

AD611432



Army

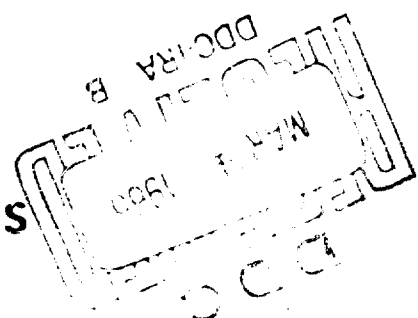
COPY	2	OF	3	UNC
HARD COPY	\$ 8.25			
MICROFICHE	\$ 2.50			

Science

526p

Conference

Proceedings



17-19 June 1964

Volume I

Principal Authors A through H

OFFICE OF RESEARCH AND DEVELOPMENT
DEPARTMENT OF THE ARMY

ARCHIVE COPY



HEADQUARTERS
DEPARTMENT OF THE ARMY
OFFICE OF THE ADJUTANT GENERAL
WASHINGTON, D.C. 20310

IN REPLY REFER TO
AGAM-P (M) (31 Dec 64) CRD/P

22 January 1965

SUBJECT: 1964 Army Science Conference Proceedings

TO: SEE DISTRIBUTION

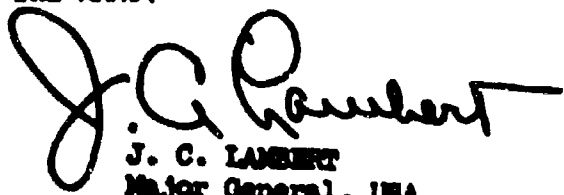
1. Inclosed for your information and use is Volume One of the 1964 Army Science Conference Proceedings, which is comprised of the unclassified papers by principal authors A thru H, which were presented at the conference.

2. Volumes Two and Three of the Proceedings are being distributed separately.

BY ORDER OF THE SECRETARY OF THE ARMY:

I Incl
Vol I, 1964 Army Science
Conference Proceedings

DISTRIBUTION:
As Indicated in Distribution
List in Volume


J. C. LAMBERT
Major General, USA
The Adjutant General

PROCEEDINGS
OF THE
1964 ARMY SCIENCE CONFERENCE
UNITED STATES MILITARY ACADEMY, WEST POINT, N. Y.

17-19 JUNE 1964

VOLUME I
(Principal Authors "A" thru "H")

INTRODUCTION

1. The proceedings of the 1964 Army Science Conference is a compilation of all papers presented at the conference. Its purposes are to:
 - a. Provide a historical record of the conference
 - b. Provide source material for scientists engaged in research
 - c. Enable conference participants to review papers in which they have a special interest
2. The report consists of three volumes as follows:

Volume I	Unclassified papers, principal authors, A thru H
Volume II	Unclassified papers, principal authors, I thru Z
Volume III	Classified papers, alphabetically by principal authors
3. Three Invited Papers, which were presented at the Opening General Session are grouped together at the beginning of Volume I. These papers are essentially different in nature inasmuch as they were oriented toward the general scientific and lay interests of a mixed audience, rather than reporting on a specific research effort within a given discipline.
4. All experiments involving live animals which are reported in the Proceedings were performed in accordance with the principles of laboratory animal care as promulgated by the National Society for Medical Research.

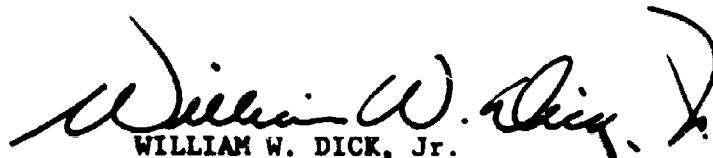

WILLIAM W. DICK, Jr.
Lieutenant General, GS
Chief of Research and Development

TABLE OF CONTENTS
 PROCEEDINGS OF 1964 ARMY SCIENCE CONFERENCE

Invited Papers

AUTHOR	TITLE	VOL	PAGE
Barila, Lt. Col, T.G.	Medical Trends in Life Support Systems	1	1
Horton, B.M.	Control, Amplification and Fluids	1	9
Johnston, Brig Gen, J.W.	Army Participation in Project SYNCOM	1	23
<hr/>			
Aaron, H.S. Ferguson, C.P., Jr.	Incapacitating Agents: Eight Stereoisomers of a Synthetic Tetrahydrocannabinol	3	1
Abbott, K.H.	Effects of Alloy Segregation on Spalling of Metallic Armor Materials	1	45
Agrios, J.P. Neudorfer, C.D. Stern, R.A.	TE ₁ O ₁ ^o Cross-Guide Transducer and Isolator	1	57
Ahearn, P.J. Quigley, F.C.	Dendritic Morphology of High-Strength Steel Castings	1	71
Ahlstrom, E.R.	SEE GERBER, E.A.	1	375
Allan, L.D. Croomes, E. Duncan, W. Murfree, J. Sandlin, B. Wharton, W.	Characterization of Nitrogen-Fluorine Containing Compounds as Propellants	3	13
Allen, H.C.	Crystallization in Solid Propellant Binders and Its Effect on Low-Temperature Missile Capability	3	27

AUTHOR	TITLE	VOL	PAGE
Alley, B.J.	Control of Propellant Composition by X-Ray Fluorescence Analysis	3	41
Anderson, V.H.	The Interpretation of Aerial Imagery of Sea Ice	3	55
Andreotti, R.	SEE KRONMAN, M.J.	2	121
Appleby, J.F. Ohmstede, W.D.	Numerical Solution of the Distribution of Wind and Turbulence in the Planetary Boundary Layer	1	85
Arendt, P.R. Papayocanou, A. Soicher, H.	Determination of the Ionospheric Electron Content Utilizing Satellite Signals	1	101
Autera, J.R.	SEE CASTORINA, T.C.	1	163
Babbitt, R.W.	SEE MALINOFSKY, W.W.	2	197
Ball, D.H. Williams, J.M. Long, L., Jr.	Potential Radiation - Protective Compounds. Synthesis of the Three Isomeric Three-Carbon Aminohydroxy Bunte Salts and Related Compounds	1	115
Balton, I.A.	SEE BOMKE, H.A.	3	99
Bauer, V.E.	SEE ROSENBLATT, D.H.	3	261
Bergdoll, M.S.	SEE SCHANTZ, E.J.	2	309
Berning, W.W.	High Altitude Ionization Associated with Nuclear Detonations	3	71
Black, Capt., J.N.	SEE CRAWLEY, G.J., 1st Lt.	1	207
Bluford, B.L., Jr.	Applications of Light-Emitting Diodes to Certain Infrared Communications and Surveillance Problems.	3	85
Blum, R.	SEE KRONMAN, M.J.	2	121

AUTHOR	TITLE	VOL	PAGE
Bomke, H.A. Balton, I.A. Grote, H.H. Harris, A.K.	Analysis of Magnetic Signals from the High-Altitude Nuclear Tests of Operation Dominic.	3	99
Brady, Lt. Col, J.V.	SEE SHARP, Capt., J.C.	2	375
Breslin, J.T.	SEE GAULE, G.K.	1	363
Brzonkala, K.J.	SEE KITTL, E.	2	105
Burke, C.L.	SEE ZEHLER, E.	2	507
Buser, R.G. Kainz, J. J. Sullivan, J.J.	The Use of Intense Pinch Discharges for Laser Illumination	1	121
Cabelli, V.J.	SEE HAYES, D.K.	1	457
Carlson, F.F.	A Paramagnetic Resonance Study of Nitrogen Atoms Trapped in X-Irradiated Alkali Azides	1	135
Cason, C.	Gas Velocity Probe for Moving Ionized Gases	1	149
Castorina, T.C. Autera, J.R.	Nitrogen-15 Tracer Studies of the Nitrolysis of Hexamethylenetetramine	1	163
Clawson, R.J.	Attenuation of ABM Radar Frequencies by High-Altitude Nuclear Fireballs	3	113
Colton, R.M. Dobbins, E.B.	Re-entry Vehicle for Radar Selectivity Evaluations	1	179
Cox, J.T. Hass, G. Ramsey, J.B.	Improved Dielectric Films for Optical and Space Applications	1	193
Crawley, 1st Lt., G.J. Black, Capt., J.N. Gray, Col., I. LeBlang, 1st Lt., W.	The Mechanism of Action of Staphylococcal Enterotoxin Poisoning	1	207
Creedon, J.E.	SEE SCHNEIDER, S.	2	331
Crim, W.M., Jr.	Closed Cycle Gas Turbine	1	221

