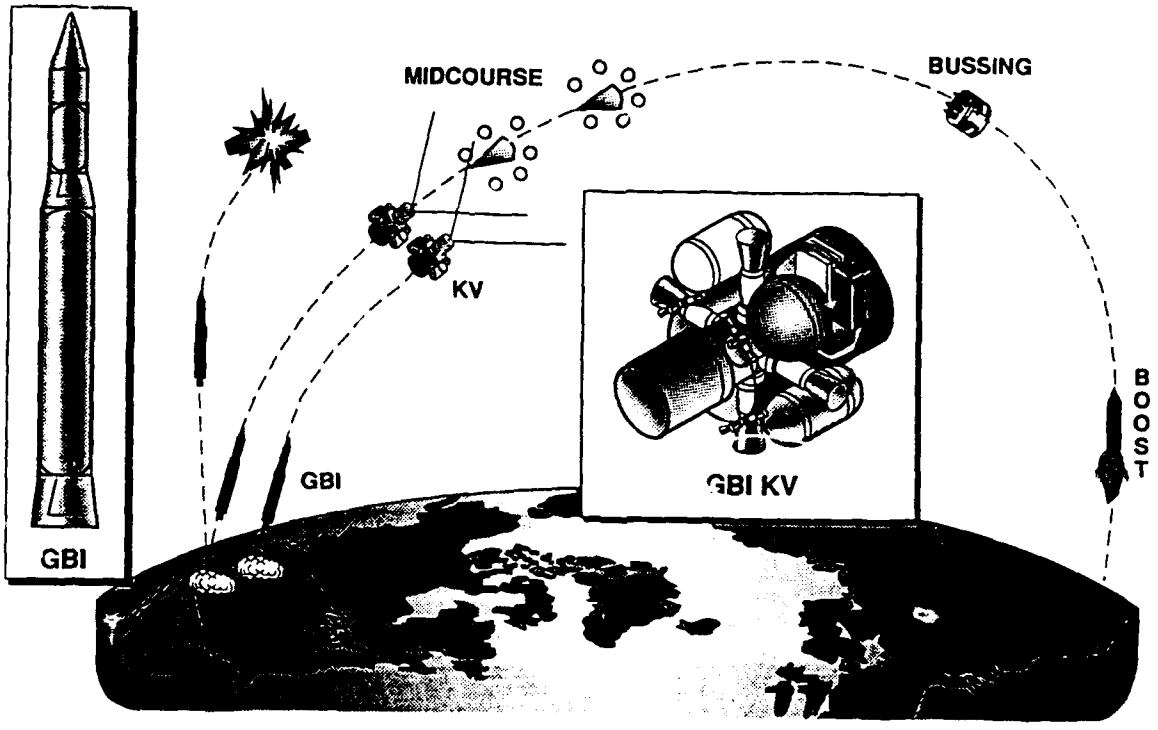


Figure 1. Ground Based Interceptor

designed to be compatible with a variety of target handover sensors whose function is to acquire, discriminate, and handover targets or target complexes to the GBI. The GBI Program is presently comprised of the Functional Technology Validation (FTV) flight tests and a series of GBI Demonstration/Validation (Dem/Val) flight tests. Together, these programs could form the foundation for the development of interceptors for the National Missile Defense Initial Deployment scheduled for the late 1990's. The purpose of the FTV flight tests is to validate low cost interceptor technologies and pursue the resolution of GBI critical issues. While minimizing weight and size, the follow-on GBI experiments develop and infuse advanced technologies with onboard endgame discrimination capability to enable transition into the Engineering and Manufacturing Development Phase (EMD) of the Defense Acquisition Process. GBI could be the foundation of the initial deployment architecture, thereby dictating that GBI use the best technologies available that have been demonstrated to ensure a low risk, robust National Missile Defense.

GBI Operational Overview

The Ground Based Interceptor (GBI) is one of the elements of the strategic defense system and consists of an interceptor, launch site, ground support equipment, associated personnel, and support services. The elements of the strategic defense system are linked through the Command and Control Element (CCE) during the defensive battle.

The GBI engagement scenario is based on sensor subsystems detecting, identifying, and designating hostile RVs (to the extent possible) and transmitting the tracking data to the CCE. The CCE determines the path of the target based on this data from the surveillance and tracking systems and plans a GBI engagement. Launch commit parameters are sent from the CCE to the GBI base, where trajectory parameters are calculated and downloaded to a specific interceptor. At the required time, the interceptor is launched using a two stage booster to propel the kill vehicle toward the on-coming threat. As the second stage burns out, the kill vehicle (KV) separates from the booster and repositions itself pointing the seeker field of view (FOV) to the predicted target position. Divert maneuvers are performed to correct the kill vehicle's trajectory for boost errors and to make any adjustments based on inflight updates of the target's trajectory. As the seeker is activated, the target and any associated objects will be acquired. The target is designated using pattern matching algorithms or onboard discrimination capability. Once the target has been designated, the KV transitions to track mode and performs its final trajectory adjustments. In the terminal homing mode, the KV selects an aimpoint on the target body in order to maximize its probability of kill. Just before impact a kill enhancement device (KED) may be deployed for added insurance of collision and lethality. Destruction of the target is accomplished by physical impact with the KV or associated KED. The intercept event can be monitored by radars and tracking systems for final kill assessment and relayed to the CCE for further battle management action if required. This operational overview is the basis of the FTV flight test program.

FTV Program Overview

The FTV interceptor subsystem is comprised of a missile and launch control system which functionally emulates a tactical

interceptor as an element in the Strategic Defense System (SDS). An operational ERIS interceptor concept was developed and optimized through extensive system engineering trades and analyses to provide an interceptor which achieved its required performance with a minimum life cycle cost. The design, functions, and specifications of the operational ERIS concept were emulated to the maximum extent possible in the FTV Kill Vehicle design while utilizing available (1986) technologies and off the shelf hardware. The FTV flight test phase of the program was to validate that the operational ERIS concept was feasible and readily achievable. The major ERIS concept design characteristics that were emulated in the FTV kill vehicle included: interceptor dormancy, cool on the rise seeker, in-flight target update capability, Threat Object Map (TOM) target designation, radiometric target selection, two color seeker for endgame CSO discrimination, aimpoint selection algorithms, kill enhancement device for added lethality, and hit-to-kill guidance capability.

Under GBI Project Office guidance, Lockheed (LMSC) and its subcontractors designed, built, and tested the FTV kill vehicles and are conducting the interceptor flight test demonstrations. The target complex is launched from Vandenberg AFB (VAFB) in California and consists of a medium RV and decoy balloons. The interceptor is launched from Meck Island at the U.S. Army Kwajalein Atoll (USAKA). The USAKA range instrumentation consists of tracking radars, telemetry receiving and recording systems, optical tracking instruments, Global Positioning System (GPS) ground stations, flight safety, and timing and communications equipment. Meck Island facilities (Figure 2) include the launch control building, missile assembly building, launch equipment room, access stand, and launch cell.

The FTV Air Vehicle (Figure 3) is broken down into three main sections which are the two stage booster, booster adapter (BA), and kill vehicle (KV). The booster is an Aires II configuration which consists of refurbished Minuteman I second and third stages with shortened nozzles and an added custom built aft skirt assembly. The BA is fabricated to interface between the kill vehicle and booster and to provide structure for the observer package (OP) and its telemetry

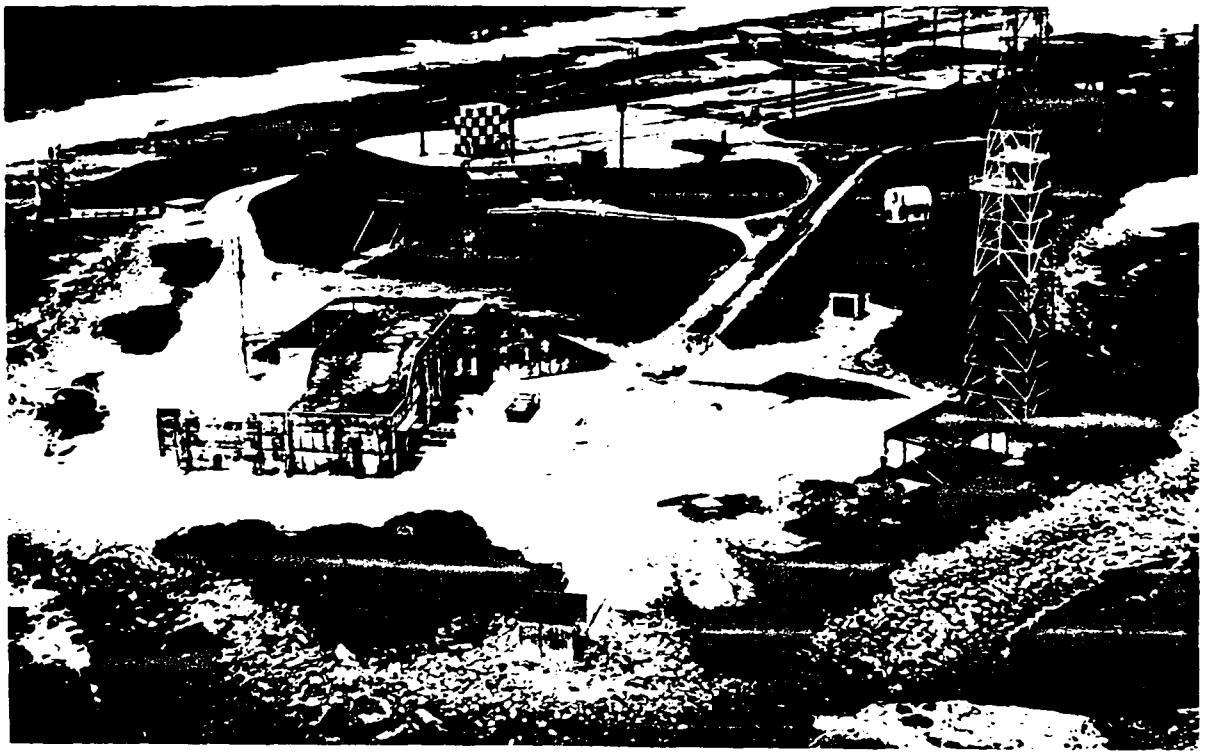


Figure 2. Meck Island Launch Facilities

and control equipment. The OP is a special collection of scientific instruments mounted in the booster adaptor designed to collect close range data on the intercept event. The BA follows a trajectory parallel to the kill vehicle and points the OP toward the expected intercept

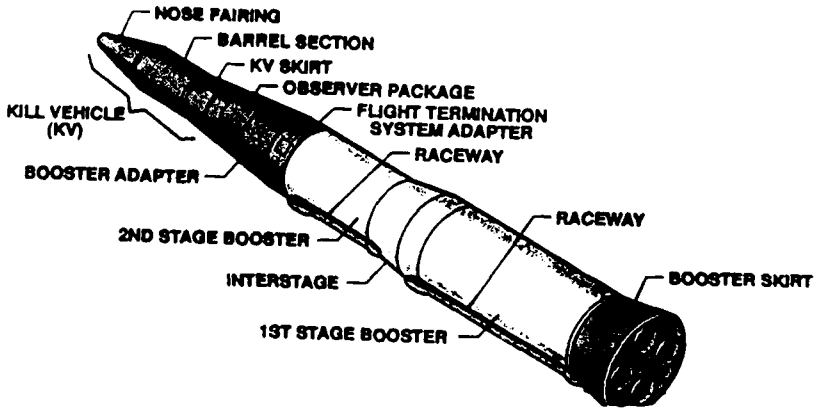


Figure 3. FTV Air Vehicle

point. Upon intercept the OP records the resulting phenomenology and transmits this data to the ground.

The Kill Vehicle (Figure 4) is designed for easy assembly and access to components which are directly mounted to the primary structure. A two color, body fixed, long wave infrared (LWIR) seeker is mounted to the end of the KV which features an in-flight cooled HgCdTe focal plane and warm optics. The propulsion and reaction control consists of a bipropellant divert system and a cold gas attitude control system. The data processor, signal processor, and inertial measurement unit (IMU) are modified off the shelf hardware. The KED is an inflatable bag used to negate the effects of target aimpoint selection countermeasures. This KV design emulates the ERIS concept and allows the FTV interceptor to remain dormant until the launch command is given. This ability to remain dormant is a significant accomplishment over previous interceptor designs and provides a very significant reduction in life cycle costs due to the decreased maintenance and supportability requirements.

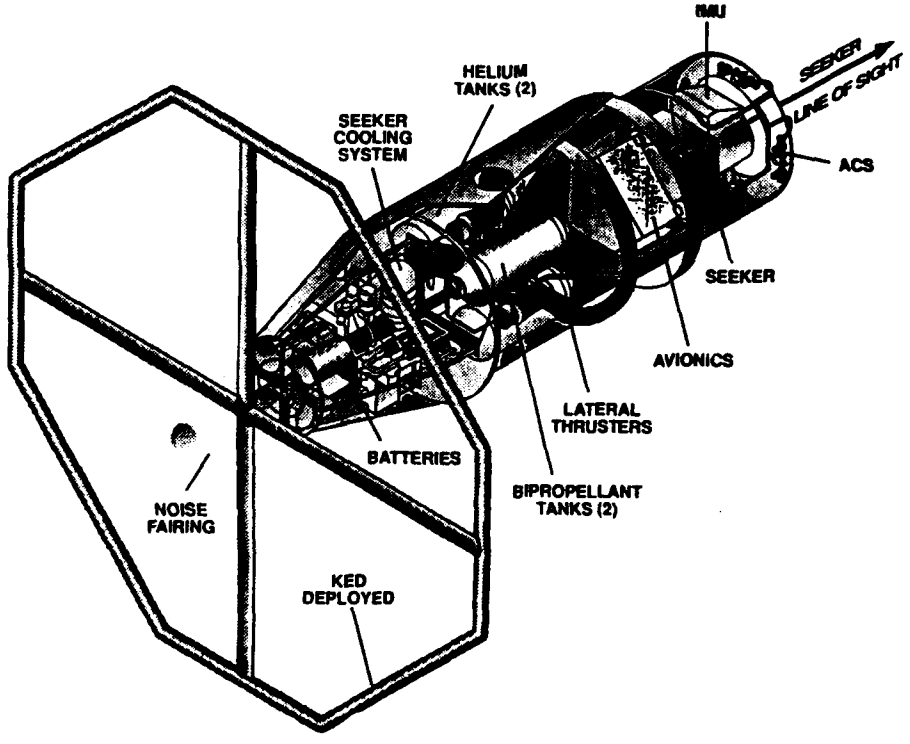


Figure 4. FTV Kill Vehicle

FTV Flight Test Description

The FTV flight tests directly emulated an operational missile defense system (Figure 5). The FTV mission began with the launch of an ICBM from Vandenburg Air Force Base in California. This Minuteman II ICBM carried a simulated threat RV, built by Sandia National Labs, and deployed the RV on a trajectory with an intended aimpoint in the mid-Pacific just north of the Kwajalein Atoll, over 5000 miles away. The target launch was monitored and tracked by radars at Vandenburg much in the same way as early warning radars and satellites would provide launch notification and initial track of a real attack. Midcourse track of the target was provided both by radars at Hawaii and USAKA and the Global Positioning System (GPS). Special equipment had been installed on the target vehicle and the interceptor to allow them to be tracked by the Kwajalein based GPS Translator Processing System (TPS). This system emulated a midcourse sensor by providing target position and velocity information to a ground battle manager. The

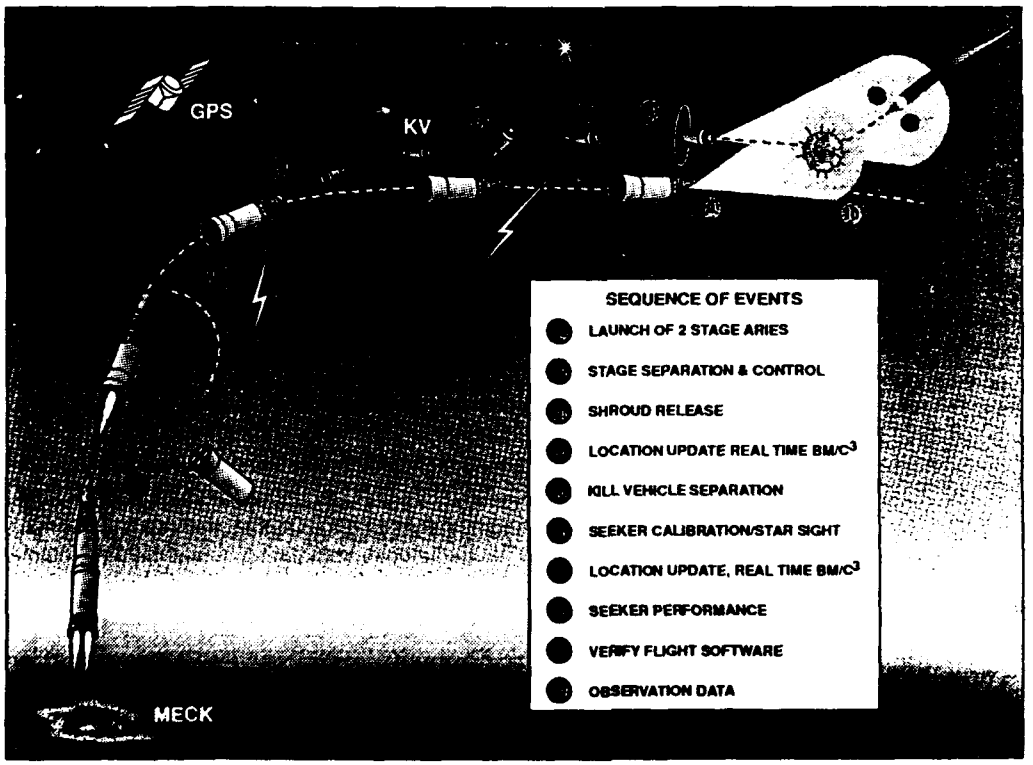


Figure 5. FTV Engagement Scenario

Mission and Launch Control Process (MLCP) software performed the role of the battle manager by tracking the target and determining the desired intercept point and the required interceptor launch and flyout parameters. Following interceptor launch, the MLCP periodically calculated and transmitted inflight target updates and a Threat Object Map (TOM) to the interceptor. Interceptor guidance used these inflight updates to maneuver the kill vehicle to an intercept course and place the target complex within the seeker's field of view. Shortly before target acquisition, the target deployed two decoy balloons as a countermeasure to the interceptor. Once the threat complex was acquired with the onboard seeker, the target was successfully designated by the TOM algorithm. During homing, the KV utilized the seeker data to close loop track the target and perform the final series of divert maneuvers. These adjustments to the trajectory placed the KV on a collision course with the target to achieve a body-to-body intercept. Kill assessment data and intercept phenomenology were collected by the KREMS radars, the observation aircraft, and the fly along Observer Package (OP).

All aspects of the FTV mission directly emulated the functions and operations of an actual defensive missile system. This ERIS FTV intercept has paved the way for more advanced hardware and software currently being developed for the GBI program.

FTV-1 Mission Accomplishments

Although success was obvious at the moment of intercept (Figure 6), most of the benefit to the GBI Program has been obtained through the post- mission analysis of the resulting data. Interceptor performance and seeker data were telemetered from the kill vehicle throughout its flight, and the fly along Observer Package provided detailed data on the intercept. Multiband infrared and visible data was also collected by the High Altitude Learjet Observatory (HALO) and Cobra Eye observation aircraft both during and following intercept. Extensive radar data was collected both pre and post intercept by the Kiernan Reentry Measurements Site (KREMS) radars located on Roi Namur. It was through this wealth of broad band multi-spectral data that the real success of the FTV 1 mission is measured.

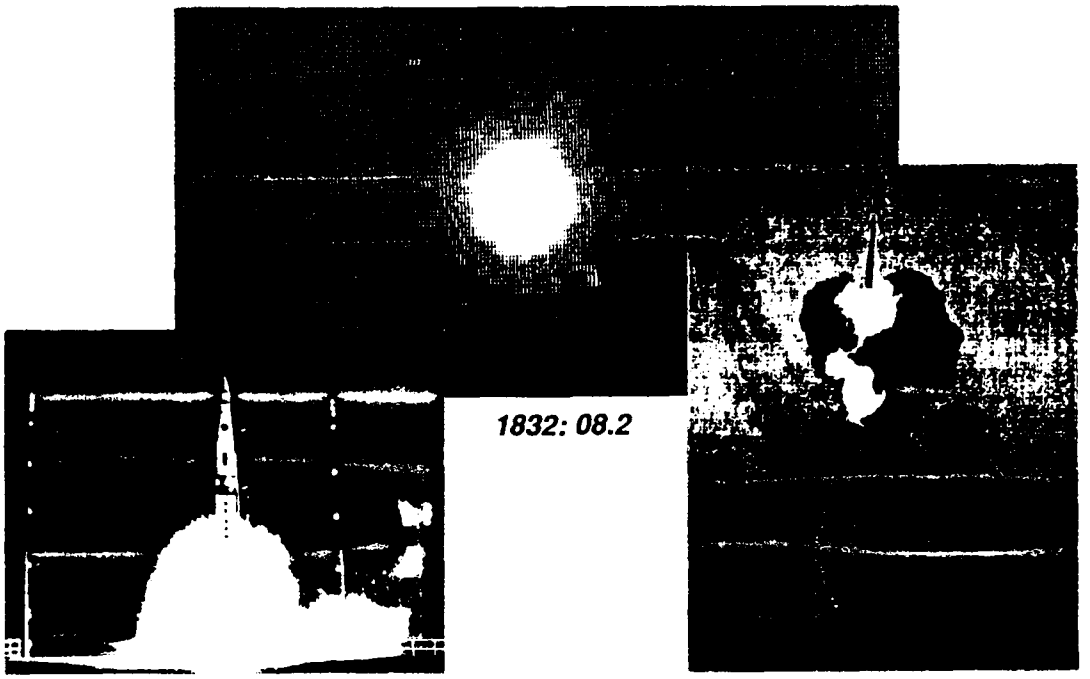


Figure 6. FTV Intercept

The highlights from the 28 January 1991 FTV-1 mission include the following:

- The countdown had no unplanned holds and the target was launched soon after the opening of the launch window.
- GPS provided accurate track of the target emulating the data from a strategic midcourse sensor. The target booster was slower than expected, but real time replanning of the intercept was performed automatically and accurately by the FTV ground software.
- Interceptor boost and flyout was nominal (considering the boosters were more than 25 years old) although some minor transients during boost indicated a need for tuning of the dynamic simulation models and adjusting the control system gains and filters.
- Inflight updates and diverts were executed properly.
- The TOM algorithm correctly selected the RV in spite of unexpectedly high seeker noise.
- Endgame homing divert maneuvers were performed correctly. A seeker boresight shift at first homing divert caused a significant

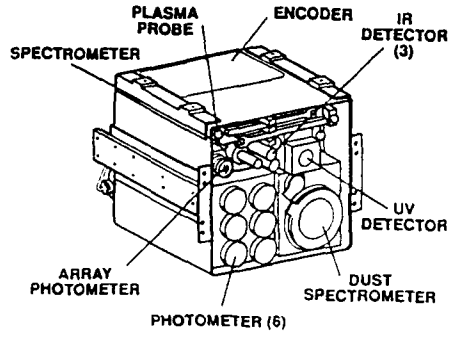
transient in the guidance filters, but the KV software design was sufficiently robust to compensate for this problem prior to the second homing divert.

- The target was completely destroyed on impact providing excellent data on impact signatures and debris.
- The test excellently represented the operation of a realistic defense system. Real time compensation for the variations from test planning was successfully accomplished.

FTV-1 provided several key contributions to the development of the strategic defense system. The important accomplishments and contributions were:

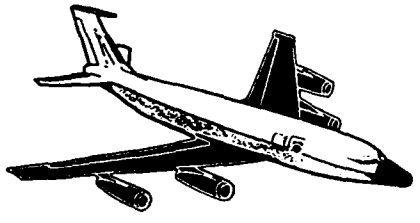
- Handover - Successfully demonstrated automated engagement planning based on real time tracking data and inflight updates of target and interceptor tracks. Conceptual algorithms for operational intercept planning and the concept of inflight updates were validated.
- Designation - The Threat Object Map (TOM) correctly selected the RV. The TOM concept and conceptual algorithms for target designation were validated for operational planning.
- Endgame Tactics - Demonstrated the effectiveness of aimpoint selection based on seeker data and the performance of the final divert maneuver. Conceptual aimpoint selection algorithms and terminal maneuver capability were validated for operational purposes.
- Immediate Kill - Body to body hit to kill was demonstrated. Validation of the lethality of a smaller KV was shown. Performance of the guidance concept, seeker approach (including body fixed mounting), P&RC miniaturization, and processing capability were shown.
- Discrimination - The limited (single color) objective was achieved. Furthermore, additional discrimination data was obtained. Specifically, the environment of the impact perturbed the trajectories of the associated objects in discernable ways.

Excellent data was also collected from the various support sensor systems (Figure 7). This data helps to support resolution of issues related to radar exo-discrimination, IR exo-atmospheric signature



FTV OBSERVER PACKAGE

HALO / IRIS



COBRA EYE

KREMS RADAR FACILITIES

Figure 7. FTV Support Sensor Systems

validation, intercept phenomenology, kill assessment, debris avoidance planning, and reentry effects on lightweight objects. The data collected by each of the sensors includes:

- KREMS Radars: balloon ejection velocity; balloon inflation; RV and balloon radar cross sections; RV and balloon dynamics; intercept flash; debris sizes and dispersion rates; debris bulk filtering techniques; and slowdown of lightweight objects during reentry.
- Cobra Eye: RV and balloon IR signatures; impact flash intensities and plasma cloud growth rate; debris IR temperatures; and reentry heating of lightweight objects.

- HALO: IR and visible data on impact flash including plasma cloud intensities, growth rates, maximum extents, and durations.
- Observer Package: IR, visible, and extreme ultraviolet data on impact flash intensities and durations. The OP also collected data on plasma electron density, variations in the electric field, and impact dust sizes, velocities, and distributions.

In addition, there are FTV accomplishments and contributions which are applicable to: future tests of this and other programs, areas of technology emphasis, and calibration of system assumptions for realism. These issues include:

- Threat realism dictates that empirical data be accommodated in signature assumptions (e.g., penaids with discernable signatures caused by imperfections) to avoid excessively conservative decisions based on superficial assumptions. The decision to deemphasize GBI should be revisited in the light of more realistic signature data.
- Several functional and operational issues were defined (blind divert, seeker boresight shift, base drag modeling values, and booster roll oscillations) which require resolution for subsequent FTV missions.
- The P&RC functioned well in the flight test environment. However, the "controlled" leakage problem requires a technology program with emphasis on modifying the design or developing new methods for controlling divert valve closure.

The FTV Program is a major step from the previously successful Homing Overlay Experiment (HOE) in the areas of air vehicle miniaturization, functional reallocation, and life cycle cost reduction. The KV propulsion and reaction control system uses a bonded rolling diaphragm fuel tank design which reduces the weight and improves expulsion efficiency with precise center of gravity control. The KV infrared seeker design is the first exoatmospheric demonstration of a fixed body seeker on a missile. Dormancy, the primary basis of operational and support cost reduction, is demonstrated through rapid cryostat operation to achieve inflight seeker cooldown. The KED design reduces the need for precise lethal aimpoint selection and increases interceptor kill capability against responsive threat countermeasures.

The FTV flight software implements velocity feedback for precise divert control, inflight measurement of thruster performance for accurate homing, scene matching target designation, closely spaced object discrimination, and target aimpoint selection in addition to basic signal processing and tracking functions. Additional advances resulting from FTV include high power density thermal batteries, and GPS translator demonstration. The FTV not only resolves functional issues but validates significant technology advances for GBI design consideration.

Remaining Issues

FTV flights address many of the issues underlying midcourse intercept. Navigation and guidance accuracy, threat object map handover, aimpoint selection, lethality, and hit to kill have been resolved during the FTV flights. Affordability and producibility have also been addressed during FTV. The remaining issues which are critical to robust interceptor capability are onboard discrimination and hardening for operation in a nuclear environment. An interceptor which can resolve and discriminate closely spaced objects not resolvable by precommit sensors will always provide a lower cost per kill. Furthermore, an interceptor must be hardened to any potential nuclear environments to ensure survivability and reliability.

Onboard discrimination is the central issue remaining on the program. How much, and how well the interceptor can do it, are questions that flight testing and simulations will answer. A very broad approach to the problem is in order. Discrimination testing will take place using a wide variety of onboard sensors. Long wave IR sensors with very long acquisition ranges will permit the time needed for temperature discrimination. Long range visible sensors, limited only by the Earth's geometry, will permit imaging and rate measurements. Ultraviolet wavebands will also be available. Active RF and radar sensors will measure doppler returns for signature modulations evidencing rotation rates, rate changes, and vibration. Concurrently, all forms of discrimination algorithms will bear on the problem. Statistical pattern recognition, expert systems, and especially real time learning techniques will be developed and implemented. The flight tests will exploit the potential in real time observation of intercept

phenomenology available to trailing interceptors and precommit sensors. The program will develop data fusion techniques to exploit data combination between interceptors and precommit sensors. At the same time, all threat objects will be modeled intensively to study their high frequency characteristics and their susceptibility to high frame rate, high resolution, and close range observation uniquely afforded by onboard interceptor discrimination.

Conclusion

The GBI Project has made significant strides in resolving the issues associated with exoatmospheric intercepts of ballistic missile RVs by ground launched interceptors. The FTV missions have provided the realistic environments and flight conditions required for true system performance verification. The extensive database resulting from the FTV missions have provided a wealth of information and will continue to be analyzed for years to come. The applications of this database span from interceptor performance, optical/radar signature evaluation to disciplines concerned with lethality, kill assessment, debris characterization, and hypervelocity impact phenomena. Due to the nature of this program, most of this database is classified. Special requests for access should be handled through the USASDC GBI Project Office in Huntsville, Alabama.

Optical Control of Resonant Tunnel Diode Devices

Dr. Robert A. Lux*, Dr. James F. Harvey[a], Mr.Peter G. Newman,
Mr. Dana J. Sturzebecher, Mr. Thomas P. Higgins, Mr. Thomas E. Koscica,
Dr. Arthur C. Paolella, Mr. Anthony Deanni[b], and Mr. Wayne H. Chang
US Army Electronics Technology and Devices Laboratory, SLCET,
Fort Monmouth, NJ 07703-5000

I. Introduction

The direct optical control of microwave semiconductor devices has received a large amount of attention because of the potential applications in phased array antennas for radar and communications, for antenna remoting, in interfacing optical fiber systems, and in isolating control and RF signals in microwave systems. A recent review article discusses these topics including optical frequency tuning of Gunn, IMPATT, TRAPATT, FET, and HEMT oscillators[1].

The direct optical control of resonant tunnel diodes (RTD's) via an optical fiber promise low noise operation of military mm-wave systems at frequencies which were previously unfeasible and with the capability of remote antenna and phased array applications. RTD's have been operated as oscillators as high as 400 GHz and as mixers above 2 THz[2,3]. GaAs RTD's operating at 10 GHz have been optimized to produce power densities of 5000 W/cm^2 of chip area with conversion efficiencies of about 10%. Power conversion efficiencies of 30% appear feasible using InGaAs RTD's[4]. These output power densities are roughly comparable to FET oscillators at this frequency. RTD's can operate at higher frequencies than FET's, HBT's, and two terminal devices such as Gunn and IMPATT diodes. In addition RTD's have very low noise characteristics. The lack of isolation between input and output in a two terminal device and the very low impedance of the RTD make the integration of these devices into conventional FET based MMIC design very awkward. Optical control of the RTD oscillators would make it possible to avoid this problem by optically isolating the RTD and antenna circuit, on a separate chip, from the rest of the system and provide the capability for remote antenna operation and for optically synchronized antenna arrays using optical fiber. In effect the optical signal acts as a third terminal, allowing isolation of input and output signals. Optical control of RTD's as

a. National Research Council research associate.
b. Employed by Vitronics, Inc.

a switching device has been reported. This paper presents the results of an ETDL investigation of optical control of these devices as microwave sources.

II. Device Structure

The material for the RTD's was fabricated at the ETDL molecular beam epitaxy (MBE) facility. The basic structure of the devices is shown in Figure 1. The thickness of the undoped spacer layers was varied from 50 nm to 250 nm. Mesas of diameters from 2 μm to 32 μm were etched. The current vs. bias curves of all devices are similar with the value of the peak current being approximately proportional to the area of the mesa. Oscillators could be made from all sizes of mesa and all spacer layers. The results reported will be for a structure with a 50 nm top spacer and a 250 nm lower spacer. Either Au-Ge or Pd-Ge contacts were made to the top of the mesa and to the bottom of the n+ substrate. No significant difference was seen in the current vs. voltage curves (I-V) between the two types of contact. The Pd-Ge contact has two advantages which make it more useful for future development: diffusion of the metal into the semiconductor is limited, and it shows better high temperature stability. A room temperature DC I-V curve for a 32μm diameter diode is shown in Figure 2. The portion of the curve labeled 'diode oscillating' does not represent a true DC

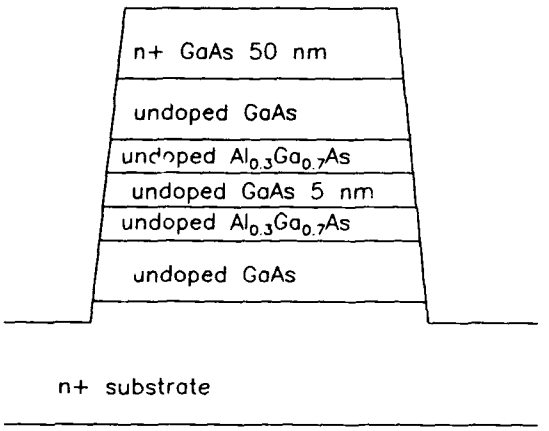


Figure 1. Layer structure of resonant tunnel diode material. Undoped GaAs layers varied from 50 nm to 250 nm.

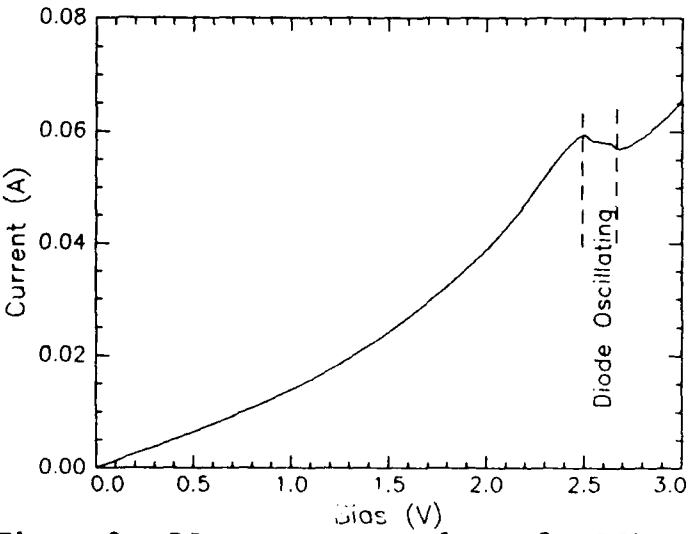


Figure 2. DC current vs. voltage for RTD

state, but is the average current in a diode that is oscillating with the parasitic impedance of the curve tracer.

III. Experimental Setup

The RTD was mounted in a 50 ohm coplanar test fixture with a wire bond (25μm Au wire me long) from the top of the mesa (Figure 3). When biased in the negative differential resistance region the RTD oscillated at a free-running frequency determined by internal and external capacitances and the external inductance, which is primarily due to the bond wire. The stub tuners were adjusted for maximum power output. An Ortel 1210C pigtail laser diode of 840 nm wavelength was used as the source of illumination. Maximum intensity at the exit of the fiber was 4 mW, with the illumination covering the edge of the diode mesa and the underlying substrate. Various laser rf generators allowed the laser to be amplitude modulated from DC to 3 GHz.

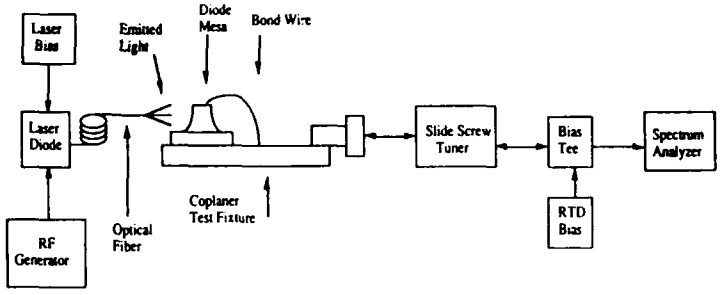


Figure 3. Schematic of RTD test fixture.

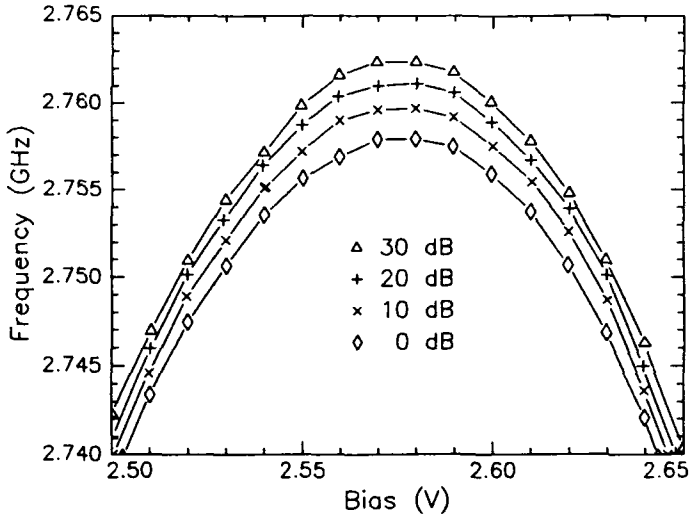


Figure 4. RTD oscillator free-running frequency as a function of bias and DC optical input. The four curves are for differing optical attenuation inserted in the fiber link.

IV. Frequency Modulation

The diode was illuminated with 2mW DC optical power through a variable attenuator. It was found that the frequency of oscillation was a function of both the bias voltage and the illumination power as shown in Figure 4. For any bias voltage, the frequency shift is a nonlinear function of the optical

power, each 10 dB increase producing approximately equal frequency shift.

The oscillator frequency can follow a high frequency modulation of the laser. Figures 5 and 6 show the frequency spectrum of an oscillator with 5 kHz square wave optical modulation and 75 MHz sine wave optical modulation respectively.

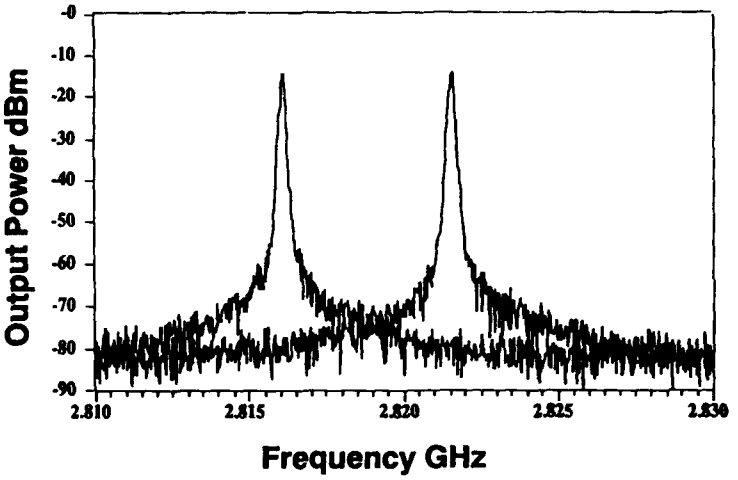


Figure 5. Spectrum of RTD oscillator modulated by 5 kHz square wave optical signal.

V. Injection Locking

Like other microwave oscillators it is possible to electrically injection-lock a free-running RTD oscillator to a rf signal. This led us to attempt injection-locking with an optical signal. With a laser source of 4 mW at the exit of the coupling fiber, it was found that when the frequency of optical modulation approached the free-running RTD oscillator frequency, the oscillator locked to the optical modulation frequency. Figure 7 shows the RTD oscillator spectrum free-running and locked to an optical signal. On locking there is the typical injection locking reduction of noise sidebands and narrowing of the central peak. The total locking bandwidth was 150 kHz.