AUTHOR	TITLE	VOL	PAGE
Croomes, E.	SEE ALLAN, B.D.	3	13
Davidson, T.E. Uy, J.C. Lee, A.P.	The Effect of Pressure on the Structure and Properties of Metals	1	237
Davis, G.T.	SEE ROSENBLATT, D.H.	3	261
Dawson, J.W. Niedenzu, K.	Hindered Rotation in Aminoboranes	1	251
Deluce, 1st Lt., D.C.	SEE ROSENBLATT, D.H.	3	261
Dienes, G.J.	SEE ROSENBLATT, D.B.	2	283
Dixon, S. Jr. Savage, R.O., Jr. Tauber, A.	High Power Microwave Device Applications for New Narrow Resonance linewidth Planar Hexagonal Ferrites	1	261
Dobbins, E.B.	SEE COLTON, R.M.	1	179
Donnert, H.J. Sasse, R.A. Klein, N.	The Response of Selected Gamma Radiation Detectors to Radia- tion Delivered of Very High Dose Rates	3	127
Dressel, R.W.	Retrofugal Electron Flux from Massive Targets Irradiated with a Monoenergetic Primary Beam	1	275
Duncan, W.	SEE ALLAN, B.D.	3	13
Dunnery, D.A.	SEE SCHANTZ, E.J.	2	309
Edwards, R.L.	Functional Dependence of the Radar Cross Section of the Wake of a Re-entry Vehicle of C_{DA}	3	141
Egli, J.J.	Electronics Counter-Counter- Measures Design Techniques for Communications	3	155
Einberg, F.	Preparation of 5-Dinitromethyl- tetrazole from Salts of Dinitroacetonitrile	1	289

AUTHOR	TITLE	VOL	PAGE
Elder, A.S.	Modal Analysis of Transient Vibration Problems in Linearly Viscoelastic Solids	1	301
Ennulat, R.D. Sherman, E.J. Schaer, F.M.	Light Scattering by Materials in the Cholesteric Mesophase	3	169
Epstein, J.	SEE ROSENBLATT, D.H.	3	261
Ferguson, C.P., Jr.	SEE AARON, H.S.	3	1
Finkelstein, R.A.	SEE NORRIS, Capt., H.T.	2	247
Flathau, W.J. Strange, J.N.	Response of Gravity Type Dams to Nuclear Weapons Effects	3	183
Frankel, J.	The Recovery of the Ultra- sonic Attenuation in Copper Single Crystals Following Small Plastic Deformation	1	317
Freitag, D.R. Powell, C.J.	Laboratory Investigation of the Mobility of Pneumatic Tires in Cohesionless Soil	1	333
Galos, G.E.	Accuracy and Information Rate Studies on Target Cross Sections Utilizing the Nike Hercules Radars	1	349
Garber, E.B.	SEE LEVINSON, H.S.	2	169
Gaule, G.K. Breslin, J.T. Loss, R.L. Logan, 2nd Lt., R.S.	Superconductors in Advanced Electronics	1	363
Gerber, B.V.	SEE MYERS, K.A.	3	223
Gerber, E.A. Ahlstrom, E.R.	Ruby Laser with Vibrating Reflector	1	375
Gisser, H. Sadjian, S.	Evaporation of Organic Compounds from Metal Surfaces at High Vacuum	1	391
Gray, Col., I.	SEE CRAWLEY, 1st Lt., G.J.	1	207

AUTHOR	TITLE	VOL	PAGE
Gray, Col, I. Hildebrandt, Capt., P.V.	The Effect of Hyperthermia on Protein Turnover in Infection	1	405
Green, E.	SEE SCHOENFELD, M.	2	345
Grote, H.H.	SEE BOMKE, H.A.	3	99
Hahn, F.E.	Molecular Mechanisms of Antimicrobial Action	1	413
Hardaway, Col, R.M., III	Influence of Trauma and Hemolysis on Hemorrhagic Shock in Dogs	1	427
Harris, A.K.	SEE BOMKE, H.A.	3	99
Hass, G.	SEE COX, J.T.	1	193
Hart, Maj, W.J., Jr. Otto, R.J. Sinclair, A.H.	Friction Hydro Pneumatic Suspension System	1	441
Hayes, D.K. Cabelli, V.J.	The Influence of Storage Aerosolization and Rehydration on the Permeability of <u>Pasteurella Tularensis</u>	1	457
Hearn, H.J., Jr.	SEE TRIBBLE, H.R., JR.	2	469
Hendron, 1st Lt., A.J.	SEE KENNEDY, T.E.	2	89
Henry, M.C.	Preparation and Uses of Some New Organolead Chemicals	1	469
Higgins, V.J.	X Band Semiconductor Switching and Limiting Using Waveguide Series Tees	1	477
Hildebrandt, Capt. P.V.	SEE GRAY, Col., I.	1	405
Holmes, L.	SEE KRONMAN, M.J.	2	121
Howard, C.D. Huskins, C.W.	A New Fuel for High Energy Rocket Propellants	3	199
Hsu, Y.C.	SEE SADOWSKY, M.A.	2	297

AUTHOR	TITLE	VOL	PAGE
Huber, W.A.	Monocycle Position Modulation	1	495
Hull, L.A.	SEE ROSENBLATT, D.H.	3	261
Huskins, C.W.	SEE HOWARD, C.D.	3	199
Hussain, M.A.	SEE SADOWSKY, M.A.	2	297
Ikrath, K. Schneider, W.	The Realization of Active Seismic Systems and Their Practical Applications	2	1
Jakubowski, E.H. Swieskowski, H.P.	Surge Waves in Automatic Weapons	2	15
Jones, G.R.	Surface Dependence of Magnetostatic Mode Line-Widths in Yttrium Iron Garnet	2	33
Kalmus, H.P.	A New Gravity Meter	2	41
Kains, J.J.	SEE BUSER, R.G.	1	121
Kant, A. Strauss, B.	The Maximum Dissociation Energies of Diatomic Molecules of Transition Elements	2	55
Kasemir, H.W.	The Relation Between Cloud Charges Precipitation and Lightning Discharges	2	67
Keithley, S.M.	Performance Evaluation of a Special Armor System	3	213
Kendall, D.P.	The Effect of Strain Rate on Yielding Characteristics of High Strength Alloy Steels	2	79
Kennedy, T.E. Hendron, Lt. A.J., Jr.	The Dynamic Stress-Strain Relation for a Sand as Deduced by Studying its Shock Wave Propagation Characteristics in a Laboratory Device.	2	89
Kittl, E. Brzonkala, K.J. Shapiro, S.J.	Performance of the Germanium-Photovoltaic Cell Under Intensive Monochromatic Light	2	105

AUTHOR	TITLE	VOL	PAGE
Klein, N.	SEE DONNERT, H.J.	3	127
Kronman, M.J. Blum, R. Andreotti, R. Vitols, R. Holmes, L.	Molecular Structure and Protein Stability; The Acid Denaturation of Alpha Lactalbumin	2	121
Laible, R.C.	High Speed Penetration of Non-woven Fabric	2	137
LeBlang, Lt., W.	SEE CRAWLEY, 1st Lt., G.L.	1	207
Lee, A.P.	SEE DAVIDSON, T.E.	1	237
Lentz, S.S.	An Optical System for Measuring the Angular Motion of the Gun Tube and the Projectile in the Bore	2	153
Levinson, H.S. Garber, E.B.	Radiation Biodosimetry and Screening for Radioprotective Compounds	2	169
Logan, 2nd Lt., R.S.	SEE GAULE, G.K.	1	363
Long, L., Jr.	SEE BALL, D.H.	1	115
Lontz, R.J.	Fluorine Hyperfine Interaction in Electron Spin Resonance	2	185
Malinofsky, W.W. Babbitt, R.W. Sands, G.C.	A Novel Process for Ultra-Fine Crystallites, and their Theory and Application in Magnetic Ferrites	2	197
Martin, D.L., Jr.	Effect of Filler Concentration on the Viscoelastic Response of a Filled Polymer System	2	211
Mather, B.	The Strength of Portland Cement Concrete as Affected by Air, Water, and Cement Content	2	227

AUTHOR	TITLE	VOL	PAGE
Mertwoy, H.E.	SEE SOLLOTT, G.P.	2	441
Miller, J.I.	SEE WILLIAMSON, C.E.	2	493
Murfree, J.	SEE ALLAN, B.D.	3	13
Myers, K.A. Gerber, B.V.	Casualty Estimations for Chemical (CB) Munitions Delivery Systems	3	223
Nagle, S.C., Jr.	SEE TRIEBLE, H.R., Jr.	2	469
Nemarich, J.	Relaxation Processes in Ferromagnetic Insulators	2	233
Neudorfer, C.D.	SEE AGRIOS, J.P.	1	57
Niedenzu, K.	SEE DAWSON, J.W.	1	251
Norris, Capt., H.T. Finkelstein, R.A. Sprinz, Col., H.	A Morphologic Study of the Pathogenesis of Experimental Cholera in the Infant Rabbit	2	247
Ohmstede, W.D.	SEE APPLEBY, J.F.	1	85
Otto, R.J.	SEE HART, Maj., W.J., Jr.	1	441
Papayocanu, A.	SEE ARENDT, P.R.	1	101
Peterson, W.R., Jr.	SEE SOLLOTT, G.P.	2	441
Portnoy, S.	SEE SOLLOTT, G.P.	2	441
Powell, C.J.	SEE FREITAG, D.R.	1	333
Prickett, M.E.	An Investigation of Aerosols of Viruses Disseminated in Very Small Particles	3	235
Quigley, F.C.	SEE AHEARN, P.J.	1	71
Ramsey, J.B.	SEE COX, J.T.	1	193
Reago, D.A.	SEE SCHOENFELD, M.	2	345
Ringel, S.	Information Assimilation from Command Systems Displays	2	261

AUTHOR	TITLE	VOL	PAGE
Roessler, W.G.	SEE SCHANTZ, E.J.	2	309
Rogers, E.B.	An Electrostatic Fuse for Antiaircraft Mines and Other Antiaircraft Weapons	3	247
Romba, J.J.	A Design for Animal-Hearing Research	2	275
Rosenblatt, D.B. Dienes, G.J.	The Effect of Reactor Irradiation Upon Hydrogen Adsorption by an Alumina Catalyst	2	283
Rosenblatt, D.H. Bauer, V.E. Epstein, J. Davis, G.T. Hull, L.A. Weglein, R.C. DeLuca, Lt., D.C.	Destruction of Toxic Agents and Model Compounds with Chlorine Dioxide	3	261
Ross, R.L.	SEE GAULE, G.K.	1	363
Sadjian, S.	SEE GISSER, H.	1	391
Sadowsky, M.A. Hsu, Y.C. Hussain M.A.	Boundary Layers in Couple-Stress Elasticity and Stiffening of Thin Layers in Shear	2	297
Sandlin, B.	SEE ALLAN, B.D.	3	13
Sands, G.C.	SEE MALINOFKY, W.W.	2	197
Sass, S.	SEE WILLIAMSON, C.E.	2	493
Sasse, R.A.	SEE DONNERT, H.J.	3	127
Savage, R.O., Jr.	SEE DIXON, S., Jr.	1	261
Schaer, F.M.	SEE ENNULAT, R.D.	3	169
Schantz, E.J. Roessler, W.G. Wagman, J. Spero, L. Stefanye, D. Dunnery, D.A. Bergdoll, M.S.	The Purification and Characterization of Staphylococcal Enterotoxin B	2	309

AUTHOR	TITLE	VOL	PAGE
Schiel, E.J.	Direct Modulation of a HE-NE Gas Laser	2	321
Schneider, S. Creedon, J.E. Yeamans, N.L.	Hydrogen System Dependence on Discharge Parameters	2	331
Schneider, W.	SEE IKRATH, K.	2	1
Schoenfeld, M. Reago, D.A. Green, E.	Optimization of the Rota- ting Reflector Q-Switch	2	345
Shapiro, S.J.	SEE KITTL, E.	2	105
Sharenow, M.	New High-Altitude Fast Rising Balloons	2	361
Sharp, Capt., J.C. Brady, Lt.Col., J.V.	The Effects of Massive Doses of Ionizing Radiation Upon Condi- tioned Avoidance Behavior of the Primate	2	375
Shear, R.E.	Predicted Blast Properties of Pentolite and Their Comparison with Experiment	2	391
Sherman, R.J.	SEE ENNULAT, R.D.	3	169
Silverstein, A.M.	Homograft Rejection in the Fetal Lamb	2	401
Sinclair, A.H.	SEE HART, Maj., W.J., Jr.	1	441
Sliney, J.L.	Biaxial Tensile Behavior of Anisotropic Titanium Sheet Materials	2	413
Snead, J.L.	SEE SOLLOTT, G.P.	2	441
Sodano, E.M.	Optical-Electronic Azimuth and Range Between Non- intervisible Distant Ground Stations	2	427
Soicher, H.	SEE ARENDT, P.R.	1	101

AUTHOR	TITLE	VOL	PAGE
Sollott, G.P. Sneed, J.L. Portnoy, S. Peterson, W.R., Jr. Mertwoy, H.E.	Phosphorus, Arsenic & Boron-Containing Ferrocene Derivatives	2	441
Soper, W.T.	SEE TRIBBLE, H.R., JR.	2	469
Spero, L.	SEE SCHANTZ, E.J.	2	309
Sprinz, Col., H.	SEE NORRIS, Capt., H.T.	2	247
Stefanye, D.	SEE SCHANTZ, E.J.	2	309
Stern, R.A.	SEE AGRIOS, J.P.	1	57
Steverding, B.	The Mechanism of Preferential Ablation	2	453
Strange, J.N.	SEE FLATHAU, W.J.	3	183
Strauss, B.	SEE KANT, A.	2	55
Sullivan, J.J.	SEE BUSER, R.G.	1	121
Swieskowski, H.P.	SEE JAKUBOWSKI, E.H.	2	15
Tauber, A.	SEE DIXON, S., JR.	1	261
Taylor, Lt., T.D.	On the Interaction of a Bow Shock of a Re- entry Body and Blast Wave	3	273
Thiele, O.W.	Ballistic Procedure for Unguided Rocket Studies of Nuclear Environments	3	289
Tribble, H.R., Jr. Hearn, H.J., Jr. Nagle, S.C., Jr. Soper, W.T.	Growth of Selected Arboviruses in Serum- Free Suspension Cell Cultures	2	469
Uy, J. C.	SEE DAVIDSON, T.E.	1	237
Vitols, R.	SEE KRONMAN	2	121
Wagman, J.	SEE SCHANTZ, E.J.	2	309

AUTHOR	TITLE	VOL	PAGE
Walton, Maj., B.C.	The Effect of Migrating Nematode Larve as a Provoking Factor in Viral Encephalitis in Mice	2	483
Weglein, R.C.	SEE ROSENELETT, D.H.	3	261
Weickmann, H.K.	Mesometeorological Applications of In- frared Instrumentation	3	299
Wharton, W.	SEE ALLAN, B.D.	3	13
Williams, J.M.	SEE BALL, D.H.	1	115
Williamson, C.E. Miller, J.I. Sass, S. Witten, B.	Design and Reaction Mechanism of Short- Lived Alkylating Agents	2	493
Witten, B.	SEE WILLIAMSON, C.E.	2	493
Yessens, N.L.	SEE SCHNEIDER, S.	2	331
Zehler, E. Burke, C.L.	Measurements of True Dark Conductivities in Ferrocene Crystals	2	507

BARILA

MEDICAL TRENDS IN LIFE SUPPORT SYSTEMS

TIMOTHY G. BARILA, LT COL, MC
WALTER REED ARMY INSTITUTE OF RESEARCH
WALTER REED ARMY MEDICAL CENTER
WASHINGTON, D. C.

Strides in medicine in the past few decades have been as great as those in other professions, probably most dramatically exemplified by the increase in life expectancy from 40 to over 70 years. Isolated examples, or even a grasp of the whole field of medical accomplishments, make one proud to be living in these times and in this country; in these times, because of these medical and other "life-improving" developments, and in this country, because many of the accomplishments are peculiar to the academic, scientific and physical environment resulting from the American way of life.

TRENDS IN MEDICAL LIFE SUPPORT SYSTEMS: In the colloquial use of the term "life support systems," the problems of maintaining life and function of an orbiting astronaut come to mind. The term "closed ecological system" is also used to denote the need for the entire system to be self-sufficient. By adding the term "medical" to the title of life support systems, I hope to explain some of the background, current research efforts and future plans of the Army Medical Service in meeting its mission to preserve the fighting strength of the Nation.

The fighting man in an extreme of environmental and time stress may have to exist in an atmosphere not unlike a closed ecological system. In my daily practice as an Army Medical Officer and an Anesthesiologist, the use of a closed ecological system does much to contribute to the successful outcome in most of today's patients undergoing elective or emergency surgery. Once the patient is rendered unconscious, he becomes part of a closed anesthetic environmental system. A skilled team controls the factors needed to maintain his vital functions and return him to full consciousness. These same factors have direct application to victims, wounded or injured in the performance of daily duty. To cope with an adverse situation with a logical, pre-planned, practiced effort, devoid of panic, is an application of these factors in a practical way. Such steps are simple enough to be learned and used by non-medical people. The immediate application of some

BARILA

simple steps of correction, such as the maintenance of an adequate airway by the proper positioning of an unconscious victim's head may meet with dramatic success. In many cases the reward is particularly great as the resuscitated victim is often a dear relative or friend.

Control of environment is important to survival prior to the need for medical care, although one cannot separate the two completely. It can be a key factor in the ability of a man to survive a stressful situation. A wound or disease occurring at a high altitude is compounded by the same environmental factors which make life for a healthy person extremely stressful at that altitude. Beside the lack of availability of medical attention, factors such as poor roads, low temperatures, low barometric pressure with consequences of problems in breathing by the personnel, and poor aerodynamics developed by a potential rescue helicopter or airplane are paramount. If the closed ecological system becomes ever more isolated, as in space flights, the problems of accident and illness may become even more acute.

Even within the boundaries of our country, a heavy snowfall in a rural area could decrease the chances for survival in, for instance, a pregnancy at term which requires a blood transfusion or operative manipulation of the infant, either of which would be rather simple procedures if the patient were in a modern hospital environment instead of her own home. The environment also becomes hostile if a person or group of soldiers are required to move abruptly from a moderate climate to one of extremes of heat or cold. Again, nature compounds these problems of physiological adaptation by the individual to the stress to which he is exposed, by making it equally difficult to logistically support him with medical or other services.