VI. Mechanism of Optical Control

Three mechanisms have been proposed to explain the observed optical effects: 1. Photoconductivity reduces the resistance of the diode external to the barrier layers; the reduced IR voltage drop changes the effective diode bias. 2. Holes generated below the barrier layer drift into the

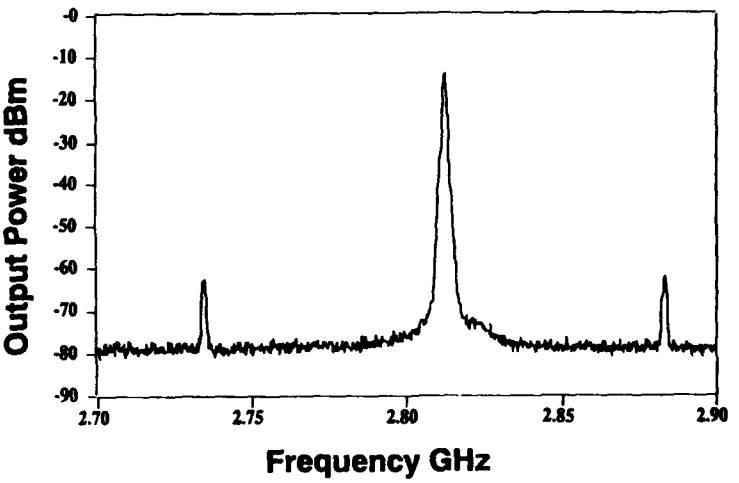


Figure 6. Spectrum of free-running RTD oscillator with 75 MHz sine wave optical modulation

barrier region and are trapped resulting in a changed electric field in the device. 3. DX center traps in the AlGaAs barriers are ionized by the light resulting in a dipole field within the barrier structure[5].

While all three mechanisms may contribute in some cases, we believe third is not likely in our diodes. All the diodes reported here have a thick metal top on the mesa, hence the only way the barrier region could be illuminated is through the edge of the mesa. This would be a very inefficient process. Further there would be a large difference in optical effects depending on the diameter of the diodes. We see no such effect.

Photoconductivity in the substrate does not seem to be a satisfactory explanation. As shown in figure 4 to produce the observed optically induced frequency shifts would require bias shifts of about 50mV. This is greater than the total voltage drop in the substrate when operating in the oscillator region, hence an increased substrate conductivity would not produce the observed effects.

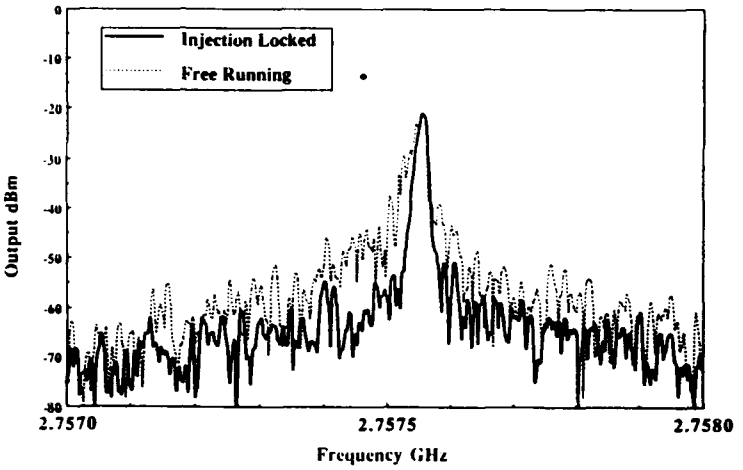


Figure 7. Injection locking of RTD oscillator by a 4 mW optical signal.

VII. Conclusions

This work at ETDL has produced the first demonstration of optical control in microwave RTD's. While these results are a proof of principle demonstration, the high frequency, low noise capability of RTD's makes them as a potential candidate for future Army systems.

VIII. References

1. A.J. Seeds and A.A.A. de Sales, "Optical Control of Microwave Semiconductor Devices", *IEEE Trans. Microwave Tech.* vol. 38, pp 577-584, (1990)

2. E.R. Brown, T.C.L.G. Sollner, C.D. Parker, W.D. Goohue, and C.L. Chen, "Oscillations up to 420 GHz in GaAs/AlAs Resonant Tunneling Diodes", *Appl. Phys. Lett.*, vol. 55, pp 1777-1779, (1989)

3. T.C.L.G. Sollner, P.E. Tannenwald, D.D. Peck, and W.D. Goodhue, *Appl. Phys. Lett.*, vol. 45, pp1319-1322, (1984)

4. V.K. Reddy, A.J. Tsao, S. Javalagi, G.K. Kumar, D.R. Miller, and D.P. Neikirk "Quantum Well Injection Time (QWITT) Diode Oscillators", *Fifteenth International Conference on Infrared and Millimeter Waves, sponsored by SPIE, R.J. Temkin ed.*, pp 88-90 (1990)

5. T.C.L.G. Solner, E.R. Brown, and H.Q. Le, "Microwave and Millimeter-wave Resonant Tunneling Devices", *Lincoln Lab. Jour.*, vol 1, pp 89-105, (1988)

MARTIN, KYLE, NUZUM, GROGL, ANDERSEN, MILHOUS

Plasmodium falciparum: Characterization of Multiple Drug Resistance Phenotypes by Reversal Modulators

CPT Rodger K. Martin[1], Ph.D.; CPT(P) Dennis E. Kyle[2], Ph.D.; MAJ Edwin O. Nuzum[1], D.V.M., Ph.D.; MAJ Max Grogl[1], Ph.D.; MAJ Steven L. Andersen[1], M.D.; and LTC Wilbur K. Milhous[1], Ph.D.

[1]Department of Parasitology, Division of Experimental Therapeutics, Walter Reed Army Institute of Research, Washington, DC 20307-5100

[2]Department of Immunology and Biochemistry, Armed Forces Research Institute of Medical Sciences, Bangkok, Thailand

Abstract

Infectious diseases historically have adversely impacted the combat readiness of United States military personnel. In military campaigns, they typically account for the majority of hospital admissions for disease-nonbattle injuries (DNBI). One of the most significant sources of morbidity has been malaria. The most lethal form, malignant tertian malaria, results from an infection by the protozoan parasite *Plasmodium falciparum*. This parasite has developed pleiotropic drug resistance to many of the drugs employed for chemoprophylaxis and chemotherapy. The drug refractory nature of the parasites is highly reminiscent of the multiple drug resistance (MDR) phenotype commonly observed in human neoplasias. In order to delineate and to define the genetic mechanisms that mediate MDR, we have developed drug resistance phenotypes based upon geography and specific reversal modulators. By employing this methodology, we have ascertained that there are at least three phenotypes. These phenotypes may translate into one or more mechanisms for resistance. These data will further assist in the description of the drug resistance phenotypes and provide impetus to lead-directed synthesis of new compounds to circumvent these resistance mechanisms.

Introduction

The annals of world military history are littered with failed military campaigns. The halting of Alexander the Great to the East, the failure of the Crusades, and in more modern times, the British Macedonian campaign of World War I failed not to armed resistance but rather, they succumbed to infectious diseases (Schmidt & Roberts, 1981). The history of the United States is no exception. From the American Revolutionary War to the Vietnam Conflict, casualties from infectious diseases far exceeded those for combat. During the Mexican-American War of 1846-1848, only one in eight deaths could be attributed to combat action among U.S. regulars and volunteers (Upton, 1881). Of the 600,000 men who died during this country's most divisive conflict, the Civil War, twice as many died from disease as were killed or mortally wounded in combat (McPherson, 1988). As recently as the Vietnam Conflict, the disease-to-combat morbidity rates were slightly more than two to one.

One of the greatest sources of morbidity and mortality throughout our history has been malaria. It was and remains the bane of the soldier and the nemesis of military health care providers. Malaria continues to represent a serious health threat for military personnel deployed in malarious endemic regions for training or operations. The World Health Organization (WHO) has estimated that approximately 40% of the world population reside in malarious regions, the prevalence may be as many as 270 million infections, and the annual incidence of 110 million with an estimated mortality figure of 1 million deaths primarily in children under the age of five years (WHO, 1990). The etiological agent of the most prevalent and lethal type of malaria is due to a parasitic protozoan known as *Plasmodium falciparum*. Given malaria's ubiquitous geographical distribution throughout temperate climates (approximately 50% of the world's land mass), malaria will adversely impact the combat readiness and effectiveness of military personnel deployed in these areas. In our most recent military expedition in Saudi Arabia,

Desert Shield & Desert Storm, malaria was considered to be a potential health care threat.

In the late 1950's, the WHO prematurely forecasted the control of falciparum malaria. This bold and optimistic prognostication was apparently justified by the development of potent insecticides for vector control and new antimalarial drugs, such as chloroquine. However, the use and the effectiveness of the insecticides were ephemeral due to environmental concerns and the manifestation of vector resistance. Furthermore, parasitic resistance to chloroquine was immediately documented in Southeast Asia and in South America. At the moment, chloroquine resistance is common to all regions where falciparum malaria is endemic. The clinical utility of chloroquine has been severely compromised due to the acute drug refractory nature of the parasites. An additional and ominous development has been the description of pleiotropic drug resistance to the current antimalarial chemotherapeutic armamentarium.

The observation of pleiotropic resistance to antimalarial drugs of diverse chemical structures was highly reminiscent of the multiple drug resistance (MDR) phenotype associated with human neoplasias. Indeed, the similarities between drug resistant malaria and human neoplasias may be extended to include a measurable extrusion of cytotoxic drugs (Krogstad *et al.*, 1987), the presence of an energy-dependent, membrane spanning p-glycoprotein (Foote *et al.*, 1989; Wilson *et al.*, 1989), and the ability to modulate the resistance phenotype by the addition of subinhibitory concentrations of drugs that do not possess any intrinsic antimalarial activity (Martin *et al.*, 1987; Bitonti *et al.*, 1989; Kyle *et al.*, 1990; Basco *et al.*, 1991; and Kyle *et al.*, 1992). It is the latter characteristic of falciparum MDR that we have tried to exploit in ascertaining the number of drug resistance phenotypes. This approach would assist in the identification of resistance mechanisms.

Methodology

Selection of Parasite Clones and Isolates

Several well-characterized clones and isolates of *P. falciparum* with differing drug susceptibility profiles were utilized for these drug assays. The clones W2 (Oduola *et al.*, 1988) and D6 (Oduola *et al.*, 1988) are reference clones from Indochina III and Sierra Leone isolates, respectively. GA3 is a clone of the GH2 isolate obtained from Thailand. 306 is a clone of the IEC 51/84 isolate from Brazil. Nigeria 60 is a patient isolate collected from Africa. Clones W2, GA3, 306, and isolate Nigeria 60 are chloroquine resistant, whereas clone D6 is chloroquine sensitive.

Drug Assays

Susceptibilities of each clone to drugs were evaluated by using a modification of the semi-automated microdilution technique (Desjardins *et al.*, 1979). The antimalarial drugs (chloroquine diphosphate, mefloquine hydrochloride, halofantrine, quinine, and artemisinin) and reversal modulating drugs (verapamil, penfluridol, desipramine, chlorpromazine, and fluoxentine hydrochloride) were obtained from an inventory of drugs maintained by the Walter Reed Army Institute of Research. Drug assays were performed in 96-well microtiter plates with 1% erythrocyte suspensions at 0.2-0.5% parasitemia in RPMI1640 supplemented with 25 mM HEPES, 32 mM $NaHCC_3$, and 10% human plasma. Microtiter plates were prepared with serial dilutions of drug to determine fifty-percent inhibitory concentrations (IC_{50}s). Each plate containing parasites and drug was incubated at 37°C in an airtight, plexiglas box in the presence of 5% oxygen, 5% carbon dioxide, and 90% nitrogen. After 24 hours of incubation, cultures were labeled with ^{3}H-hypoxanthine (NEN Dupont) and incubated for an additional 18-20 hours prior to harvesting. Incorporation of radiolabel in each well was determined by scintillation spectrophotometry, and computer-generated

concentration response curves were analyzed by non-linear regression. Susceptibilities of the parasites to the antimalarial drugs alone and in combination with subinhibitory concentrations of reversal agents were compared to ascertain the degree of modulation. IC_{50}s were determined for each drug alone and for various drug combinations.

Modulation of the susceptibility of the parasite to the antimalarial drugs singularly and in the presence of the resistance modulator at predetermined concentrations was expressed as the response modification index (RMI) (Kyle *et al.*, 1990). RMI's were calculated by the following formula:

$$RMI = IC_{50(A,B)}/IC_{50(A)}$$

where A is the antimalarial IC_{50} and B represents the resistance modulator IC_{50}. A RMI equivalent to 1.0 represents no change in the IC_{50} upon addition of the resistance modulator at predetermined levels. A RMI greater than 1.0 demonstrates an antagonistic effect between the resistance modulator and the antimalarial drug, whereas a RMI of less than 1.0 represents potentiation or synergism.

Results

Table 1 is a summary of *in vitro* IC_{50} data for parasite isolates from three different geographical regions of the world. This table shows how reversal phenotypes can be defined by the parasite's origin, the antimalarial drug to which resistance develops, and the chemotherapeutic agents that modulate that resistance.

Table 1. Table of reversal phenotypes based on reversal modulators that decrease parasite resistance to the antimalarial drug listed at the top of each respective column.

	Reversal Phenotype I	Reversal Phenotype II			Reversal Phenotype III		
Origin of Parasite Isolates	Chloroquine and Quinine	Mefloquine	Halofantrine	Artemisinin	Mefloquine	Halofantrine	Artemisinin
West Africa	Verapamil Chlorpromazine Desipramine Amitriptyline Prozac Cyproheptadine Promethazine	Penfluridol[1]	Penfluridol[1]	Penfluridol[1]			
Southeast Asia	Verapamil Chlorpromazine Desipramine Amitriptyline Prozac Cyproheptadine Promethazine				Penfluridol Verapamil Chlorpromazine Desipramine Amitriptyline	Penfluridol Verapamil Chlorpromazine Desipramine	Penfluridol[2]
South America	Verapamil Chlorpromazine Desipramine	NA[3]	NA[3]	NA[3]	NA[3]	NA[3]	NA[3]

[1]Penfluridol is the only reversal modifier found to decrease resistance to mefloquine, halofantrine and artemisinin in West Africa.

[2]Reversal of artemisinin resistance by chlorpromazine and desipramine is suspected but not confirmed.

[3]There is no known resistance to mefloquine, halofantrine and artemisinin in South American *P. falciparum* parasites.

Discussion

Since 1963, the United States Army Medical Research and Development Command has undertaken the most ambitious drug screening program in the history of malaria chemotherapy. Over 350,000 compounds have been evaluated for their antimalarial activity. Due to the complex nature of drug development and licensing, only a handful of drugs have actually received FDA approval and are available in most formularies. The prolonged clinical utility of current antimalarial drugs is of paramount concern. Therefore, the emergence of multiple drug resistant *Plasmodium falciparum* represents a significant health concern and challenge for military health care providers. The reduced efficacy of currently approved antimalarial drugs for chemoprophylaxis and chemotherapy is alarming. The diminished clinical utility of chloroquine has resulted in a concerted effort to rush through the approval process the dissemination of mefloquine, halofantrine, and artemisinin as alternative drugs. However, in specific geographical areas (most notably Southeast Asia), resistance to these drugs has been encountered and the longevity of their future effectiveness remains questionable. Until the phenotypes and mechanisms involved in mediating resistance to these drugs are described and characterized the resistance phenomenon will continue to be a prominent health care problem. In order to ascertain the number of phenotypes and mechanisms, we have attempted to delineate the drug resistance phenotypes by geography and by the ability to modulate resistance to antimalarial drugs with specific reversal drugs.

The rationale for this approach was as follows. Mediation of resistance by a single gene was plausible. Yet, the cumulative molecular, biochemical, and drug efficacy data seemed to suggest otherwise. Indeed, it has been thought that resistance is multigenetic. However, only one putative gene has been identified. The primary candidate gene that has been the subject of intense investigation has been the *pfmdr*1 gene that encodes the falciparum

equivalent of the human P-170 glycoprotein. The P-170 glycoprotein has been implicated in human neoplasias as the "pump" that reduces the intracellular accumulation of antineoplastic drugs in cancer cells (Gottesman & Pastan, 1988). The falciparum p-glycoprotein homolog shared amino acid sequence and structural identity with the human P-170 glycoprotein (Wilson *et al.*, 1989; Foote *et al.*, 1989). Initially, it was anticipated that this particular mechanism would account for the observed efflux of antimalarial drugs, especially chloroquine, and reduced accumulation within resistant parasites. Another piece of data linking the falciparum p-glycoprotein to drug resistance was the localization of the p-glycoprotein to the parasite's food vacuole membrane, the postulated target of chloroquine's cytotoxic action (Cowman *et al.*, 1991). However, the mere presence of the p-glycoprotein on the food vacuole was not sufficient to explain the broad cross-resistance especially since the target of the "new" antimalarials such as mefloquine and halofantrine have yet to be identified. Whereas an association was made between amplification of the p-glycoprotein gene and recent mefloquine-resistant falciparum isolates and clones from Thailand (Wilson, personal communication), no conclusive data has been published that directly correlates chloroquine, mefloquine, or multiple drug resistance with *pfmdr*1 amplification, over-expression, or mutations (Foote *et al.*, 1989; Wilson *et al.*, 1989; Foote *et al.*, 1990). Therefore, a single gene defect seemed unlikely to sufficiently account for the resistance observed.

Previously, it was observed that falciparum malaria's susceptibility to various antimalarial drugs differed from one geographic region to another (Milhous & Kyle, personal observation). In addition, the ability to modulate resistance seemed to be contingent upon geography , the *antimalarial drug,* and the reversal *agent.* Therefore, we decided to determine the number of mechanisms based upon the ability to modulate resistance with reversal agents in resistant parasites from diverse geographical areas of the world. By employing this strategy, we have determined that there at least

three distinct reversal phenotypes and quite possibly three mechanisms for the mediation of resistance.

In Table 1, the malarious regions of the world have been segregated into three distinct areas: South/Central America, Africa, and Southeast Asia. Although this breakdown is not inclusive of all malarious regions, it does include most countries monitoring falciparum drug resistance. We also elected to define the parasites by the degree of modulation of drug resistance tc the chinchona alkaloids (chloroquine, desethylchloroquine, quinine, and quinidine), mefloquine, halofantrine, and artemisinin. The drugs that have been identified as increasing the parasitic sensitivity to chloroquine were found to be common to all geographical areas. In other words, the reversal agents that increase chloroquine accumulation in resistance parasites from South America are equally effective in resistance parasites from Africa or Southeast Asia. However, some subtle differences have been noted in the degree of reversal of quinine resistance in South American parasites when compared to parasites studied from Africa and Southeast Asia. Nevertheless, this constitutes the first drug resistance phenotype and is consistent throughout the world.

The second drug resistance phenotype is unique to Africa. Resistance to mefloquine has been documented *in vitro* from African patient isolates (Salako & Aderounmu, 1987; Oduola *et al.*, 1987; and Oduola *et al.*, 1992). This particular phenotype exclusively involves the ability to modulate mefloquine resistance with the neuroleptic drug, penfluridol (Peters & Robinson, 1991; Kyle *et al.*, 1992). No other reversal modulator was effective at increasing the accumulation of mefloquine in African drug refractory parasites. Unlike Southeast Asian parasites, mefloquine resistant, African *P. falciparum* parasites are presently susceptible tc the drugs halofantrine and artemisinin, although transient increases in their IC_{50}s have been reported (Oduola *et al.*, 1992). These data from African parasites make them truly unique when compared to

mefloquine resistant isolates obtained from Southeast Asia.

The final drug reversal phenotype is also the most complex. Parasites from Southeast Asia constitute the most mutidrug resistant populations. These parasites exhibit cross-resistance to chloroquine, mefloquine, halofantrine, and artemisinin. Whereas Southeast Asian parasites are comparable to African parasites in their reversal phenotypes to chloroquine and the other chinchona alkaloids, they differ in their mefloquine resistance phenotype. In addition to penfluridol, these Southeast Asian parasites may have their resistance to mefloquine reversed by many of the drugs that modulate resistance to chloroquine. Furthermore, these same reversal modulator drugs are effective at increasing *in vitro* sensitivity to halofantrine and artemisinin (Kyle *et al.*, 1992). The *in vivo* linkage of mefloquine and halofantrine resistance has been observed in Thai soldiers who failed mefloquine prophylactically and subsequent chemotherapy with halofantrine (Shanks *et al.*, 1991). Halofantrine and mefloquine resistance may be genetically linked in Southeast Asia, whereas there is disagreement over a similar linkage in African parasites.

The description and characterization of these drug resistant phenotypes as determined by their *in vitro* modulation by tricyclic antidepressants, calcium antagonists, and phenothiozines presents a unique approach to the elucidation of the MDR phenomenon. These observations should accelerate the search for the mechanisms that mediate falciparum drug resistance.

References

Basco, L.K., P. Ringwald, and J. Le Bras. 1991. Chloroquine-potentiating action of antihistaminics in *Plasmodium falciparum in vitro. Ann. Trop. Med. Parasit.* 85: 223-228.

Bitonti, A.J., A. Sjoerdsma, P.P. McCann, D.E. Kyle, A.M.J. Oduola, R.N. Rossan, W.K. Milhous, and D.E. Davidson, Jr. 1988. Reversal of chloroquine resistance in malaria parasite *Plasmodium falciparum* by desipramine. *Science* 242: 1301-1303.

Cowman, A.F., S. Karcz, D. Galatis, and J.G. Culvenor. 1991. A p-glycoprotein homologue of *Plasmodium falciparum* is localized on the digestive vacuole. *J. Biol. Chem.* 113: 1033-1042.

Desjardins, R.E., C.J. Canfield, J.D. Haynes, and J.D. Chulay. 1979. Quantitative assessment of antimalarial activity *in vitro* by a semiautomated microdilution technique. *Antimicrob. Agents Chemother.* 16: 710-718.

Foote, S.J., J.K. Thompson, A.F. Cowman, and D.J. Kemp. 1989. Amplification of the multidrug resistance gene in some chloroquine-resistant isolates of *P. falciparum. Cell*: 57: 921-930.

Foote, S.J., D.E. Kyle, R.K. Martin, A.M.J. Oduola, K. Forsyth, D.J. Kemp, and A.F. Cowman. 1990. Several alleles of the multidrug resistance gene are closely linked to chloroquine resistance in *Plasmodium resistance. Nature* 345: 255-258.

Gottesman, M.M., & I. Pastan. 1988. The multiple transporter, a double-edged sword. *J. Biol. Chem.* 263: 12163-12166.

Krogstad, D.J., I.V. Gluzman, D.E. Kyle, A.M.J. Oduola, S.K. Martin, W.K. Milhous, and P.H. Schlessinger. 1987. Efflux of chloroquine from

Plasmodium falciparum: mechanism of chloroquine resistance. *Science* 238: 1283-1285.

Kyle, D.E., A.M.J. Oduola, S.K. Martin, and W.K. Milhous. 1990. *Plasmodium falciparum*: Modulation by calcium antgonists of resistance to chloroquine, desethylchloroquine, quinine, and quinidine *in vitro*. *Trans. R. Soc. Trop. Med. Hyg.* 84: 474-478.

Kyle, D.E., A.M.J. Oduola, H.K. Webster, K. Pavanand, G.D. Shanks, and W.K. Milhous. 1992. *Plasmodium falciparum: in vitro* reversal of resistance to mefloquine, halofantrine, chloroquine, and artemisinin. *Exp. Parasitol.* (submitted).

Martin, S.K., A.M.J. Oduola, and W.K. Milhous. 1987. Reversal of chloroquine resistance in *Plasmodium falciparum* by verapamil. *Science* 235: 899-901.

McPherson, J. M. 1988. *Battle Cry of Freedom: The Civil War Era.* Ballatine Books: New York, p. 485-488.

Oduola, A.M.J., W.K. Milhous, L.A. Salako, O. Walker, and R.E. Desjardins. 1987. Reduced *in vitro* susceptibility to mefloquine in West African isolates of *Plasmodium falciparum. Lancet* ii: 1304-1305.

Oduola, A.M.J., W.K. Milhous, N.F. Weatherly, J.H. Bowdre, and R.E. Desjardins. 1988. *Plasmodium falciparum*: cloning by single-erythrocyte micromanipulation and heterogeneity *in vitro. Exp. Parasitol.* 66: 86-95.

Oduola, A.M.J., O. Sowunmi, W.K. Milhous, D.E. Kyle, R.K. Martin, O Wallker, and L.A. Salako. 1992. Innate resistance to new antimalarial drugs in *Plasmodium falciparum* from Nigeria. *Trans. R. Soc. Trop. Med. Hyg.* (in press).

Peters, W. & B.L. Robinson. 1991. The chemotherapy of rodent malaria. XLVI. Reversal of of mefloquine resistance in rodent *Plasmodium. Ann. Trop. Med. Parasit.* 85: 5-10.

Salako, L.A., & A.F. Adernounmu. 1987. *In vitro* chloroquine and mefloquine resistant *Plasmodium falciparum* in Nigeria. *Lancet* i: 572-573.

Schmidt, G.D. & L.S. Roberts. 1981. *Foundations of Parasitology.* C.V. Mosby Co.: ST. Louis, p. 152.

Shanks, G.D., G. Watt, M.D. Edstein, H.K. Webster, V. Suriyamongkol, C. Watanasook, S. Panpunnung, and W. Kowinwiphat. 1991. Halofantrine for the treatment of mefloquine chemoprophylaxis failures in *Plasmodium falciparum* infections. *Am. J. Trop. Med. Hyg.* 45(4): 488-491.

Upton, E. 1881.*The Military Policy of the United States.* Fourth Impression. Government Printing Office: Washington, D.C., p. 216-218.

WHO Report. 1990. *"World Report on Tropical Diseases."*

Wilson, C.M., A.E. Serrano, A. Walsey, M.P. Bogenschutz, A.H. Shankar, and D.F. Wirth. 1989. Amplification of a gene related to mammalian *mdr* genes in drug-resistant *Plasmodium falciparum. Science* : 244-1184-1186.

MATTHEW

An Animal Model of Drug-Induced Thermoregulatory and Endurance Decrements

Ms. Candace B. Matthew
US Army Research Institute of Environmental Medicine
Natick, MA 01760-5007

INTRODUCTION

Effects of the neurotransmitter acetylcholine are divided into muscarinic or nicotinic depending on the receptors that are activated. Stimulation of muscarinic receptors results in vasodilation, decreased heart rate and cardiac contractility, increased secretions of exocrine glands including sweat and salivary, increased intestinal and gastric contractions, and increased bronchial secretions. Agonists at nicotinic receptors stimulate autonomic ganglia and neuromuscular junctions. Over-stimulation results in asynchronous excitation, fatigability, and involuntary twitching[1]. Anticholinergic drugs are commonly used as antihistamines, tranquilizers, cold medications, antidiarrheal medications, as well as treatments for motion sickness and anticholinesterase poisoning. Use of anticholinergic drugs has detrimental effects on thermoregulation particularly in hot environments due to sweat suppression[2]. Anticholinesterase drugs are used clinically for treatment of myasthenia gravis, Alzheimer's disease, anticholinergic syndrome, and as a pretreatment against potential organophosphorus exposure[1].

Earlier work from this laboratory established the exercising rat and the sedentary heat-stressed rat as models of human heat-induced injury[3,4,5,6]. More recently, interest centered on the effects of anticholinergic and anticholinesterase drugs on physical, physiological, and thermoregulatory responses to hot environments and exercise in moderate to hot environments. Atropine, the prototypical anticholinergic drug, inhibits evaporative cooling by suppressing sweating in man[1,7,8] and by suppressing saliva secretion in rats which is behaviorally spread for evaporative cooling[9,10]. Clubley *et al.*[11] reported that sweat and saliva production are both analogously regulated in man. A 2 mg dose of atropine in a 70 kg man has been shown to suppress the sweat rate by about

40%[12,13,14,15]. Our work demonstrated that a 200 ug/kg dose of atropine (equivalent to 2 mg in man[16]) inhibits salivary water loss by the same 40%.

Sweat in man[17] and saliva in rats[18] contain Na^+ and K^+ but are hypotonic with respect to serum. Thus, both man[19] and rats[20] lose electrolytes, are dehydrated, and have increased serum osmolarities as a result of prolonged evaporative cooling. It has long been known that thirst is an inadequate stimulus for rehydration in man[21]. Recently, Barney and West[20] reported that thermally-induced salivary loss in rats increases plasma osmolality and that thirst is stimulated only to the point of replacing 50% of the water when the osmolality returns to control levels. Thus, mechanisms of evaporative cooling in man and rats are similar in cooling fluid composition, neural control, and resultant effects on dehydration and thirst.

MODEL DESCRIPTION

Experimental Animals: Adult male Sprague-Dawley (Charles River CD strain) 510-530 g rats were used for all of the studies unless otherwise indicated. The animals were caged individually and housed in an environmental chamber maintained at 26°C and 50% rh. Lighting was controlled automatically (on, 0600-1800 hours) and rat chow and water were available ad lib. except during experimental intervals. Drug administration and blood sampling, when required, were accomplished using a lateral tail vein or a jugular cannula surgically implanted in rats under sodium pentobarbital anesthesia.

Sedentary Heat-Stressed Rats: Except as noted below, rats were unrestrained during heat stress to facilitate saliva spreading activity. Rats, in their own cages, were placed in an environmental chamber maintained at 41.5°C until the desired end point core temperature (Tc) was attained (usually 42.6°C), when the animals were removed to a 26°C chamber. During heat stress, Tc, weight loss (as a measure of fluid spread for evaporative cooling), and the extent of saliva or urine spread were monitored. Fecal pellet production was monitored both to correct weight loss to more accurately reflect water loss and as an indication of intestinal motility.

When saliva is unavailable as in the case of surgical desalivation[10] or inhibition of salivation by anticholinergic drugs[10,22,23], rats will spread urine if no other fluid is available. We reported[22] (Fig. 1) that the extent of saliva or urine spread for evaporative cooling was a useful measure of the surface available for

evaporative cooling. Since, saliva-spreading requires both physiological and grooming activity, this provides a behavioral as well as physiological measure.

Exercising Rats: To identify and quantify effects on exercise endurance and thermoregulation, rats were run on a motor driven treadmill (11 m/min, 6° incline, shock avoidance contingency). Unless otherwise indicated, they were run to exhaustion, defined as the point at which rats fail to keep pace on the treadmill and do not immediately right themselves when placed on their backs. Rats were run for 20 min then allowed a 2 min rest, run for an additional 20 min followed by a 2 min rest and then run to exhaustion without any further rest (control animals ran for about 60 min). During running and subsequent recovery Tc (thermistor probe inserted 6.5 cm beyond the anal sphincter) as well as tail temperature (Tt, surface thermistor taped to the dorsal mid point of the tail) were monitored with a computerized data acquisition system.

For studies of anticholinesterase drug effects, a drug symptom check list[24] was developed to quantitate tremors, salivation, exophthalmus, defecation, run performance, and overall behavior. These indices proved useful in identifying the efficacy, dosages, and specificity of anticholinergic drugs[24,25].

RESULTS

Sedentary Heat-Stressed Rats: Following are descriptions of results reported using the sedentary heat-stressed rat model in our laboratory[10,22,23,30]:

Prior to using the heat-stressed rat as a model for examining the thermoregulatory effects of anticholinergic drugs, it was necessary to compare the response of heat-stressed rats that were unable to spread saliva for evaporative cooling with the response of heat-stressed control animals. In this study,[10] rats were surgically desalivated (DESAL) by ligation and transection of the salivary ducts. Since surgically desalivated rats were found to spread urine for evaporative cooling when no other fluid was available, it was necessary to restrain (RES) these animals to prevent spread of any fluid for evaporative cooling. Unrestrained (UNRES) control animals (Table 1) spread saliva on their ventral surface losing 8% of their body weight (Tc final= 42.6°C) with a rate of rise of core temperature (heating rate) of 0.02°C/min. Restraint (CONTROL RES) inhibited, but did not completely prevent, behavioral spread of saliva; some saliva was groomed into the fur and some dripped out of the mouth and was lost. DESAL UNRES rats spread some urine resulting in similar % weight loss and heating rate to those of the

CONTROL RES rats. A combination of desalivation and restraint was required to achieve heat stress with no evaporative cooling; DESAL RES rats had the minimum % weight loss and maximum heating rate. Pharmacological desalivation via administration of 1 mg/kg of atropine prior to the start of heat stress resulted in inhibition of both salivation and urine spreading with % weight loss and heating rate similar to those of the DESAL RES rats. Thus, Hubbard, *et al.*[10] concluded that the heat-stressed rat was a potentially useful model of anticholinergic-induced thermoregulatory decrements.

TABLE 1

Weight Loss and Heating Rate of Surgically and Chemically Desalivated Rats

Group [a]	% Weight Loss	Heating Rate, °C/min
CONTROL UNRES	8.0 ± 0.4	0.02 ± 0.002
CONTROL RES	2.2 ± 0.2	0.08 ± 0.01
DESAL UNRES	1.9 ± 0.4	0.07 ± 0.01
DESAL RES	0.7 ± 0.1	0.16 ± 0.01
ATROPINE 1mg/kg	0.7 ± 0.1	0.12 ± 0.004

a Data from Hubbard, *et al.*[10]; see text for group definitions.

In a subsequent study[22], the dose-response and pharmacokinetic effects of atropine were examined in the sedentary heat-stressed rat. Doses of atropine ranging from 10-4000 ug/kg were administered via tail vein 15 min prior to the start of heat stress. In the range of 25 to 1000 ug/kg atropine was found to increase heating rate and decrease % weight loss and fecal pellet production (anticholinergics decrease intestinal motility) in a dose-response manner. A 200 ug/kg dose (equivalent to a 2 mg dose in man[16]) resulted in a 38% decrease in water loss rate which was similar to the 42%[12] and 43%[15] seen in man. The pattern of saliva or urine spread was also found to have a dose-response relationship to atropine administration. Despite some individual variability, the patterns in Fig 1 were found to be consistent with the following: S-1, 200 ug/kg atropine; S-2, 50 ug/kg; S-3, 10 ug/kg or saline controls; U-1, 500 ug/kg; U-2 or U-3, surgically desalivated saline controls (from the previous study[10]). In this same

study, 250g rats were found to be less sensitive to dose-response effects of atropine administration than 500 g rats. Thus, the larger rats were used in all studies. This study also examined the effect of increasing the time interval between atropine administration and the start of heat stress. While intestinal motility returned to normal levels after a 60 min delay, and % weight loss returned to control levels after a 120 min delay, heating rate required a 210 min interval to return to control levels. These times correspond closely with the 2-4 hour time required for atropine effects to subside in man[26,27,28,29]. Based on the similarity of the rat and human responses to atropine, it was determined that the sedentary heat-stressed rat was a promising model with which to examine relative anticholinergic drug effects.

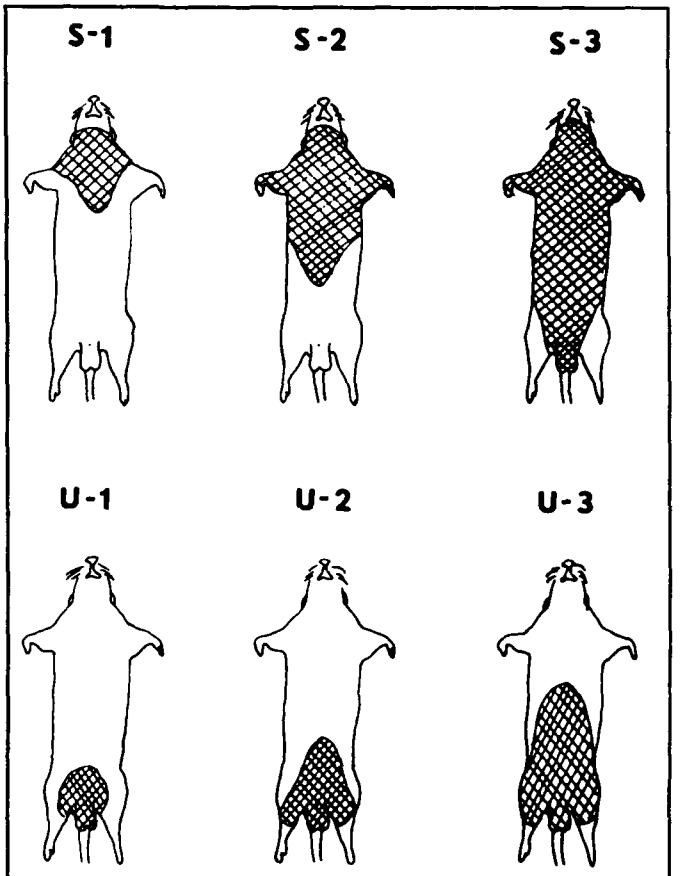


Fig. 1 Typical progressive saliva (S-1, S-2, and S-3) spread pattern of control rats and urine (U-1, U-2, and U-3) spread of atropinized or surgical desalivated rats. (Matthew *et al.*[21])

Since atropine is frequently administered intramuscularly (im) rather than intravenously (iv), a comparison of iv and im routes of administration in the heat-stressed rat was undertaken[30]. The effects of im or iv atropine were quantified on the following variables: heating rate, % weight loss, fecal loss (intestinal motility), and mydriasis (in rats that were not heat stressed.) Both iv and im atropine had similar effects on weight loss (decreasing with increasing atropine dose) and mydriasis (pupil dilation to a greater extent and longer duration with increasing dose.) The range of doses of atropine (10-1000 ug/kg) over which there was a dose-response increase in heating rate with iv administration was truncated with im administration (10-50 ug/kg). Therefore, we concluded that the iv route was the better route for examining anticholinergic-induced inhibition of evaporative cooling.

TABLE 2

ANTICHOLINERGIC POTENCY RELATIVE TO ATROPINE

Drug	Dose (ug/kg)	Heating Rate (°C/min)[a]	Potency Ratio[b]	AT=Dose (ug/kg)[c]
Saline[d]	----	0.022 ± 0.001	----	----
Atropine[d]	200	0.087 ± 0.004	1	200
Atropine Methyl Nitrate[d]	25	0.067 ± 0.006	4	50
Scopolamine[d]	10	0.079 ± 0.005	16	12
Chlorpromazine[d]	1000	0.067 ± 0.004	0.1	2000
L-Hyoscyamine[d]	100	0.089 ± 0.004	2	100
Aprophen[e]	2000	0.072 ± 0.004	0.067	3000
Trihexyphenidyl[e]	400	0.033 ± 0.004	0.061	3200

a Values are mean ± standard error.
b Calculated potency relative to atropine from Matthew, *et al.*[23]
c 200 ug/kg AT/potency ratio; drug dose required to elicit the same heating rate in the heat-stressed rat as 200 ug/kg of AT.
d Values taken from Matthew *et al.*[23]
e Values taken from Matthew[24].

Since the heating rate of sedentary heat-stressed rats was the most sensitive index of anticholinergic activity, the heating rates of animals given other anticholinergic drugs were compared to the dose-response increase in heating rate induced by atropine administration[23]. An atropine equivalent dose and an anticholinergic potency relative to a value of 1 for atropine were determined as indicated in Table 2. Additionally, the relative abilities of carbamates to reverse the atropine-induced increase in heating rate as a measure of relative anticholinesterase potency were compared. Neostigmine was determined to have a value of 8 relative to 2 for physostigmine and 1 for pyridostigmine[23]. Thus, the core temperature response of sedentary heat-stressed rats can be used as a sensitive bioassay for both anticholinergic and anticholinesterase efficacy.

Exercising Rats: In the sedentary heat-stressed rat, the cholinergic drug effects examined were primarily muscarinic (salivation, intestinal motility, vasodilation, etc.). However, the exercising rat is unable to spread secreted saliva while exercising; thus, nicotinic cholinergic effects (primarily at the neuromuscular junction) can be evaluated. Following are descriptions of cholinergic drug effects using an exercising rat model[24,25,32,33,34,38,39,40,41].