Finally, the longer life span achieved does not lead many of us to a feeling of complacency nor to say that most of the problems preventing a long, happy, healthful life have been eliminated. The first day of life, the first month, and the first year of life, all still carry high and similar mortality rates to those recorded 25 or 50 years ago. At the other end of the life span, cancer and heart disease have increased simply because people are living long enough to acquire these diseases. Further, the American heritage has led to a strange paradox which permits thousands to be killed in traffic accidents, and more thousands to die by accidents in their homes, and the events noted as little more than statistic, while a prominent person, in a similarly sudden, shocking, unexpected instant, dies of a heart attack or an assassin's bullet and becomes a national tragedy. Which person is the more dead or the greater loss to his family becomes somewhat academic if considered in a somewhat metaphysical detachment by claiming that most of these deaths are preventable.

I will not consider the epidemiology of automobile accidents in this presentation. It is better done by others. But it is indeed strange that seat belts or chest harnesses are not more widely used nor are obvious design changes forthcoming for safety in the

BARILA

automobile itself for fear, perhaps, that the car will suffer esthetically in the market place. It is equally strange that methods are known today for the reversal of death, or the maintenance of life, in accident or disease victims, which will permit these victims to reach more sophisticated medical care; but these methods are not used, or worse, are used improperly for various reasons, not the least of which is an attitude of "it can't happen here." Yet when such methods as mouth-to-mouth breathing to resuscitate drowned or asphyxiated victims were shown to be so effective when taught by a simple daily lesson on the front page of a Scandinavian state-controlled newspaper, the mortality from such accidents was reduced by almost 75%. How many of us consider ourselves competent to resuscitate a friend or relative who has drowned or stopped breathing from any of many other causes?

In a similar way, I believe that perhaps half of the 500,000 persons who die each year of heart disease might not have had to die at that particular time. When the heart of such a person is examined by a pathologist, he is unable to detect by the finest of our modern instrumentation any new disease in that heart. It appears to be the same heart that hours or days before was functioning well in a living human being. To say it in another way, an examination of the heart of a person who dies in his seventies after long, various and sometimes severe episodes of heart disease, shows the heart to be anatomically more severely involved by the disease processes which apparently killed the person 20 or 30 years his junior.

It is true that death could be so sudden that changes would not have time to occur. In refutation, studies such as that conducted in Framingham have shown that an adult population can be examined and some of the population will be found to have heart disease. When followed over a period of years, some of these patients with heart disease die. Yet over 70% of these fatalities never reach medical attention even within as long a time as three weeks from the onset of new or severe symptoms. In our excellent Army Medical Service, similar records show that about 60% of our active duty personnel die without medical attention. Incidentally, the major causes of death in civilians and in our citizen army are identical: heart disease and accidents.

When the victim of the myocardial infarction or "heart attack" lives to enter the hospital, his chances of dying are reduced to about 25%. But even in this area, there are good reasons to believe many of these 25% had hearts too good to die. About 18 of these 25 deaths are due to ventricular fibrillation, a condition in which the thousands of heart muscle fibers attempt to contract out of synchrony with their neighbors, resulting in an ineffective output of blood from the heart and the death of the patient. This condition is likely due to an electrocution phenomenon not unlike a death from an electric shock incurred when the victim's feet were wet, permitting the electric charge to be conducted easily through the body- a not uncommon

BARILLA

home accident.

In the heart attack victim, an abnormal pattern of electrical activity is caused by the lack of blood nourishing the area normally supplied by the vessel which has been blocked by the disease process. This area does not have to be large since the key feature is a difference in electrical potential in the diseased area compared to the remaining normal heart muscle. This same area of difference can then become an irritable focus to a degree sufficient to permit random discharges of electrical potential great enough to cause abnormalities in the rhythm of the heart, continuing to the end point of the completely disorganized rhythm of ventricular fibrillation just mentioned. The nearer this area of infarction or dead muscle lies to the normal area of nervous pathways for heart stimulation, the more likely it is to become an irritable focus to precipitate the fatal muscle contractions. In contrast, if it is far removed from this area of nervous conduction, the heart will be less irritable and the victim is able to develop collateral channels over a period of time to nourish the areas of the heart previously supplied with blood by way of the blocked artery.

But even if this abnormal rhythm of ventricular fibrillation does occur, you may surmise from what I have said that there are good reasons to expect this death to be reversed if the treatment is prompt and proper.

The output of blood from the heart can be diminished or stopped in other ways such as the inadvertent overdosage and depression of the heart, as well as the rest of the body, by anesthetic agents, so that it no longer can respond to normal nervous impulses. Or it can stop because of lack of oxygen or a wound or accident causing such severe loss of blood that an insufficient quantity is being returned to the heart to "load" or fill it. Regardless of the cause, the effect can be catastrophic. Prompt treatment, similar in both mechanisms, may permit successful resuscitation and complete recovery. These mechanisms for death were not described and understood until early in this century, and probably are not yet fully explained. Yet, in 1921, and perhaps earlier in undocumented cases, a surgeon was able to take over artificially the function of the heart by rhythmically squeezing it with his hand while adequate oxygenation could be restored, permitting the patient to survive. In 1946, the first successful reversal of the electrical mechanism for death was reported. In this instance, the surgeon reversed the fatal outcome by applying an even greater electrical charge to the surface of the heart by means of two electrodes applied to opposite sides of the heart, causing all the heart muscle fibers to contract at once. After the heart fibers recovered from this effort of contraction, they were ready to respond normally and in unison to the next nervous impulse reaching the heart through its usual pathways. In 1956, the placement of the electrodes directly to the heart was proved to be unnecessary if sufficient voltage was delivered to two large electrodes placed on

BARILA

opposite sides of the chest. The need to manually squeeze the heart with the hand to force blood to be circulated, while the cause of the trouble was being eliminated, was solved by noting the same effect could be achieved by pressing forcibly on the chest, trapping the heart between the chest, the backbone and the lungs on either side of the heart. When this method is combined with the mouth-to-mouth breathing technic mentioned above, a victim of an otherwise fatal heart attack might be brought to a medical facility by a trained non-medical rescuer sufficiently alive to permit his basic mechanism of death to be combatted and reversed. And indeed, this is the favorable outcome of more and more heart attack victims today. It is true that these simple methods for first aid are not without danger, not universally taught or learned, nor even fully accepted by the medical profession. But the concept is sound and seems to need only a concerted effort by teaching sessions to have it mean life for many people otherwise doomed to an unexpected and often early death.

Of course, the complete story of reversal of death or the support of life has not been outlined above. The complete picture must also include time. The rescuer needs time to act. The victim's reparative processes need time to come into play. The timing of the catastrophe itself is unexpected and unknown, often delaying the onset of resuscitative efforts. The damage of the disease or injury may be so great that prolonged periods of time are required to permit the body to repair or recover. For these and other reasons, additional concepts of life support are being investigated. These fall into two broad categories of tissue or organ transplantation and artificial internal organs. These are arbitrary delineations but serve to divide attempts to substitute healthy tissue or organs for that which is diseased or destroyed from those who attempt to devise mechanical substitutes to temporarily or permanently take over the function of such damaged tissue or body organs. Again, there is some precedent for each. Transfusions of blood cells and plasma, transplantation of corneas, bone and cartilage have been successful with proper donors. Artificial tissue and organs, such as heart valves, a cardiac pacemaker or the artificial kidney, work well for varying periods of time. Both groups tend to fail when whole organs become involved: the organ transplants are rejected by immune mechanisms developed by the host, and artificial or mechanical organs tend to be too large to meet many logistical requirements.

The unique power system called "fluid amplification" mentioned by Mr. Horton in the previous address, and by Mr. Woodward in the 1962 session of this Conference, has continued to lend itself to miniaturization and to overcome some of the other objections to conventional power sources for artificial internal organs. The medical equipment development effort by the Harry Diamond Laboratories and the Walter Reed Army Institute of Research has been quite rewarding. This effort can be used as an example of the increasing role of the engineer and other scientists in assisting the physician to achieve better patient care, a deeper understanding and explanation of many

BARILA

complex body processes and to realize the value of a multiple discipline approach to the solution of medical military problems.

Such miniaturization may lead to the availability of a small (one cubic foot) heart-lung substitute which could be quickly connected to a casualty's circulation, to take over its vital function, and "buy time" for the victim until conditions are optimal for definitive repair of his injury.

Some suggestion that this is a reasonable concept follows from some animal experiments with the Army Artificial Heart Pump. Two groups of ten dogs were artificially kept alive by heart-lung machines delivering the same volumes of blood flow at the same mean input pressures, the only difference being in the quality of the blood flow. Ten dogs were so perfused by a conventional steady flow pump and the other ten by the Army Heart Pump set to deliver a pulsatile blood flow quite similar to normal pulsed blood flow, but still at the same mean pressure and minute flow delivered by the conventional pump. The difference in survival was remarkable. In the group of dogs pumped with a steady blood flow for two hours who were then taken off this artificial circulation, none survived longer than 21 hours, and the mean survival time was 4 hours. In marked contrast 7 of the 10 dogs perfused with a pulsatile blood flow, again at the same mean flow and pressure as the first group, survived the two hours of artificial circulation. Actually, there is some bias here as deaths from any explainable cause are still reported. Two of the three dogs on the pulsatile blood flow died of air embolism, a technical area rectified by the addition of an air trap to the artificial blood circuit. The third dog had inadvertently been fed on the morning of the procedure, regurgitated stomach contents and could not be adequately ventilated prior to the artificial circulation experiment.

Apparently similar benefit of pulsatile blood flow has been reported by civilian investigators using the Army Heart Pump. But as is often the case, many new questions arise. Why should a pulsating heart and an elastic arterial system be the optimal system for maintenance of man's life within the bounds of his body's environment? What is the best experimental design available to prove that a pulsatile blood flow will offer ever greater chances of surviving a period of interruption of normal blood flow when thousands of heart disease victims have been successfully cured while their circulatory systems have been taken over by a steady flow pump? We do not have the answers to these and other more basic questions - only speculations most of which involve, again, a factor of time. The human body has great reserves which enable it to withstand severe physiological insults if they are of short duration. Patients who require artificial circulation for more than two hours, with conventional steady flow of blood, have a much higher mortality than those whose disease can be repaired in less than two hours. This may reflect no more than the severity of the disease, but the possibility that even greater achievements in medicine will follow the modification of the

BARILA

character of the artificial blood flow is more than exciting.

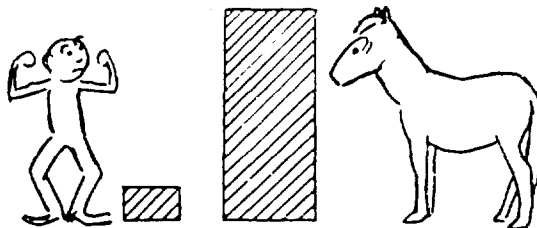
Departing from specifics of medical research, rewarded by collaboration with other scientists, one may ask other general questions. One of paramount interest to me also involves time and can be posed: "Whither Army Medical Research?" or even: Whither medical research? A physician treads his path of treatment of the sick while maintaining his drive and ambition to learn and teach more of the mysteries of the body and is at first unaware that his path is not one, but four or more. As time passes, his clinical path of healing the sick meets a fork which requires him to teach other less experienced physicians what he has learned. He is still able to travel both paths with some expenditure of time, but a third path branches off, requiring him to study or do research experiments to determine for himself even better ways of curing the ill. Again, his agility may permit him to progress along all three paths, albeit at a slower rate than if all his energies could be concentrated on a single goal. At a time when he might be reconciled to such gymnastics, a fourth path looms and must also be travelled, that of administrative duties almost divorced from one or more of the other paths. With more time, he finds he is spending more and more time on one or two paths, working little or none on the others. How then is he to be the complete physician, demanded by his four-in-one "career"? I do not know the answer, but part of it lies in the help and encouragement of other professional people who find themselves on similar paths requiring some agility. If the travel along some of these paths coincide between the scientific professions, all of the paths can be savored and the goal still reached expeditiously. I earnestly request your help and pledge my own.

CONTROL, AMPLIFICATION, AND FLUIDS

BILLY M. HORTON
 HARRY DIAMOND LABORATORIES
 WASHINGTON, D. C.

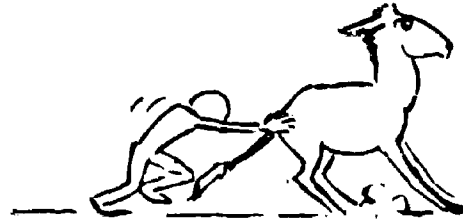
When General Lotz invited me to come and speak before you, I was first tempted to parade: a list of functions which can be accomplished by a fluid amplification system; a projection of things to come; and some discussion of their applicability to military (particularly Army) problems. Then I decided that those of you who have a strong or special interest in this field may have attended, or been represented at, the Second Fluid Amplification Symposium, held just three weeks ago in Washington; and that in a short presentation such as this, anything like adequate coverage of activity in this field would be impossible. I do believe, however, that it is possible to present the "why" of fluid amplification, that is: why is there such a thing; and why is there such current high interest in this field? (The current R&D budget is several million dollars annually.) The short answer is that it is a new, conceptually simple, method of control, the potentialities of which are being determined and exploited. But of course I shan't stop with the short answer. Instead I'll say that the control of energy lies at the root of man's standard of living and use that as the starting point.

For most of recorded history, the bulk of the energy which man has controlled has been that which he processes with his body. At 2500 calories per day, this turns out to be about 3 kilowatt hours, or about 4 horsepower hours per day (not a whole lot), and most of this is used to keep him warm. A 200-pound man walking up a 5% grade for a 75-hour week accomplishes about 1 horsepower hour of work. Now man has always been able to think of more things to do than he's had the energy to do, so it isn't too surprising that he began looking for means of controlling energy beyond that which he



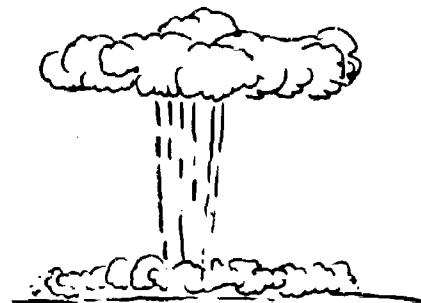
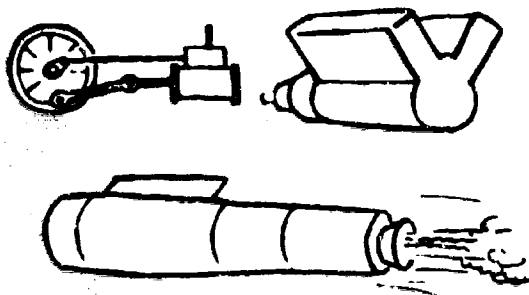
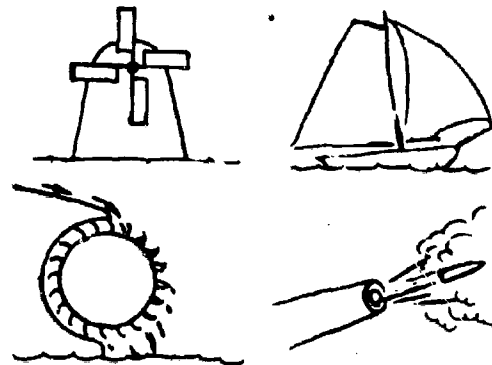
HORTON

himself ingested. He has domesticated animals. A man with a horse can raise this "energy" standard of living by an order of magnitude, even assuming the horse works only a 40-hour week. (By the way, an ordinary horse would wear out if you insisted that he perform according to a one-horsepower specification (a 125-pound pull at three miles per hour)). Wild animals expend large amounts of energy, but neither they nor certain stubborn mules contribute to this "energy" standard of living, because the element of control is missing.



Directed energy is what man needs and is what he wants in order to achieve many of his goals.

Now he has been a prolific inventor of ways to do this: He has used windmills, sail boats, water wheels, gunpowder, steam engines, gasoline and diesel engines, jet engines, nuclear fission, and nuclear fusion as means of bringing various sources of energy under his control. He has also been an innovator of institutions to control and direct the energy of other human beings, such as slavery, serfdom, and the totalitarian state. Of course these institutions are much more complex than the physical systems I just mentioned, but I mention them because they represent another part of the long history of man's effort to achieve a standard of living by augmentation of what he alone can accomplish.



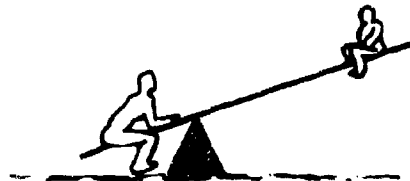
This brings us to the next question: Under what conditions is the control and direction of some source of energy worthwhile? Well, it's pretty obvious that the controlled energy ought to be greater than the energy required for the controlling process, or, to be a little more exact, the result accomplished by the controlled energy should be greater than the results which could be accomplished by the controlling energy acting directly. If this condition is not met, the system hardly seems worthwhile. If it is met, then the process has amplification. There are, of course, many necessary and useful controls which do not have amplification, but those with amplification have received much more attention.

If a string of controls, each having amplification, were connected in series, then a small amount of energy injected as a con-



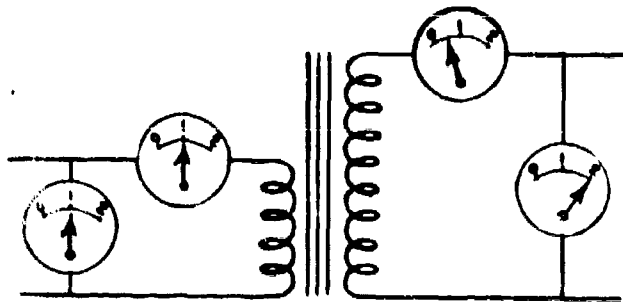
trol to the first could effect the control of larger and larger amounts of energy down the line. In connecting them together, however, we get into the problem of the "medium" or the "media" involved in the process. A windmill is a system for directing the energy of the wind so as to turn the millstone, but is not particularly well suited to controlling the flow of energy through another windmill system. In other words, you can't connect windmills in cascade very well because a rotating shaft doesn't control wind. This input-output problem is one which greatly hampers any system when applied to any complex problem (I'll have more to say about this later).

But before we get too deeply involved in a discussion of the cascading of amplifiers, their use as controllers, and into such things as feedback systems, I'd like to explore a little further just what is meant by the term "amplification". We can say that a simple lever "amplifies" force. Or, by moving the pivot, we can make it "amplify" displacement. In this case, as with all other simple machines, the fact that no "other" source of energy is used means that the energy amplification is unity or less, i.e., no more work comes out than goes in. Now no one considers a seesaw, a wedge, nor a screw an amplifier, but an imprecise definition of amplification might well qualify them as such.