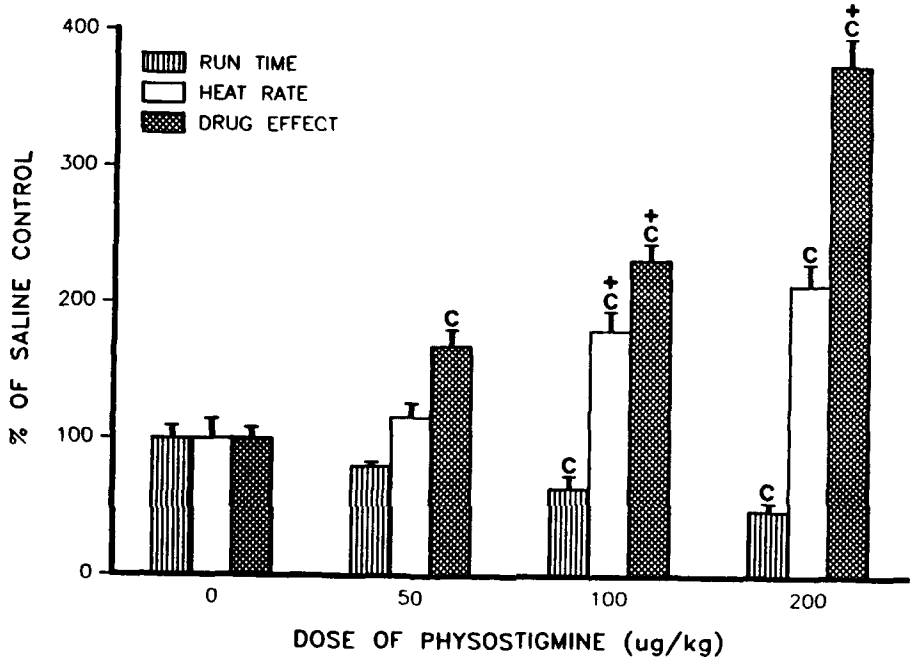


Fig. 2 Run time, heating rate, and cholinergic drug effect score for rats receiving saline (0), 50, 100, 200 ug/kg of PH are plotted as a % of saline control values. Significant difference from control value = "C", from previous lower dose = "+". (From Matthew, *et al.*[25]).

Physostigmine, (PH) a reversibly acting carbamate anticholinesterase, was initially chosen to establish dose-response drug effects in the running rat. PH was selected because it is a drug in clinical use, has central as well as peripheral sites of action[1], and has been shown to be protective against organophosphorus poisoning[31]. In the results illustrated in Fig. 2[25], PH was administered via tail vein 15 min prior to the start of run. Control animals had a run time of 67 $\pm$ 6 min (mean $\pm$ S.E.), a heating rate of 0.051 $\pm$ 0.007°C/min, and cholinergic symptom

score of 2.5 $\pm$ 0.2. Endurance declined as heating rate and cholinergic symptoms increased with increasing doses of PH. Since endurance, thermoregulatory effects, and cholinergic symptoms all vary in a dose-dependent manner with PH, this exercising rats model represents a potentially useful model with which to examine other anticholinesterases or anticholinergics.

In order to alleviate PH-induced thermoregulatory and endurance decrements in running rats, a combination of atropine (muscarinic anticholinergic) and diazepam (anticonvulsant) was administered with PH to running rats[32]. The administration of atropine (200 ug/kg) and diazepam (500 ug/kg) with PH (200 ug/kg) resulted in run times and heating rates which were not significantly different from those of saline controls. In a separate study[33], subchronic (2 week) administration of PH (60% inhibition of cholinesterase (ChE)) resulted in endurances and heating rates which were not significantly different from controls. In the same study[33] subchronic PH administration resulted in attenuation of ultrastructural changes commonly seen with acute PH administration. Accommodation to excess acetylcholine at receptors was noticeable after one week of PH administration via osmotic pump and complete after 2 weeks.

Subsequently, the sedentary heat-stressed rat model and the exercising rat model were used to determine optimal doses of 4 anticholinergic drugs to neutralize the PH-induced decrements in exercising rats[24]. The optimum dose of atropine (200 ug/kg) was determined by administering a range of doses of AT with PH prior to exercise. Doses of AT either higher or lower than 200 ug/kg elicited higher heating rates and cholinergic symptom scores and lower endurance times. Doses of scopolamine (S), aprophen (AP), and trihexyphenidyl (THP) equivalent to 200 ug/kg of AT **in the sedentary heat-stressed rat** were determined by the method[23] indicated above (Table 2) to be S- 12, AP- 3000, and THP- 3200 ug/kg. In the **exercising rat** model, the optimum dose of S (12 ug/kg) and AP (3000 ug/kg) as adjuncts to PH were determined to be the same doses that were equivalent **in the sedentary heat-stressed rat** to the dose of AT (200 ug/kg) found to be optimum in the **exercising rat**. AT, S, and AP are primarily muscarinic anticholinergic drugs. The optimal dose of THP (800 ug/kg) in the **exercising rat** is only 1/4 of the dose equivalent to the standard AT dosage in the **sedentary heat-stressed rat**. This could be explained because THP selectively blocks only M-1 (central, neuronal) cholinergic receptors, while AT, S, and AP all block both M-1 and M-2 (secretory, smooth muscle, and cardiac) receptors[1,34,35].

The peripheral vasodilatory effect[36,37] of PH and other centrally acting anticholinesterases may result in decreased core temperature[25,32,36,37]. In sedentary

animals this peripheral vasodilation results in a decrease in core temperature at or below an ambient temperature of 25°C[25,32,37] which should confer a thermoregulatory advantage. However, when PH-treated animals are exercised at 26°C, they have an increased heating rate[24,25,32,33]. The cause of this increased heating was postulated to be increased acetylcholine and metabolic activity at the neuromuscular junction. To elucidate the mechanism involved in PH-induced thermoregulatory decrements, rats were exercised after treatment with PH at ambient temperatures of 10, 15, 26, and 30°C[38]. Results indicated that 10°C may be low enough to allow sufficient radiant cooling in both saline and PH-treated groups, that 30°C is too high for effective cooling in either group, and that at 15 and 26°C the PH-treated animals had significantly increased rates of rise of core temperature. Thus, the PH-treated rat model is useful over a temperature range of 15-26°C.

The quaternary carbamate pyridostigmine (PY) was fielded and used by the military in Desert Storm as a nerve agent pretreatment drug. At low to moderate clinical doses PY does not cross the blood-brain barrier in pharmacologically significant amounts[1]; therefore, its effects can be attributed to peripheral sites of action. PY acutely administered to exercising rats (Tamb= 35°C, 40-60% ChE inhibition) decreased endurance and increased heating rates[39]. Chronic oral PY administration (20-40% ChE inhibition), also in rats exercising in a warm environment, did not decrement endurance or thermoregulation[40] indicating that, as was the case with PH, there may be accommodation with chronicity. Additionally, the endurance and thermoregulatory decrements induced by acute PY administration (40% ChE inhibition, 26°C) were attenuated by AT and diazepam administration[41].

CURRENT AND POTENTIAL APPLICATIONS

If PY is given prophylactically, it will be given subchronically which will result in consistently elevated acetylcholine levels at receptors. This chronic elevation may result in accommodation and eliminate adverse effects seen with acute administration as was the case with PH[33]. However, if there is an alteration in receptor sensitivity following subchronic PY, acute administration of atropine subsequent to toxic exposure may not result in anticipated levels of anticholinergic activity[23,42]. An evaluation of the effects of subchronic PY with acute AT administration is currently underway in our laboratory.

The acute administration of either PH or PY resulted in endurance and thermoregulatory decrements in exercising rats[32,39,41]. This is hypothesized to be due to increased metabolic activity at the neuromuscular junction[24] for the following reasons: 1) The decrements are peripheral in origin because both the PH and PY elicit the same decrements. 2) Preliminary results from this laboratory indicate that PH administration to sedentary rats results in an increased oxygen consumption despite the drop in core temperature subsequent to tail vasodilation. 3) PH- or PY-stimulated salivary secretion does not contribute to cooling while rats are running on the treadmill. Since none of the anticholinergic drugs used to block PH-induced decrements is a specific nicotinic blocker[24], conclusive evidence for this hypothesis awaits use of ganglionic and neuromuscular junction nicotinic blockers with PH in the exercising rat.

Because of the similarities between sweat loss dehydration in man and salivary loss dehydration in the rat, the sedentary heat-stressed rat model is currently being used to examine treatment of hyperthermic dehydration with hypertonic saline in dextran (HSD). This new use of the model increases its potential usefulness for examining problems in military medicine where adequate hydration is a chronic problem. With the existing data base on the exercising and sedentary heat-stressed rat models and current work to identify specific sites of anticholinesterase-induced decrements, these models are available for evaluating prospective new therapies as they are identified.

The views, opinions, and/or findings contained in this report are those of the author and should not be construed as official Department of the Army position, policy or decision, unless so designated by other official documentation. In conducting the research described in this report, the investigators adhered to the "Guide for the Care and Use of Laboratory Animals," as prepared by the committee on Care and Use of Laboratory Animals of the Institute of Laboratory Animals Resources, National Research Council. Citations of commercial organizations and trade names do not constitute an official Department of the Army endorsement or approval of the products or services of these organizations.

REFERENCES

1 A.G. Gilman, *et al.* Ed., *Goodman and Gilman's The Pharmacological Basis of Therapeutics 8th ed.* (Pergamon, New York, 1990), pp. 122-165.

2 S. Shibolet, M.C. Lancaster, and Y. Danon, *Aviat. Space Environ. Med.* **47**,280 (1976).

3 R.W. Hubbard, W.T. Matthew, J.D. Linduska, F.C. Curtis, W.D. Bowers, I. Leav, and M. Mager, *Am. J. Physiol.* **231**,1119 (1976).

4 R.W. Hubbard, W.D. Bowers, W.T. Matthew, F.C. Curtis, and R.E.L. Criss, G.M. Sheldon, and J.W. Ratteree, *J. Appl. Physiol.: Respirat. Environ. Exercise Physiol.* **42**,809 (1977).

5 R.W. Hubbard, W.T. Matthew, R.E.L. Criss, C. Kelly, I. Sils, M. Mager, and W.D. Bowers, and D. Wolfe, *J. Appl. Physiol.: Respirat. Environ. Exercise Physiol.* **45**,463 (1978).

6 R.W. Hubbard, *Medicine and Science in Sports.* **11**,66 (1979).

7 F.N. Craig, *J. Appl. Physiol.* **28**,779 (1970).

8 M.A. Kolka, L. Levine, B.S. Cadarette, P.B. Rock, M.N. Sawka, and K.B.Pandolf, *Aviat. Space Environ. Med.* **55** 1107 (1984).

9 F.R. Hainsworth, *Am. J. Physiol.* **212**,1288 (1967).

10 R.W. Hubbard, C.B. Matthew, and R.P. Francesconi, *J. Appl. Physioi.* **53**,1174 (1982).

11 M. Clubley, C.E. Bye, T. Henson, A.W. Peck, and C.A. Riddington, *Eur. J. Clin. Pharmacol.* **14**,221 (1978).

12 S. Robinson, *Chemical Corps Medical Laboratory Contract Report* **15**,1 (1953).

13 W.C. Randall and K.K. Kimura, *Pharmacol. Rev.* **7**,365 (1955).

14 H. Cullumbine and S. Miles, *Qtr. J. Exp. Physiol.* **41**,162 (1956).

15 A.C. Custance and C.A. DeCandole, *Defense Research Chem. Lab.* **427**,1 (1964).

16 E.J. Freireich, E.A. Gehan, D.P. Rall, L.H. Schmidt, H.E. Skipper, *Cancer Chemo. Rpts.* **50**,219 (1966).

17 S. Robinson and A.H. Robinson, *Physiol. Rev.* **34**,202 (1954).

18 C.A. Schneyer and L.H. Schneyer, *Proc. Soc. Exp. Bio.* **101**,568 (1959).

19 C.B. Wenger, in *Human Performance Physiology and Environmental Medicine at Terrestrial Extremes*, K.B. Pandolf, M.N. Sawka, and R.R. Gonzalez, Ed. (Benchmark, Indianapolis, 1988), pp.153-193.

20 C.C. Barney and D.R. West, *Physiol. Behav.* **48**,387 (1990).

21 C.S. Leithead, *Fed. Proc.* **22**,910 (1963).

22 C.B. Matthew, R.W. Hubbard, R.P. Francesconi, and P.C. Szlyk, *Aviat. Space Environ. Med.* **57**,659 (1986).

23 C.B. Matthew, R.W. Hubbard, and R.P. Francesconi, *Aviat. Space Environ. Med.* **57**,1061 (1986).

24 C.B. Matthew, *Neurosci. Biobehav. Rev.* **15**,141 (1990).

25 C.B. Matthew, R.P. Francesconi, and R.W. Hubbard, *Life Sci.* **50**,39 (1992).

26 J.S. Kerchum, F.R. Sidell, E.B. Crowell Jr., G.K. Aghajanian, and A.H. Hayes Jr., *Psychopharmacologia* **28**,121 (1973).

27 R. Virtanen, J. Kanto, E. Iisalo, E.U. Iisalo, M. Salo, and S. Sjovall, *Acta Pharmacol. Toxicol.* **47**,206 (1980).
28 J. Kanto, R. Virtanen, E. Iisalo, K. Maenpaa, and P. Liukko, *Acta Anesth. Scand.* **25**,85 (1981).
29 R.G. Adams, P. Verma, A.J. Jackson, and R.L. Miller, *J. Clin. Pharmacol.* **22**,477 (1982).
30 C.B. Matthew, G.J. Thomas, R.W. Hubbard, and R.P. Francesconi, *Aviat. Space Environ. Med.* **59**,367 (1988).
31 S.S. Deshpande, G.B. Viana, F.C. Kauffman, D.L. Rickett, and E.X. Albuquerque, *Fund. Appl. Toxicol.* **6**,566 (1986).
32 C.B. Matthew, R.W. Hubbard, R.P. Francesconi, and G.J. Thomas, *Aviat. Space Environ. Med.* **58**,1183 (1987).
33 C.B. Matthew, R.P. Francesconi, W.D. Bowers Jr., and R.W. Hubbard, *Life Sci.***47**,335 (1990).
34 R. Blumenson, G. Razoni, A. Shalev, and H. Munitz, *Drug Intell. Clin. Pharm.* **20**,786 (1986).
35 E.G. Giachetti, H. Ladinsky, and E. Montagna, *Br. J. Pharmacol.* **89**,83 (1986).
36 E. Meeter, *Arch. Int. Pharmacodyn. Ther.* **182**,416 (1969).
37 K.S. Fehliner and C.J. Gordon, *Neuropharmacology* **24**,993 (1985).
38 C.B. Matthew, *FASEB J.* **5**,A1401 (1991).
39 R.P. Francesconi, R.W. Hubbard, and M. Mager, *J. Appl. Physiol.* **56**,891 (1984).
40 R. Francesconi, R. Hubbard, C. Matthew, N. Leva, J. Young, and V. Pease, *Pharmacology Biochemistry & Behavior.* **25**,1071 (1986).
41 C.B. Matthew, R.W. Hubbard, R.P. Francesconi, and G.J. Thomas, *Life Sci.* **42**,1925 (1988).
42 J.R. Keeler, *Military Med.* **61**,430 (1990).

Nonlinear Optical Properties of Layered Composites

Joseph K. McDonald, Dr.
Weapons Sciences Directorate, AMSMI-RD-WS-CM
U.S. Army Missile Command, Redstone Arsenal, AL 35898-5248

Earl F. Pearson, Dr.
Department of Chemistry, University of Western Kentucky
Bowling Green, KY 42101

Kenneth LaiHing, Dr.
Department of Chemistry, Oakwood College, Huntsville, AL 35896

INTRODUCTION

At ordinary intensities, light interactions are independent of the intensity of the light. However, when high intensity laser sources are used, the interaction may depend upon the intensity of the light beam as well. Such effects are said to be "nonlinear" optical effects. Nonlinear optical effects may give rise to optical analogues of gates and switches used in digital computers and signal processors. If the appropriate nonlinear optical materials can be discovered[1], it could be possible to generate, transmit and process information in all-optical circuits.

Almost all materials show some degree of nonlinearity. However, symmetry may restrict the magnitude of some of the effects. For example, symmetry requires that second order nonlinear materials not contain a center of inversion. Third order nonlinear effects have no symmetry restrictions. Optical phase conjugation is a third order nonlinear effect whereby a time-reversed replica of an input wave is generated. The phase conjugate beam has some very unusual and useful properties. If an input beam is degraded by passing through a distorting medium and the phase conjugate is created in a nonlinear material, the phase conjugate will travel backwards toward the source (backwards in time for the input beam) and retrace its path. Upon returning to the original point of origin of the input beam, the phase conjugate will be near perfect, with all the distortions corrected[2-5]. This property of phase conjugate "mirrors" has

obvious applications such as pointing and tracking moving objects through a turbulent atmosphere containing smoke, fog, or other scattering aerosols.

The two common methods of generating an optical phase conjugate are stimulated Brillouin scattering and four-wave mixing in nonlinear materials. In degenerate four-wave mixing (DFMW), three laser beams of the same frequency are mixed in a nonlinear medium. The sample is illuminated by two counter-propagating pump beams with time-varying electric fields, given by E_f and E_b. The forward and backward pump beams are of equal intensity and co-linear. A weaker probe beam with electric field E_p is propagating in the same direction as the forward pump beam but intersects with the forward and backward pump beams in the sample at an angle of about 5 degrees. If the sample is nonlinear, a phase conjugate replica of the probe beam will be generated. This beam retraces the path of the probe beam, regardless of the angle between the probe beam and the forward pump beam. The intensity of the phase conjugate signal is determined by the nonlinearity of the sample and the intensity of the other three beams.

The nonlinear optical properties of metal colloids have been studied extensively and found to have large third order nonlinearities[6]. Theoretical calculations on nanoparticle layered colloid materials have predicted large enhancements in the third order effects[7,8]. To determine if these predictions could be obtained, a series of metal coated particles have been prepared in our lab and studied using DFWM.

NONLINEAR RESPONSE THEORY

In a Kerr type medium the dielectric function for the material is given as

$$\varepsilon = \chi^{(1)} + \chi^{(2)}E + \chi^{(3)}E^2 \ , \qquad (1)$$

where the χ's are the susceptibility tensors and E is the electric field at the sample. The tensor, $\chi^{(1)}$, gives rise to the ordinary absorption, dispersion, and refraction of the sample. The tensor, $\chi^{(2)}$, is responsible for three-wave mixing (second harmonic generation) and $\chi^{(3)}$ produces the four-wave mixing, (phase conjugate) and third harmonic generation. The dependence of the phase conjugate signal on the intensity of the beams which produce it can be explained from the polarization produced in the

sample by the input beams. The polarization, P, produced by an electric field may be written as the dielectric function times the electric field intensity

$$P = \varepsilon E \ . \tag{2}$$

Substituting for ε gives

$$P = \chi^{(1)}E \ + \ \chi^{(2)}E^2 \ + \ \chi^{(3)}E^3 \ . \tag{3}$$

The electric field of the phase conjugate signal is proportional to the third-order term, $\chi^{(3)}E^3$. Since the electric fields of each beam may be different, the phase conjugate electric field may be expressed as:

$$E_c \ = \ K\chi^{(3)}E_fE_bE_p \ , \tag{4}$$

where c, f, b, and p represent the electric fields of the phase conjugate, forward pump, back pump, and probe beams, respectively, and K is a proportionality constant. The intensity of a light wave is proportional to the square of the electric field. So the intensity of the phase conjugate is given as:

$$I_c \ = \ kK^2\chi^{(3)}\,2I_fI_bI_p \ . \tag{5}$$

The effective intensity of each of the two counter-propagating beams, I_f and I_b, is the geometric mean of the two, $\sqrt{I_fI_b}$. Due to the experimental arrangement (all beams split from the same input beam), the probe beam intensity will also be proportional to the effective intensity of the forward and backward pump beams.

$$I_p \ = \ k'\sqrt{I_fI_b} \ . \tag{6}$$

Thus we have:

$$I_c \ = \ k'kK^2(\chi^{(3)})^2(\sqrt{I_fI_b})^3 \ . \tag{7}$$

Taking the logarithm of both sides of equation (5) we have:

$$\log(I_c) \ = \ \log(k'kK^2(\chi^{(3)})^2) \ + \ 3\log(\sqrt{I_fI_b}) \ . \tag{8}$$

Thus a plot of the log of the conjugate signal intensity vs. the log of the geometric mean of the forward and backward pump beam intensities should be linear, with a slope of 3. This slope is expected for homogeneous samples or suspensions of small metallic spherical particles (colloids). Any slope other than 3 indicates a thermal or other effect contributing to the phase conjugate signal.

The above approach must be modified for particles composed of layers of material[7,8]. The local field, E_L, at the center of a spherical particle is given by

$$E_L = \frac{9\varepsilon_h\varepsilon_s}{(\varepsilon_c+2\varepsilon_s)(\varepsilon_s+2\varepsilon_h)+2(a_c/a_s)^3(\varepsilon_c-\varepsilon_s)(\varepsilon_s-\varepsilon_h)} E_0 . \qquad (9)$$

The ε's are the complex dielectric constants of the host(h), the core(c) and the shell(s) and E_0 is the electric field far from the particle. The a's are the radii of the core and shell. The local electric field, E_L, at the center of the particle may be substituted for E in equation (3) and the derivation continued as before to determine the intensity dependence of the phase conjugate signal. There are two important predictions implied by the derivation based on equation (7). First, if the dielectric constants are chosen so that the denominator is close to zero, a large enhancement of the phase conjugate would be observed. Also, a slope of 3 for a plot of the log of the phase conjugate signal intensity vs. the log of the laser intensity is no longer required. Instead, a higher order intensity power may be observed, since the signal is dependent upon a product of the dielectric functions. If $\chi^{(2)}$ is not allowed, then only odd powers could be observed. It is expected that a slope of 3 would be observed at low laser intensities, with higher slopes occurring as the laser intensity is increased.

EXPERIMENTAL RESULTS

Various coated particles which included metal on metal and metal on semiconductor cores were prepared in our lab. Using standard procedures[9-12], the metal shells and cores were prepared by reducing the metal salt. By varying the amount of starting materials, the thickness of the composite layer could be changed. Samples studied included gold on silver, gold on platinum, and gold on alumina.

In the DFWM experiments an excimer (XeCl, 308 nm) pumped dye laser was used. Most of the experiments were run at *480 nm*; however, the same results were observed when other wavelengths were used. The experimental setup is shown in Figure 1.

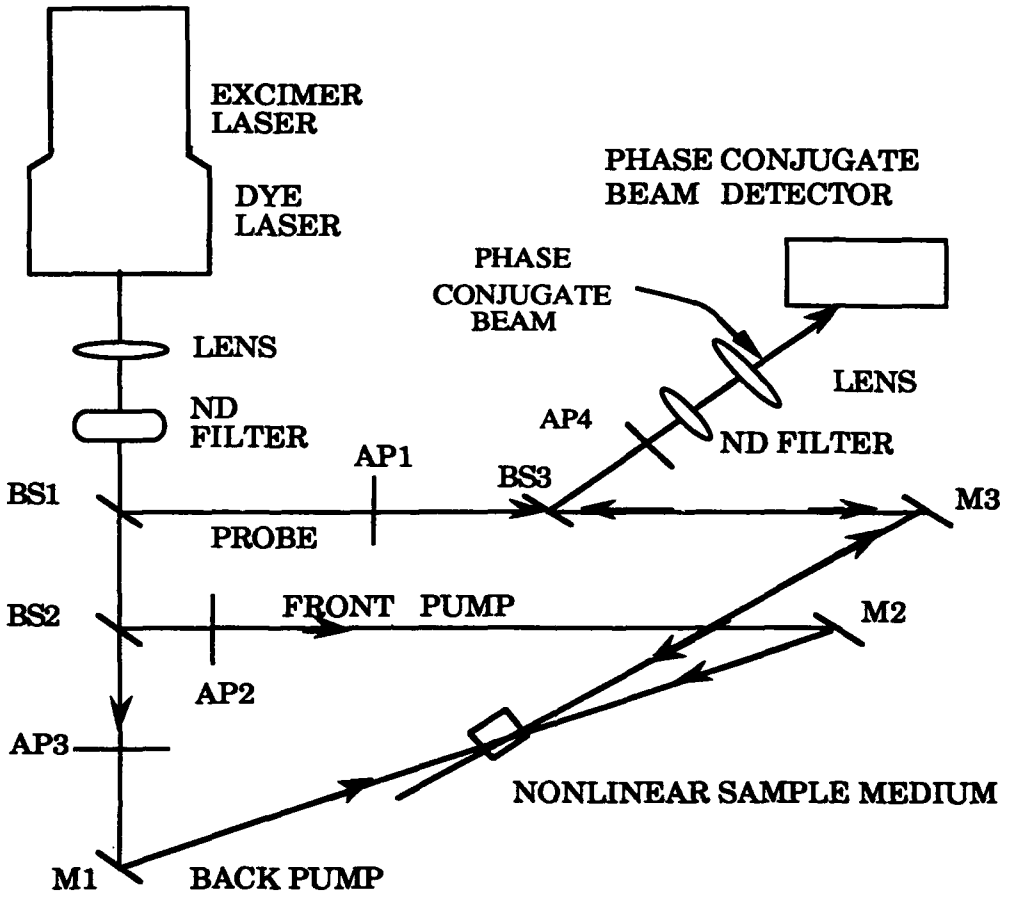


Figure 1. DFWM experimental setup

The phase conjugate signals measured for the first samples showed only a normal response, with no enhancement of the signal when compared to pure metal colloids. When samples of thinly gold coated

silver and platinum were measured, an anomalous phase conjugate was observed. In the log-log plots the slope obtained was 4, instead of the expected 3 (see Figure 2). Because of the higher intensity dependence, an order of magnitude enhancement of the phase conjugate signal was observed. A thinly gold coated alumina also produced the same anomalous results (see Figure 3). Transmission electron microscopy of the gold on silver sample showed that the gold coat was *1.5-2.0 nm* thick for the anomalous sample. The sample that showed a normal phase conjugate response was approximately twice this thickness. The core diameters were not a determining factor in the anomalous signals observed since they varied from *5 to 20 nm* for the different samples.

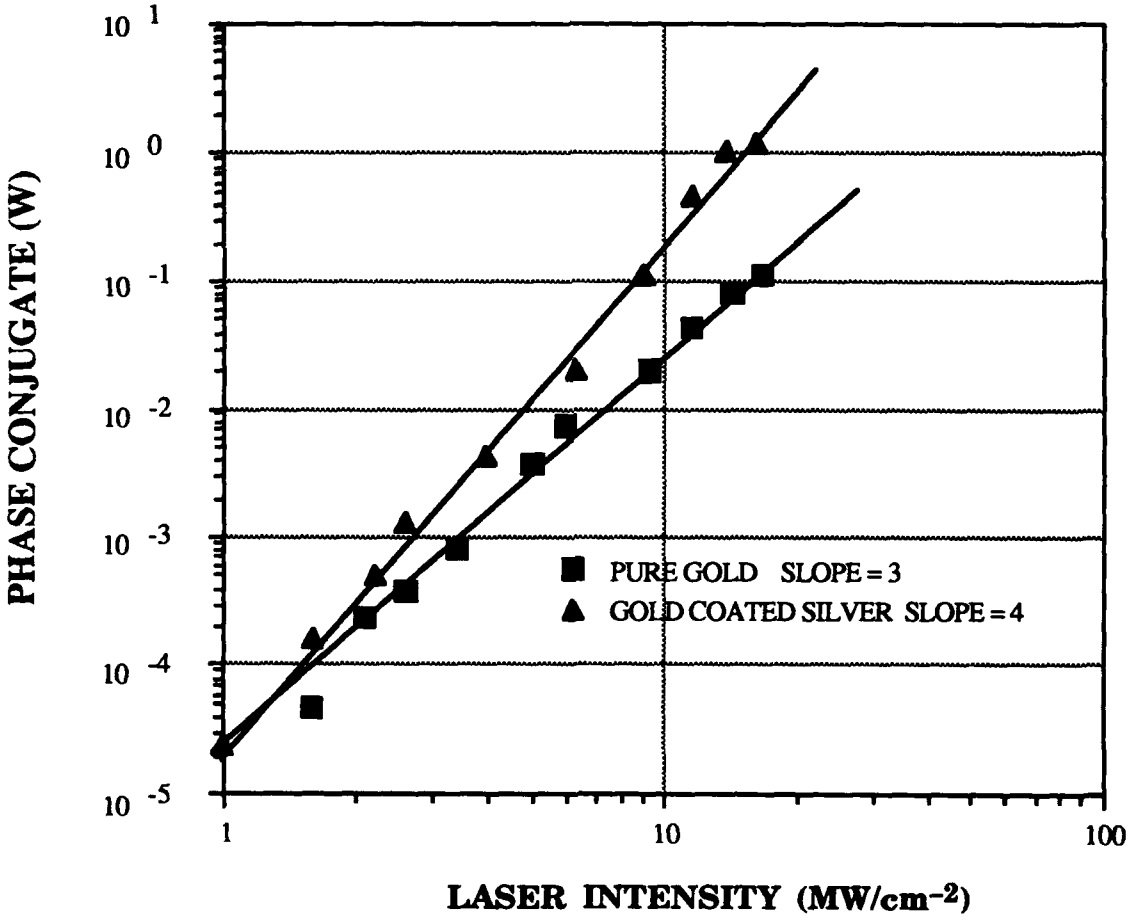


Figure 2. DFWM of pure gold and gold–coated silver colloids

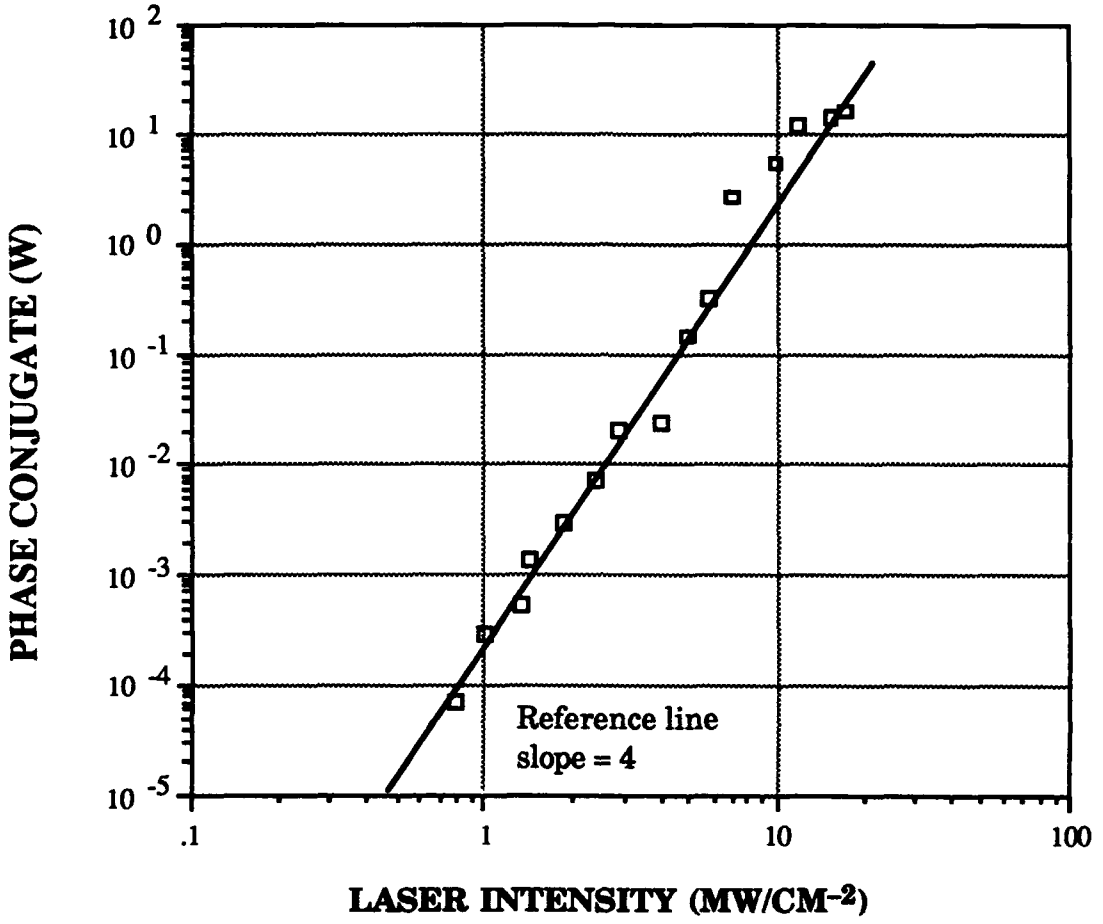


Figure 3. DFWM of gold on alumina

Since only odd power dependence on the laser intensity for the phase conjugate was expected, we thought it possible that the slope 4 was just the transitional stage from slope 3 to 5. To determine if this was the case, higher laser intensities were used by adding a dye amplifier stage. Changes in the slope were observed but they appeared to be shifts instead of a change to slope 5 (see Figure 4). Color changes were also noted after irradiation at the higher laser intensities (above *20 MW/cm²*). The higher laser intensities produced larger thermal effects which could be observed as a rapid decrease in the phase conjugate signal after the initial irradiation. This decrease in signal could be explained as a decrease in particle concentration due to thermal diffusion. To eliminate this effect

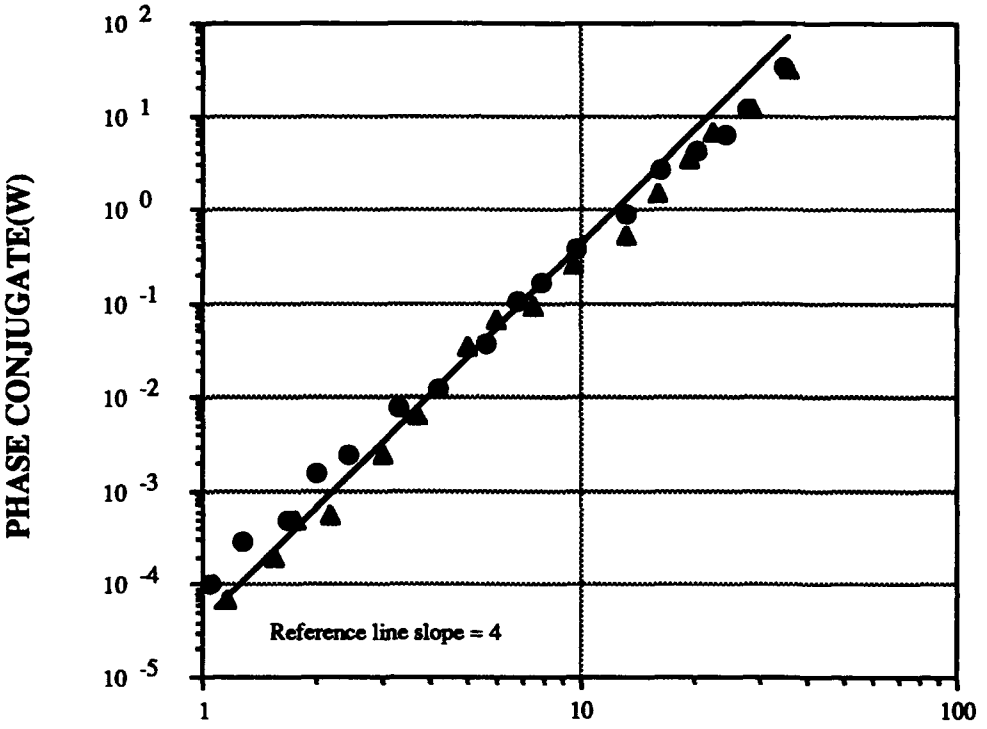


Figure 4. DFWM of gold on platinum

the solutions were stirred during the experiments.

While stirring the solution, the slopes remained the same at lower laser intensities, as had been previously observed; however, in most of the samples above *15 MW/cm²* (see Figure 5), a discontinuation was observed. At the higher laser intensities, an increase in phase conjugate signal was observed and, as in the case shown, by as much as two orders of magnitude. The discontinuity occurred at approximately the same intensity regardless of the shell or core material.

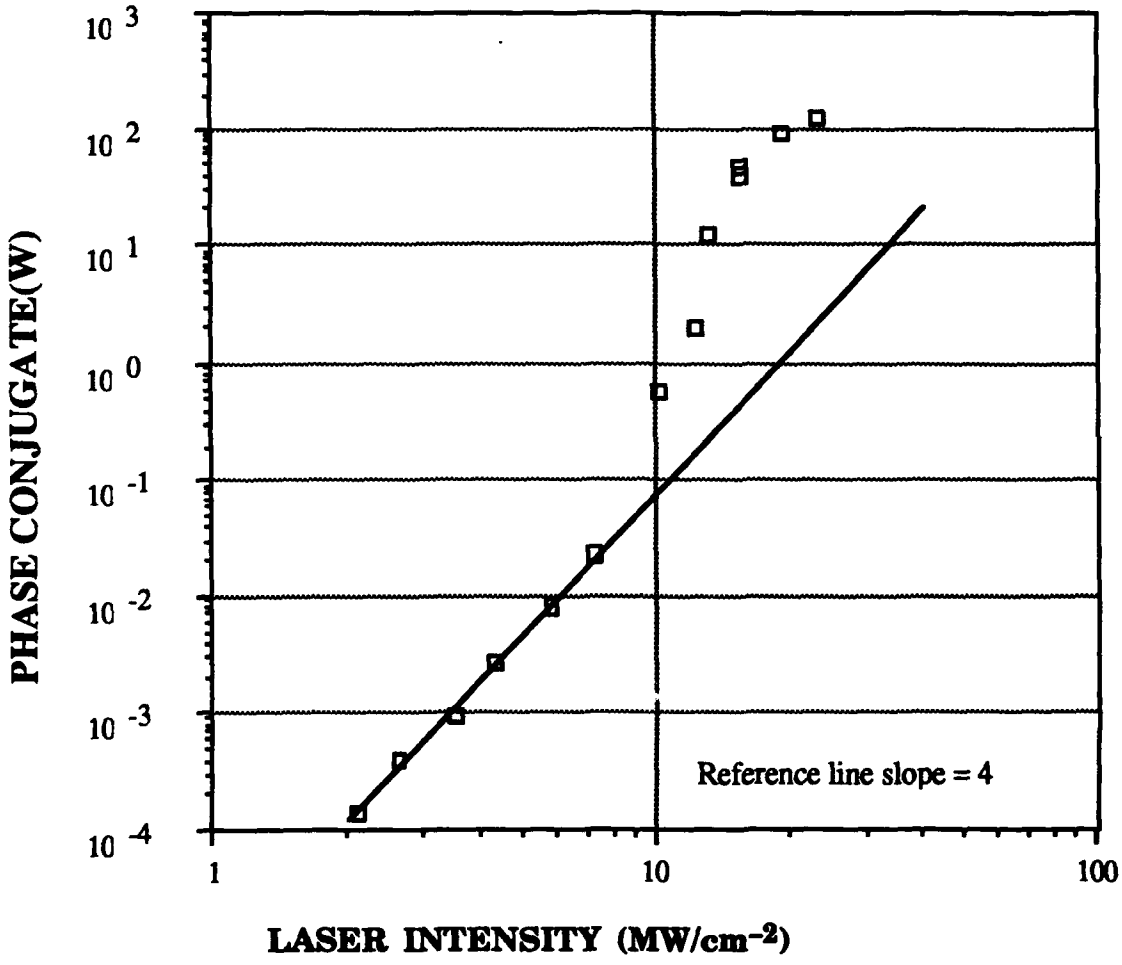


Figure 5. DFWM of gold on alumina while being stirred

It is believed that laser ablation of the shell material is the primary cause of this sudden increase in phase conjugate signal. Laser ablation of the shell also explains the color changes observed with the gold coated composites. The absorption spectra of these samples before and after irradiation showed that the mixture became more homogeneous with a narrowing of the bandwidth, and smaller with a shift in the maximum to higher energy (see Figure 6). It has not been determined if the gold is removed permanently from the composite or if the surface becomes more uniform. The samples still showed the anomalous slope and there was no indication of any slope 5 dependence.

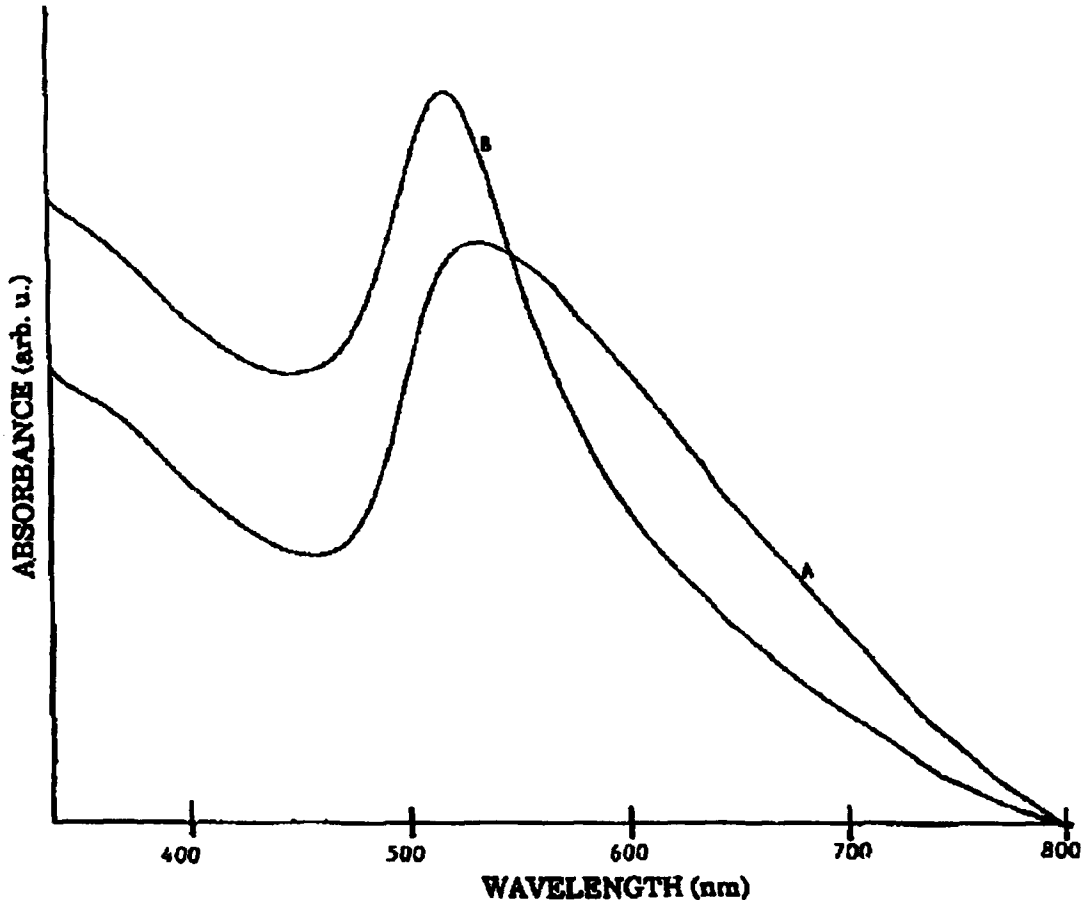


Figure 6. Gold–coated alumina before (A) and after (B) DFWM

Coated composites obtained from other labs also showed the same anomalous phase conjugate signal if the shell was less than *3 nm* thick. These samples included silver on silica and silver sulfide on silver.

CONCLUSIONS

An anomalous fourth power intensity dependence in the phase conjugate signal has been observed in certain gold and silver coated microparticles. An enhancement of up to one hundred times the phase conjugate signal in an uncoated particle is observed. The results were unaffected by the type of particle core. The anomalous signals were seen only with thinly coated particles ($< 3\ nm$). If $\chi^{(2)}$ effects are ignored, current coated particle theory does not adequately explain the anomalous results observed. Optical second-harmonic generation are known to occur at surface interfaces[13]. Because of the two interfaces involved with the layered composites, it is probable that $\chi^{(2)}$ cannot be eliminated from the dielectric function. Theoretical calculations are currently being conducted to determine possible explanations for the anomalous signals and to reproduce the experimental results.

REFERENCES

1. B. I. Greene, J. Orenstein, S. Schmitt-Rink, Science **247**, 679 (1990).
2. R. W. Hellwarth, J. Opt. Soc. Am. **67**, 1 (1977).
3. D. M. Bloom, G.C. Bjorklund, Appl. Phys. Lett. **31**, 592 (1977).
4. C. R. Guiliano, Phys. Today **34**, 27 (1981).
5. A. Yariv, D. Fekete, D.M. Pepper, Opt. Lett. **4**, 52 (1979).
6. F. Hache, D. Ricard, C. Flytzanis, and U. Kreibig, Appl. Phys. A **47**, 347 (1988); F. Hache, D. Ricard, and C. Flytzanis, J. Opt. Soc. Am. B **3**, 1647 (1986); D. Ricard, P. Roussignol, and C. Flytzanis, Opt. Lett. **10**, 511 (1985).
7. A. E. Neeves, and M.H. Birnboim, J. Opt. Soc. Am. B. **6**(4), 787 (1989).
8. N. Kalyaniwalla, J. W. Haus, R. Inguva, and M. H. Birnboim, Phys. Rev. A **42**(9), 5613 (1990).
9. R. M. Wilenzick, D.C. Russell, R.H. Morriss, and S.W. Marshall, J. Chem. Phys. **47**, 533 (1967).
10. G. Frens and J.T.G. Overbeck, Kolloid- Z. und Z. fur Poly. **233**, 922 (1969).
11. J. Turkevich, P.C. Stevenson, and J. Hiller, Discuss. Faraday Soc. **11**, 55 (1951).
12. J. A. A. J. Perenboom, P. Wyder, and F. Meier, Phys. Repts. **78**, 173 (1981).
13. Y. R. Shen, Nature **337**, 519 (1989).

TITLE: Magnetron Ion Etching for GaAs Device Processing

*Dr. G.F. McLane, Ms. M.W. Cole, Dr. H.S. Lee,
Dr. A. Lepore, Mr. D.W. Eckart and Dr. R. T. Lareau
U.S. Army Electronics Technology and Devices Laboratory
LABCOM, Ft. Monmouth, NJ 07703-5601
Dr. M. Meyyappan
Scientific Research Associates, Inc., Glastonbury, CT 06033-6058
Mr. M. Namaroff and Dr. J. Sasserath
Materials Research Corporation, Orangeburg, NY 10962

INTRODUCTION

For many microelectronic device applications plasma assisted (dry) etching has replaced wet chemical etching as the preferred method for providing fine line submicron patterns and selective etch fabrication steps. For device fabrication it is necessary that the plasma etching process provide smooth etched surfaces with minimal damage and degradation of surface region electrical properties, along with anisotropic etching characteristics for the desired etched sidewall angle. Etch-induced damage may be either in the form of etch gas surface residues or subsurface crystal defects.

Reactive ion etching (RIE) is a plasma processing technique which has been used for etching silicon and various III-V compound semiconductors. In this process, reactive species are created in the plasma and combine with the wafer surface atoms to form volatile chemical species which can then be removed. Positive ions which are also generated in the plasma impinge anistropically upon the negatively biased wafer mounted on the cathode and provide energy to promote gas-surface chemical reactions and help remove surface etch products. This anisotropic characteristic of RIE results in faithful transferral of the mask pattern to the semiconductor wafer, with good fidelity of the pattern being reproduced throughout the etch. RIE eliminates mask undercut, which is a major disadvantage of wet etching. However, ion bombardment of the wafer surface may lead to etch-induced damage which can be detrimental to device operation.

Magnetron reactive ion etching (MIE) is similar to RIE, with the addition of a magnetic field to confine the plasma electrons close to the wafer. This minimizes electron loss to the chamber walls and increases ionization efficiency and reactive species generation. Compared to RIE, the magnetic field of MIE allows a higher density plasma discharge to be created which is able to sustain itself at relatively low cathode bias voltages, resulting in low ion bombardment energies and therefore minimal etch-induced wafer damage. In addition, the relatively high

density discharge of MIE allows it to operate at lower etch gas pressure than RIE. This leads to anisotropic etching with smaller feature size capability since lower pressures provide reduced ion scattering and associated lateral etching, and allow reactive species and wafer etch products to be transported more easily to and from high aspect ratio, small dimension etched wells.

MIE is a relatively new technique and although it is a modification of RIE with a magnetic field added, the knowledge base previously developed for RIE is not directly transferable since the extent of ionization, electron density, and energy distribution is altered considerably by the magnetic field. Hence, there is a strong incentive to investigate the plasma reaction mechanisms and wafer sample etching characterisatics of the MIE technique. While MIE has been investigated for etching Si and SiO_2, there has been relatively little work done on exploiting this technique for GaAs processing. Since the chlorides of Ga and As are relatively volatile, chlorine based chemistry is the choice for etching GaAs. We report here the results of a study of MIE using silicon tetrachloride ($SiCl_4$) and Freon-12 (CCl_2F_2) etch gases. Although some work has previously been reported for CCl_2F_2[1], this is the first report of magnetron ion etching of GaAs using $SiCl_4$.

We have investigated MIE to determine its effectivness as a processing technique for the fabrication of GaAs devices. Etched wafers are characterized in terms of surface smoothness, sidewall angle, etch product surface residues, subsurface defects, and surface region electrical properties, since these are all important factors in terms of device performance. Wherever possible, we compare our MIE results with results reported by others for RIE.

EXPERIMENTAL

The etch experiments were performed in a magnetron reactive ion etch chamber having a powered cathode (13.56 MHz) containing two bar magnets, with two additional magnets located outside the chamber at the top of the chamber cover. The rf power source is capacitively coupled to the cathode resulting in a negative dc voltage bias being developed on the cathode during etching. The cathode is water cooled and covered with an aluminum pedestal which supports the wafer sample. Wafers were thermally heat sunk to the cathode with high vacuum grease. The magnets provide a uniform magnetic field at the wafer location. The chamber is typically evacuated to base pressures in the range 10^{-6} Torr.

Etch depth of patterned wafers was measured using a Dektak profilometer after removal of the mask. Etched surface characterization was performed using scanning electron microscope (SEM), transmission electron microscopy (TEM), Auger electron spectroscopy (AES), and Schottky barrier diode I-V measurements. GaAs wafer samples with (100) surface orientation and $1 \times 10^{17}/\text{cm}^3$ n-type carrier concentrations were used for Schottky barrier diode measurements. Semi-insulating GaAs wafers with the same surface orientation were used for all other measurements.

RESULTS

a) $SiCl_4$ etch gas

Figure 1 shows the etch rate and cathode self-bias voltage as a function of power density for 2 mTorr pressure and 15 standard cubic cm/min (sccm) $SiCl_4$ flow rate. The etch rate initially increases with power and then levels off at higher power densities, eventually reaching an etch rate of approximately 1 micron/minute. An increase in power density usually results in an increase in electron density and consequently in Cl atom production, which contributes to the observed increase in etch rate at low powers. Additionally, the cathode self-bias voltage increases linearly with power as shown in Fig. 1, resulting in increased ion bombardment energies. Increased ion bombardment energies can result in higher etch rates by increasing sputter removal of etch products or by providing additional activation energy necessary for the formation of volatile chlorides of Ga and As in the first place. At the higher power levels of Fig. 1, however, the increased ion bombardment may interfere with Cl atom adsorption on the surface as pointed out by Pearton, et al [2], and prevent the etch rate from increasing further. Figure 2 shows etch rate as a function of pressure for 0.16 watts/cm^2 power density and 15 sccm $SiCl_4$ flow rate The observed increase in etch rate with pressure is due to an increase in reactant species production resulting from an increase in density of available etch gas. The etch rates reported here are generally higher by a factor of five or more when compared with results reported on $SiCl_4$ RIE at comparable power densities and self bias levels.[3,4] High etch rates allow short processing times in the hostile plasma environment; since the eventual number of defects in a dry etched sample is a balance between defect generation due to ion bombardment and the removal of wafer material and surface defects due to etching, high etch rates may reduce damage accumulation. [5]

Figure 3 shows a scanning electron micrograph of a patterned GaAs sample etched for 2 minutes at 2 mTorr pressure and 0.24 watts/cm^2 power density with 15 sccm $SiCl_4$ flow rate. The 200 nm wide line masks were fabricated by direct electron beam writing on SAL-601, a novolac-based electron beam resist which was used directly as the etch mask. The etched surface is clean and smooth and the sidewalls are vertical, with an etch depth of approximately 2 microns. Striations on the sidewalls are due to non-uniformities of the mask edge being replicated into the sample. Figure 4 shows a scanning electron micrograph of the same sample at higher magnification, showing more clearly the quantum wire-like etched structure. Figure 5 shows another patterned sample etched under the same conditions in $SiCl_4$. These SEM results show that $SiCl_4$ etching of patterned GaAs wafers results in a surface morphology which is appropriate for fabrication of small dimension, high speed devices, in addition to quantum wire and quantum dot structures.

Etched surface profile and subsurface lattice defects were investigated using TEM. Figure 6 shows a TEM micrograph of a sample etched under the same conditions as above.The surface is extremely smooth, with a peak to valley distance along the surface of only 4 nm.

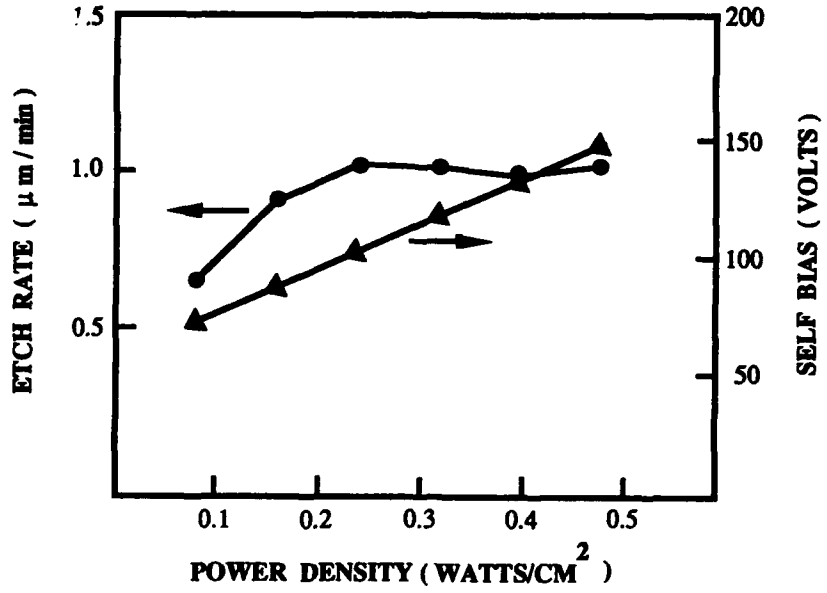


Figure 1. GaAs etch rate and cathode bias voltage as a function of power density for 2 mTorr pressure and 15 sccm $SiCl_4$ flow rate.

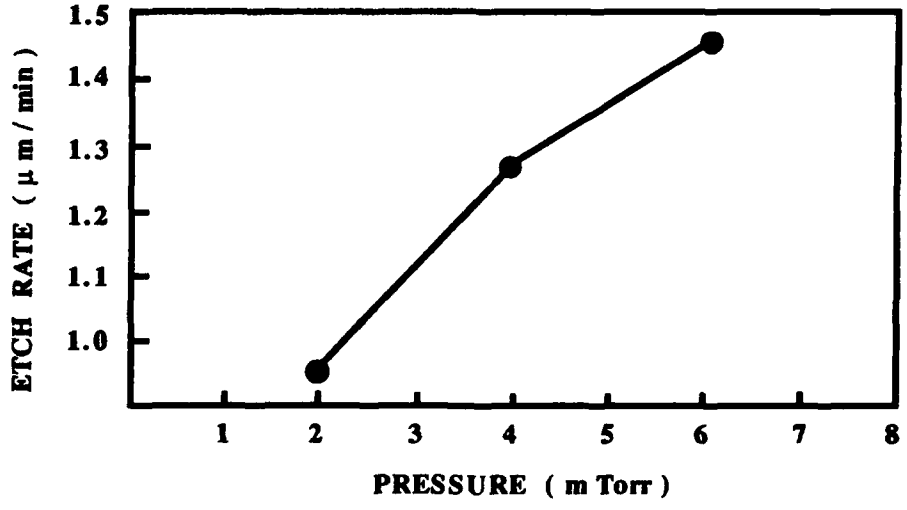


Figure 2. GaAs etch rate as a function of pressure for 0.16 watts/cm^2 power density and 15 sccm $SiCl_4$ flow rate.

Figure 3. Scanning electron micrograph of a patterned GaAs sample etched for 2 minutes at 2 mTorr pressure and 0.24 watts/cm^2 power density with 15 sccm $SiCl_4$ flow rate.

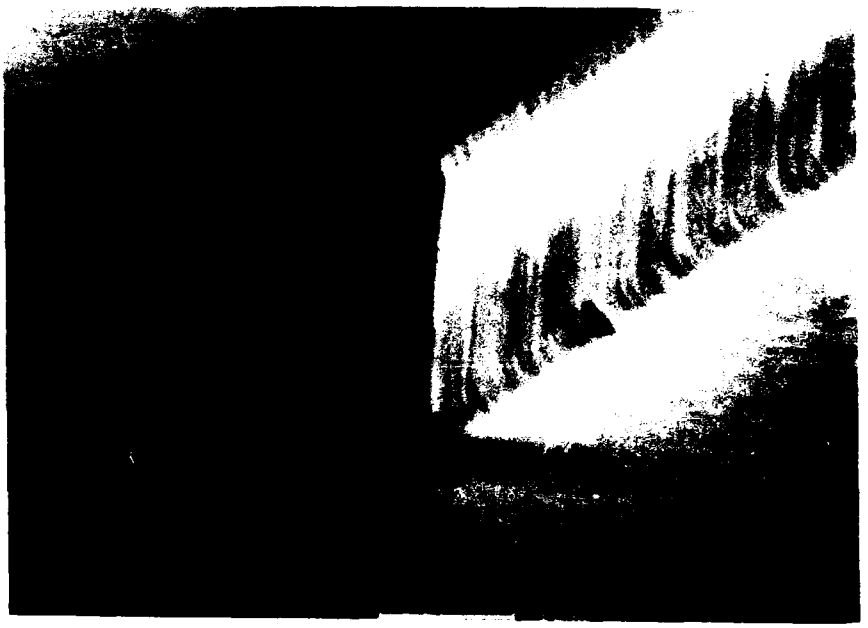


Figure 4. High magnification scanning electron micrograph of Figure 3 .

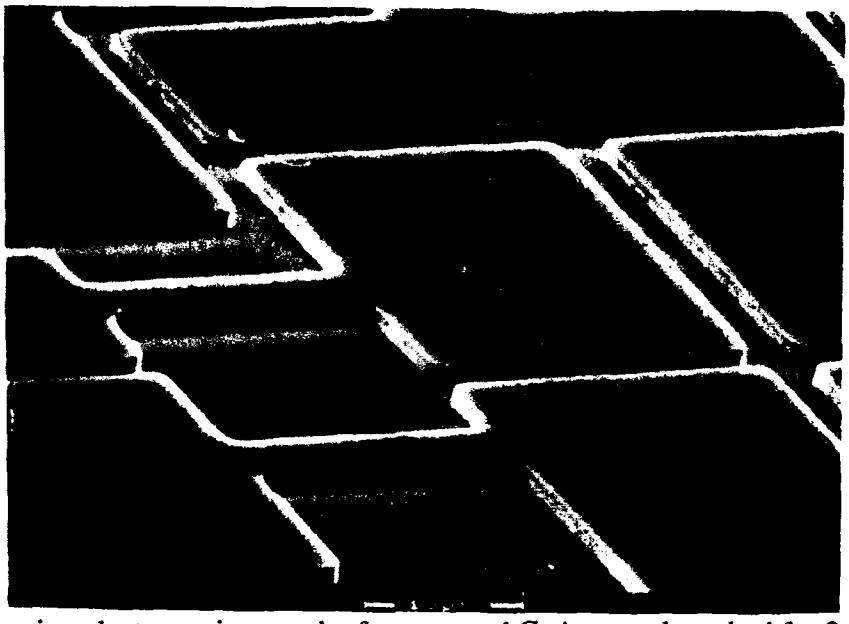


Figure 5. Scanning electron micrograph of a patterned GaAs sample etched for 2 minutes at 2 mTorr pressure and 0.24 watts/cm^2 power density with 15 sccm $SiCl_4$ flow rate.

Figure 6. Transmission electron micrograph of GaAs etched for 2 minutes at 2 mTorr , 0.24 watts/cm^2 , 15 sccm $SiCl_4$.

The defects are in the form of small dislocation loops 0.7 to 2.0 nm in diameter, extending a depth of 300 nm below the surface, with a density of $10^{10}/cm^2$. This is an order of magnitude lower than defect densities reported for RIE under similar conditions. [6]

To determine the effect of etch-induced damage on the surface region electrical characteristics, Schottky barrier diodes were fabricated on unpatterned etched surfaces and diode I-V characteristics are compared with those of a Schottky diode formed on a control (unetched) sample. Schottky contacts were formed by electron beam deposition of a Ti/Pt/Au structure. Schottky diode measurements for GaAs samples etched in $SiCl_4$ are shown in Table 1 for several values of pressure and applied power density along with results for a control sample. Ideality factor (n) and Schottky barrier height ($\emptyset_B$) are obtained from the slope and intercept, respectively, of the measured forward bias ln I vs V curve.[7] Ideality factor is normally close to unity, whereas Schottky barrier height can depend upon semiconductor material, semiconductor surface states, and metal contact material. These parameters are of particular interest because they are sensitive to Schottky barrier surface region carrier concentration, defect levels and traps. Etch induced surface region defects which can affect device performance are thus manifest in changes in the values of n and $\emptyset_B$. Minimal changes in these parameters indicate low etch induced wafer damage. Table 1 shows that n increases and $\emptyset_B$ decreases for the etched samples relative to the control sample, but most important is the fact that the deviations from the control sample are small. This indicates minimal etch-induced degradation of surface region electrical properties. Schottky diode measurements performed on GaAs samples etched under similar conditions using RIE also showed an increase in n and a decrease in $\emptyset_B$, but the increases in n were much larger than the results reported here.[2,8,9] This indicates that MIE results in less surface degradation than RIE.

Auger electron spectroscopy measurements were performed on $SiCl_4$ etched samples to determine the possible presence of etch gas residues. Figure 7 shows AES measurements on a horizontal etched surface. The surface is essentially residue free, with only the constituent Ga and As atoms being observed. These residue free etched surfaces make magnetron reactive ion etching of GaAs with $SiCl_4$ an attractive etching process for device fabrication. In contrast, Stern, et al[4], found several monolayers of chlorinated or oxidized Si present on GaAs layers reactively ion etched in $SiCl_4$ Figure 8 shows Auger measurements performed on the vertical sidewalls of the magnetron ion etched sample of Figure 7. This data indicates the presence of Si, C and O on the sidewalls, which could be responsible for the excellent anisotropic etching properties of $SiCl_4$ since this residue may prevent lateral chemical etching of sidewalls.

Power Density (watts / cm 2)	Pressure (m Torr)	Ideality Factor, n	Schottky Barrier height, Φ_B (eV)
Control	—	1.07	0.786
0.16	2	1.12	0.766
0.16	4	1.13	0.775
0.16	6	1.17	0.764
0.24	2	1.11	0.778
0.40	2	1.16	0.769

Table 1. Schottky diode measurements for GaAs samples etched for 2 minutes at 15 sccm $SiCl_4$ flow rate.

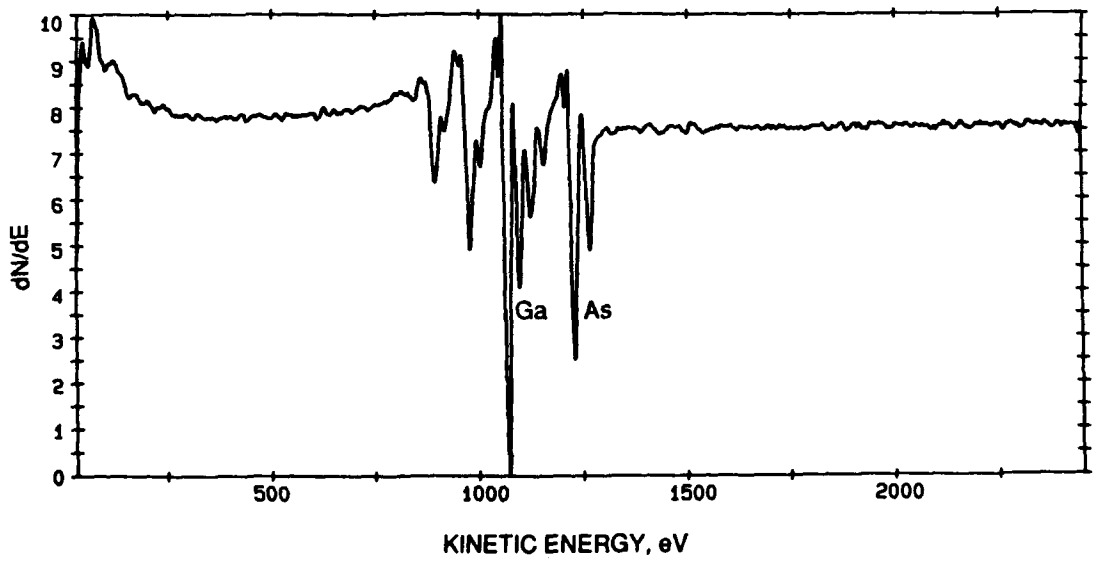


Figure 7. Auger electron spectroscopy measurements on GaAs horizontal surface etched for 2 minutes at 2 mTorr, 0.24 watts/cm^2, 15 sccm

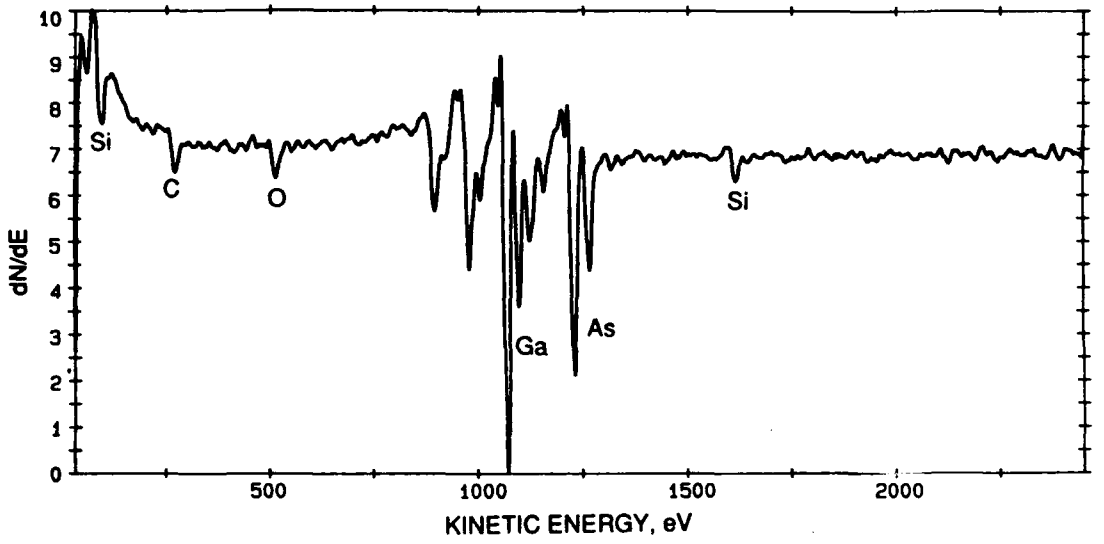


Figure 8. Auger electron spectroscopy measurement on GaAs vertical sidewall etched for 2 minutes at 2 mTorr, 0.24 watts/cm^2, 15 sccm.

b) CCl_2F_2 etch gas

Figure 9 shows the etch rate as a function of power density and pressure for CCl_2F_2 with a flow rate of 20 sccm. The etch rate increases with an increase in applied power density and a decrease in pressure. An increase in power density results in an increase in Cl atom production similar to $SiCl_4$ etching, leading to increased etch rates. Unlike $SiCl_4$, however, the etch rate for CCl_2F_2 increases as the pressure is lowered. In the case of pure chemical etching, a decrease in pressure would result in an etch rate decrease, since the production of active chemical species diminishes due to less etch gas being available.

In Figure 9 at 0.4 watts/cm^2 , cathode bias voltage increased from 95 volts to 105 volts as pressure was lowered from 20 mTorr to 10 mTorr. This results in an increase in ion bombardment energies. Under these conditions, it appears that as pressure is lowered the increase in the ion bombardment component of etching is a more dominant etch rate mechanism than a possible decrease in the available reactant species. The net result is an increase in GaAs etch rate with CCl_2F_2 as pressure is lowered. We have reported previously[10] that x-ray photoelectron spectroscopy (XPS) measurements reveal the presence of GaF_3 and other non-stoichiometric (Ga_xF_y) gallium fluorides on the surfaces of samples etched with CCl_2F_2.

Since gallium fluorides are relatively non-volatile, these etch product species may require enhanced ion bombardment for their removal. Indeed, Klinger and Greene [11] have reported that the removal of Ga in discharges containing fluorine occurs primarily by ion-assisted desorption of GaF_x radicals where x=1 or 2. The etch rates reported here for MIE are generally much higher than reported for RIE under similar conditions.[12]

Figure 10 shows a scanning electron micrograph of a pattern etched for 30 minutes in 20 sccm of CCl_2F_2 at 4 mTorr and a power density of 0.52 watts/cm^2. The etch mask consisted of an array of nickel/titanium layers in the shape of squares 10 microns on a side. The etch depth is approximately 5 microns. The etched surface is clean and very smooth. The sidewalls are straight and exhibit negative undercutting, i.e., they slope away from the mask edge. This negative undercut profile has been previously observed in RIE of GaAs with CCl_2F_2,[11] and could be due to the redeposition of surface material or formation of polymers on the sidewall as the sample is etched. Negative undercut characteristics are desirable for the fabrication of via hole connections for MIMIC applications, since they facilitate the subsequent sidewall metallization step. For via hole connection applications, etch rates can be increased by using higher power densities and through the addition of Cl_2 to the etch gas.

Figure 11 shows a TEM micrograph for an unpatterned GaAs sample etched for 4 minutes under the conditions of Figure 10. The surface morphology is extremely good, with peak to valley distances along the surface of only about 9 nm. The defects are due to small dislocation loops 2-4 nm in diameter which extend to a depth of about 70 nm below the etched surface. Most of the defects are concentrated in the top 10 nm where the density is about $10^{10}/cm^2$; total defect count is lower than for RIE under similar etching conditions.[13] Also, GaAs samples etched in CCl_2F_2 using the RIE technique under similar conditions exhibit rough surfaces[13] with peak to valley distances along the surface of 170-200 nm, more than an order of magnitude rougher than for MIE. The defects extend to a distance of 220 nm into the sample, much deeper than for MIE. The low damage in magnetron ion etching is due to the reduced cathode dc bias. For example, the dc bias voltage for MIE under the etching conditions of Figure 10 is 148 volts compared to 380 volts reported for RIE under similar conditions.[13]

Schottky diode measurements for GaAs samples etched in CCl_2F_2 are shown in Table 2 for several values of pressure along with results for a control sample. Both the ideality factor and Schottky barrier height increased slightly for the etched samples compared to the control sample, indicating minimal etch-induced degradation of surface region electrical properties. For RIE etched samples, n increases and $\emptyset_B$ decreases under etching, but the increase in n is somewhat higher than the results reported here for MIE.[14]

The increase in $\emptyset_B$ for the CCl_2F_2 etched samples reported here is atypical for etched GaAs samples, since $\emptyset_B$ normally decreases upon etching. An explanation for this result is

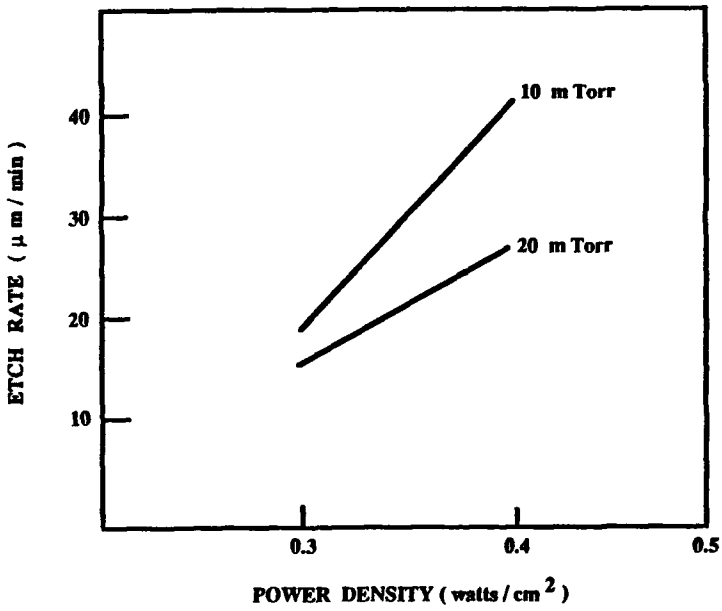


Figure 9. GaAs etch rate as a function of power density and pressure for 20 sccm CCl_2F_2 flow rate.

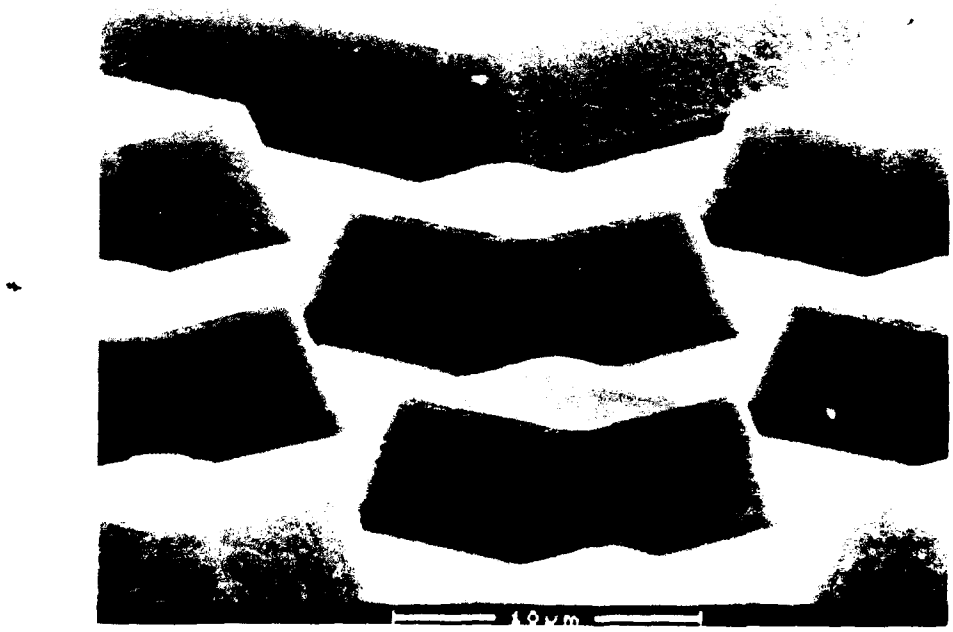


Figure 10. Scanning electron micrograph of a patterned GaAs sample etched for 30 minutes at 4 mTorr ,0.52watts/cm^2 , 20 sccm CCl_2F_2.

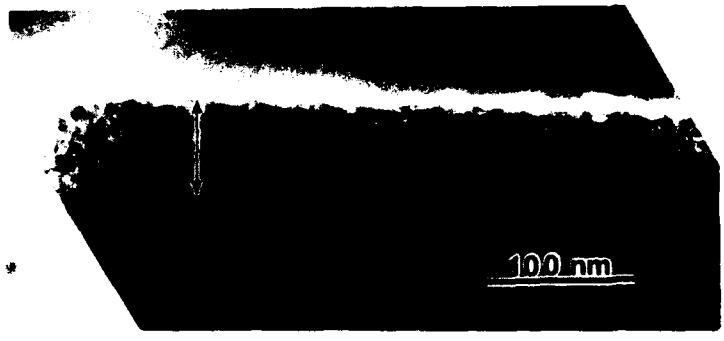


Figure 11. Transmission electron micrograph of GaAs sample etched for 4 minutes at 4 mTorr, 0.52 watts/cm^2, 20 sccm CCl_2F_2.

Pressure (m Torr)	Ideality Factor, n	Schottky Barrier height, Φ_B (eV)
CONTROL	1.12	0.747
2	1.12	0.816
4	1.14	0.792
10	1.12	0.772

Table 2. Schottky diode measurements for GaAs samples etched for 4 minutes, 0.31 watts/cm^2, 20 sccm CCl_2F_2.

obtained from Auger electron spectroscopy measurements performed on a CCl_2F_2 etched surface as shown in Figure 12. Carbon, fluorine, chlorine, and oxygen are seen on the surface in addition to gallium and arsenic. These residual etch gas contaminants formed on the surface could effectively be producing a metal-insulating layer-semiconductor structure instead of the intended metal-semiconductor Schottky barrier diode. This thin insulating residue acts as an additional tunneling barrier, resulting in an apparent increase in $\emptyset_B$. Similar effects have been seen for GaAs reactive ion etched in CF_4 and CHF_3,[15] and also in CCl_2F_2:O_2 mixtures.[13] These residues are strictly surface impurities, however, since they were easily removed by a light sputtering step.

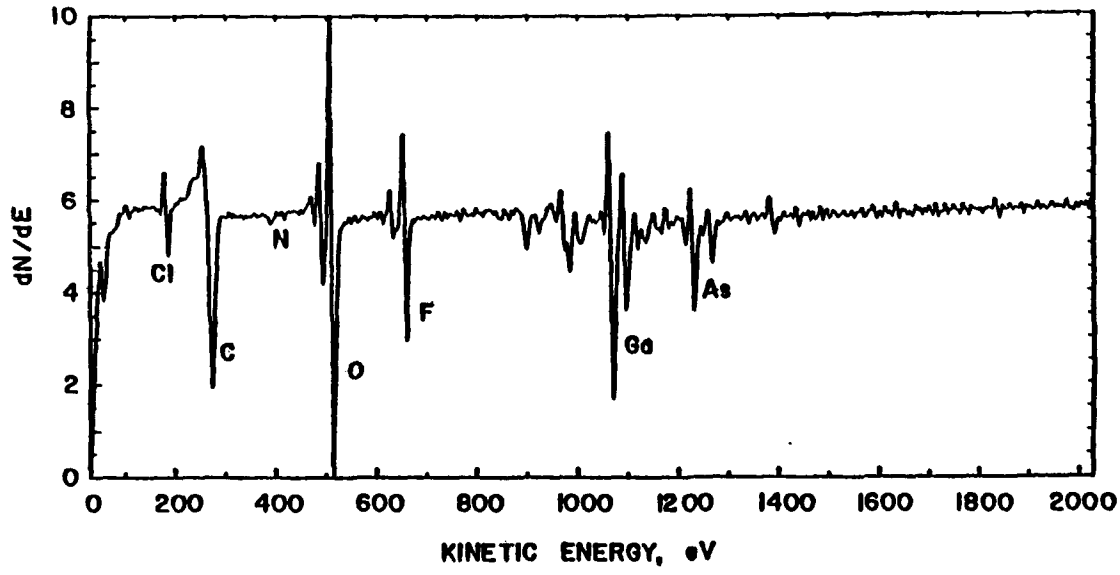


Figure 12. Auger electron spectroscopy measurement on GaAs horizontal surfaces etched for 4 minutes at 4 mTorr, 0.52 watts/cm^2, 20 sccm CCl_2F_2.

CONCLUSIONS

Magnetron reactive ion etching has been found to produce higher etch rates, smoother surfaces, and less wafer damage and electrical property degradation than standard (unmagnetized) reactive ion etching. This relatively new plasma etching method has the potential for being an important processing technique for the fabrication of GaAs devices for Army applications.

REFERENCES

1. EJ. Contolini, L.A. D'Asaro, J. Vac. Sci. Technol. **B 4**,706 (1986).
2. SJ. Pearton, U.K. Chakrabarti, W.S. Hobson, A.P. Kinsella, J. Vac. Sci. Technol. **B 5**, 607 (1990).
3. S. Salimian, C. Yuen, C. Shih, C.B. Cooper, J. Vac. Sci. Technol. **B 9**, 114 (1991).
4. M.B. Stern, P.F. Liao, J. Vac. Sci. Technol. **B1**, 1053 (1983).
5. G.F. McLane, M. Meyyappan, MW. Cole, C. Wrenn, J. Appl. Phys. **69**, 695 (1991).
6. M.W. Cole, S.Salimian, C.B. Cooper, H.S. Lee, M. Dutta, Scanning, The Journal of Scanning Microscopy **14**, 31 (1992).
7. S.M. Sze, "Physics of Semiconductor Devices", John Wiley and Sons, Inc., 1969.
8. R. Van Daele, D. Lootens, P. Demeester, Vacuum **41**, 906 (1990).
9. T. Mantei, J.J. Ibara, J Appl. Phys. **61**, 4885 (1987).
10. M. Meyyappan, G.F. McLane, M.W. Cole, R. Lareau, M. Namaroff, J. Sasserath, C. Sundararaman, 38th Annual AVS Symposium and Topical Conference, Seattle, WA, Nov. 11-15, 1991.
11. R.E. Klinger, J.E. Greene, J. Appl. Phys. **54**, 1595 (1983).
12. S.W. Pang, J. Electrochem. Soc. **133**, 784 (1986).
13. S.J. Pearton, M.J. Vasile, K.S. Jones, K.T. Short, E. Lane, T.R. Fullowan, A.E. VonNeida, N.M. Haegel, J. Appl. Phys. **65**, 1281 (1989).
14. S.J.Pearton, A.B. Emerson, U.K. Chakrabarti, E. Lane, K.S. Jones, K. T. Short, A.E. White, T.R. Fullowan, J. Appl. Phys. **66**, 3839 (1989).
15. S.W. Pang, G.A. Lincoln, R.W. McClelland, P.D. DeGraff, M.W. Geis, M.J. Piacentini, J. Vac. Sci. Technol. **B1**, 1334 (1983).