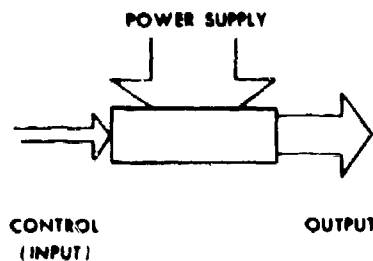


HORTON

Similarly, in the electrical field, a transformer can provide a very high voltage gain. Is it an amplifier? Since the voltage gain occurs at the expense of a reduced current, it is no more an amplifier than a seesaw.



We need a concept of amplification which will distinguish between such things as an electrical transformer and a vacuum tube, between a hydraulic jack and a power steering system. This concept must deal with the ability of input energy to control output energy, more of this output energy than comes in as control at the "input". This means, of course, that some sort of a power supply must be available to provide this "greater amount" of energy. This power source can be anything from a fuel cell to an atom bomb.



Now in continuous processes, people are very often interested in energy rates, so we more often than not talk about input power, output power, and power amplification, or, more commonly, power gain of an amplifying device or of a system.

I remember once when a friend seriously questioned the validity of talking about "power gain" when the output power was considerably less than that coming in from the power supply. There was some confusion in his mind between the concepts of "gain" and "efficiency". While the power gain of an amplifier needs to be greater than one to attract much attention, its efficiency will surely be less than one, since it is just about impossible to modulate or control the flow of energy by any means without some losses in the process.

We've been discussing an area of thought which has occupied man's thoughts for centuries. And he has been very successful in achieving the kind of control of energy he desires. For example, hydraulic, pneumatic, steam, and electrical controls act as amplifying devices, some giving a continuous range, and some giving only discrete levels of output power. Today we would class them as either "analog" or "digital" systems. Up until about a half century ago, the operation of all these systems required the motion of mechanical

HORTON

parts to accomplish their purposes. Now there is no natural disadvantage to moving parts. A steering wheel serves very well, and the valves it controls do a magnificent job, thanks to a little oil pump attached to the engine. Similarly, the exciter generator in a power generating station delivers a large excitation current to the alternator with high efficiency, and in this capacity, it is acting as an electrical amplifier. This kind of electrical amplifier is, however, neither sensitive enough nor fast enough to handle very small radio signals (often less than a billionth of a watt). This is, of course, where the vacuum tube came in, and, as you know, for a half a century has dominated the high-speed electrical control of energy for purposes too numerous to mention. In the vacuum tube, De Forest combined the electric field effects of Faraday and others with the Edison effect to achieve a control of energy which gave a power gain, and it had no moving parts. The same medium was used at input and output; that is, electron flow controlled electron flow. This is a very happy state of affairs as one contemplates connecting together many elements into complex systems.

The transistor has augmented the capabilities of, and in many cases replaced, the vacuum tube but has retained the two important characteristics of: no moving parts; and single medium control.

But this is enough about amplification as an idea, and enough about electrical systems. Let's see what has been happening to fluid systems for a few hundred years.

The control valve or gate in a water wheel control system is a very neat and high-gain fluid controller. The sliding valve on a steam locomotive and regulating valve on a steam turbine are effective controllers of a large power. The flapper valve on a pneumatic jack hammer serves as an amplifier as it controls its supply of fluid power. The capabilities of these controls and many others are such that if electrical systems had not been developed around the turn of the century, there is little doubt that fluid distributing systems would have taken over a considerable burden. In London, Manchester, and Glasgow, hydraulic power transmission systems were installed. Pressures from 700 to 1600 pounds per square inch were employed. The fluid was water. Cast iron pipe up to six inches in diameter permitted delivery of 140 horsepower over a ten-mile stretch with a 10% line loss. The maximum horsepower which can be transmitted over a one-mile-long, six-inch-diameter line at 1120 pounds per square inch is 570 horsepower. While such systems were expensive to install and maintain, they did operate. Some work was also done on wave transmission hydraulic systems (i.e., alternating current, or alternating flow systems), some of them were even three-phase systems.

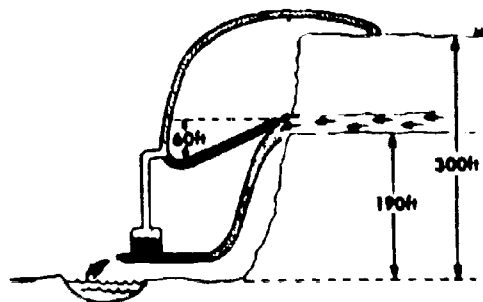
During this same period, pneumatic systems were being exploited mostly in France. In the 17th Century, Denis Papin compressed air with power from a water wheel and transmitted it through tubes to be used elsewhere. From 1800 to 1900, many improvements in

HORTON

compression, distribution and use of air were made as the thermodynamics involved became better understood. Pneumatic systems became popular in mines for safety reasons. The ready availability of air made pneumatics popular for things ranging from rock drills and pneumatic wrenches to railroad brake systems. These systems all require moving parts to control the air flow. The various valves are acting as amplifiers, since the controlled energy, or power, is much greater than the controlling energy.

But there was one man who sought to operate a pneumatic system without moving parts. I'd like to quote from the 1947 issue of the Britannica:

"A remarkable pneumatic transmission system was installed in 1890 by Priestly in Snake River desert, Idaho. On the north side of the river is a cliff, nearly perpendicular, about 300 ft. high. One hundred and ninety feet above the river, for a considerable distance along the cliff,

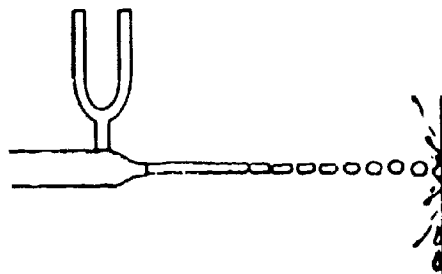


streams of water gush out from between the bottom of the great lava bed and the hardened clay of the old lake bottom. Priestly, . . . , built a pipe line down the bluff and trained the water into it in such a way that it carried a very considerable quantity of air in the form of bubbles along with it down the pipe, compressing it on the way. The air was collected at the bottom in a covered reservoir, and taken up the cliff again to the lower part of an inverted siphon pipe, one side of which reached down from the water-supply about 60 ft. and the other side reached up and over the bluff. Allowing the water to fill both sides of the pipe to the level of the water-supply, he admitted his compressed air at about 75 lb. pressure into the long side of the pipe near the bottom, and soon had water flowing upwards over the cliff and irrigating a large tract of rich lava land. He had made a power, a transmission and a motor plant without a moving part. A similar compressor was installed near Montreal, Canada, in 1896; another at Ainsworth, British Columbia, in 1898; and another at Norwich, Conn., in 1902. These are called hydraulic air compressors and show an efficiency of about 70%. They are particularly adapted to positions with a large flow of water with a slight fall or head."

HORTON

Some later editions of the Britannica don't have this account, displaced, presumably, by more modern technological developments. But it would be difficult to find a more elegant system than Priestly's, especially if you imagine how badly he must have wanted that water up on top of that plateau. Although Priestly's system did not amplify, it did show a strong appreciation for the simplicity and reliability of a no-moving-parts system.

Meanwhile, back in the physics laboratory, Edison, in the United States, and C. V. Boys, in England, performed some very interesting experiments in which they achieved sound amplification by vibrating an extremely fine jet of water, then allowing the jet to fall on a sounding board.



What happens is that the jet breaks up into droplets whose size is determined by the frequency of the sound. In this case, the first source of energy is the surface tension energy being continuously supplied to the surface of the jet by the sources of pressure. The jet is going to break up onto droplets anyway, as you know, and the sound waves simply determine exactly where the breaks occur. The amount of surface tension energy available is small, and the sound does not have "full control", being able to only slightly retard or speed up the stream break-up process. This subdivision into droplets then controls the timing of the delivery of the kinetic energy of the droplets to the diaphragm, and the whole apparatus had a power gain.

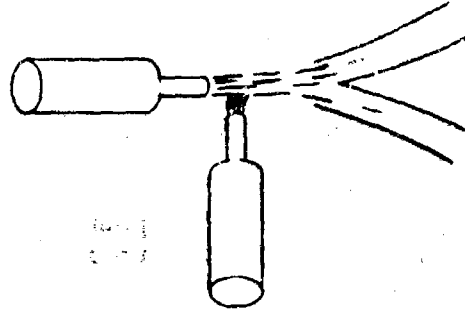
The various ways of controlling energy which I have referred to so far are, except for the transistor, more than a half century old. The tools available for discussing and analyzing them today are, however, rapidly improving. For example, the concepts of impedance, gain, available power, and the whole field of network theory permit rapid analysis of most linear systems. Our understanding of system stability (and what to do about a "hunting" governor), our ability to cope with distributed systems, our "black box" concepts, our treatment of noise, interference, nonlinearity, and parametric changes, and above all our tendency to look at all systems from an "information transfer" viewpoint--all of these have warranted taking a fresh view of many old systems. In the particular field of fluids, rapid advances in aerodynamics and hydrodynamics, spurred on by airplanes, bullets, torpedoes, rockets, turbines and jets, having provided a much improved understanding of fluid flow in recent years, though I would hasten to point out that the basic properties in fluids pretty well assure that analysis on fluid systems will always be more complex than the analysis of similar electrical systems.