Imaging and Analysis of Combusting Liquid Propellant Sprays

*Michael J. McQuaid, Dr.; Avi Birk, Dr.; Gary Bliesener, Mr.
U.S. Army Ballistic Research Laboratory
Aberdeen Proving Ground, MD 21005-5066

Introduction

Numerous power and propulsion systems involve liquid spray combustion. They range in character from the ubiquitous residential oil furnace, to diesel, turbine and powerful rocket engines. Perhaps the most exotic propulsion system utilizing liquid injection and spray combustion is the regenerative liquid propellant gun (RLPG) which has been selected for incorporation into the Army's Advanced Field Artillery System. The RLPG operates at pressures at least an order of magnitude higher than other propulsion systems and achieves combustion intensities as high as those produced by the Saturn V rocket. RLPG technology has matured to the point where a fieldable 155 mm howitzer developed by General Electric is routinely fired successfully, but technical issues remain.

One of these issues is the large, high-frequency pressure fluctuations produced in virtually all RLPG systems developed to date.[1] Performance is little affected by these fluctuations, but there is concern that they may have an adverse effect on projectile components. Analysis of pressure fluctuations reveal that they have both coherent and incoherent wave structure, with most of the energy stored in the incoherent waves. The coherent waves, in most cases, relate to acoustic modes which depend on the global geometry of the particular combustion chamber. In RLPGs the coherent waves have been associated with transverse acoustic modes. (This is in contrast to the coherent waves encountered in solid propellant guns which have been associated with lower frequency longitudinal modes.[2]) Engineering solutions, such as the use of a jet dispersion device in an experimental 30 mm RLPG at the BRL,[3] have been implemented which reduce coherent fluctuations.[4] However, the energy content of coherent modes is so small relative to the incoherent content that overall pressure v. time traces are essentially unchanged.[5]

Considerable research effort has been devoted to identifying the mechanisms driving incoherent pressure fluctuations in RLPGs, but conlusive answers have yet to be obtained. It appears certain, however, that the mechanisms are to liquid atomization and gas dynamic interactions in the highly turbulent combusting spray. This paper summarizes experiments conducted at the BRL Spray Research Facility to elucidate jet break-up, atomization, and combustion of liquid monopropellants through the use of imaging and spectroscopic techniques. These studies demonstrate the importance of particulate and vortex burning for

hydroxylammonium nitrate (HAN) based propellants at low gun pressures. Experiments with LGP1898, a recently formulated HAN-based propellant,[5] are the first to provide information regarding its performance during actual spray combustion. Finally, experiments with nitromethane show that the dynamical behavior of a propellant spray will change as the pressure and temperature in a gun chamber exceed supercritical values for the propellant.

Experimental Setup

The experimental facility employs techniques designed to overcome the high optical density and particulate generation which prevents adequate visualization of spray phenomena in a real gun chamber.[6,7,8,9,10] While studies of liquid atomization at high pressures are routinely done elsewhere at room temperature, the BRL facility is unique in enabling such studies at gas temperatures as high as 550°C. Thus, the facility provides controlled test conditions for investigating liquid spray reactivity; i.e. evaporation, ignition, and combustion.

The experimental setup is shown schematically in Figure 1. Propellant is injected upward into hot nitrogen delivered to the test chamber from a high pressure gas reservoir via a particle bed heater. The test chamber is cylindrical with a 75 mm inside diameter. It has four opposing rectangular sapphire windows with 97 mm x 32 mm apertures. Chamber pressure is measured with a pressure transducer (Kistler, Model 607C) and the temperature at the center of the chamber is measured with a 0.05 mm chromel/alumel thermocouple. For the experiments reported here the propellant was typically injected into 500°C, 33 MPa nitrogen. These temperature and pressure conditions were the highest that could be reliably sustained because the sapphire windows, though rated for 65 MPa service, were prone to failure above 33 MPa.

The injector is a patented design based on a regenerative concept.[8] For the experiments to be discussed the injector was loaded with 35 cc of propellant. Among the propellants studied were LGP1845, LGP1898, the subcomponents HAN and TEAN, and nitromethane. A small amount of water (0.5 to 0.7 cc) was loaded between the propellant and the injection orifice to prevent the propellant from being ignited prior to injection by heat transfer from the ambient gas. (Thus the injection of the buffer water preceded the actual propellant injection.) Forty seven cc of water was also loaded "behind" the propellant to flush the system after the propellant was injected. The two (water) buffer zones in the injector were isolated from the propellant by traveling seals. In most tests propellants were seeded with approximately 0.5 cc of a saturated nitrate salt solution to produce flame enhancement and spectroscopic markers. Aqueous $Sr(NO_3)_2$, $Ba(NO_3)_2$, $LiNO_3$, and $NaNO_3$ solutions were tested for these purposes. Each is fully miscible with the HAN-based propellants and it is considered that their impact on the combustion process is minimal. Injector piston motion was monitored by a Wolff Hall-effect magnetic sensor, and by a photodiode-based device from which the injection velocity could be calculated. In an attempt

to achieve complete burning of the spray within the field of view, and to provide maximum test duration, the propellant was injected through a small circular (0.5 mm diameter) orifice. Tests utilizing a 1.0 mm diameter orifice were also conducted.

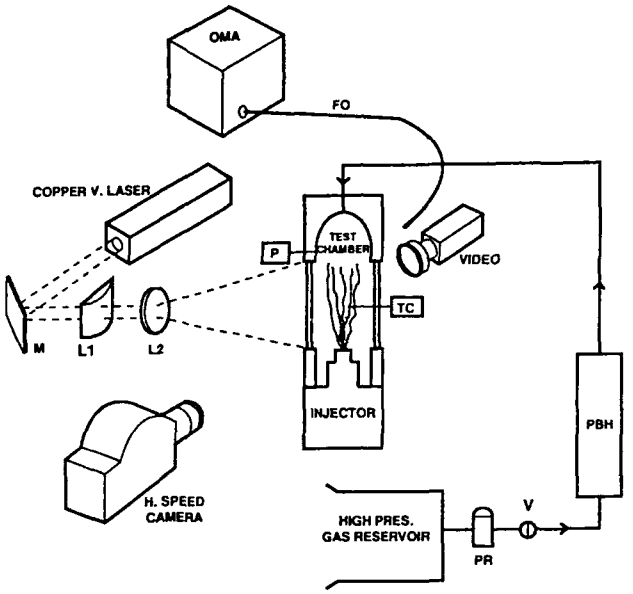


Figure 1: Schematic of the Experimental Setup.

For visualization, both a color video camcorder (Quasar, Model VM-21) and a high-speed framing camera (Photonic Systems, Photec IV) were used. A 10 watt copper vapor laser (Oxford Lasers, Model CU10) served as a strobe. The laser pulses, which have a duration of 25 ns, were synchronized to the framing rate of the Photec. The framing rate was typically set to either 2000 or 5000 fps. The laser's 25 mm diameter beam was optically steered to provide a 3 mm thick, planar light sheet that sliced the chamber volume symmetrically along the cylinder axis and at 90° to the cameras.

Spectroscopic measurements were facilitated with an optical multichannel analyzer (Princeton Instruments, Model ST-1000) coupled to a monochromator (ISA Instruments, Model HR-20). Measurements were limited to recording emission spectra of the combustion process. Scan rate (integration) cycles of 4 or 5 msec were used. A 150 groove/mm grating was used to observe spectral ranges of approximately 500 nm. A 600 groove/mm grating, which resulted in a 125 nm spectral range, was used to obtain better wavelength resolution of features identified in the 500 nm scans. The spectra were collected from a region subtending a 25 mm circle of the spray approximately 50 mm downstream of the injector - the region of highest flame intensity. A one meter long quartz fiber optic was used for light collection.

Image Digitization and Processing. Image analysis techniques were applied to better interpret the cinematic data. To employ image analysis techniques it is first necessary to digitize the cinematic records. This was accomplished by projecting the photographic images onto a screen and recording the projection to a Super-VHS formatted tape via a professional video camera (JVC, Model HZ-811) and a digital VCR (Panasonic, Model AG-1830). The white balance feature of the video camera was used to correct the color bias of the projected image caused by the deviation of the projector lamp output from "true white". ("True white" is defined as I(R)=I(G)=I(B), where I(R), I(G), and I(B) are the intensity values of the NTSC "red", "green", and "blue", respectively.) The video tape records were digitized with PC-AT based frame grabbers to obtain 8-bit monochrome (Truevision, TARGA M8) or 24-bit color (Truevision, ATVISTA) images. A PC-based software package (Jandel Scientific, JAVA) was used for monochrome image analysis. The color image files were ported via a local area network (Ethernet) to a Silicon Graphics Personal Iris 4D20E workstation for analysis. Color image analysis was performed with a UNIX-based software package (G.W. Hannaway and Associates, WHIP) which was further developed in house for the specific needs of this study.

Results

Temperature and Pressure Profiles. A representative sample of chamber pressures and temperatures versus time for HAN-based propellant tests is shown in Figure 2. The chamber

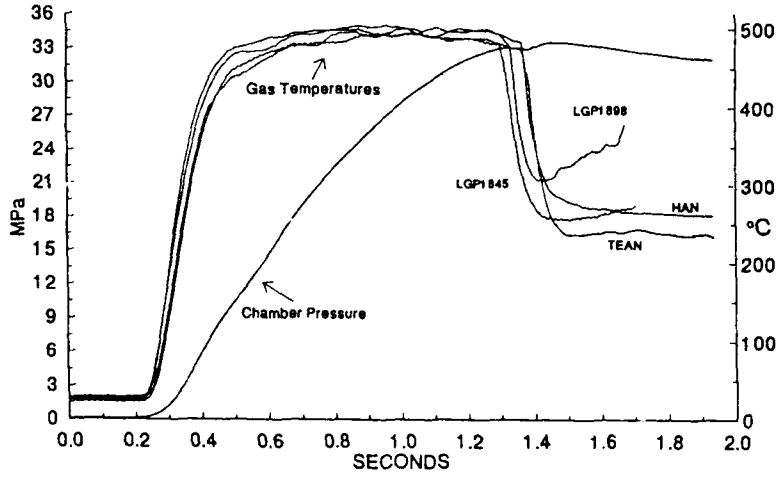


Figure 2. Chamber Pressure and Temperatures v. Time for Selected Experiments.

pressure shown is an average from three LGP1845 tests. The LGP1845 gas temperature is also an average of these three tests. In all tests, the thermocouple failed shortly after injection. Injection takes place about one second after chamber pressurization. Of interest

are the post-injection minimum temperatures and the subsequent rate of temperature rise. This rate is a direct reflection of the reactivity of the given propellant and also indicates flame propagation upstream and inward on the spray. The minimum temperatures are influenced by a propellant's reactivity and its thermal properties. The quality of phase mixing during the early injection period is also reflected in the minimum temperatures obtained. The flameless HAN and TEAN sprays exhibit flat post-injection temperature profiles. The post-injection temperature measured in the HAN test is well above the temperature at which HAN starts to decompose. The post-injection temperature measured in the TEAN test is relatively low, indicating that TEAN provides a greater heat sink to the chamber gas. Both HAN and TEAN, used as monopropellants, release only a small fraction of the energy produced by LGP1845 combustion. The temperature records of the LGP1845 versus that of the LGP1898 sprays clearly reveal the greater reactivity of the LGP1898 spray.

Spectroscopic Results. Emission spectra observed during an unseeded LGP1845 test are shown in Figure 3. Only OH A-X emission at 310 nm and Na $3s^2S$-$3p^2P$ (D-line) emission at 589 nm are identifiable. (Sodium is a propellant contaminant.) These transitions were observed in all LGP1845 and LGP1898 tests. Line broadening of the Na D-line is pronounced. The scan to scan variation in OH A-X emission intensity is attributable to background level variations. A broad band of emission in the range of 350-550 nm is apparent, but carriers for the signature could not be identified. No emission spectra were detected in the HAN tests and no NO_2 fluorescence, induced via excitation by the copper vapor laser, was detected in any of the experiments.

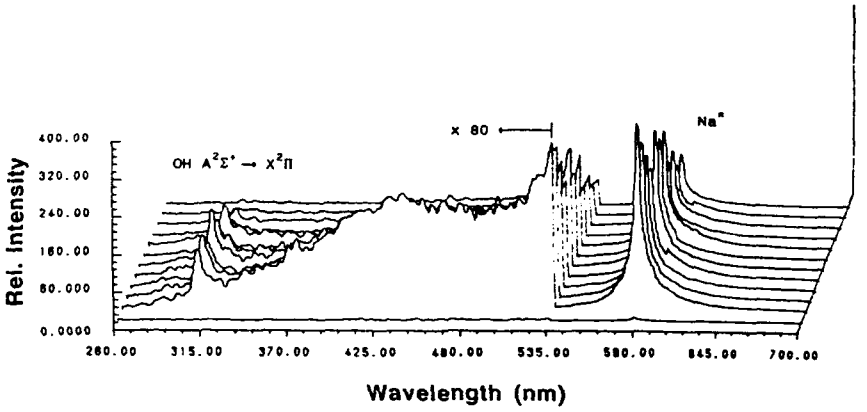


Figure 3. Emission Spectra from Unseeded LGP1845 Spray Combustion.

It was hoped that seeding the propellants with $Sr(NO_3)_2$ or $Ba(NO_3)_2$ would produce an excited state population distribution of the metal atoms which could be determined from its visible emission spectra. This, coupled with color image analysis, would permit mapping of flame temperatures. However, emission from these species was not observed. In addition

to Na $3s^2S$-$3p^2P$ transitions observed in all experiments, Li $2s^2S$-$2p^2P$ transitions were observed in experiments seeded with $LiNO_3$. Unfortunately, the energy splitting in the upper state in these transitions is insufficient for calculating flame temperature. The dominance of these transitions in the emission spectra is used in processing the image though. as will be discussed below.

Cinematic Observations. Sidelighting with the copper vapor laser strobe and seeding the propellants with nitrate salts provided breakthroughs in the imaging of the sprays. Sidelit photography produces better images than the more commonly used back-lit photography because it reduces the problem of image degradation due to index of refraction variations. Figure 4 shows a cinematic image of LGP1845 injection through a 1 mm diameter orifice into 33 MPa, 500°C nitrogen. Liquid boundaries appear clearly against the black background in this figure. Seeding the propellants with nitrate salts created sufficient additional combustion luminosity for flame to be recorded at the high framing rates necessary to capture the spray dynamics. Figure 5 shows a cinematic image of LGP1898 injection through a 0.5 mm diameter orifice into 33 MPA, 500°C nitrogen. The detail of the combustion process shown in the figure would not have been possible without seeding.

Figure 4. Combusting LGP1845 Spray with Laser Illumination.

The cinematic records reveal higher degrees of liquid turbulence, flame penetration, and intensity in the LGP1898 spray than in the LGP1845 spray. The slow accumulation of opaque gases in the chamber indicates the LGP1845 spray did not burn with optimum efficiency. In contrast, decomposition products were not observed in the LGP1898 tests. During the test flame occasionally advanced upstream in the spray as close as 10 mm from the injector. The mechanisms of flame advance are of interest because such phenomena could cause piston reversal in a RLPG, a potentially serious problem.

Figure 6 shows a cinematic image of the nitromethane inmjection through a 0.5 mm diameter orifice into 33 MPa, 500°C nitrogen. These values are above nitromethane's critical point (6 MPa, 314.8°C). Nitromethane did not ignite at these conditions, but the dynamics of the spray are distinctly different from those observed for HAN-based propellants. One

noticeable difference is that there is little evidence for the generation of large droplets. The global characteristics of the spray are also different. The divergence angle and penetration distance are much smaller, and the liquid core and boundary fluctuate to a larger degree. These results may have implications for HAN-based propellant spray dynamics in an RLPG since the temperature and pressure reached during the injection cycle will presumably exceed supercritical values for the propellant.

Figure 5. Combusting LGP1898 Spray without Laser Illumination.

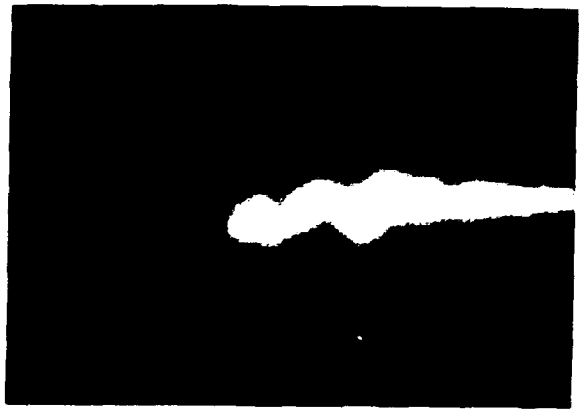


Figure 6. Nitromethane Spray with Laser Illumination.

Image Processing and Analysis

Color image processing was useful for enhancing the details of images involving laser illumination. The spectroscopic data show that the flame can be considered to have just two wavelength components, the 589 nm emission line of sodium and the 670.7 nm emission line of lithium. The copper vapor laser radiates at 510.6 nm and 578.2 nm with the 510.6 nm line approximately twice as intense as the 578.2 nm line. Thus, the color images effectively result from the combination of just four different wavelengths. From the published spectral sensitivity curves for the high speed film (Kodak, Ektochrome 2253), the relative amounts of I(R), I(G), and I(B) for unit exposure at these four wavelengths is shown in Table I. These values indicate that the 510.6 nm laser line is the only color source which will contribute to I(B). Likewise, I(R) will be predominantly due to emission from lithium and sodium. Thus the color content provides a means of identifying contributions due to flame and laser scatter from liquid in regions where these two processes are coincident.

Table I

RGB Sensitivity v. Selected Wavelengths for Unit Exposure
(Kodak Ektochrome 2253 High Speed Daylight Film)

Wavelength (nm)	R	G	B
510.6	0.5	6.1	5.9
578.2	1.6	15.1	<0.1
589.0	1.8	8.4	<0.1
670.7	27.5	<0.1	<0.1

It should be noted that the film color sensitivity data has only been established for exposure times ranging from 10^{0} to 10^{-4} seconds. Exposure durations outside this range necessitate corrections for color biasing and/or f-stop. Since the film's exposure to flame is controlled by the framing rate of the camera (<10^{4} fps), this data should provide reasonable values for analysis of flame. The film's exposure to laser scatter, however, is determined by the length of laser pulses, i.e., 2.5×10^{-8} seconds. In order to establish sensitivity data for the 2.5×10^{-8} second exposure to laser scatter, images from experiments involving the injection of H_2O were analyzed. For these experiments the image is produced solely by laser scatter. It was found that the ratio I(B)/I(G) was close to 1:1.25, versus the 1:2 ratio calculated from Table I.

The deviation from the ratio calculated from the sensitivity data may also be due to light scattering phenomena. The effective cross-section for light scattering from particles change with their size and index of refraction in accordance with the Mie theory. For particles of submicron size (Rayleigh) scattering efficiency is greater at shorter wavelengths. Hence , for finely atomized sprays undergoing significant evaporation it is conceivable that the 510.6 nm laser line is scattered more effectively than the 578.2 laser line. This would result in an image with higher than expected blue content. On the other hand, the known LGP1845 combustion intermediate NO_2 has an absorption cross-section which is much higher at 510.6 nm than at 578.2 nm. Preferential absorption of the 510.6 nm line would tend to lower the I(B)/I(G) ratio. Since it is not possible to quantitatively assess all the factors which could affect the color balance, the current analysis focused on noise reduction and image enhancement based on the simple process identification criteria noted above. That is, I(B) is associated with laser scatter and I(R) is associated with flame emission.

Criteria to threshold the images for noise (reduction) was established by noting that both emission from the combustion process and laser scatter produce images with I(G) values. (See Table I.) By sampling background regions which should have been black (I(R)=I(G)=I(B)=0), a threshold for true signal could be determined. Based on this approach a pixel element was considered noise if I(G) was less than 50(/256), and I(R), I(G), and I(B) were set equal to 0.

The next aspect of the analysis involved enhancing the details of gas-liquid interactions. Because both flame emission and laser scatter produce images with I(G) values, I(G) values do not provide information which delineates the combustion process from scatter from the droplets. Therefore, it is not considered in processing the image. The ratio I(R):I(B) is calculated to show relative contributions from flame emission and laser scatter. The ratio is associated with a color in a user defined set with typically up to 8 or 9 colors. The I(R), I(G), and I(B) values of each color in this set was chosen in a systematic way to provide the best detail of the gas (flame) - liquid interactions. Figure 7 shows a processed version of Figure 4. The dynamic processes are analyzed by running the processed images in an animation. The results support observations from the raw cinematic images regarding the importance of vortex and particulate burning.

Figure 7. Image Processed Version of Figure 4.

Test sequences taken in the absence of external (laser) illumination, such as shown in Figure 5, were analyzed as monochrome images to quantitatively establish the intensity and extent of combustion. Intensity histograms of consecutive cinematic frames (Figure 7) provide a temporal profile of this information. The gray level in the histograms correlates with flame intensity and has a dynamic range from 0 (total darkness) to 256 (CCD element saturation). The number of pixels per gray level range is proportional to the projected area of the corresponding luminosity level. Such histograms provide the potential to correlate "combustion noise" with pressure fluctuations, but fluctuations were not observed during any of the experiments. It has been observed that pressure fluctuations in RLPGs tend to onset at pressures greater than 60 MPa. Thus the current experimental apparatus may be incapable of achieving the conditions which lead to pressure fluctuations in a RLPG. The experiments with nitromethane were the first attempt to use a simulant which would behave in the experimental chamber like HAN-based propellants behave in a RLPG. Though the nitromethane experiment was informative, a simulant which will ignite is desired. Other approaches along this line are being investigated.

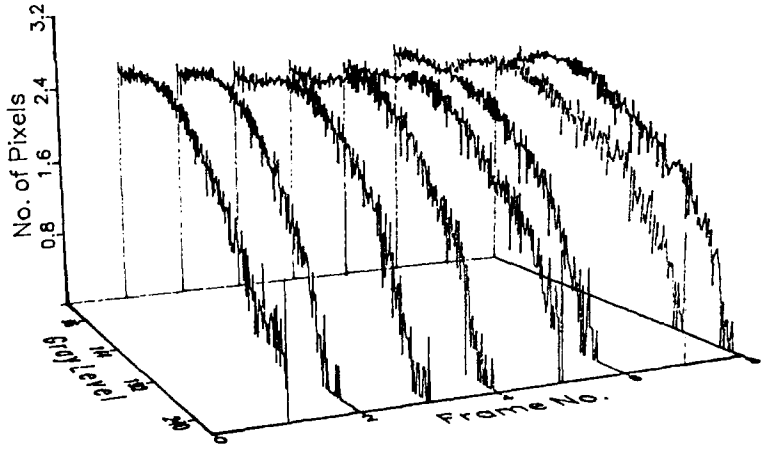


Figure 8. Intensity Histogram of Combusting LGP1898 Spray

Conclusions

Experimental techniques have been developed which facilitate the imaging of combusting liquid propellant sprays at low gun pressures. Novel chamber design, laser-strobed sidelit photography, and propellant seeding have resulted in excellent visualization of reacting sprays. The experimental results, coupled with image processing and analysis techniques confirm the assertion that particulate and vortex burning are salient mechanisms

in (R)LPG spray combustion. The results also support the proposition that the use of LGP1898 may reduce reactant accumulation during the early stage of the combustion process. This could result in reduced pressure fluctuations. Finally, experiments with nitromethane show that the dynamical behavior of the spray will change as the ambient temperature and pressure exceed the critical point of the propellant.

References

1. J.D. Knapton, C.A. Watson, and N.E. Boyer, BRL-TR-3322, March 1992.

2. J.D. Knapton and T.C. Minor, 27th JANNAF Combustion Subcommittee Meeting, (CPIA Pub.557, 1990), Vol.II, pp 81-94.

3. J. DeSpirito, N.E. Boyer, J.D. Knapton, A.W. Johnson, and R.E. Rychnovsky, 28th JANNAF Subcommittee Meeting, (CPIA Pub. 573, 1991) Vol.II, pp.179-188.

4. R.E. Rychnovsky, R.W. Carling, S.K. Griffiths, S.R. Vosen, and R.F. Renzi, 27th JANNAF Subcommittee Meeting, (CPIA Pub. 557, 1990), Vol.II, pp. 71-79.

5. N. Klein, T.P Coffee, and C.S. Leveritt, 28th JANNAF Combustion Subcommittee Meeting, (CPIA Pub. 573, 1991), Vol.I, pp. 227-234.

6. A. Birk, M. McQuaid, and G. Bliesener, 28th JANNAF Subcommittee Meeting, (CPIA Pub. 573, 1991), Vol.I, pp. 235-243.

7. A. Birk and G. Bliesener, 27th JANNAF Subcommittee Meeting, (CPIA Pub. 557, 1990), Vol.I, pp. 425-434.

8. A. Birk and G. Bliesener, BRL-TR-3209, February 1991.

9. A. Birk and P. Reeves, BRL-TR-3684, July 1988.

10. A. Birk and P. Reeves, BRL-TR-2780, March 1987.

Multifractal Analysis of Chaotic Point Sets

Dr. L.V. Meisel and Mr. M.A. Johnson*
U.S. Army ARDEC
Close Combat Armaments Center
Benet Laboratories
Watervliet, NY 12189-4050

INTRODUCTION

Many important physical systems[1] exhibit highly unstable or chaotic behavior. Traditional methods for treating such systems have been linearization and stability analysis. It is generally acknowledged that non-linear systems exhibit important features that cannot be understood, controlled, or exploited without addressing their behavior in a non-linear form. Thus, techniques are being developed to model and analyze such complex systems. The approach entails the analysis of complicated structures that exhibit fractal geometry. Fractal parameters describe how the length, area, and/or density of an object varies with scale and may provide quantitative measures of the complex structures that occur in nature. To optimally exploit fractal perspectives, quantitative techniques for measurement of fractal properties are evolving. In this paper, qualitative and quantitative techniques are applied to elucidate the fractal structure of the strange attractors that govern the evolution of symmetric chaos in some model systems.

BACKGROUND

The qualitative features of the time evolution of many physical systems are embodied in the solutions to equations of the form:

$$\dot{x}(t) = F(x(t), a, b, \ldots) \qquad (1a)$$

where a, b, etc. are parameters and $x \in R^n$. The function F determines, in principle, the non-linear evolution of the system. In practice, the behavior of non-linear dynamical systems is quite complex, and the solutions to Eq. (1a) (i.e., the trajectories x(t)) are frequently extremely sensitive to small changes in initial conditions or the system parameters. Equation (1a) is a continuous time

description of the dynamical system, however, such systems are generally treated in terms of difference equations or discrete time mappings of the form

$$x(n+1) = F(x(n), x(n-1), \ldots, a, b, \ldots) \qquad (1b)$$

The motion of non-linear systems, governed by Eq. (1b), may be unbounded or may settle down after an initial "transient" to a subset of R^n called an attractor. The motion of such systems can be characterized by their attractors rather than by specific trajectories. The attractors of non-linear dynamical systems fall into two categories:

i. Periodic attractors. When an attractor comprises a finite set of points (that are visited in sequence), the motion of the system is periodic and one refers to a periodic attractor. If the number of points is N, it is called a period N attractor.

ii. Chaotic attractors. When an attractor contains an infinite set of points, the motion of the system is called "chaos," trajectories are called "chaotic," and one refers to "strange attractors" or "chaotic attractors." Strange attractors have fractal structure.

The attractors studied here describe the dynamics of an important subset of chaotic systems: symmetric chaotic mappings in E^2, which were qualitatively characterized by Field and Golubitsky[2]. In particular, we focus our efforts on the quantitative determination of the fractal parameters that characterize the attractors of mappings, which transform from D_n to Z_n symmetry as a parameter is changed. Field and Golubitsky[2] referred to the motion of non-linear systems, whose attractors were invariant with respect to such symmetry groups, as "symmetric chaos."

Box-counting algorithms for fractal analysis are based on a statistical description of the variation with scale of the distribution of "mass" of a fractal set. We employ a box-counting technique to extract the fractal parameters that characterize the attractors of four systems exhibiting symmetric chaos.

APPROACH

The attractors of the polynomial mappings determined by iteration of

$$z \rightarrow (\lambda + \alpha |z|^2 + B\, Re(z^n) + i\, \omega)\, z + \gamma\, (z^*)^{n-1} \qquad (2)$$

where $z \epsilon \acute{C}$, and λ, α, B, γ, and n are real constants, have D_n symmetry for $\omega=0$, and Z_n symmetry for $|\omega| > 0$. Field and Golubitsky[2] recently discussed qualitative techniques for analyzing the attractors of polynomial mappings.

The nature of the attractors of Eq. (2) is elucidated by generating graphical representations of their structure, similar to those of Ref. 2, and by fractal analysis.

The purpose of the graphical representations is to visually illustrate the chaotic structure and symmetry of the attractors studied. Although a number of subtle decisions must be made in such representations, only standard techniques are employed.

The fractal analysis provides quantitative measures of the attractor structures. The fractal analysis employs the agglomeration box-counting (ABC) algorithm of Meisel, Johnson, and Cote[3]. ABC is especially suitable for analysis of large point sets.

Since the ABC algorithm is not a standard procedure, we outline the technique here. ABC employs a set of hypercubes (boxes) of linear dimension E to cover a subset of the fractal set. For each E, the partition function

$$Z(q;E) = \sum_i p_i^q(E) \qquad (3)$$

is computed for a range of the descriptive index q, and $p_i(E)$ is the occupation probability for the i^{th} box of side E. The partition function is related to the Hentschel and Procaccia[4,5] fractal dimension D(q) by the scaling relationship:

$$Z(q;E) \sim E^{(q-1)D(q)} \qquad (4)$$

where D(q) is a generalization of the Mandelbrot fractal dimension[6]. At q=0, D(q) reduces to the standard Mandelbrot fractal dimension. If D(q) is non-integer and invariant, the set is monofractal. If D(q) varies with q, it is said to be multifractal.

The essential steps in applying ABC are:

i. Define a set of elementary boxes for the given point set S, such that the elementary boxes have edge lengths,

$$E(H,K,M,\ldots)=E_0/(2^H 3^K 5^M \ldots)$$

where H, K, M, ... are integers, and E_0 is to be the largest box edge length in the set.

ii. Compute (measure, etc.) and store the occupation numbers (integers) $n_i(E(H,K,M,...))$ for the elementary boxes. (The probabilities $p_i(E)=n_i(E)/N$ need not be stored.)

iii. For each $h \in \{0,1,...,H\}$, $k \in \{0,1,...,K\}$, $m \in \{0,1,...,M\}$, ... define sets of boxes of edge length

$$E(h,k,m,...)=E_0/(2^h 3^k 5^m...),$$

which also cover the point set and contain integer numbers of elementary boxes.

a. Compute the occupation numbers $n_i(E(h,k,m,...))$ for each box by summing the occupation numbers in the elementary boxes contained.

b. For a representative set of q values, compute $Z(q,E(h,k,m,...))$ from Eq. (3).

iv. For each q value, compute D(q) by least squares fitting of a straight line to $\ln(Z(q,E(h,k,m,...))/(q-1)$ vs $\ln(E(h,k,m,...))$ as follows from Eq. (4). In applications, the elementary box size must be large enough to contain a "substantial number" of members of the set, and the largest boxes included in the fit must be smaller than the extent of the point set. (Special techniques are required for $q=1$, but D(q) is a smooth function of q in any case.)

ABC has been tested on Euclidean point sets and simple monofractal and multifractal constructs for which analytic values of D(q) could be computed. Worst case results were within 5 percent of the analytic values for $q \geq 0$ and $N=10^7$ for the cases studied in Ref. 3. Reliable values of D(q) were not obtained employing ABC (or conventional box-counting[4]) for $q<0$; therefore, D(q) for $q<0$ are not reported here.

RESULTS

Results are presented for four classes of symmetric chaotic mappings based on Eq. (2): Figures 1 and 2 pertain to attractors having 3-fold symmetry (i.e., D_3 and Z_3 symmetry), Fig. 3 to attractors having 6-fold symmetry, and Fig. 4 to attractors having 16-fold symmetry.

For each class of mappings defined by a specific choice of λ, α, B, γ, and n, attractors were generated based on 10^7 iterations of Eq. (2) for ω values starting from 0 and stepping by 0.01 over the range for which the motion is bounded. Smaller changes of ω were taken in "interesting" ranges. The parameter set defining each class of mappings is given in the captions of the a-parts of each figure. (The λ, α, B, γ, and n parameter sets were chosen to match the Field and Golubitsky[2] D_n mappings.)

Two-dimensional images of subsets of the attractors (generated for $\omega=0$ and selected other values of ω), which were formed using the density of the attractor mapping as an index to a gray scale using a technique similar to Field and Golubitsky[2], are shown for each case. The attractors are depicted in square regions whose corners are at (-2, -2) and (2, 2) for the 3-fold mappings and (-1,-1) and (1,1) for the others. Each image is based upon 768x768 pixels shaded according to number of hits modulo 512.

Three-dimensional images, in which the smoothed density of the attractor is the third coordinate, are shown for the 6- and 16-fold symmetric attractors to enhance visualization. Although a more faithful representation was obtained prior to smoothing, unsmoothed density is not shown because the rapid (chaotic) density variations were difficult to interpret.

Graphs of the Hentschel-Procaccia[5] generalized fractal dimension D(q) (for q=0, 2, 5, and 9) vs ω over the range for which chaotic attractors exist and D(q) vs q for the ω values corresponding to the chaotic attractor images are shown for all cases.

For all the mappings presented:

i. For $\omega>1$, unbounded motion results.

ii. For $\omega<1$,

a. a range of values of ω for which periodic orbits, possibly comprising parts of period doubling cascades with decreasing ω, was found;

b. chaotic attractors were found for a range of positive ω extending from 0;

c. dramatic changes in the structure of the attractors with varying ω were apparent in their two- and three-dimensional images.

D(0) is essentially invariant under changes in ω except for the mappings of Fig. 1; however, D(q) exhibits substantial non-monotonic variation with ω over the chaotic range for all classes of mappings studied.

The 3-fold symmetric mappings of Figs. 1 and 2 are most interesting:

The mappings of Fig. 1 yield a single basin of attraction at small ω that fissions into three symmetrically disposed sub-basins of attraction for ω near 0.33. Near the "critical point" one passes through a "periodic window." An enhanced rendering (i.e., each point is rendered as a small circle) of one of the three period 8 sub-attractors in the periodic window at ω=0.338500... is shown. An image of a corresponding chaotic sub-attractor at $\omega \approx 0.39$ is also shown. The full attractors in the three sub-basin range have Z_3 symmetry. However, the motion of the dynamical system would be confined to one of the sub-basins of attraction and would not have Z_3 symmetry. Searches based on changes in ω of 0.001 did not yield periodic orbits with a period other than 8 ($=2^3$) in the window. (Although the ABC algorithm yielded values as large as 0.06, D(q) values are shown as 0.00 for ω in the strictly periodic windows in Figs. 1 and 2.)

The mappings of Fig. 2 yield a single basin of attraction or unbounded orbits for almost all ω. Two strictly periodic windows were found, and their attractors are shown. The window at ω=0.06 contains a period 2 attractor in the midst of a one basin chaotic range of ω. The window at ω=0.12 contains a period 39 (=3*13) attractor, which has Z_3 symmetry, but the factor of 13 is puzzling. The window at ω=0.28 contains a quasi-period 12 attractor, which has Z_3 symmetry, but is not strictly periodic even after 10^6 iterations. The ABC algorithm yields values of D(q) greater than 0.30 in accord with the apparent chaotic nature of the attractor for ω=0.28.

The mappings of Figs. 3 and 4 yield substantial variations in structure and multifractal D(q) with ω. Although prominent dips were apparent in D(q) near specific values of ω, no periodic windows were discovered searching via steps of 0.001 in ω near those dips.

CONCLUSIONS

The results presented in Figs. 1 through 4 comprise a qualitative and

quantitative solution (except for the possible omission of some narrow periodic windows) of the non-linear problems defined. One needs both approaches to characterize the dynamics.

Even the simplest non-linear dynamical systems exhibit unexpectedly complex behavior. Indeed, even in retrospect, it seems impossible to deduce the complex form of the symmetric chaotic attractors or the diversity of D(q) vs q for representative ω-values, such as those illustrated in Figs. 1 through 4, by analysis of the form of the mappings of Eq. (2).

The nature of the change in symmetry from D_n to Z_n is generally not reflected in the Mandelbrot fractal dimension D(0), but it is manifested in changes of the Hentschel and Procaccia[3,4,5] generalized fractal dimension D(q). Thus, both qualitative and multifractal characterization of the attractors of non-linear systems of interest are recommended.

REFERENCES

1. For example, see C. Grebogi, E. Ott, and J.A. Yorke, *Science*, **238**, 632 (1987); J.-P. Eckman and D. Ruelle, *Rev. Mod Phys.* **57**, 617 (1985); A. Davidson, *Computers in Physics*, **5**, 574 (1991); or E. Hornbogen, *International Materials Reviews*, **34**, 277 (1989).

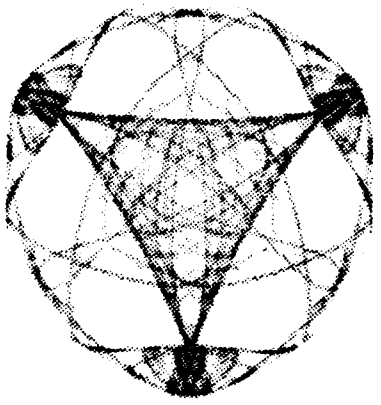
2. Mike Field and Martin Golubitsky, *Computers in Physics*, **4**, 470 (1990).

3. L. V. Meisel, Mark Johnson, and P. J. Cote, "Box-Counting Multifractal Analysis," *Phys. Rev. A* (in press).

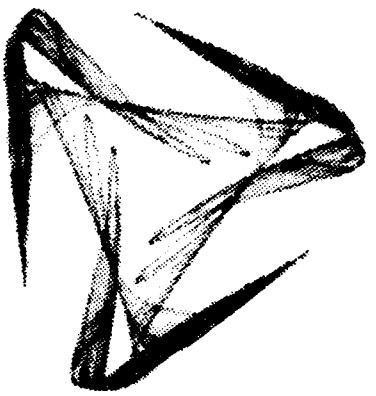
4. For example, see Harvey Gould and Jan Tobochnik, *Computers in Physics*, **4**, 202 (1990) or A. Block, W. von Bloh, and H.J. Schellnhuber, *Phys. Rev.*, **A 42**, 1869 (1990).

5. H. G. E. Hentschel and I. Procaccia, *Physica D*, **8**, 435 (1983). See also, Thomas C. Halsey, Mogens H. Jensen, Leo P. Kadanoff, Itamar Procaccia, and Boris I. Shraiman, *Phys. Rev.* **A33**, 1141 (1986).

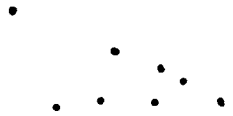
6. For example, see B.B. Mandelbrot, *The Fractal Geometry of Nature*, Freeman, New York (1982).



ω=0.00



ω=0.26



ω=0.33854

ω=0.39

Figure 1a. Attractor with n=3; λ=1.56; α=-1.0; B=0.1; γ=-0.82.

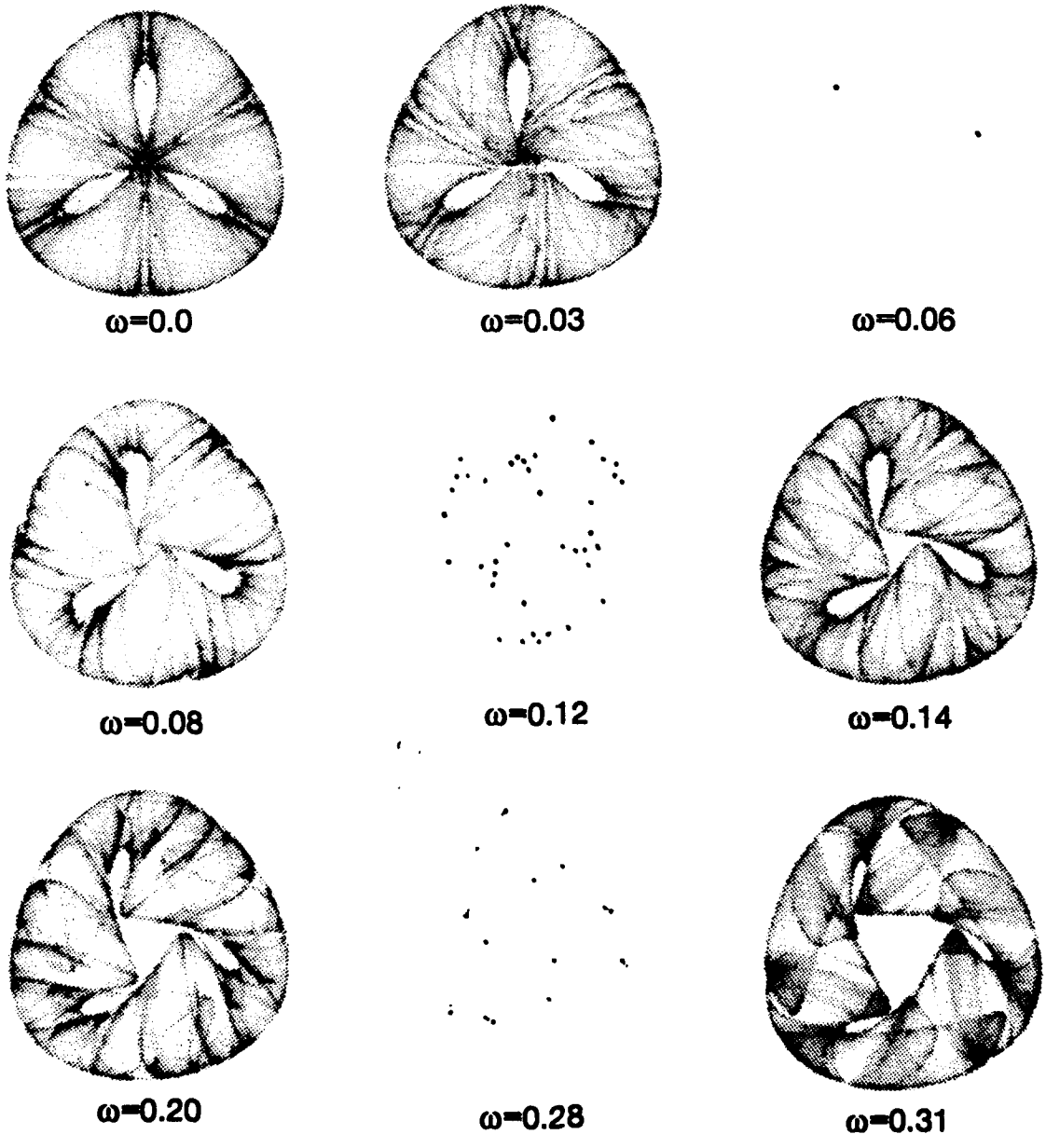


Figure 2a. Attractor with n=3; λ=-2.38; α=1.0; B=0.0; γ=0.1.

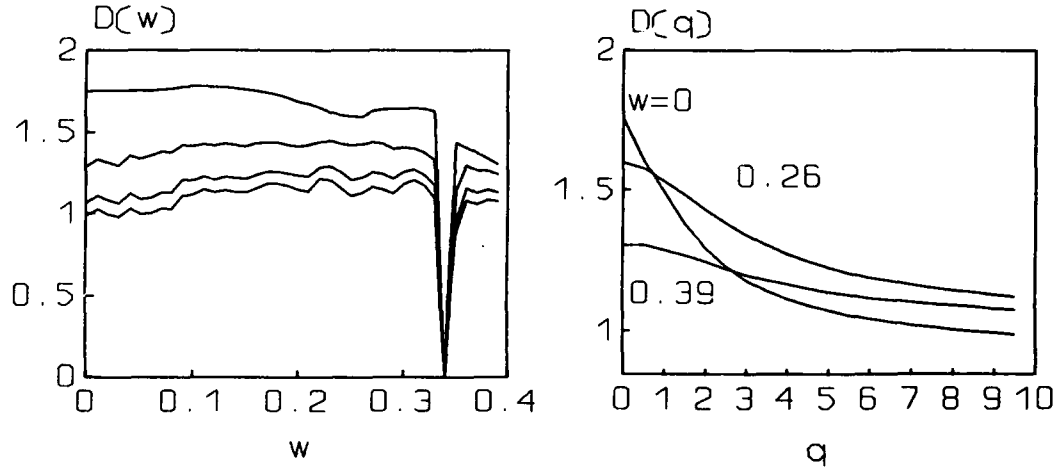


Figure 1b. Fractal analysis of the attractors depicted in Fig. 1a. (D is non-increasing in q and for all attractors D(w) are shown for q={0,2,5,9}.)

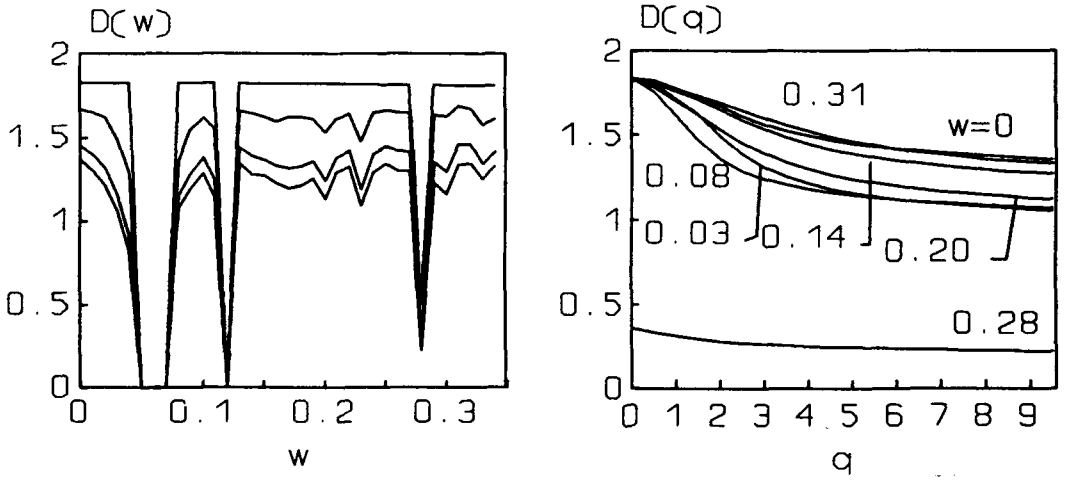


Figure 2b. Fractal analysis of the attractors depicted in Fig. 2a.

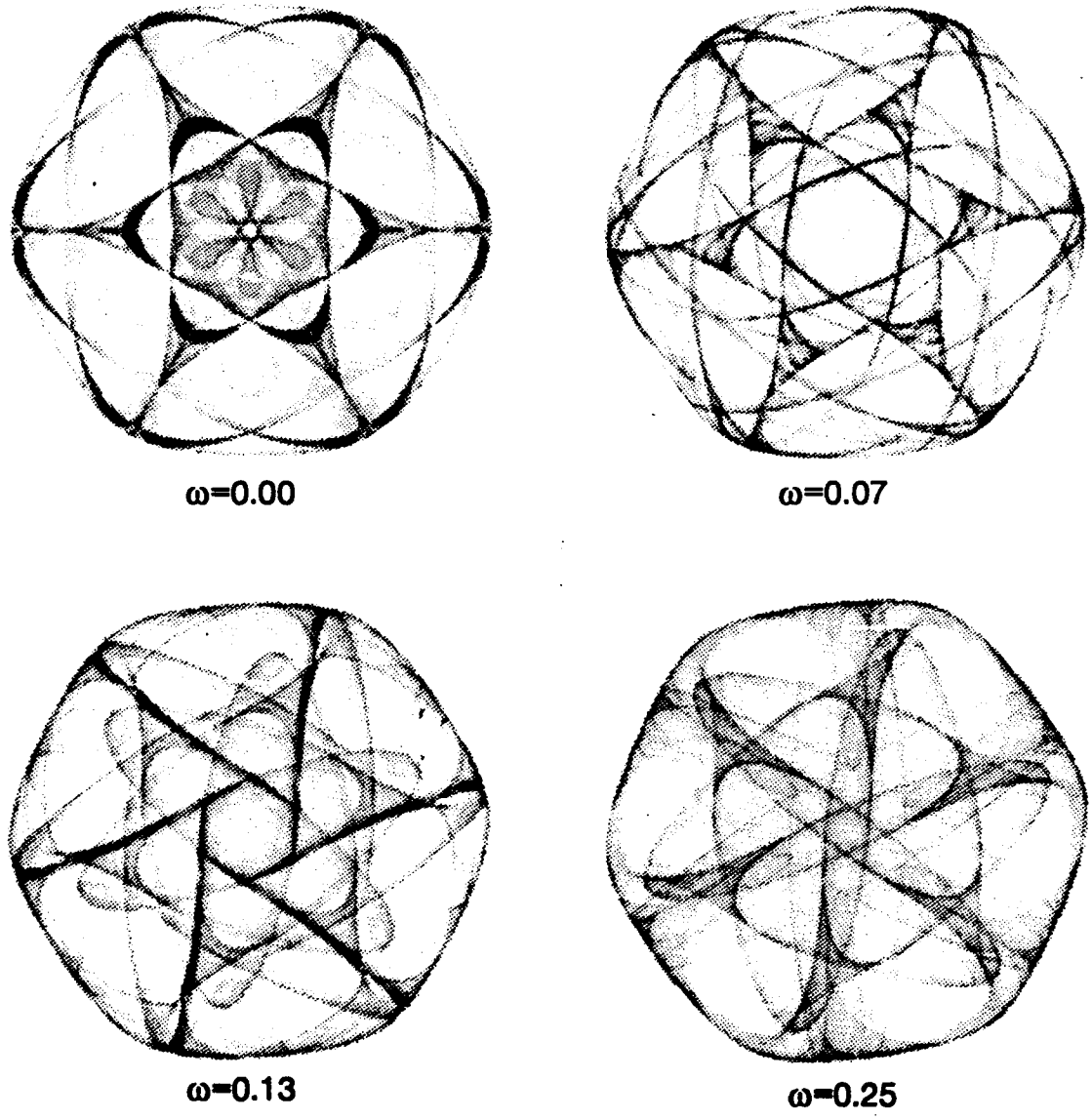


Figure 3a. Attractor with n=6; λ-2.7; α=5.0; B=2.0; γ=1.0.

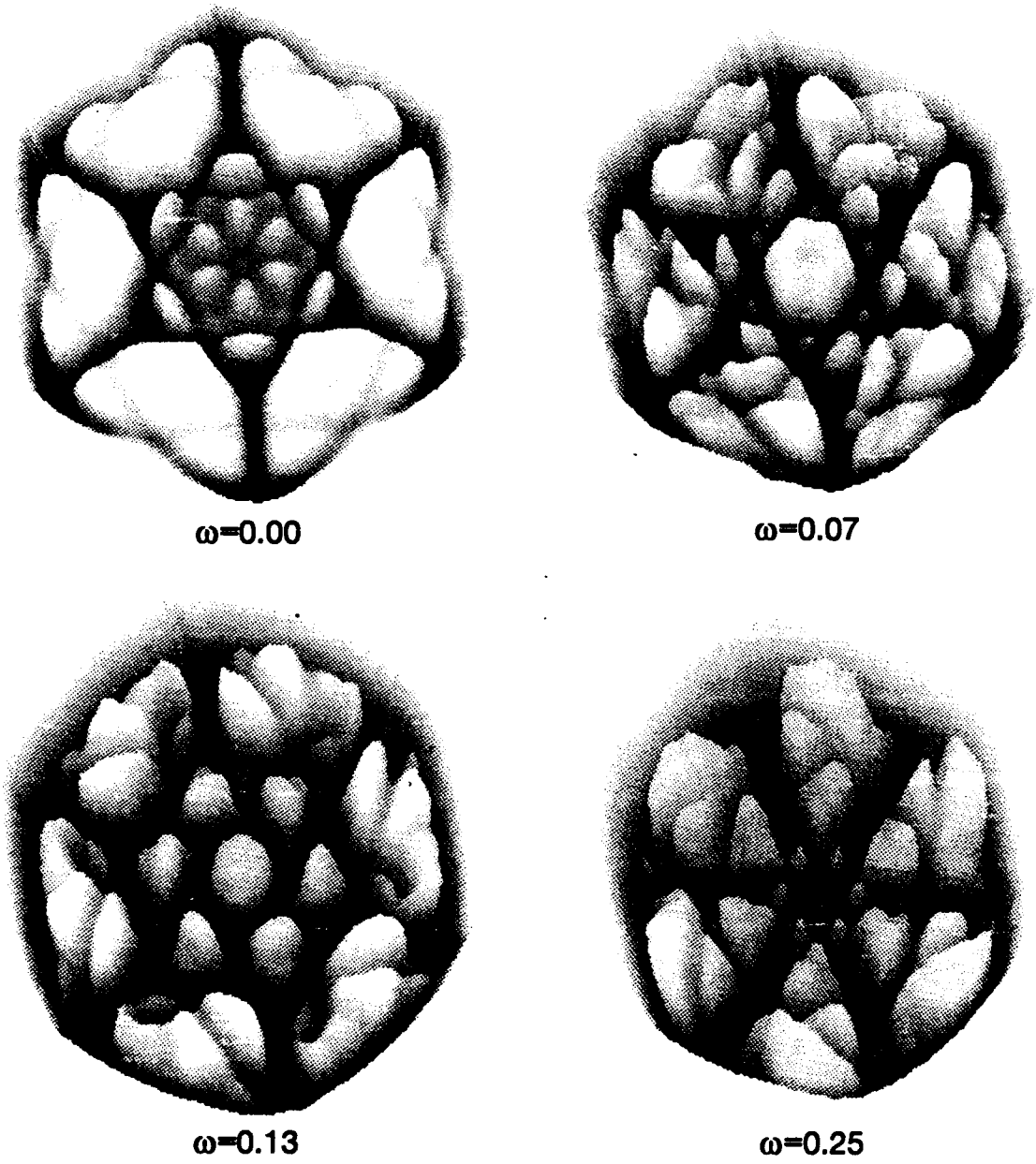


Figure 3b. Three-dimensional representation of attractor in 3a.

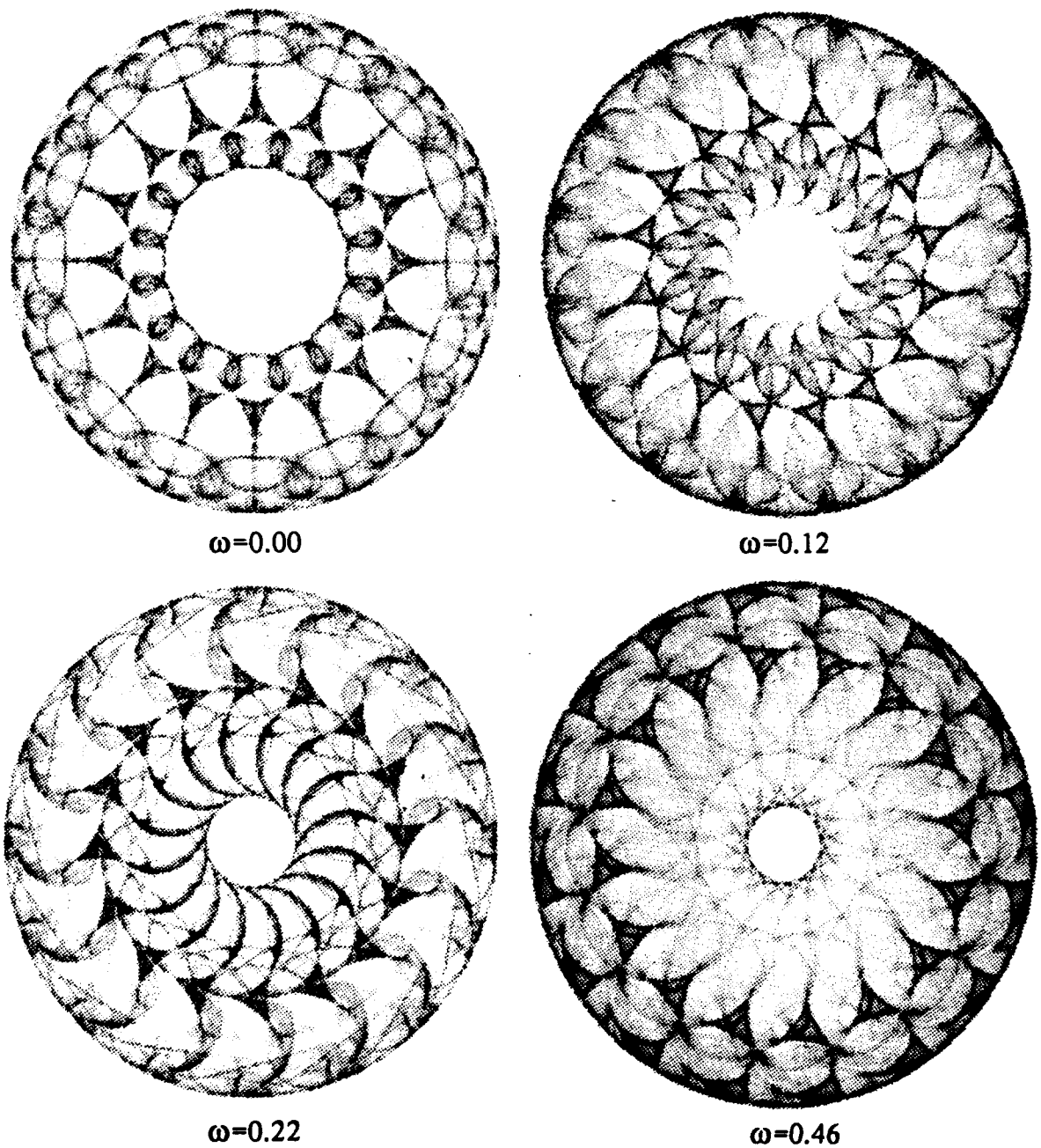


Figure 4a. Attractor with n=16; λ=2.39; α=-2.5; B=-0.1; γ=0.9.

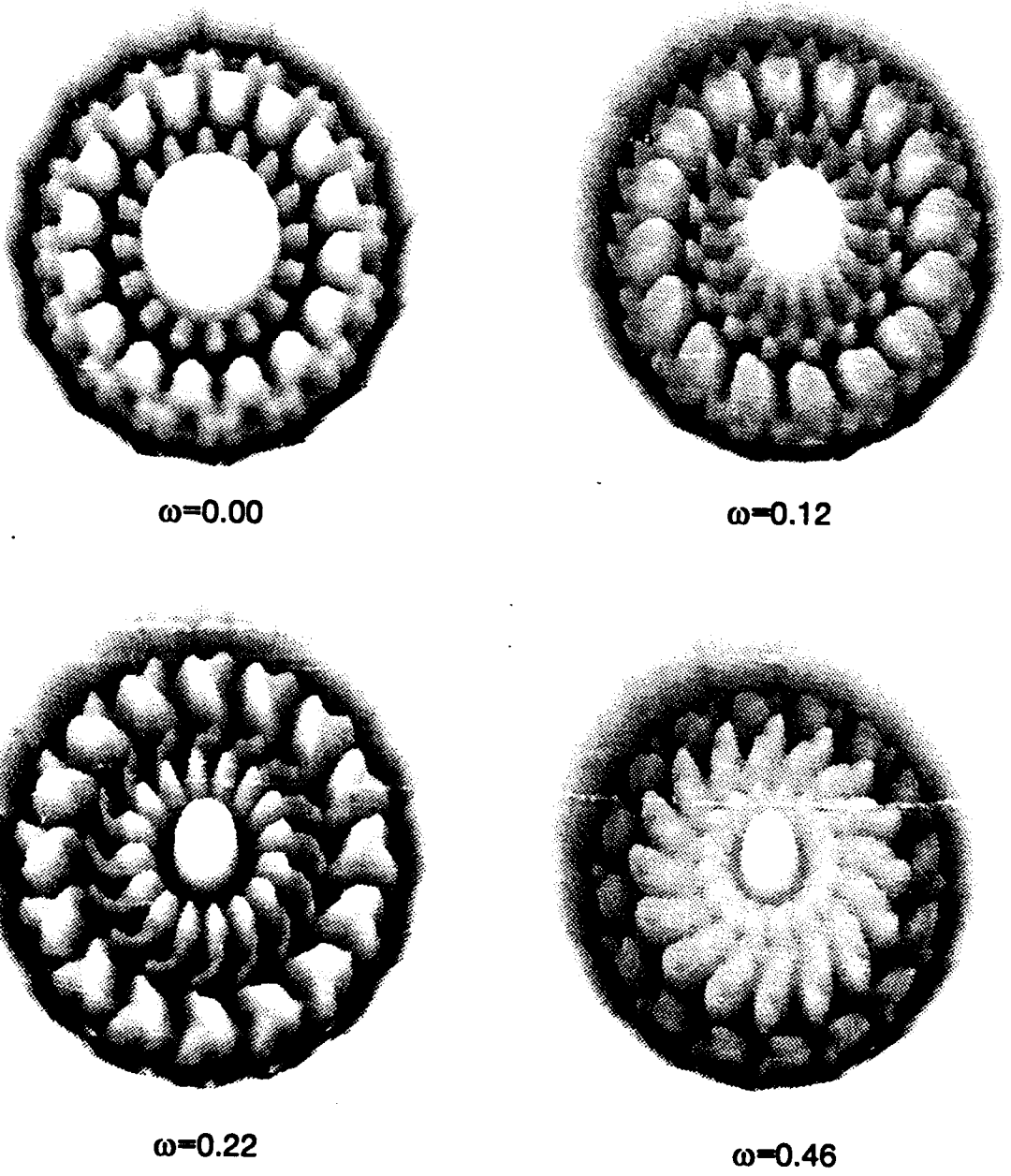


Figure 4b. Three-dimensional representation of attractor in 4a.

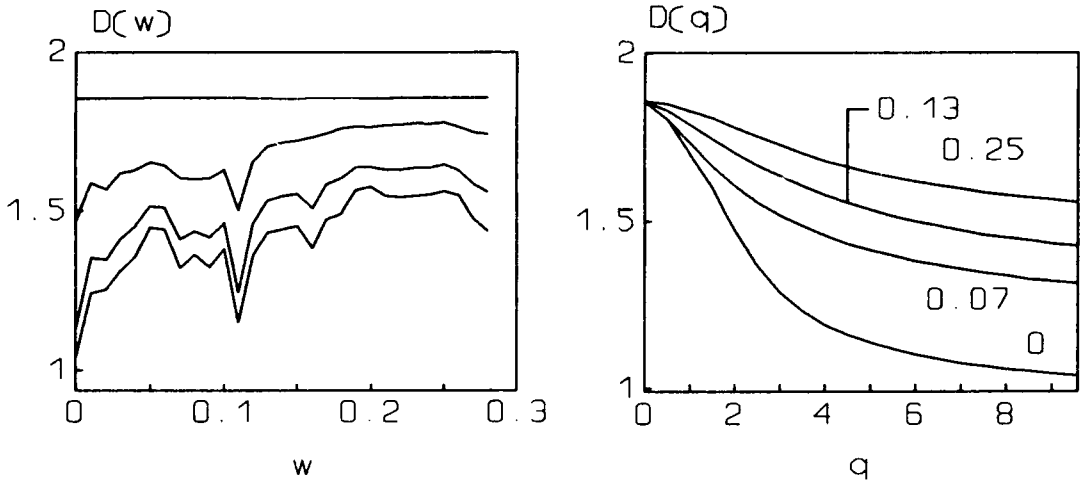


Figure 3c. Fractal analysis of the attractors depicted in Figs. 3a and 3b.

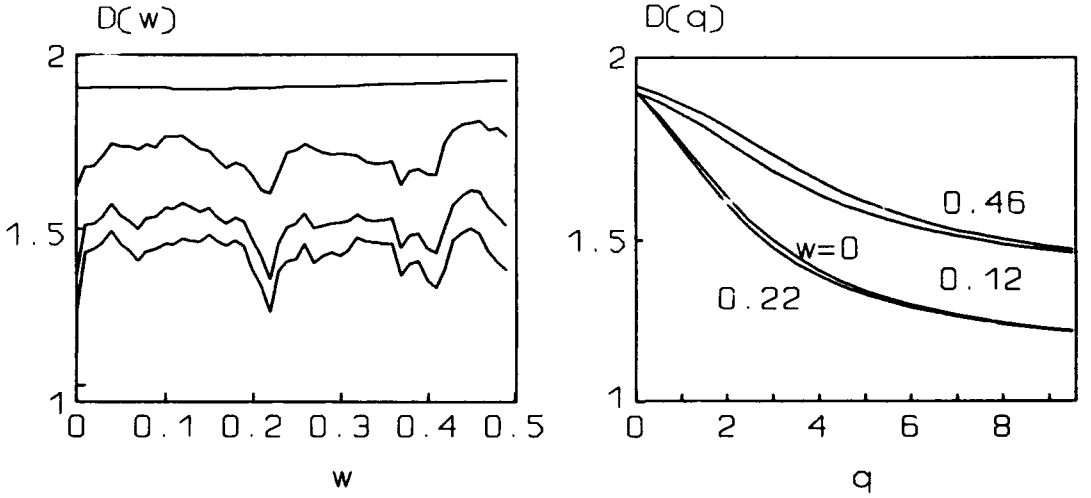


Figure 4c. Fractal analysis of the attractors depicted in Figs. 4a and 4b.

River and Lake Ice Conditions as Interpreted from Microwave Imagery

Ms. Rae Melloh* and Mr. Lawrence Gatto
U.S. Army Cold Regions Research and Engineering Laboratory
Hanover, New Hampshire 03755

ABSTRACT

Passive microwave (33.6 GHz) and synthetic aperture radar (C-,L- and P-band) images of ice covered rivers and lakes can provide information on types and distribution of river and lake ice needed for studying river ice processes and dynamics, and monitoring ice during winter. Visible and IR remote sensing systems do not provide comparable data and are hindered by darkness and cloud cover often encountered in northern regions. Our study has resulted in several findings that will be useful for imagery analysts: active radar systems provide the capability to distinguish rough-textured from smooth-ice covers, zones of variable ice surface roughness within rough ice fields can be differentiated, C- and L-band data are more sensitive than P-band to the range of surface roughnesses encountered, and polarization synthesis and signatures provide a way to improve contrast of ice features on radar imagery. K_a-band passive microwave systems allow the differentiation of wet versus dry snow conditions; open water in ice covered rivers and lakes; and compacted areas in depth hoar snowpacks. Possible applications of these results include location of river bridging sites, dangerous thin-ice zones, packed-snow trails, and preferential water supply points.

INTRODUCTION

Passive and active microwave sensors provide a tool for monitoring river and lake ice conditions. The larger area coverage afforded by airborne and spaceborne sensors provides a means of extrapolating known site conditions over long river reaches or lake surfaces. Unlike visible and short wavelength infrared, microwave systems are capable of sensing both through cloud cover and in darkness. Microwave sensitivity to surface roughness, volume inhomogeneities and dielectric properties of materials provides advantages for sensing conditions in ice/water environments. Active radar systems, for example, have been used to distinguish areas of freshwater lakes that were frozen completely to the bed from those that were not based on the strength of radar backscatter (1). Investigations

using passive microwave systems over freshwater lake ice have shown that shorter wavelength (.81 cm) radiometers are better suited for gathering information about the upper layers of ice and snow on the surface, while longer wavelengths (6 cm) may sense conditions involving the entire ice thickness (2).

The study objective is to obtain a better understanding of active and passive microwave signatures created by freshwater ice conditions. This information is needed in order to utilize current and future data sources fully. Current airborne passive microwave systems can provide adequate spatial resolution needed in river and small lake environments; however, this type of data is not widely available and its applications to freshwater ice studies have not been fully explored. The airborne synthetic aperture radar (AIRSAR) system used in this study is being developed and tested by NASA and the Jet Propulsion Laboratory (JPL) (3). The AIRSAR provides an opportunity to compare river and lake ice signatures on coincident images of three frequencies (C-, L- and P-band) and four polarizations (HH, HV, VV and VH); the first H or V corresponding to send polarization and the second H or V the receive polarization. Images can be synthesized from the data for any desired polarization orientation and ellipticity (3). Recent launches of SAR systems on space platforms include the European Space Agency's ERS-1 and Japan's ERS-1; proposed launches include Canada's Radarsat, the United States' Shuttle Imaging Radar and EOS platforms. The AIRSAR's polarimetric capability places it among the more advanced radar systems such as those planned for the United States' Shuttle Imaging Radar and EOS platforms. Our study is timely considering the increasing availability of SAR imagery, and to a larger user community.

BACKGROUND AND METHODS

During March 1988, images from both passive and active microwave systems were acquired over the Tanana River Valley in and near Fairbanks, Alaska. The two airborne systems flown were the NASA/JPL AIRSAR aboard a NASA DC-8 and the Naval Ocean Research and Development Activity (NORDA) K_a-band Radiometric Mapping System (KRMS) aboard a Navy RP-3. CRREL field teams obtained extensive ground and aerial observations and measurements at a variety of river sites and at Harding Lake for use in interpreting image patterns. Aerial observations, photographs and videotapes were taken to document regional conditions. Field observation and measurements were made to determine site-specific conditions.

The study area lies within the discontinuous permafrost zone of the interior of Alaska, about 160 km south of the Arctic Circle. The average annual temperature in Fairbanks is -3° C and precipitation averages 25.4 to 55.9 cm, including 76 to 152 cm of snow per year (4). The Tanana River changes from a braided river with unstable bars, multiple channels and a predominantly gravel bed upstream of Fairbanks to an anastomosing river with several main channels separated by stable vegetated islands and a sand

bed downstream of Fairbanks (5). Width of the active floodplain ranges from 610 to 2130 m. Mean monthly discharge in March was 161 m^3/s, the lowest discharge during the 1988 water year. Harding Lake is roughly circular in shape has a surface area of about 10 km^2 and an elevation of 220 km above sea level.

SENSOR CHARACTERISTICS

The ability to distinguish freshwater ice and snow conditions on microwave imagery depends on the characteristics of the sensor system and the physical properties of the materials being sensed. Passive systems measure naturally emitted radiation in terms of brightness temperature (T_b), which is the product of the physical temperature (T_o) of an object and its emissivity (E) at the wavelength being sensed, ($T_b = ET_o$). The ability to distinguish freshwater ice conditions in an environment, where physical temperatures vary little depends largely on differences in emissivity. For example, smooth, open water surfaces have a very low emissivity, resulting in the coldest brightness temperatures in a scene while wet snow is highly emissive, resulting in a much warmer signature. Active radar systems do not rely on naturally emitted radiation, but transmit energy of a given frequency and then record the magnitude of radar backscatter. Where surfaces are rough, relative to the wavelength, strong backscatter results. In contrast, low backscatter occurs where transmitted energy reflects specularly off of smooth air-water, air-ice and ice-water interfaces. The images constructed from brightness temperature or backscatter measurements display patterns and tonal variations that may be correlated with known site conditions.

The K_a-band Radiometric Mapping System (KRMS) is an airborne passive microwave imager that operates at a center frequency of 33.6 GHz. KRMS image data has been extensively used in polar ice research to distinguish ocean surface types such as open water, first year ice and multi-year ice (6,7). The imagery and analyses presented here summarize the first use of the KRMS for analysis of freshwater ice conditions (8). The sensor was flown at an altitude of 1524 m, resulting in an image width of 3.6 km. The angle of incidence along a scanline varies from 50° at the outer edges of the image to 0° at nadir, along the center-line of the image.

The NASA/JPL AIRSAR is capable of collecting three frequencies and four polarizations simultaneously, yielding twelve co-registered radar images during each flight. The image data were processed by JPL to a standard compressed four-look format. The compression algorithm spatially averages four adjacent pixels to reduce speckle variance (3, 9). Data were collected over a swath width of roughly 8 km, with the incidence angle across the swath varying between 22.5 to 52.5 degrees. Pixel resolutions are approximately 10 m.

RESULTS

Image interpretations are organized by those observed on rivers and those observed on lakes and by passive versus active systems. The passive microwave system, operating at 33.6 GHz (.89 cm) was sensitive to characteristics of the snow cover. KRMS image interpretations are thus primarily related to the presence or absence of snow, variation in snow grain size, snow wetness and occurrence of open water. The active radar system at longer wavelengths (5.7, 24, and 68.4 cm) was less affected by snow conditions and more affected by smooth and level versus rough or inclined surfaces.

K_a-band Radiometric Mapping System (passive microwave)

River Patterns

Ice jams: Freeze-up ice jams on the Tanana River are not readily distinguished from smooth, level ice on the KRMS images (Fig. 1). Active radar provided a superior ability to distinguish ice jams.

Hardpacked snow trails: Radiometrically warmer (dark) lines, faint across the ice jams, and more pronounced along the river (Fig. 1) are snowmobile, dogsled and our SUSV (Small Unit Support Vehicle) trails. Detection of the trails may be explained by compaction of the depth hoar into a hardpacked fine grained snow. The hardpacked snow appears radiometrically warm (dark) compared to undisturbed depth hoar snow. The large hoar crystals scatter microwave emissions (10,11) resulting in radiometrically colder (lighter) signatures for uncompacted depth hoar. Depth hoar crystals ranged generally from 1 to 12 mm in size, with an occasional columnar crystal as long as 20 mm near the base of the pack. Snow densities were as low as 0.15 g/cm^3 in undisturbed depth hoar. The density of compacted snow on a 4 m wide trail traveled by snowmobiles was 0.46 g/cm^3 in the top 7 cm and 0.33 g/cm^3 in the bottom 6 cm.

Open leads: The low emissivity of the smooth water surface compared to the snow or ice surfaces makes open water readily detectable on KRMS imagery. Radiometrically cold (light tone) open water is seen on the Chena River in Fairbanks, and in open leads on the Tanana River. Observations from bridges at Cushman Street and Old Steese Highway showed the radiometrically warm (dark) reach on the Chena River (Fig. 2) to be well traveled hardpacked snow over ice. The radiometrically cold (light tone) reach at Grant Street, Peger Road and University Avenue was open water bordered by snow-covered ice. The open lead on the Tanana River (Fig. 3) is in the vicinity of Delta and Clear Creeks southeast of Harding Lake. The open water signature is so distinct on KRMS images that automated classification using simple level slicing routines could be used to map it.

Wet Overflows: The low emissivity of a water surface may also result in radiometrically cold signatures for puddles of water on the ice. The second flight of the KRMS on March 11, followed two days of unseasonably warm weather with daily mean temperatures above freezing. Air temperatures

Fig. 1. KRMS image of ice jam at the Tanana and Chena rivers confluence.

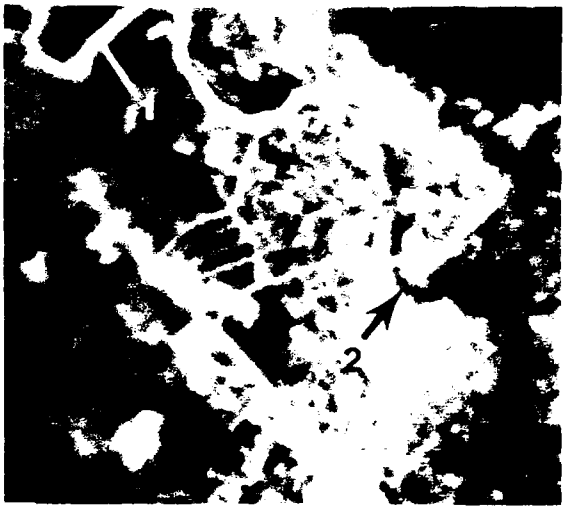


Fig. 2. Open water (1) and hardpacked snow over ice (2) on the Chena R. in Fairbanks.

Fig. 3. Open lead on the Tanana River southeast of Harding Lake near Delta Creek.

Fig. 4. Potential overflow on the Tanana River southeast of Harding Lake, near Little Delta R.

at the time of the March 11 flight were around 6.7° C. A feature that appears on the March 11 image (Fig. 4) but is absent on the March 8 image may be related to overflow of tributary streams or upflow through cracks due to a rise in river stage. Brightness temperatures of this feature approach that of open water, suggesting water overflow onto the ice surface. Unfortunately, we have no ground verification or video coverage of this feature.

Lake Patterns

Snow wetness: A comparison of images of Harding Lake acquired on March 8 and 11 show an overall increase of brightness temperature and decrease in image detail between the two dates (Fig. 5). Snow wetness was low or absent on March 8 and very high on March 11 in response to air temperatures at the time of overflight of 0° C and 6.7° C, respectively. Increased emission due to moisture in the snow results in increased brightness temperature and a darker image tone on the second flight date.

Fractures: The number and arrangement of the radiometrically warm (dark) linear features on the March 8 image of Harding Lake (Fig. 5) suggest that they originated over fractures in the ice cover. Ground observations and aerial videos were used to correlate these linear features with white snow ice superimposed on lake ice. The observations suggested that water had seeped through the fractures and wetted the overlying snow. A linear snow ice feature observed in the field was roughly 60 m wide, and lacked significant snow cover. The linear snow ice features are radiometrically warmer (darker), presumably because the melted, refrozen snow scatters less energy than the depth hoar over the surrounding snow-covered ice (8). The fracture patterns are less prominent on the second flight date when increased emission by the wet snow more closely matched the emission of the linear snow ice above the fractures (8).

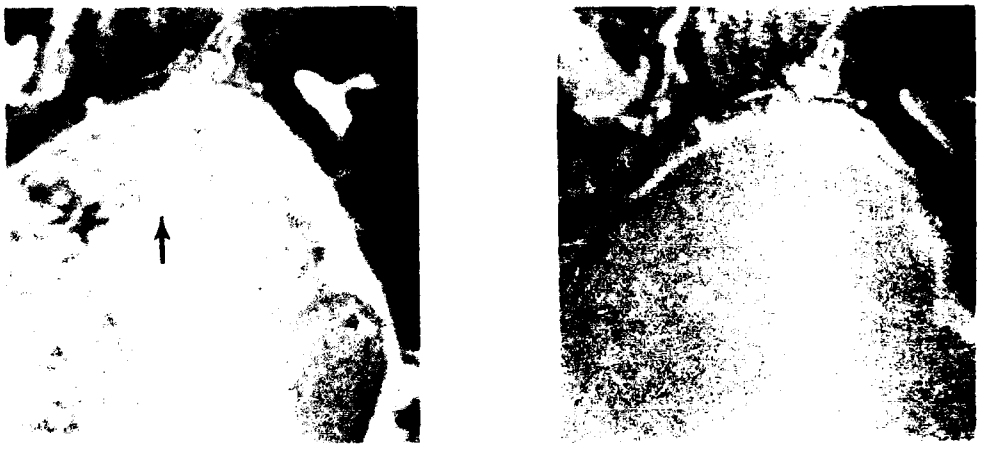


Fig. 5. Harding Lake KRMS images of March 8 (left) when fracture patterns are apparent and March 11 (right) when the snow is wet.

Fig. 6. NASA/JPL AIRSAR image (L- and C-band composite), 13 March 1988; ice jams (1), longitudinal patterns (2), smooth river ice (3), river bed and bars (4), open lead (5), brash ice (6), vegetated islands (7), and smooth lake ice (8). The Fairbanks International Airport is the large dark feature in the upper right portion of the image.

NASA/JPL AIRSAR (active radar)

River patterns

Ice jams: Obvious patterns on the SAR images of the Tanana River (Fig. 6, items 1a and 1b) are light (high backscatter) channel reaches within the relatively dark toned (low backscatter) floodplain. The high backscatter areas correspond to frazil ice floe accumulations that developed during freeze-up (Fig. 7); we refer to these frozen and snow covered frazil floe accumulations as ice jams. Within the ice jams shown on the composite C- and L-band image (Fig. 6, items 1a and 1b), the lightest (highest backscatter) portions are the downstream ends. The tone along the jams darkens in the upstream direction because the surface roughness (due to packing density, size of ice fragments or concentration) within the jams typically decreased upstream, resulting in less backscatter. Typical root-mean-square roughnesses in the ice jams were 4 cm for the snow surface and 10 cm for the ice surface. Root-mean-square roughness measurements at the downstream end of the southern ice jam (Fig. 6, item 1b) were 7 cm for the snow and 19 cm for the ice. An aerial video frame of the northern ice jam on 3 November, 1988 (Fig. 7) shows a gradation in ice fragment concentration along the accumulation prior to freeze-up. A photograph of this northern ice jam in March (Fig. 8) shows a moderately rough surface.