Today a fresh view is being taken of fluid-actuated systems, triggered by a question I had the good fortune to ask and to answer.

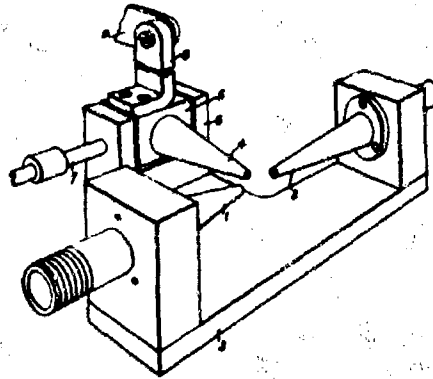
HORTON

That question was: Is it possible to make an amplifier (one with a power gain), for use in a fluid-actuated system, without the use of any moving parts? Not a very impressive sounding question. For some reason it doesn't seem to have been asked before, or, if asked, not answered before.

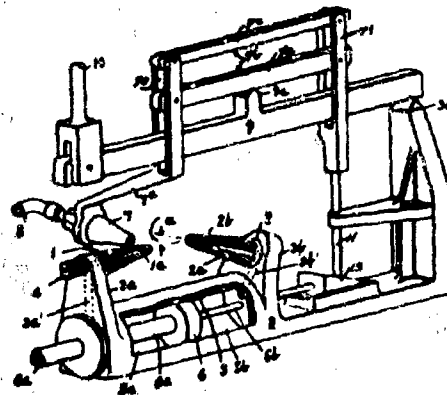
The answer is, of course, "yes", and such an amplifier is shown here. All of you are familiar with the work that can be accomplished by a high-speed jet of fluid. A fluid amplifier has what we call the power nozzle to provide such a high-speed jet or power stream. The source of energy is the pump which supplies the nozzle. In the jet, the energy takes the form of kinetic energy of translation of the fluid. The control nozzle(s) determine where the kinetic energy of this stream is delivered and hence control the flow and pressure in each of the two output apertures. An amplifier such as this can have a power gain of from 10 to 100, depending upon what other characteristics (such a pressure gain and efficiency) are desired.



Jet relays which had been used in the past are shown in the 1946 patents of Braithwaite

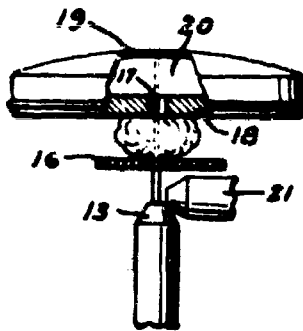


and of Todd. Both of these permit relative position determination, readout, and control, but neither of them uses the deflected stream and neither is designed to conserve stream energy nor to achieve a power gain.

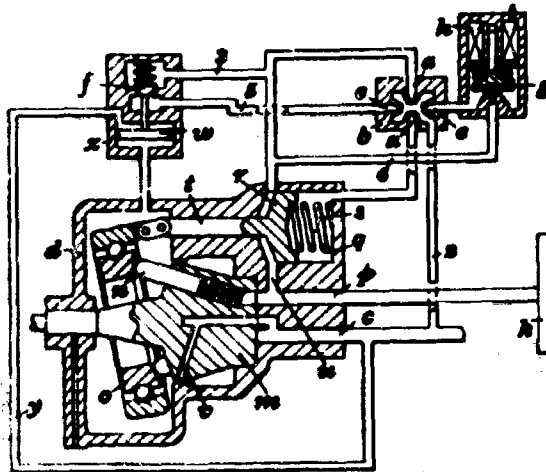


HORTON

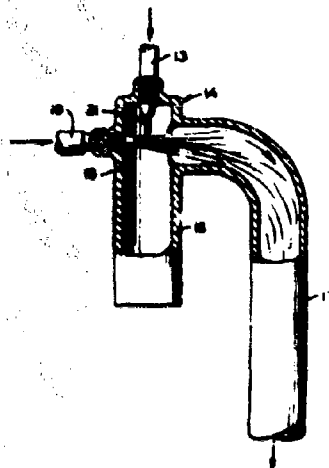
Hall in 1953,



Harris in 1955,



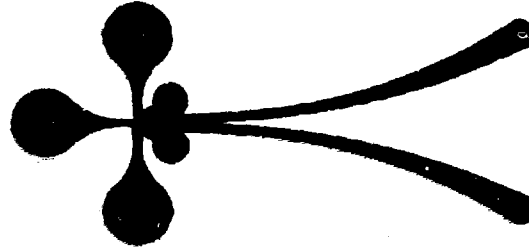
and Magnuson in 1959



show jets deflecting jets for sound generation, feedback signal interruption and bottle filling, but again none uses the energy of the deflected stream, none is designed for good power gain, and none discusses gain or amplification.

HORTON

Now taking another look at a fluid amplifier, we see that it is sensible to connect them in push-pull, since each is a "beam deflecting" device rather than a "beam stopping" device, as is a triode. But connecting fluid amplifiers is just about the biggest problem in this field. There are good reasons for the difficulties: first, the basic forces in fluids in motion, as shown by the Navier-Stokes differential equations, are nonlinear (the inertial forces depend upon velocity squared and the viscous or frictional forces depend upon velocity); second, the continuity equation must hold (what goes in must come out).



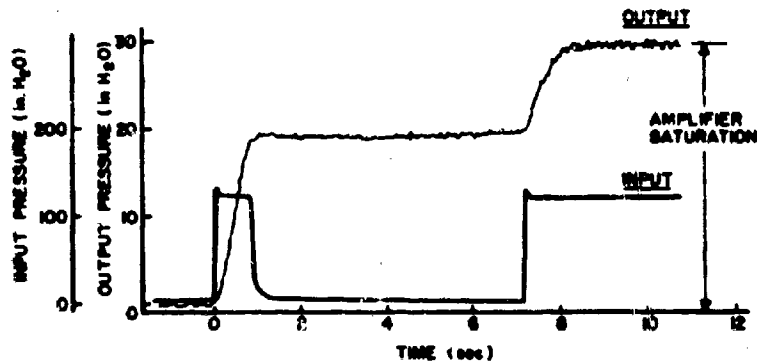
Now there's nothing new or exciting about either of these, but they do have special significance in fluid amplification systems. With moving parts, it is easy enough to shut off fluid flow. Without moving parts, the continuing flow must be accommodated. Without diaphragms (the equivalent of blocking capacitance in electrical systems), this flow must be either properly used or properly discarded after its energy has been extracted. The fluid nonlinearity makes it a ready converter of its steady (or DC) flow energy into fluctuating forms. You have all heard the noises and whistles of fluid jets. Of course electrical systems have their noises and parasitic oscillation too, but nowadays these are under pretty good control. Fluid systems need the same discipline and more, because of nonlinear effects.

Lest you get the impression that nonlinearity is, per se, a disadvantage, I'd like to point out that without nonlinearities appropriately used, neither present day telephonic nor radio communication could operate. In a fluid amplifier, the extreme simplicity is a direct consequence of these nonlinearities and is very closely associated with the vector properties of a jet of fluid, without which a fluid amplifier would not work.

Much more work is needed to clear up this area. More attention needs to be given to the processes by which stream energy gets degraded by the parasite, turbulence, not only because of our interest in conserving that energy, but, more importantly, because we want to avoid its showing up as noise. We need to adopt and adapt techniques from the communications field which arrange for most of the noise to be generated in frequency bands not used by the signal. We have hardly scratched the surface in this area, nor in problems of matching passive components and transmission lines to fluid signals consisting of combined acoustic and mass transport positions.

HORTON

We have been working on a fluid integrator, the basic element, for an analogue computer. How well it is doing is shown in this integration. The range is about fifty times the noise level. So you see we have much more work to do.



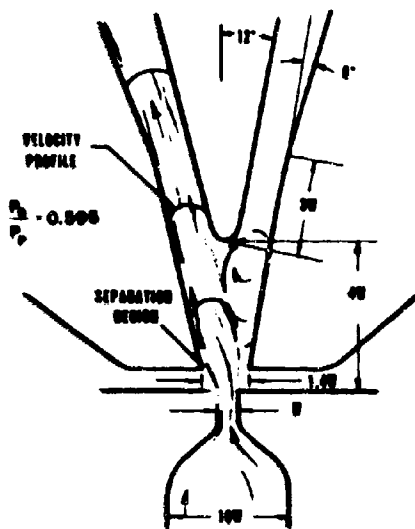
INTEGRATOR RESPONSE TO A STEP INPUT-TWO STAGE AMPLIFIER

You know that proportional systems can be made to operate in a digital fashion by the use of feedback. Here we see such a flip-flop. By using two logic elements, this can be made into a scaler stage.



But Ray Bowles and Ray Warren discovered a much more elegant way to accomplish bistability and digital operation. They did this by positioning the sidewalls of a fluid amplifier so as to cause the stream to attach to one side or the other by the Coanda effect, known since about 1932. This basic bistable element has formed the foundation for a whole rash of digital computer and controller work throughout the country, and, in fact, in a number of foreign countries.