The longitudinal, linear, light (high backscatter) features within the jams (Fig. 6, item 2) probably delineate the main flow channel below the ice cover. The ice surface here may be rougher than along the channel margins because higher forces in the middle of the channel after 3 November (Fig. 7), may have resulted in thicker and rougher accumulations there. Ground electrical conductivity measurements (12) we made across the river channel just downstream of and through the middle of the northern ice jam suggest that the main channel is in the location indicated by the linear, high backscatter features.

Fig. 7. An ice floe accumulation at the northern jam site on 3 November 1987.

Fig. 8. The ice jam surface at the northern jam site in early March 1988.

Ice thickness in the smooth ice section downstream of the southern jam ranged from 51 to 76 cm and consisted predominantly of columnar ice over congealed frazil. The smooth ice downstream of the northern ice jam ranged from 35 to 69 cm and consisted primarily of columnar ice. Thickness of the rough ice at the southern ice jam, which consisted of congealed frazil with some columnar ice, ranged from 99 to 112 cm.

Smooth ice: Dark image patterns (low backscatter) within the Tanana and Chena rivers are formed by unjammed ice with a smooth, level surface (Fig. 6, item 3) such as occurs downstream of the two ice jams. Many of the lakes adjacent to the Tanana and Chena rivers occur in old gravel borrow pits and have surface areas of less than 10 acres. Most of these lakes appear dark on the SAR images (Fig. 6, item 8) and are difficult to discern from the surrounding abandoned river bed.

Open water: A dark linear pattern (low backscatter) within the Tanana River (Fig. 6, item 5) corresponds to an open water lead. Open leads were observed at numerous points along the river between Fairbanks and Harding Lake during the aerial reconnaissance.

Snow: Although the snow was substantially wetter on 13 March than on 19 March, differences due to the snow patterns on the images are subtle. Channel patterns and drainage ways in undeveloped, tree-covered areas of the Tanana River floodplain are more distinct in images acquired during thaw conditions. The contrast between the vegetated islands and the unvegetated bars of the Tanana River is also greater during thaw conditions. Dielectric changes in the tree trunks and canopy are thought to be the primary cause of the increased contrast between vegetated and unvegetated areas during the thaw; a concurrent study in the same flight path (13) described substantial increase in backscatter from forested areas during the warm flight date.

River bed and bars: The river bars and exposed river bed show up as mottled dark areas with variable light and dark tones caused by the irregular surface, scattered vegetation and ephemeral channels (Fig. 6, item 4). Some of the bars and exposed bed have light-toned lines within them on the SAR images that are likely scarps less than 0.3 to 1 m high which border these old channels.

Band and Polarization Comparisons

Composite multi-band and multi-polarization images (Fig. 6) will not always be available to the image analyst. The relative sensitivities of single band (C, L and P) co-polarized (HH,VV) and cross-polarized (HV,VH) radar configurations were compared (14) and are summarized here for discerning ice jam features. The single band and polarization images (Fig. 9) were synthesized from compressed Stokes matrix data provided by JPL (9, 16) and are uncalibrated.

Contrast with snow-covered river bed: The ice jams could adequately be distinguished from the surrounding snow-covered river floodplain on C-

and L-band co- and cross-polarization images but not on P-band images. The ice jam surfaces were not rough relative to the longer wavelength (68 cm), resulting in low contrast between the jams and other floodplain features on the P-band images. This is especially the case for the P-band cross-polarization image.

Contrast with tree-covered areas: The ability to distinguish the ice jams from the tree-covered areas along the river varies with wavelength, polarization and freeze or thaw conditions. Here we compare P-and L-band data (Fig. 9a) on the warm day (13 March), L- and C-band data (Fig. 9b) on the cold day (19 March), and warm and cold day L-band data.

On the cold day delineation of the boundary between tree-covered areas and ice jams was easiest on C-band co-polarization images, on which backscatter from the ice jams was stronger than from the tree-covered areas. On the L-band images backscatter from the tree-covered areas and ice jams was more similar; the VV image provided the best contrasting polarization.

On the warm day delineation of the boundary between the ice jams and the tree-covered areas was most satisfying on cross-polarization L-band (VH) images where backscatter from the tree-covered areas was stronger than from the jammed ice and the ice jams were discernible. On co- and cross-polarization P-band images high backscatter from the tree-covered areas contrasts well with low backscatter from the entire floodplain but the ice jams were not discernible.

Increased L-band backscatter from the tree-covered areas on the warm day (13) had a significant effect on our ability to delineate the trees from the ice jams. On co-polarization images the delineation was easier on the cold day; on cross-polarization images the delineation was easier on the warm day.

Linear longitudinal features: Ability to distinguish the linear longitudinal features within the ice jams varies with wavelength and radar look direction. The features are not apparent on the P- or C-band images but are apparent on L-band images on both the warm and cold days.

The orientation of a longitudinal feature relative to the radar look angle is more important for the cross-polarization than for the co-polarization L-band images. The longitudinal feature in the northern ice jam is nearly parallel with the radar look direction while the longitudinal feature in the southern jam is at an oblique angle (45°). Both longitudinal features are apparent on co-polarization images but only the obliquely angled feature is apparent on the cross-polarization image.

Gradation in Backscatter: The C-band cross-polarization (HV) image (Fig. 9b) shows the most sensitivity to the gradation in roughness along the ice jams. The decrease in backscatter in the upstream direction may be related to a gradation in fragment concentration or packing density (Fig. 7) resulting in reduced surface roughness in the upstream direction.

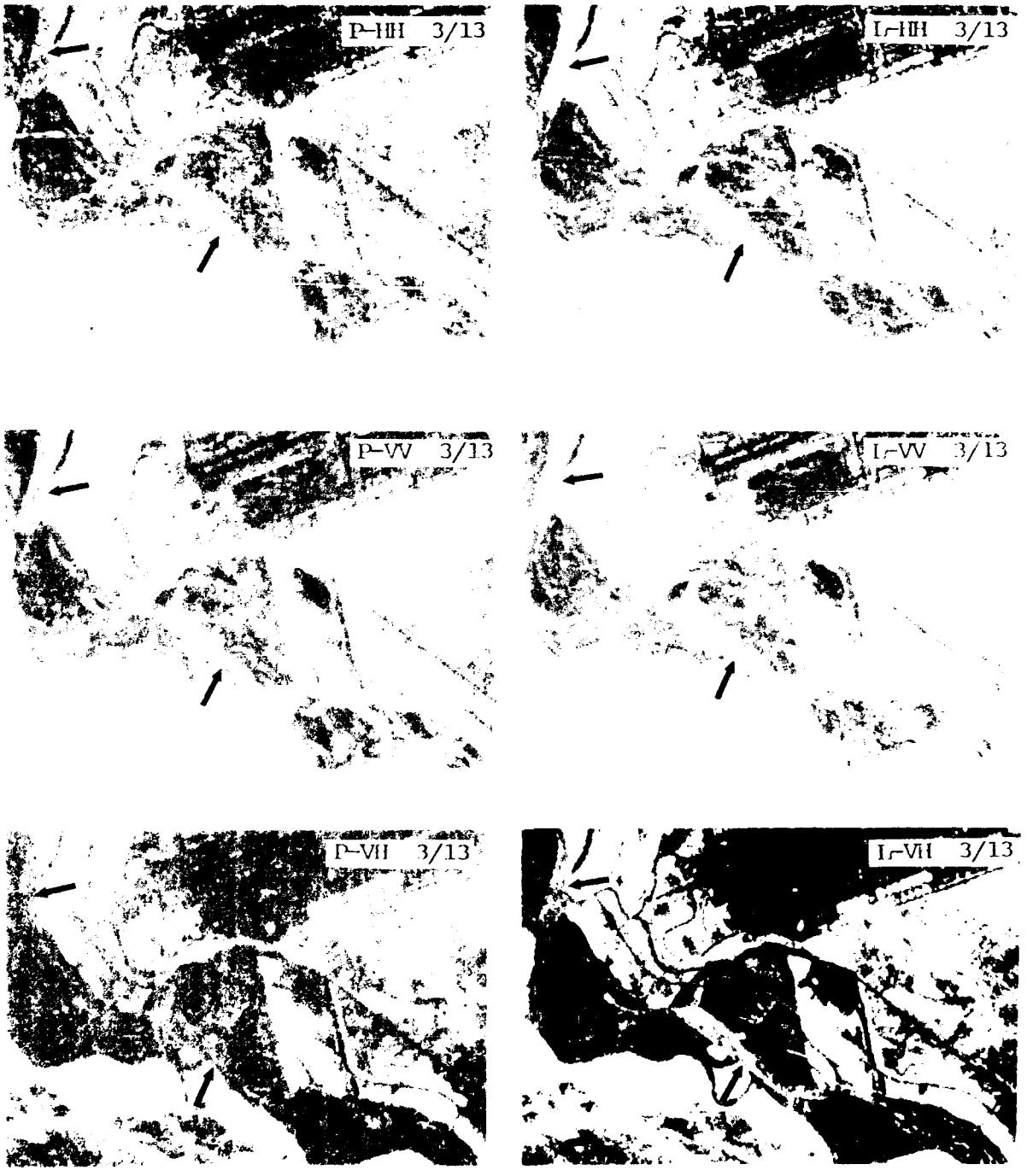


Fig. 9a. P- and L-band, HH-, VV- and VH-polarization images of two ice jams near the Tanana and Chena rivers confluence on 13 March.

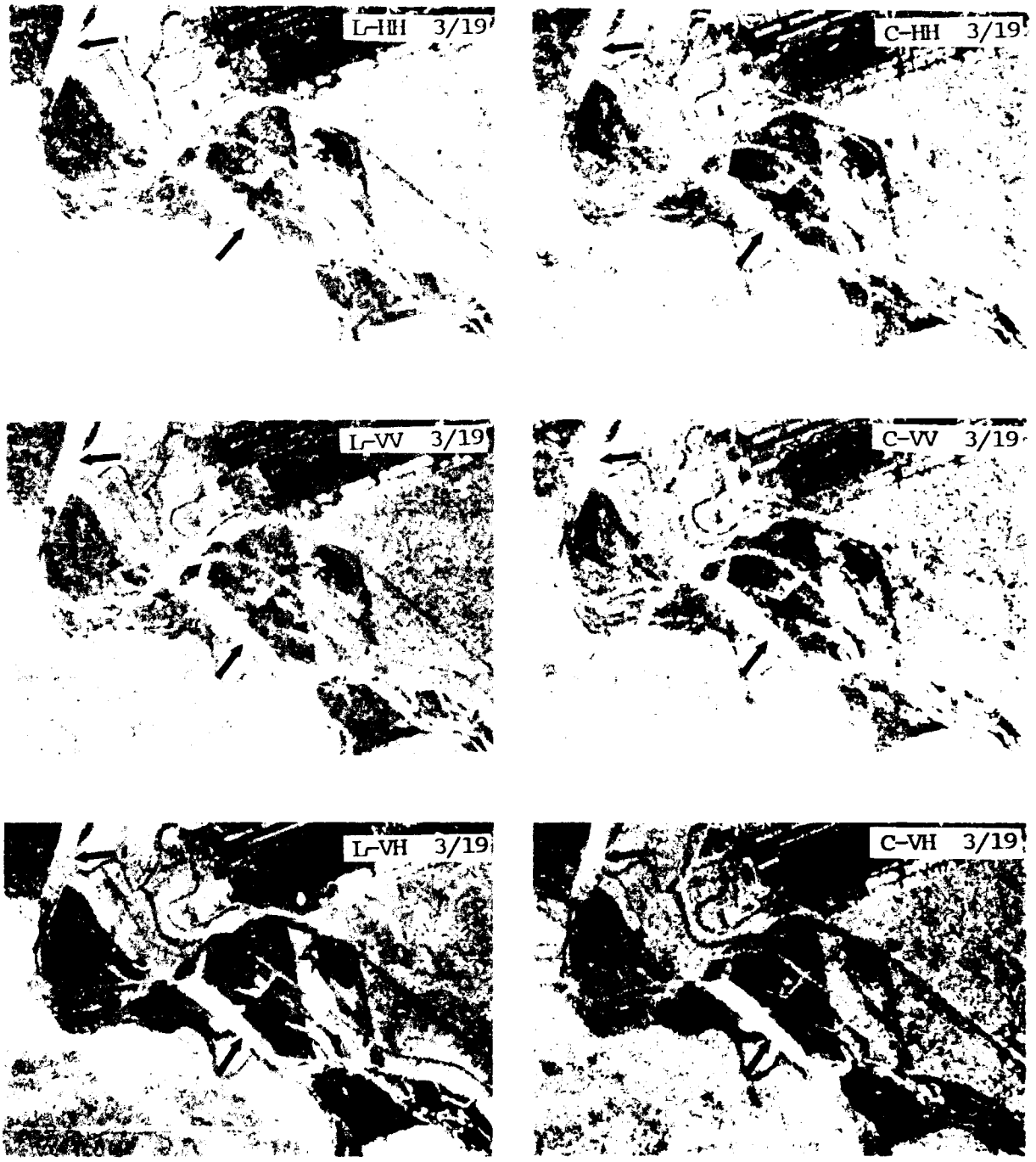


Fig. 9b. L- and C-band, HH-, VV- and VH-polarization images of two ice jams near the Tanana and Chena rivers confluence on 19 March.

Using Polarimetry to Increase Contrast

Polarimetric processing of AIRSAR data allows the synthesis of images that maximize contrast of surface types by chose of ellipticities and orientations for which polarimetric signatures (15) of two surface types differ the most. In order to study river ice on SAR images it is important to distinguish the ice flow accumulations from the surrounding tree covered islands and river banks. On the HH,HV,VV, VH polarization images we found this was not always possible.

L-band co-polarization signatures of the calibrated 13 March data were computed (16) for a tree covered area and rough ice (Fig. 10). The scattering mechanism in tree covered areas includes a double bounce from the ground and tree trunks and diffuse scattering in the canopy (13). The

Fig. 10. Co-polarization signatures for a tree covered (left) and rough ice (right).(15,16)

Fig. 11. A synthesized L-band image at 45° ellipticity and 0° orientation.

relatively high pedestal resulting from diffuse scattering (15) is evident in the tree signature. The co-polarization signatures for ice have a shorter pedestal indicating less diffuse scattering (Fig. 10). Synthesis of an image with an ellipticity of 45° (right-hand circular) dramatically improved the contrast between the ice jam and surrounding tree covered areas (Fig. 11) compared to the L-HH and L-VV images (Fig. 9).

SUMMARY OF POTENTIAL APPLICATIONS

Passive microwave and synthetic aperture radar sensors provide a data source that will be useful to the image analyst. Possible applications are suggested here.

Passive Microwave:

1. In areas where depth hoar snows occur, foot and vehicle paths on the ice cover can be detected.
2. Passive microwave surveys of lake and river surfaces in early spring could provide a map of open water and overflows that may indicate dangerous thin-ice zones.
3. Observations of water saturated snow cover where tributaries overflow onto the ice cover could be used as an early indication of the onset of snowmelt runoff.
4. Open leads and wetted snow cover can be detected; these could be used to monitor progressive features of river ice melt or breakup.

Synthetic aperture radar:

1. SAR images can distinguish rough ice surfaces from smooth ones. The rough ice surfaces would be difficult to traverse.
2. SAR images detect ice jam locations. Preferable field water supply points would be downstream of ice jams where subsurface frazil ice deposits are less likely to occur and ice is thinner.
3. Historical records of SAR images could be assembled to establish normal freeze-up conditions on a river including location and extent of ice jams.
4. Open leads on rivers can be identified and represent zones of potentially weak ice and dangerous locations for obtaining water or traversing the river.

References and Notes

1. W. Weeks, A. Fountain, M. Br[illegible]n, and C. Elachi, J. Geophys. Rsch., 83(C8), 4069 (1978).
2. D. Hall, J. Foster, A. Chang, and A. Rango, IEEE Trans. Geosc. and Rmt. Sen., GE-19(4), 189 (1981).

3. D. Held, W. Brown, A. Freeman, J. Klein, H. Zebker, T. Sato, T. Miller, Q. Nguyen and Y. Lou, The NASA/JPL multifrequency, multipolarisation Airborne SAR system. In Proceedings IGARSS '88, Edinburgh, Scotland, September 1988, Vol. 1, pp. 345-347.
4. L. Selkregg, Alaska regional profiles. Vol. 6, Yukon Region, (University of Alaska, Anchorage, 1974).
5. C. Neill, J. Buska, E. Chacho, C. Collins and L. Gatto, Overview of Tanana River monitoring and research studies near Fairbanks, Alaska. Special Report 84-37, U.S. Army CRREL, Hanover, NH. 03755-1290, (1984).
6. D. Eppler, L. Farmer and A. Lohanick, Digital Processing of passive K_a-band microwave images for sea-ice classification, NORDA, NSTL, Mississippi, Rpt. 51, (1984).
7. D. Eppler, L. Farmer, A. Lohanick and M. Hoover, Freshwater ice thickness observations using passive microwave sensors, IEEE Trans. Geos. Rmt. Sen., 91(C9), 10661 (1986).
8. R. Melloh, D. Eppler, L. Farmer, L. Gatto and E. Chacho, Interpretation of Passive Microwave Imagery of Surface Snow and Ice: Harding Lake, Alaska. CRREL Rpt. 91-11, U.S. Army CRREL, Hanover, NH. 03755-1290, (1991).
9. P. Dubois and L. Norikane, Data volume reduction for imaging radar polarimetry, Proceedings IGARSS, Michigan, USA, May 1987, Vol. 1, pp.691-696.
10. J. Shiue, A. Chang, H. Boyne and D. Ellerbruch, Remote sensing of snowpack with microwave radiometers for hydrologic application, In Proceedings of the 12th International Symposium on the Remote sensing of the Environment, Ann Arbor, Michigan, pp. 887-896 (1978).
11. M. Hallikainen, F. Ulaby and M. Abdelrazik, The dielectric behavior of snow in the 3- to 37- GHz range. In Remote Sensing From Research Towards Operational Use. International Geoscience and Remote Sensing Symposium (IGARSS '84), Strasbourg, August 27-30. European Space Agency, ESA SP-236, Vol.2, pp. 169-174 (1984).
12. S. Arcone, and A. Delaney, J. Glaciology, 33(115), 330 (1986).
13. J.Way, J. Paris, E. Kasischke, C. Slaughter, L. Viereck, N. Christensen, M. Dobson, F. Ulaby, J. Richards, A. Milne, A. Sieber, F. Ahern, D. Simonett, R. Hoffer, M. Imhoff, and J. Weber, Int. J. Remote Sensing 11(7) 1119 (1990).
14. R. Melloh and L. Gatto, Interpretation of synthetic aperture radar imagery of snow-covered river ice, In Proceedings of the USACE Seventh Remote Sensing Symposium, Portland, Oregon, May 1990, pp. 429-441.
15. T. Ray, T. Farr and J. van Zyl, Polarization Signatures for Abandoned Agricultural Fields in the Manix Basin Area of the Mojave Desert, In Proceedings of the Third Airborne Synthetic Aperture Radar (AIRSAR) Workshop, May 1991, pp.117-125.
16. We acknowledge E. Rignot and R. Kwok of the Jet Propulsion Laboratory for providing the polarimetric signatures, and F. Burnette for providing the software to synthesize images from the Stokes matrix data.

Microwave Solid-State Transmitter for Miniature Multiband Beacon Transponder

*Muhammad Mizan, Mr., Leo DiDomenico, Mr.,
Ted Lukaszek, Mr., Ken Klohn, Mr.,
US Army LABCOM, Electronics Technology & Devices Laboratory,
SLCET-MP-C, Fort Monmouth, NJ 07703-5601

ABSTRACT

An X-band solid-state transmitter has been designed to satisfy the frequency stability and the power output requirements of a Miniature Multiband Beacon (MMB). The transmitter consist of a highly stable dielectric resonator oscillator (DRO) and a high power GaAs FET amplifier module. The DRO exhibited a frequency stability of 65 ppm over the temperature range of -50°C to +50°C. The DRO also exhibited a single-sideband (SSB) phase noise of -65 dBc/Hz and -93 dBc/Hz at 100 Hz and 1 kHz carrier offset frequencies, respectively. This phase noise and frequency vs temperature data represent state-of-the-art performance for a 9 GHz DRO. The high power, high efficiency, three-stage amplifier module was designed to amplify the DRO output power to 10 watts. The 10 watt output power is then fed into a 16 element microstrip patch antenna having a gain of 21 dB and a half power beamwidth of 14°. The total radiated power was calculated to be about 1200 watt which is sufficient to meet MMB's range requirement.

INTRODUCTION

The availability of sophisticated and reliable electronic systems to support Army troops can often make the difference between victory and defeat. One such system, an Army radar transponder AN/PPN-19, is essential to our Special Operating Forces (SOF), but the current technology employed relies on components which do not provide optimum performance by today's standards. The unit is large in size, heavy and some of its components are power consuming and not as frequency stable as required.

The man-portable radar transponder is deployed by long-range SOF patrols for use as a forward air controller/strike aircraft rendezvous

aid. It provides the location of friendly troops and/or can mark an offset aiming point for close air support. In order to operate in a covert/secure manner without risk of interception, the unit must be capable of immediate on/off signal transmissions when interrogated by incoming aircraft. Current units require a substantial warm up period before attaining the required operating frequency to assure transmission within the bandwidth of the aircraft receiver. This additional power drain mandates the use of a battery larger than its operating requirement thus contributing significantly to the beacon's overall size and weight.

The U.S. Army Training and Doctrine Command (TRADOC) has recently established a ROC for its replacement, referred to as the Miniature Multiband Beacon (MMB). An analysis of the unit determined that most of the size, weight and power consumption is associated with the battery and the transmitter portion of the system. The ETDL team had designed and demonstrated significant candidate improvements to the system through their in-house research and development efforts with stable, highly efficient dielectric resonator oscillator and solid-state FET amplifier circuits. This in turn reduced the power consumption requirements allowing the use of a smaller battery.

Furthermore, according to Army doctrine, the duration of future wars can no longer be thought of in terms of years, but in terms of weeks or months. In the spirit of the Army doctrine, the in-house developments reported herein promise to significantly reduce the time required to field an improved version of the PPN-19.

DIELECTRIC RESONATOR OSCILLATOR (DRO) AND HIGH POWER AMPLIFIER DESIGN AND PERFORMANCE

The two most critical elements of a solid-state transmitter are the frequency source and the high power amplifier module. A DRO was chosen as the frequency source for its ultra-stable frequency vs temperature characteristic, simple construction, spurious-free RF output spectrum, fundamental frequency of operation from 2 GHz to 30 GHz, low cost, power efficiency and extremely low phase noise.[1-4] The amplifier module, capable of 10 watts of output power, was designed using commercially available GaAs FETs procured from Fujitsu, Inc. The DRO and the high power amplifier module design is described in the following paragraph.

Initially, a single-stage feedback-loop DRO was designed with a Murata Erie zero ppm/^{o}C dielectric resonator (DR) and the GaAs FET amplifier constructed on 10 mil thick 5880 RT/duroid board. The microstrip circuit board layout of the single-stage DRO is shown in

Figure 1. The oscillator was designed to operate at 9.3 GHz with 15 dBm of RF output power. The frequency vs temperature characteristic of the DRO is shown in Figure 2. The center frequency shift from -45°C to +55°C is less than 100 ppm (less than 1 MHz at 9.3 GHz).

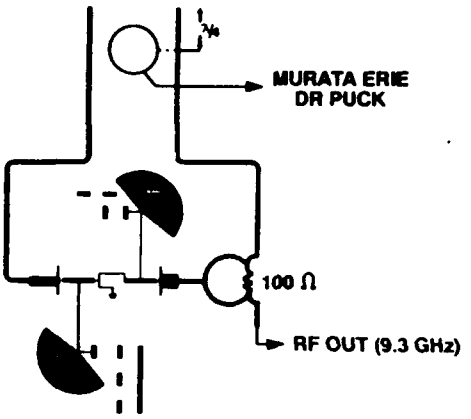


Figure 1 Dielectric Resonator Stabilized Single-Stage Feedback-Loop Microstrip Circuit Board Layout

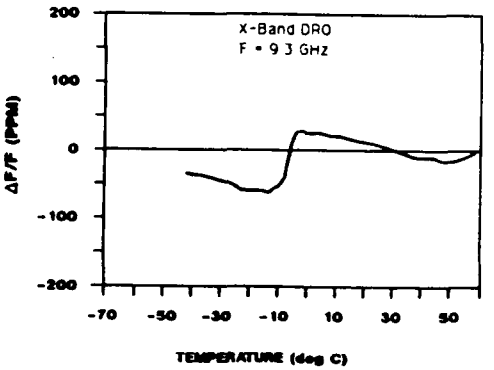


Figure 2 Frequency vs. Temperature Plot of Single-Stage DRO

Although the frequency stability of the single-stage DRO was good enough to satisfy the MMB's requirements, a second effort was made to further improve the frequency stability of our initial DRO. A two-stage DR stabilized feedback-loop oscillator was designed, fabricated and tested. The circuit board layout of this second oscillator is similar to the first one shown in Figure 1, except in this case two amplifiers were used to double the loop gain. Because of the additional loop gain we were able to weakly couple the DR and realize a very high loaded Q, on the order of 13,000 at 9 GHz. A plot of frequency vs temperature measurements for the two-stage DRO is shown in Figure 3. The total frequency drift from -50°C to +50°C is only 65 ppm. This data is about two times better than the single-stage amplifier DRO design and about five times better than commercially available production DROs. In addition to having an excellent frequency stability, the two-stage DRO also exhibited extremely low phase noise. The absolute phase noise was measured by using a Hewlett Packard (HP) 3047A phase noise measurement system. The oscillator exhibited a single-sideband (SSB) phase noise of -65 dBc/Hz at 100 Hz offset frequency. The phase noise measurement test setup and the phase noise plot are shown in Figures 4 and 5, respectively. The measured phase noise data is significantly better than previously published data.

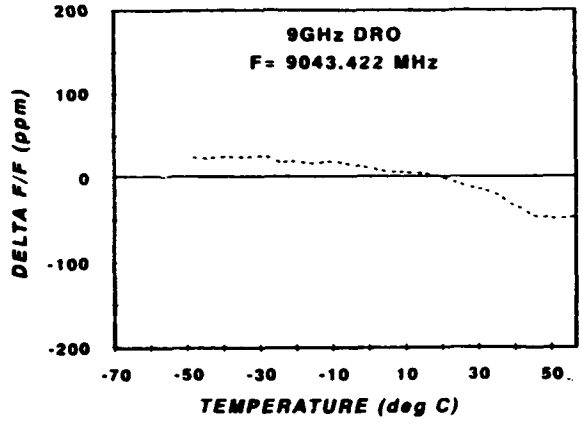


Figure 3 Frequency vs. Temperature Plot of Two-Stage DRO

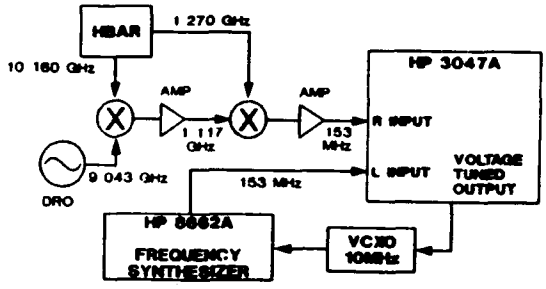


Figure 4 9GHz DRO Phase Noise Measurement Test Setup

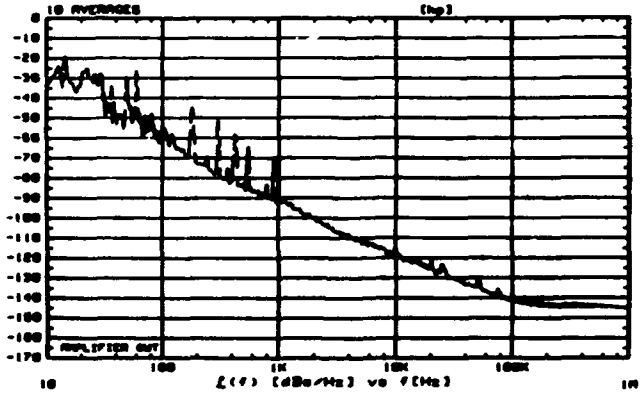


Figure 5 Absolute Phase Noise of 9GHz DRO

The high power amplifier module consists of three stages of amplification and required 21 dBm of input power to provide 40.13 dBm of output power. The microstrip circuit board layout of this amplifier is shown in Figure 6. The measured small signal gain is provided in Figure 7 indicating greater than 27 dB. The input and the output VSWR, much better than 1.5:1, are shown in Figures 8 and 9, respectively. An HP 8510B Network Analyzer was used to measure the above parameters. A single-stage preamplifier was designed to provide the power amplifier module with the necessary drive level. The block diagram of our X-band transmitter is given in Figure 10. The RF output power was measured by turning the gate bias on and by switching the drain supply on and off manually. The output power measurement test setup is shown in Figure 11. Total line loss between the transmitter output port and the spectrum analyzer including the 20 dB pad is about 20.9 dB. The RF output power as measured on the spectrum analyzer is shown in Figure 12 and the cw output after taking the line loss into account is 10.3 watts. For pulsed power operation, the output device can be driven harder and output power of 41 dBm is easily attainable. A photo of the X-band DRO/High power amplifier circuit development is shown in Figure 13.

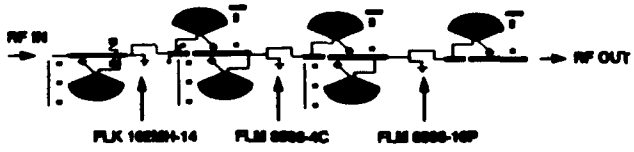


Figure 6 Circuit Components of GaAs FET Amplifier Module

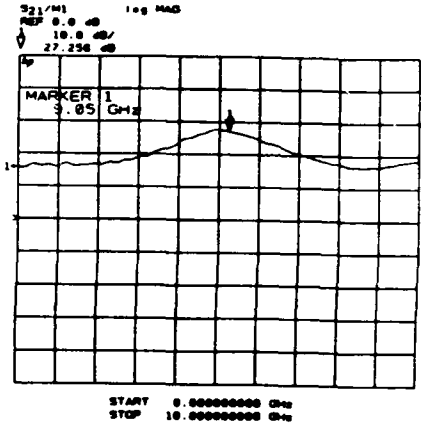


Figure 7 Small Signal Gain Performance of 10 Watt Amplifier Module

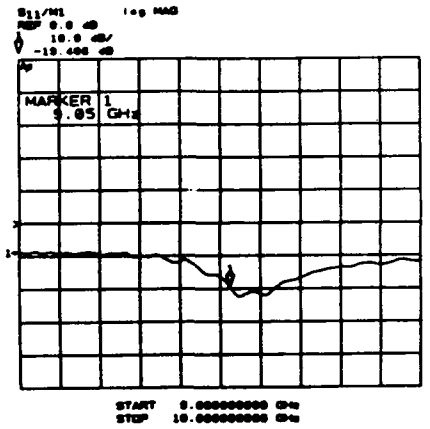


Figure 8 Input VSWR of 10 Watt Amplifier Module

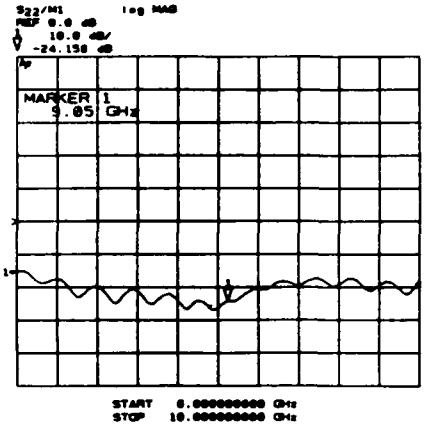


Figure 9 Output VSWR of 10 Watt Amplifier Module

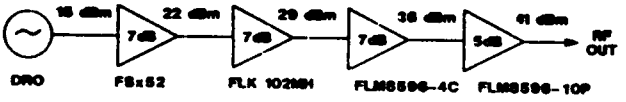


Figure 10 Solid-State X-Band Transmitter

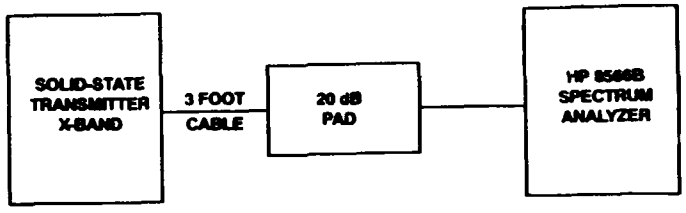


Figure 11 Transmitter RF Output Power Measurement
Test Setup

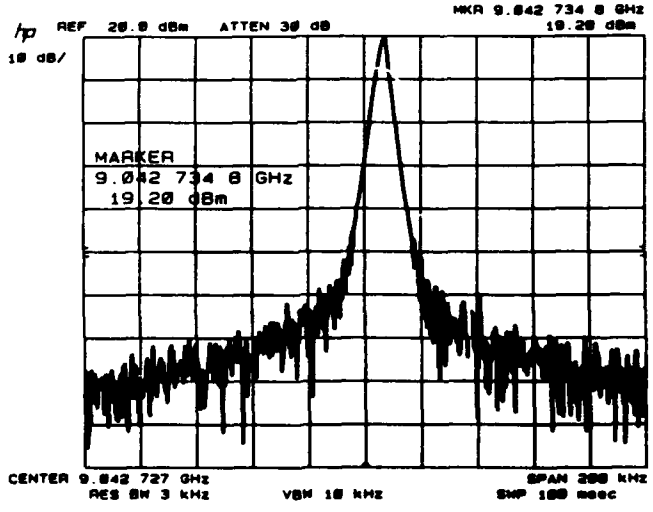


Figure 12 Transmitter RF Output Power

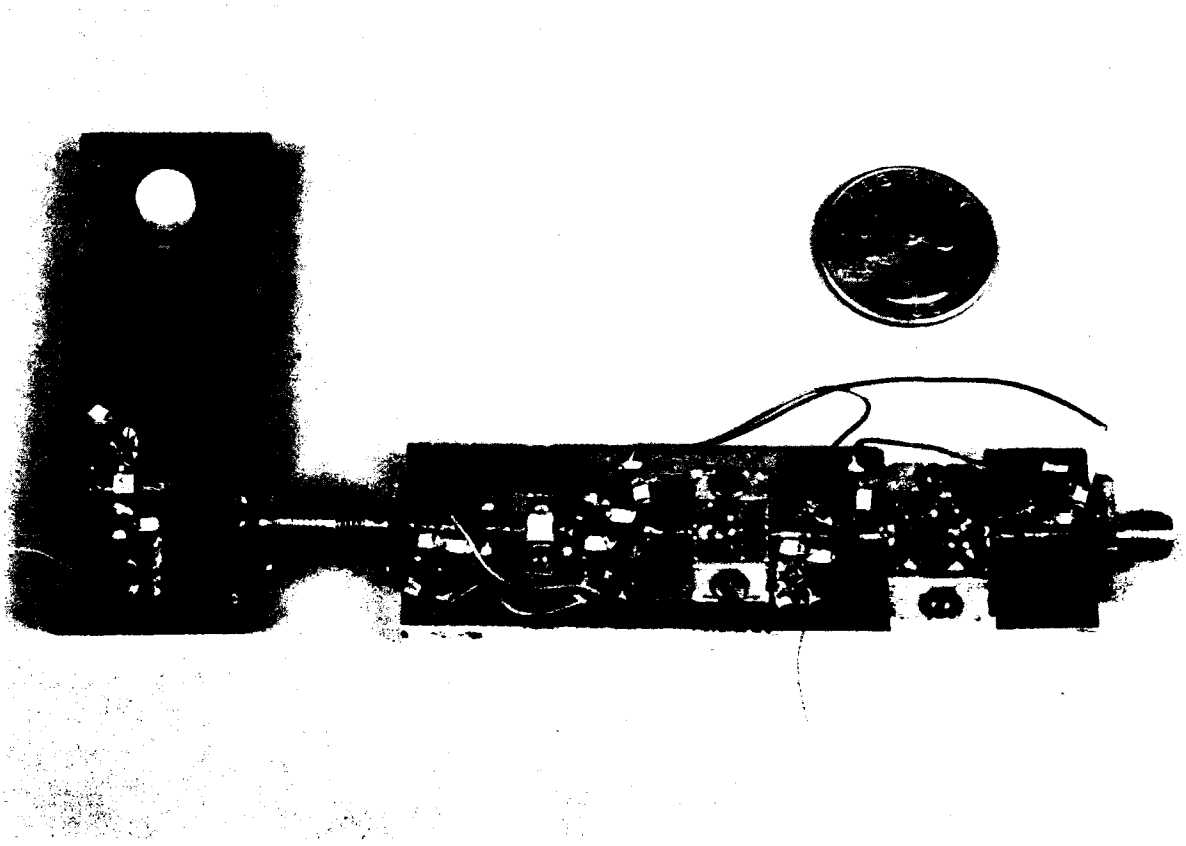


Figure 13. DRO/High Power Amplifier Circuit

PLANAR ARRAY ANTENNA

The initial signal radiation tests of the high power low phase noise oscillator required an antenna. The first order requirements for the antenna are that it be of low profile, of compact design and capable of producing a gain comparable to horn antennas and easy to construct. The initial antenna type chosen was a simple 4x4 corporate feed microstrip patch array antenna. In the initial experiments polarization diversity and minimization of sidelobe levels were not a major driving factor and hence these aspects were deferred to a future date.

The basic antenna layout is shown in Figure 14. The entire structure is on 508 micron (20 mil) thick duroid with a relative dielectric constant of 2.2 and a cladding of copper as the field supporting medium. The overall input characteristic impedance of the antenna is 50 ohm. This is achieved by the use of a standard power division and quarter wave transformation circuit on the same side of the antenna structure as the actual radiators. The exact dimensions and the characteristic impedances are shown in Figure 15. The antenna gain was empirically found to be about 21 dB above isotropic radiation, this is in good agreement with the theory for the given frequency of operationand dimensions. In addition, the antenna has input SWR of 1.4 and an antenna gain pattern as given in Figure 16.

CONCLUSIONS

As a result of this developmental effort, ETDL's technical capability in the area of frequency stable microwave oscillators and their accompanied high power solid state amplifiers circuits have been significantly enhanced. The achievement not only satisfied a real and present Army requirement but established the scientific basis for obtaining even greater technical improvements with the potential for impacting other current and/or future military systems.

REFERENCES

(1) M. Mizan and R. McGowan, "Ultra-Low -Noise, 8.3 GHz Dielectric Resonator Oscillator", Proceeding of 45th Annual Symposium on Frequency Control 1991, pp. 693-699. IEEE Catalog No. 91CH2965-2.

(2) M. Mizan and R. McGowan,"Extremely Low-Phase Noise X-Band Field Effect Transistor Dielectric Resonator Oscillator", 1991 IEEE MTT-S Digest,AA-5, pp. 891-894. IEEE Catalog No. 91CH2870-4.

(3) M. Mizan, R. McGowan, T. Lukaszek and A. Ballato,"Temperature Stable,Low-Phase Noise 2 GHz Dielectric Resonator Oscillator", 1991 IEEE MTT-S Digest, LL-4, pp. 1183-1186. IEEE Catalog No. 91CH2870-4.

(4) M. Mizan, R. McGowan, T. Lukaszek and A. Ballato,"Determination of the Limiting Factors in the Absolute Phase Noise of an L-Band Dielectric Resonator Oscillator", Proceeding of 45th Annual Symposium on Frequency Control 1991, pp. 687-692. IEEE Catalog No. 91CH2965-2.

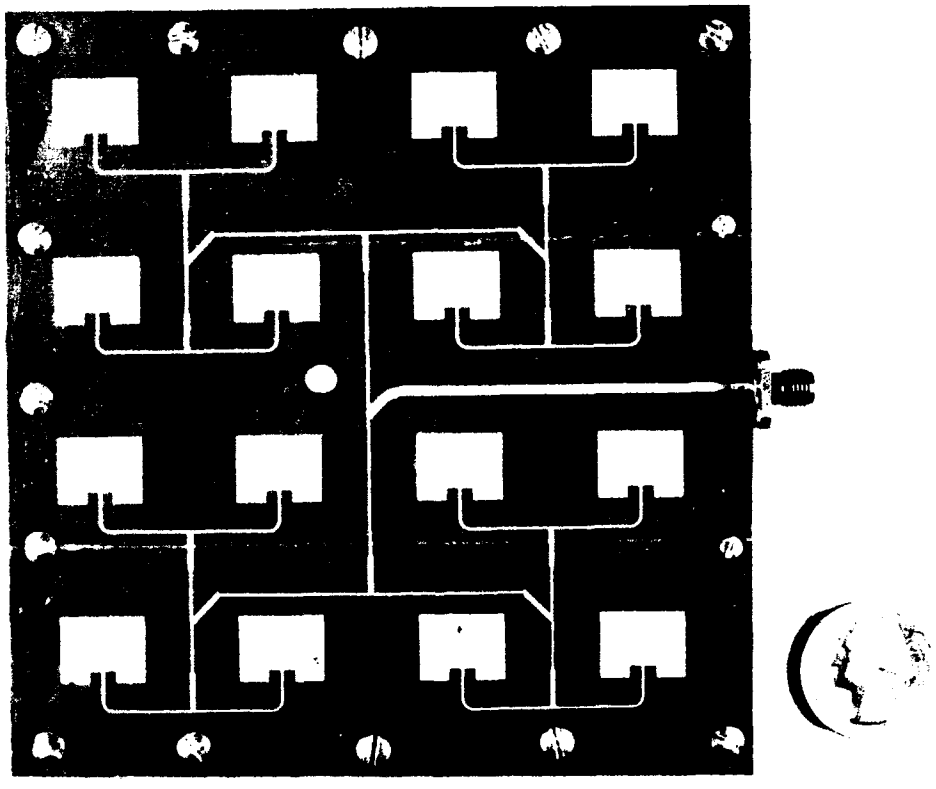


FIGURE 14 ANTENNA LAYOUT

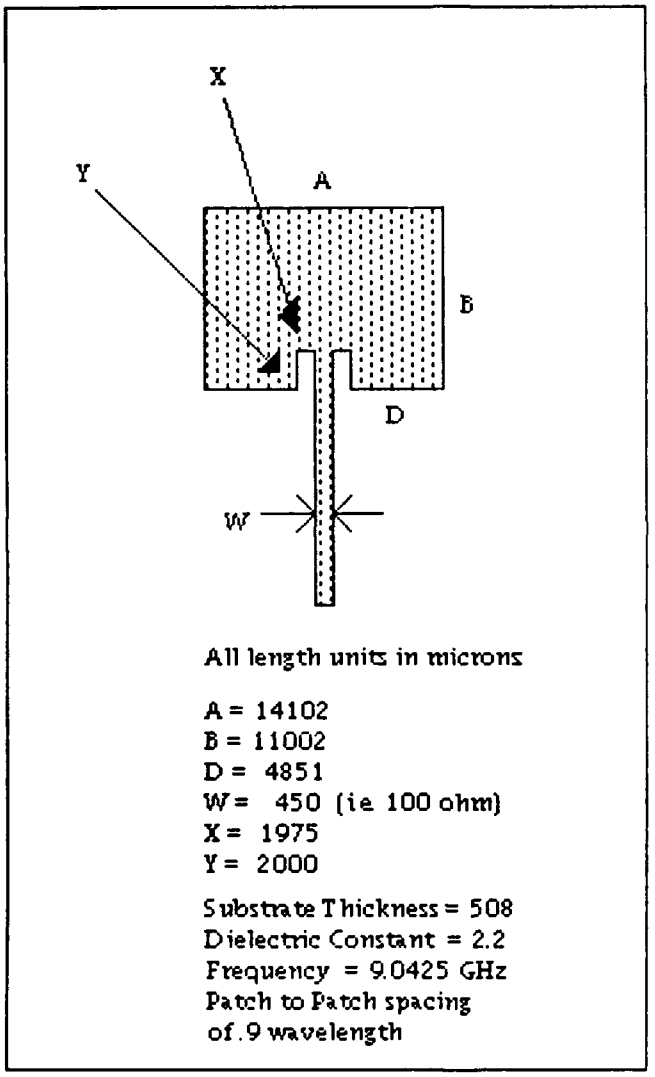
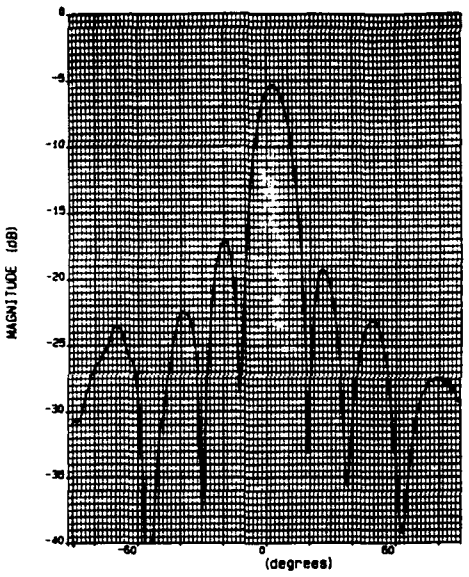
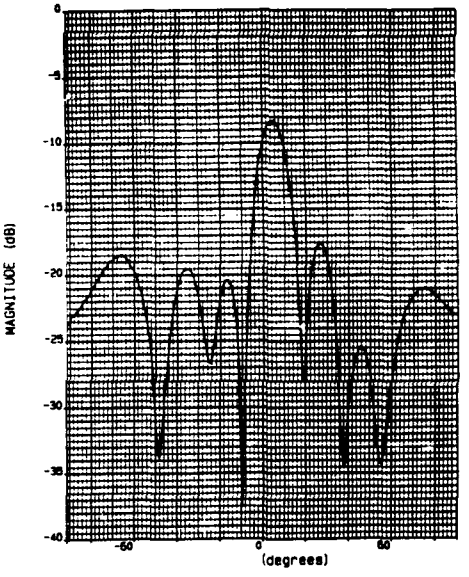


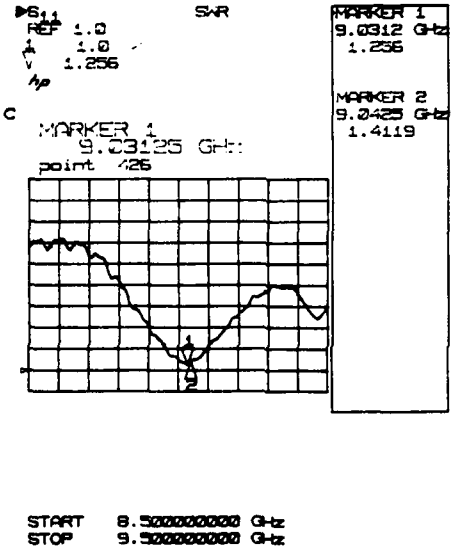
Figure 15 Patch radiator dimensions



A



B



C

Figure 16, A) H-Plane, B) E-Plane, C) SWR

Chromium-free Carbon, a New Adsorbent for Chemical Warfare Agent Filters

Mr. Robert W. Morrison*
U.S. Army Chemical Research, Development and Engineering Center
Aberdeen Proving Ground, MD 21010-5423

Dr. David T. Doughty
Calgon Carbon Corporation
Pittsburgh, PA 15230-0717

INTRODUCTION

Ever since the first use of toxic chemicals as a weapon during World War I, activated carbon impregnated with reactive chemicals has been used to provide respiratory protection for military personnel.[1] One of these early impregnated carbons used copper oxide as the reactant to serve as an acid gas sink. As new chemical threats were identified, additional reactants were added to the carbon to detoxify them. During World War II hexavalent chromium was added as a reactant to provide protection against hydrogen cyanide and cyanogen chloride. Since that time some forms of hexavalent chromium have been determined to be human carcinogens, and as a result materials which contain this chemical fall under stringent environmental regulation. Because this hexavalent chromium impregnated adsorbent (ASC Carbon) is used in all U.S. military chemical warfare agent filters, the Department of Defense spends over $2,000,000 per year to dispose of spent filters in accordance with environmental regulations. This paper describes the development of an environmentally acceptable adsorbent for use in military chemical warfare agent filters.

BACKGROUND

In the early 1980s, with the increasing attention to environmental matters, the presence of hexavalent chromium on ASC Carbon became an important issue of concern. ASC Carbon is activated carbon which has been impregnated with copper, silver and chromium compounds to provide high levels of filtration performance against all chemical warfare agents. The activated carbon is effective in removing low vapor pressure chemical warfare agents such as the nerve and mustard agents by physical adsorption. However, the high vapor pressure agents such as the blood agents, cyanogen chloride and hydrogen cyanide, are not retained well on activated carbon by physical adsorption, and thus, chemical reactants (usually called impregnants) are applied to the surface of the larger pores of the activated carbon. These chemical reactants convert the toxic chemical agents to reaction products (such as copper chloride) that are retained by

the carbon and to nontoxic volatile chemicals (such as carbon dioxide) that pass out of the filter in the purified air.

Each year as much as a million pounds of ASC Carbon are produced for the manufacture of U.S. military filters used to replace spent filters requiring change-out. These filters are located world-wide and must be disposed of in accordance with environmental regulations. Because of the hexavalent chromium, ASC Carbon has been shown to be a hazardous waste under Resource Conservation and Recovery Act (RCRA) regulations.[2] As a RCRA hazardous waste, ASC Carbon filled filters are subject to extremely expensive disposal procedures which include not only the costs associated with the actual disposal, but also those required to retrieve the filters from their use location.

In addition to the high disposal costs associated with filters containing ASC Carbon, the carcinogenic chromium also poses potential physiological risks to personnel who could potentially come in contact with the carbon.[3] These personnel include those involved in the manufacture of the carbon and filters that contain the carbon, in the event that appropriate protective measures are not used to keep ASC Carbon dust out of the breathing air. Even the soldiers, sailors and airmen, who rely on the filters for chemical protection, could be at slight risk of physiological harm in the event that damage to the filter occurs during use and dust from the chromium impregnated carbon exits the canister in the breathing air.

BEGINNING OF THE CHROMIUM-FREE CARBON DEVELOPMENT PROGRAM

To address these concerns resulting from the hexavalent chromium, the U.S. Army Chemical Research, Development and Engineering Center (CRDEC) began to develop a nonhazardous adsorbent for the filtration of chemical warfare agent vapors. Beginning in 1982, a preliminary screening of options for a chromium-free impregnant formulation was conducted under contract with Westvaco, Inc.[4] Work under this contract demonstrated that alternative nonhazardous impregnant formulations could provide sufficiently high levels of filtration performance to meet filter requirements. Under this preliminary contract, little effort was undertaken to find the optimum impregnant formulation.

In 1986 the Chromium-free Carbon development program intensified with the letting of a contract to Calgon Carbon Corporation with the primary objective of developing a new chemical warfare agent adsorbent which does not contain any hazardous materials while maintaining the current high level of agent filtration performance. The program had the secondary objective of reducing the concentration of ammonia which off-gases from filters when first placed in service. Ammonia, a residual product of the production process, can be evolved at levels that are irritating to the user and can result in reduced confidence in the effectiveness of the protection being provided. The contractual effort consisted of the following phases: impregnant screening, performance optimization, process development, and performance qualification of the carbon.

SCREENING OF CHROMIUM REPLACEMENT OPTIONS

During Phase 1 of the Chromium-free Carbon Development Program, Impregnant Screening, an assessment of the total role of the hexavalent chromium in removing chemical warfare agents was undertaken. This assessment determined that the chromium was only important in providing protection against cyanide containing blood agents such as hydrogen cyanide and cyanogen chloride.[5] The role of the hexavalent chromium in the filtration of hydrogen cyanide by ASC Carbon is to react with the toxic product cyanogen produced by the reaction of hydrogen cyanide with the cupric oxide impregnant. Evidence indicates that the hexavalent chromium plays a role in converting the cyanogen to oxamide which is retained by the activated carbon by physical adsorption.[6] The role of the hexavalent chromium in cyanogen chloride filtration is not as well understood. Pytleweski has reported that hexavalent chromium complexed with divalent copper results in the decomposition of cyanogen chloride by ASC Carbon.[7] For no other chemical warfare agent was a significant role identified for hexavalent chromium in the removal of that agent by ASC Carbon. Based on this determination, a literature search was conducted to identify reactants for hydrogen cyanide and cyanogen chloride. The field of potential candidates was initially reduced based on toxicological considerations. Those candidate impregnants for which carcinogenicity, mutagenicity or high acute toxicity had been reported were eliminated from consideration. This criteria, of course, was essential, but it did greatly add to the difficulty of the task, because those chemicals which do exhibit high levels of reactivity tend also to be toxic. Several very promising candidates (e.g. nickel and cobalt), for which cyanide reactivity had been reported in the literature, were excluded based on toxicity considerations.[8,9] Over 100 different impregnated carbons, of sufficiently low toxicity, were prepared and tested for cyanide filtration performance. These carbons varied in the actual chemical impregnants used, the loading levels of the impregnants and specific treatments for the carbon surface. The impregnants consisted of those which stoichiometrically reacted with the agents and those that acted catalytically to decompose the agents.

As a result of the Phase 1 work, the five most promising candidate impregnant formulations were identified. All of these formulations were prepared on the identical activated carbon substrate as is used for ASC Carbon. By so doing and by limiting the total loading of metal impregnants to 12 percent, the very high physical adsorption capacity that exists for ASC Carbon for nerve and mustard agents would not be compromised by the Chromium-free formulation. All of these candidate formulations included copper and silver to retain established reactivity toward numerous reactively removed agents (e.g. phosgene and arsine). The potential chromium replacements in the five formulations were molybdenum, vanadium, zinc, manganese and tungsten. All five of these formulations were prepared with and without triethylenediamine as a potential aid to cyanogen chloride removal.

DETERMINATION OF THE OPTIMUM FORMULATION

Under Phase 2 of the program, Performance Optimization, the five candidate impregnant formulations were optimized for overall filtration performance against all agents of concern.[5] The most effective loading

level for each impregnant in each of the five formulations was determined by statistical contour mapping of agent removal performance as a function of impregnant loading level. As a result of this formulation optimization, an impregnant composition of copper oxide, zinc oxide, silver and triethylenediamine was initially identified as the best replacement for ASC Carbon. This Chromium-free Carbon formulation was named ASZ-TEDA Carbon. During subsequent testing of this ASZ-TEDA Carbon to determine the effect of long term exposure to a variety of temperature and humidity conditions on agent filtration performance, the hydrogen cyanide filtration performance of the carbon was found to fall to unacceptably low levels after exposure to high humidity environments. This poor performance was manifested by premature breakthrough of cyanogen from beds of the carbon when challenged with hydrogen cyanide. The cupric oxide on the ASZ-TEDA Carbon was converting the hydrogen cyanide to cyanogen, without an effective reaction path for the decomposition of the cyanogen being provided by the sorbent. The reason for the degradation of the initially high level of hydrogen cyanide performance for fresh carbon to very poor performance after exposure to humidity was not immediately apparent. Spectroscopic analyses (X-ray photoelectron spectroscopy and energy dispersive X-ray spectroscopy) along with scanning electron microscopy were performed to identify differences in the chemical state and distribution of the impregnants before and after humidity exposure.[10] These analyses determined that while the chemical state of the impregnants did not change, the location of the impregnants was altered by the moisture. For the fresh carbon, zinc was heavily concentrated on the external surface of the granule while the cupric oxide was dispersed throughout the granule. Figure 1 is a scanning electron microscopy image of the cross-section of an unaged ASZ-TEDA Carbon granule. The figure shows a high concentration of impregnants on the exterior surface. Energy dispersive X-ray spectroscopy determined this heavy impregnant concentration to be primarily zinc. Thus, the hydrogen cyanide challenge to the carbon finds the zinc to be more accessible than copper, and reaction to zinc cyanide is favored. No cyanogen is produced from this reaction.[11] As the carbon is exposed to moisture, the cupric oxide was found to migrate to the external surface of the granule. With this increased accessibility of the cupric oxide, the hydrogen cyanide reacts with it to form cuprous cyanide and cyanogen according to the reaction:[12]

$$4\ HCN + 2\ CuO = 2\ CuCN + (CN)_2 + 2\ H_2O$$

To address this problem of premature cyanogen breakthrough, a modification to the ASZ-TEDA impregnant formulation was sought. The impregnants considered earlier in the program were rescreened for hydrogen cyanide reactivity. As a result, the addition of molybdenum, in the form of ammonium dimolybdate, to the ASZ-TEDA formulation was found to stabilize the hydrogen cyanide filtration performance. The impregnant loadings on this new Chromium-free Carbon, named ASZM-TEDA Carbon, are listed in Table 1 along with those for the precursor ASZ-TEDA Carbon.

The ASZM-TEDA Carbon was evaluated spectrometrically to acquire an understanding of the role that the molybdenum played in the removal of hydrogen cyanide.[13] This analysis determined that oxidation state of the molybdenum was not changed by exposure to hydrogen cyanide, indicating that the ammonium dimolybdate was apparently not involved in an oxidation-reduction reaction. However, the dispersion of the copper was noted to

Figure 1

Zinc Impregnant Distribution (Light Areas) on
a Granule of ASZ-TEDA Carbon

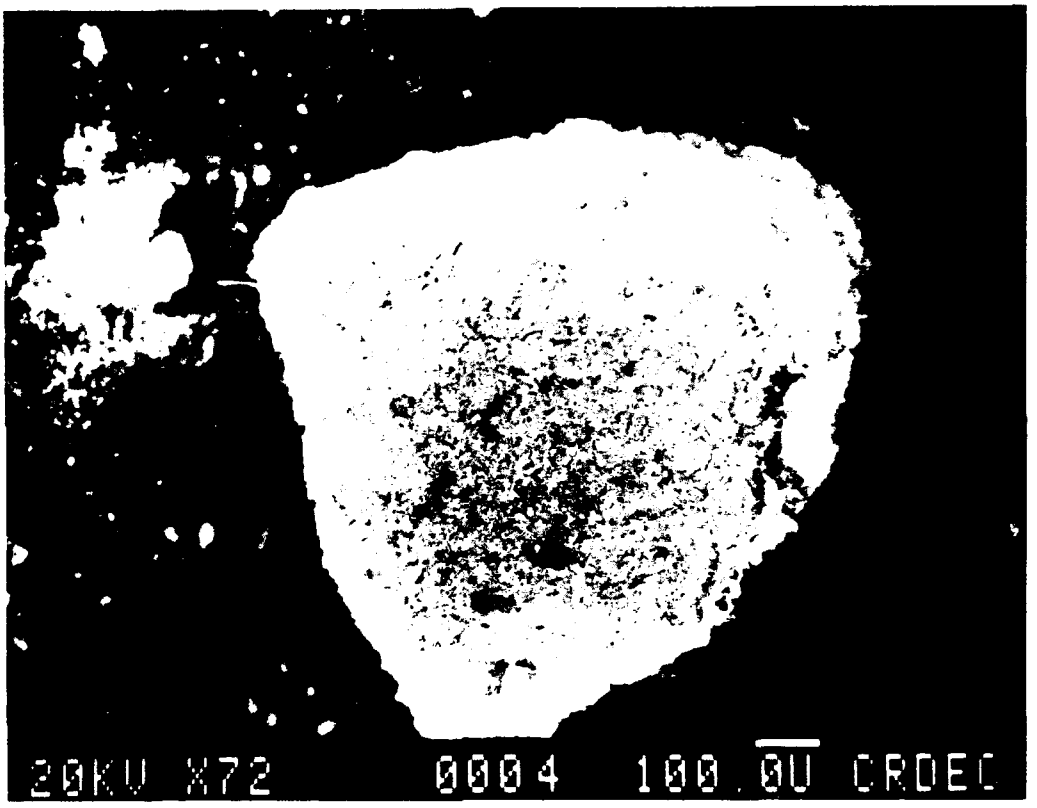


Table 1

Impregnant Loadings on ASZ-TEDA and ASZM-TEDA Carbons

Impregnant	Loading in Percent	
	ASZ-TEDA	ASZM-TEDA
Copper as Cupric Oxide	6	5
Silver in the Elemental Form	0.05	0.05
Zinc as Zinc Oxide	6	5
Molybdenum an Ammonium Dimolybdate	0	2
Triethylenediamine	3	3

Table 2

Specific Mechanisms of Agent Removal Effected by ASZM-TEDA Carbon

AGENT	MECHANISM OF REMOVAL
Nerve	Strong Physical Adsorption Followed by Hydrolysis.
Blister	Strong Physical Adsorption Followed by Hydrolysis.
Cyanogen Chloride	Weak Physical Adsorption Followed by Reaction with Copper, Zinc and TEDA Impregnants.
Hydrogen Cyanide	Weak Physical Adsorption Followed by Reaction with the Copper and Zinc Impregnants. Reaction with Copper Forms Cyanogen; Molybdenum Minimizes Production of Cyanogen.
Arsine	Weak Physical Adsorption Followed by Oxidation by Copper at Low Relative Humidity, and/or Catalytic Oxidation by Silver at High Humidity.
Phosgene	Weak Physical Adsorption Followed by Hydrolysis to HCl and CO_2. HCl Then Reacts with the Copper and Zinc Impregnants.

have been stabilized, thus maintaining the zinc as the most accessible reactant. Figure 2 shows the effect of humid exposure on the copper to carbon ratio on the surface of an ASZ-TEDA Carbon granule as compared to that for a ASZM-TEDA Carbon granule, as determined by X-ray photoelectron spectroscopy. An attempt was made to identify the nature of any complex of copper with molybdenum that may form on the carbon surface to account for this stabilization of copper dispersion, but without success. Table 2 provides a description of the specific mechanisms by which the ASZM-TEDA Carbon removes a number of key chemical warfare agents. From the table, the importance of physical adsorption performance and the wide diversity of chemical reactions provided by the adsorbent is apparent.

Once the final chromium-free carbon formulation had been established, Industrial Hygiene Office of CRDEC obtained approval from the medical community for the use of the ASZM-TEDA Carbon in military filters. The adsorbent was approved for use by the Department of Army Office of the Surgeon General in August 1990. In addition, a CRDEC environmental assessment determined the Chromium-free Carbon, when discarded, would not be classified as a hazardous waste material under EPA regulations.

DEVELOPMENT OF THE BEST PROCESSES FOR APPLYING THE IMPREGNANTS

The next phase of the Chromium-free Carbon Development Program was the determination of the optimum processes for applying the impregnants to the pore structure of the carbon. These processes are critical to achieve the desired impregnant loading level on the carbon and chemical state of the impregnants. The processes were optimized such that maximum reactivity was achieved from the impregnant formulation. The carbon impregnation process studies included determination of the optimum impregnant solution concentration, contact time between the impregnant solution and the carbon, drying time and temperature, and heat treatment conditions. The optimum impregnation processes were developed, first through laboratory experimentation, and then through pilot plant studies. The loading of the impregnants on the carbon is controlled primarily through the concentration of the impregnants in the solution into which the activated carbon is dipped and the contact time between the carbon and the impregnant solution. The chemical state of the impregnants is controlled by the specific form of the impregnants dissolved in the solution and the temperature profile used to dry and heat-treat the carbon. The temperature for these steps is ramped up from relatively modest values under which the water is slowly evaporated to much higher values under which the impregnants convert to their final chemical state. As a result of improvements in drying and heat-treating the carbon, the amount of ammonia which remains on the final product was greatly reduced over that on ASC Carbon.

A pilot plant scale lot consisting of 3,000 pounds of ASZM-TEDA Carbon was produced using the optimum process conditions. The carbon was produced over a one week period by a continuous operation. The carbon was tested for physical properties and chemical agent filtration performance and was determined to comply with specification requirements. Table 3 compares the properties and performance of the ASZM-TEDA Carbon with the specification requirements for the carbon. The specification requirements had been set at values to insure that filters containing the carbon would meet their performance requirements.

Figure 2

Effect of Humid Air on the Copper to Carbon Ratio on the Surface of ASZ-TEDA Carbon and ASZM-TEDA Carbon Granules

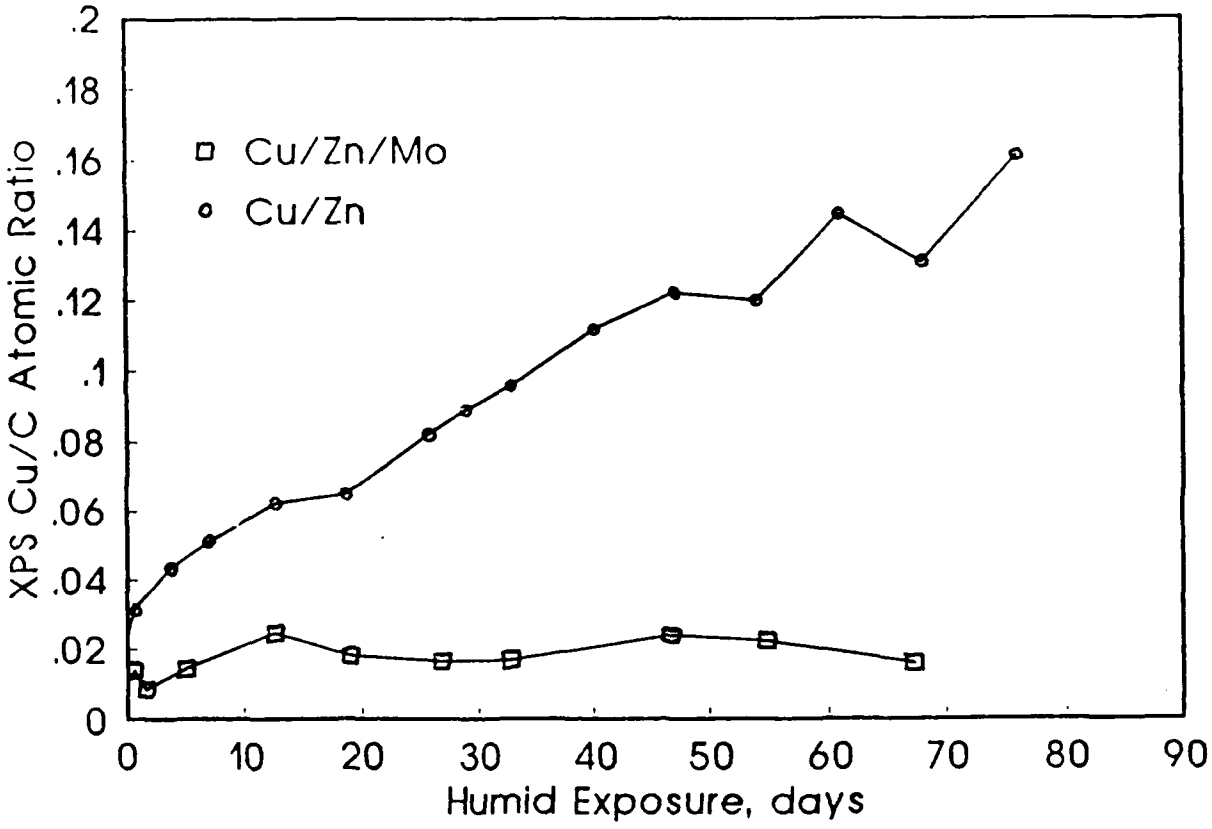


Table 3

Properties and Performance of ASZM-TEDA Carbon as Compared to Specification Requirements

Property Specified	Specification Value		ASZM-TEDA Performance
	Minimum	Maximum	
Cyanogen Chloride Life (minutes)	55	NA	67
Cyanogen Chloride Life After Aging (minutes)	45	NA	58
Hydrogen Cyanide Life (minutes)	28	NA	35
Phosgene Life (minutes)	25	NA	38
DMMP Life (minutes) (Nerve Agent Simulant)	175	NA	210
Copper Content (%)	NA	6.0	5.0
Zinc Content (%)	NA	6.0	5.1
Silver Content (%)	0.03	0.1	0.048
Molybdenum Content (%)	NA	2.5	1.8
TEDA Content (%)	NA	3.5	3.1
Ammonia Off-gassing (milligrams per liter)	NA	0.0010	0.0006
Moisture Content (%)	NA	2.5	1.7
Density (g/cm^3)	NA	0.68	0.61
Hardness (relative value)	85	NA	95
Particle Size (millimeter)	0.6	1.7	0.6-1.7

EVALUATION OF THE FILTRATION PERFORMANCE OF ASZM-TEDA CARBON

The performance evaluation testing of the carbon was concentrated on establishing the level of hydrogen cyanide and cyanogen chloride filtration performance, since as described earlier, the filtration of these agents would be the primary impact from the replacement of the chromium impregnant. Samples of the pilot plant production lot of ASZM-TEDA Carbon were subjected to long term exposure in test chambers under a variety of temperature and humidity conditions representing field use. These exposure environments were as follows:

Environmental Condition	Temperature °F	Relative Humidity (%)
Desert	120	10
Tropic	86	88
Temperate	60	70
Arctic	-25	90

The environmentally exposed carbon was then agent tested at a wide range of bed depths and airflow velocities so that the performance of the carbon in all military filter designs could be assessed. Figure 3 shows a comparison of the effect of tropic exposure on the cyanogen chloride filtration performance of the ASZM-TEDA Carbon with that of ASC Carbon. The ASZM-TEDA Carbon is noted to be much less affected by exposure to this tropic condition. Under the other temperature/humidity conditions (desert, temperate, and arctic), the cyanogen chloride performance of both ASZM-TEDA and ASC Carbon was relatively stable. Hydrogen cyanide challenge testing of the exposed carbons showed relatively stable performance under all environments of the study. The level of performance against this agent for the ASZM-TEDA Carbon was about three-quarters of that of ASC Carbon, but is still at an acceptable level. Figure 4 shows an example of the plot obtained from cyanogen chloride testing conducted at various bed depths and airflow velocities. From such a plot estimates of the performance of specific filters can be obtained from the airflow velocity and carbon bed depth for that filter. The data from this testing indicated that the agent filtration performance of all military filters filled with ASZM-TEDA Carbon would meet their performance requirements. Under a follow-on Engineering Study Project, testing of full scale filters substantiated this conclusion. Currently, the ASZM-TEDA Carbon in being implemented into the Technical Data Packages of several of the chemical warfare agent filters. The implementation process will continue until all military filter systems are converted to Chromium-free Carbon.

RESULTS OF THE DEVELOPMENT EFFORT

All goals of the Chromium-free Carbon Development Program have been met. The environmental and physiological hazards associated with the current ASC Carbon have been eliminated. The ASZM-TEDA Carbon is not a RCRA hazardous material. The U.S. Army Office of the Surgeon General has approved the material for use in all military filters. Because of its environmentally compatible characteristics, over $2,000,000 in military filter disposal costs will be saved. In addition, potential health risks associated with the chromium have been eliminated for personnel involved in

Figure 3
Effect of Tropic Aging of ASZM-TEDA and ASC Carbons on Cyanogen Chloride Filtration Performance

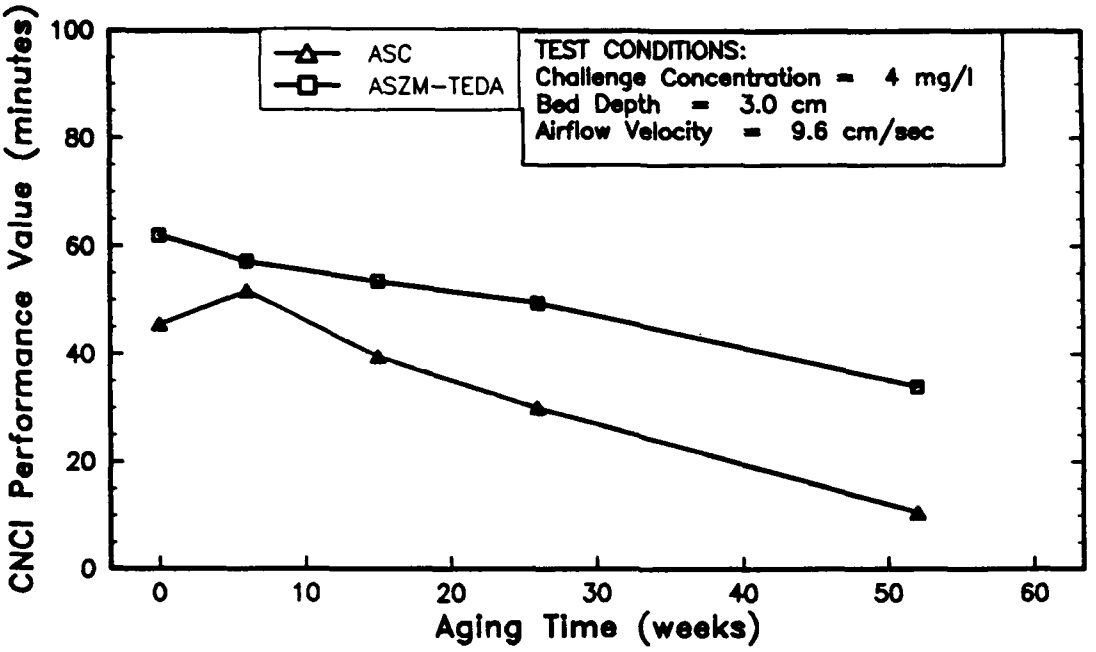
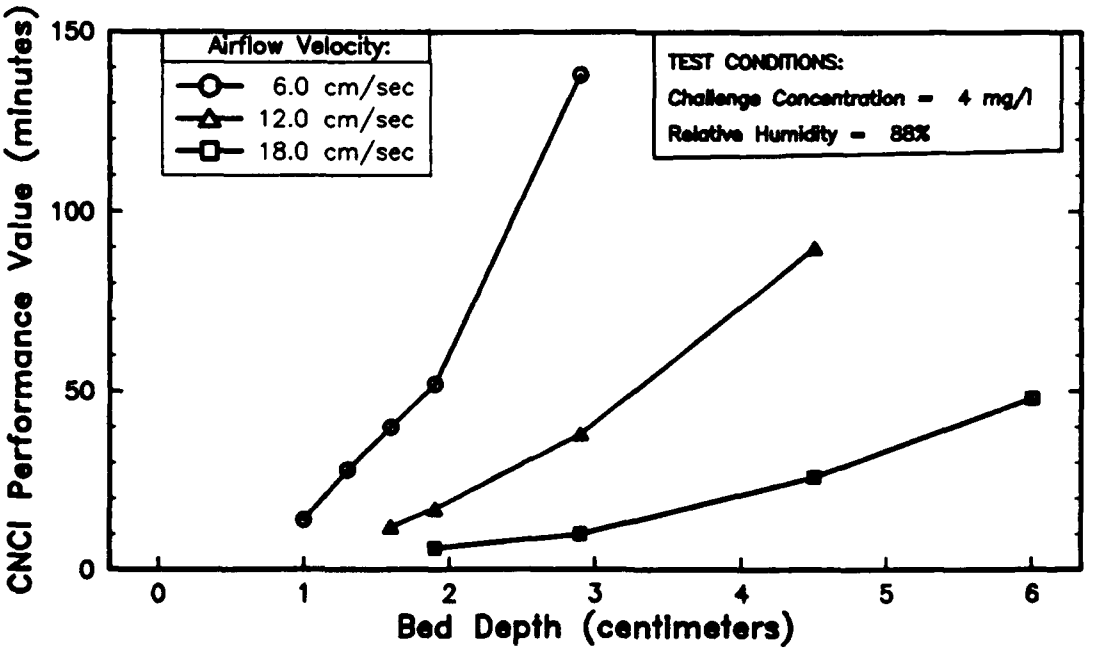


Figure 4
Effect of Tropic Aging of ASZM-TEDA and ASC Carbons on Cyanogen Chloride Filtration Performance



the manufacture of the carbon and military filters as well as users of military filters in the event that the filter is damaged.

Extensive agent challenge testing has been conducted to insure that the performance of filter systems containing this adsorbent are adequate to meet their requirements. As described above, this testing was concentrated on those agents that, for ASC Carbon, had required hexavalent chromium for their removal. The filtration performance of ASZM-TEDA Carbon was determined to be at a level generally comparable to that of ASC Carbon under the wide variety of environmental conditions investigated. Under some conditions the performance of the ASZM-TEDA carbon is a somewhat better than ASC Carbon (cyanogen chloride under tropic conditions) and in some cases the new carbon is somewhat worse (cyanogen chloride under desert conditions), but overall the filtration performance of two carbons is comparable. Because of the use of the identical activated carbon substrate, the physical adsorption of nerve and mustard agents by the ASZM-TEDA Carbon is nearly the same as by ASC Carbon. The filtration performance of the new carbon against these agent is slightly less (about 10 percent) because the impregnant TEDA occupies adsorption space in the micropores. This reduction in performance is not a concern since all military filters have excess performance against nerve and mustard agents. The filtration performance of the ASZM-TEDA Carbon against the acid gas producers such as phosgene is somewhat greater than that of ASC Carbon because of the higher metal oxide impregnant loading for the new carbon.

The objective of reducing the level of ammonia concentration which off-gases from military filters has also been met. While much effort was spent in striving to completely eliminate the need for the ammonia in the impregnation solution, such an approach was not identified. However, the reduction achieved is very large (about 70 percent), such that the highest levels of ammonia off-gassing from the ASZM-TEDA Carbon are less than five parts per million. At this level, the ammonia that off-gases would not be offensive to the filter user.

An additional benefit achieved from the development of the new carbon is improved stability from exposure to heat and humidity over that of ASC Carbon. This benefit offers the opportunity to reduce filter change-out intervals, which will further reduce the production and logistical costs associated with providing chemical warfare protection to military personnel and equipment.

CONCLUSION

This development effort has resulted in the most significant improvement in the adsorbent used in military chemical warfare agent filters since World War II. The elimination of potential environmental and health hazards associated with military filters will have a tremendous benefit to the environment, to the well-being of personnel potentially exposed to the filter adsorbent, and to the life cycle costs associated with the filters. In addition, the ASZM-TEDA Carbon will provide a high level of stable chemical warfare agent protection for military personnel and equipment in a chemical environment.

ACKNOWLEDGEMENT

The authors wish to acknowledge the important contributions of Joseph Rossin of GEO-CENTERS, Inc. for the spectrometric analyses of the chemical state and dispersion of the impregnants on the carbon surface and for his insight into the significance of the results of these analyses.

REFERENCES

1. W. C. Pierce, "U.S. Gas Masks", Chapter 1 in *Summary Technical Report of Division 10, NDRC, Military Problems with Aerosols and Nonpersistent Gases*, National Defense Research Committee, Washington, D.C. (1946).

2. J. R. Conlisk, R. J. Puhala, and L. M. Sturdivan, *Determination of ASC Carbon as a Resource Conservation and Recovery Act (RCRA) Hazardous Waste*, U.S. Army CRDEC Report No. CRDEC-TR-109 (1989).

3. G. A. Lutz, *Physiological and Environmental Effects Relating to the Preparation of Impregnated Charcoal*, U. S. Army CRDEC Contractor Report No. ARCSL-CR-80072, Battelle Columbus Laboratories (1981).

4. E. D. Tolles, *Improved Charcoal Impregnation Technology*, U.S. Army CRDEC Contractor Report No. CRDEC-CR-86066, Westvaco Corporation (1986).

5. D. T. Doughty, *Development of a Chromium-free Impregnated Carbon for the Adsorption of Toxic Agents*, U.S. Army CRDEC Contractor Report No. CRDEC-CR-118, Calgon Carbon Corporation (1991).

6. B. R. Alves and A. J. Clark, *Carbon*, 24, 287-294 (1986).

7. L. L. Pytleweski, *Studies of ASC Whetlerite Reactivity*, U.S. Army CRDEC Contractor Report No. ARCSL-CR-79008, Drexel University (1979).

8. J. F. Alder, P. R. Fielden, and S. J. Smith, *Carbon*, 26, 701-711 (1988).

9. R. J. Grabenstetter and F. E. Blacet, "Impregnation of Charcoal", Chapter 4 in *Summary Technical Report of Division 10, NDRC, Military Problems with Aerosols and Nonpersistent Gases*, National Defense Research Committee, Washington, D.C. (1946).

10. J. A. Rossin and R. W. Morrison, *Carbon*, 29, 887-892 (1991).

11. M. Smisek and S. Cerny, *Active Carbon*, (Elsevier, New York, 1970), p. 188.

12. P. N. Brown, G. G. Jayson, G. Thompson, and M. C. Wilkinson, *Carbon*, 27, 829 (1989).

13. J. A. Rossin and R. W. Morrison, paper presented at the Annual Meeting of the American Institute of Chemical Engineers, Los Angeles, CA, November 1991 (unpublished data).