Shown here is one such computer in breadboard form.



HORTON

Fluid amplification has its share of zealots (one journal said "Their strange devices can do what is now done by gears, valves, and vacuum tubes, yet they consist only of weirdly shaped conduits for streams of liquid or gas").

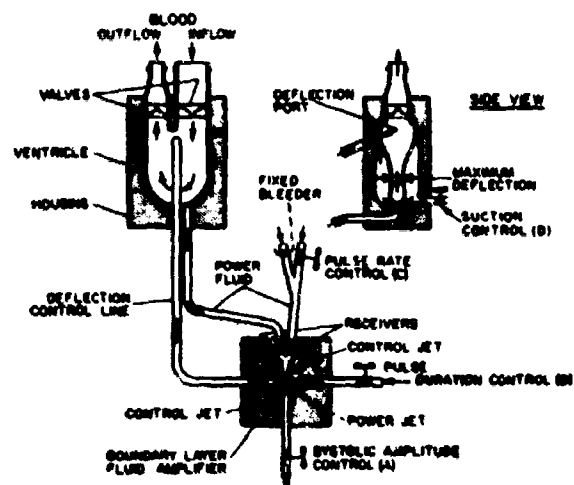
Then there are others whose reference to a diode as a "fluid control" element and Coanda's very fine work as though it were a usable amplifier, seem to miss the whole point in achieving the control of energy flow. But any field is bound to have a spectrum of viewpoints. Fluid amplification certainly does, due in part to the mixture of disciplines which it employs.

Current rate of activity is high. At the Second Fluid Amplification Symposium in May, over fifty papers were presented, touching such topics as:

"...Dissipation of Energy in Free Turbulent Flows..."
"...Curved Turbulent Mixing..." "...Underexpanded Jets..."
"Noise Reduction..." "Wall Attachment..."
"...Stability of Fluid Systems" "Acoustic Control of..."
"Response...to Step Input" "...Characteristics of Vortex Valves"
"Application...to On-Off Control Systems"
"Development...for Multistage Operations"
"...Analog Computation with Fluids" "...Fluid Encoding..."
"Application...to Speed Control" "...Hydrofoil Control..."
"...Pure Fluid Timers" "Transient Behavior..."
"Interconnection of Fluid Amplification Elements"
"Staging of Closed Proportional Fluid Amplifiers"

So you see people have a variety of interests and objectives.

Probably the most interesting application of all is the Army Heart Pump which began at the Harry Diamond Laboratories (then DOFL) under Mr. Woodward with sponsorship and medical guidance provided by the Walter Reed Army Institute of Research. Col. Barila will have more to say about this.



HEART PUMP SCHEMATIC

HORIZON

Military applications of fluid amplification currently being investigated include: logic and timing circuits for ordnance; rocket thrust vectoring, missile attitude control; automatic piloting of aircraft; hydrofoil control; jet engine control; heart massaging; and artificial respiration. It is a little early to say just how far this new medium will go in each of these fields, but the inherent advantages of simplicity and ruggedness give strong impetus to the work. The term "simplicity" above does not necessarily refer to the state of understanding necessary to cause an all-fluid system to work properly. In fact, each device having significantly different characteristics usually points out the need for additional fundamental knowledge of fluid flow. The resulting practical impetus to research activity is aided by the realization that all-fluid systems can "think" in the manner of electronic systems (though not nearly as fast) and that full exploitation of this capacity requires better analysis and synthesis of fluid flow phenomena in order to permit more orderly design of fluid systems.

The rate of development of fluid amplification will depend heavily upon this increase in fundamental knowledge, as well as upon inventiveness and ingenuity. Its ultimate role in the control and computation field will not be known for years.

ARMY PARTICIPATION IN PROJECT SYNCOM

Brigadier General J. Wilson Johnston, USA
Commanding General
United States Army Satellite Communications Agency
Fort Monmouth, New Jersey

The SYNCOM-SATCOM network we shall discuss today is the newest frontier in a series of U.S. experiments using communications satellites to transmit instantaneously messages sent between inter-continental ground stations. Both the Department of Defense and private industry have long been interested in the development of such a satellite communication system to improve long distance communications, because the present communications systems are subject to disruption by natural or man-made phenomena. Satellites offer a means of overcoming these serious limitations.

To provide some background for a description of the current SYNCOM-SATCOM network, I will first review the several types of communications satellites. Satellites fall basically into two categories: active and passive. Passive satellites merely act as reflectors of radio energy received from ground stations, while active satellites contain electronic equipment to retransmit signals.

SLIDE 1 ON. "SATELLITE COMMUNICATIONS CAPABILITY"

This slide shows both passive and active communications satellites. Three types of passive satellites are shown: the earth's natural satellite, the moon, which has been used by the Navy in their MOON BOUNCE experiments between Maryland and Hawaii; the ECHO balloon satellite; and, finally, the Project WESTFORD dipole belt.

The active category can be subdivided into medium altitude, random orbit satellites; and high altitude, synchronous orbit satellites. The Army's COURIER, NASA's RELAY and the commercial TELSTAR are examples of the former. NASA's SYNCOM satellite, which we will discuss today, is the world's first high altitude, synchronous orbit satellite.

SLIDE 2 ON. "24-HOUR SATELLITE COVERAGE"

JOHNSTON

The next slide shows the communications coverage obtained using a stationary satellite stationed at 55 degrees west longitude, approximately the original position of SYCOM II, which was in an inclined synchronous orbit, from mid-August 1963 until 17 March 1964, when the satellite was commanded to move to its present position in the Pacific. You can see that with one satellite at an orbital altitude of 22,300 miles, links can be provided to all the capitals in North and South America, Europe and Africa.

SLIDE 3 ON. "MEDIUM ALTITUDE ORBIT EARTH COVERAGE"

In contrast, the coverage of a random orbit satellite at 5,000 miles is shown on Slide 3. Because of the lower altitude, coverage is slightly reduced. More important, since this type of satellite is not stationary, the pattern of coverage is continuously moving in the orbital direction. This means that to provide highly reliable, continuous, world-wide communications, as many as 30 such satellites would be required. (This disadvantage is, I must add, partially offset by the relative simplicity of a medium altitude satellite.)

The Army has had a long background of experience in communications satellites, and I am going to show some examples of earlier experiments.

SLIDE 4 ON. "SATELLITE BACKGROUND"

The first satellite communications experiment of the Army, conducted by the Signal Laboratories at Fort Monmouth, N.J. in 1946, was a project called DIANA, wherein Signal Corps scientists bounced radio signals off the moon.

SCORE represents a later experiment in December 1958, when the Army Signal Laboratories participated in Project SCORE, which was the world's first active repeater type communications satellite. This was the satellite that broadcast President Eisenhower's Christmas message.

More recently, in September 1960, the Army Signal Laboratories at Fort Monmouth further advanced the state-of-the-art in Project COURIER, which was a highly sophisticated successor to SCORE. This was designed to be a delayed repeater but could also be used in a real-time mode, and operated with ground stations in New Jersey and Puerto Rico.

ADVENT was a program initiated at about the same time as COURIER, but for a far more complex stationary satellite. The management of this tri-service program was assigned to the Army, but because of technical problems in the satellite and rocket booster, the program was reoriented in early 1962 to the present military medium altitude program.

JOHNSTON

SLIDE 4 OFF.

With this as a background, let us look at the Army support of NASA in the SYNOOM program. Under a joint DOD/NASA arrangement, the U.S. Army Satellite Communications Agency, or SATCOM Agency as it is often called, was charged with providing this support using assets derived from the DOD ADVENT Project. The SATCOM Agency was created in 1960, and many of its members are "old hands" who have participated in those earlier satellite communications projects that I have mentioned.

The Agency was formed to provide the communications stations and to plan, conduct, and evaluate satellite communications tests for the Defense Communications Satellite Program. As a result, it was prepared to perform these functions for NASA's SYNOOM Project, and we have enjoyed very fine working relations.

In describing the Army participation in SYNOOM, I should like to first tell you about the spacecraft, which was built for and launched by the National Aeronautics and Space Administration.

SLIDE 5 ON. "PROJECT SYNOOM"

The SYNOOM spacecraft is a synchronous altitude satellite. It is shown on this slide. The satellite was built by the Hughes Aircraft Corporation. It is about 28 inches in diameter and weighs about 85 pounds in orbit. It carries, in addition to communications gear, a small "on board" propulsion system for orbital corrections.

The ground communications support for this project was developed, engineered, installed and operated by the Army. The complex consists of fixed terminals at Fort Dix, New Jersey and Camp Roberts, California, several transportable terminals and one seaborne terminal, the U.S. Naval Ship Kingsport. I shall describe these in some detail.

SLIDE 6 ON. "FORT DIX"

This slide shows the large, fixed, high precision research terminal AN/FSC-9 located at Fort Dix, N.J. The Fort Dix station has a large, parabolic antenna with a tracking accuracy of plus or minus 0.024 degrees, and an associated operations building. The antenna weighs 190 tons and is 80 feet high. The diameter of the dish is 60 feet. This antenna stands on a massive concrete foundation 30 feet deep and 84 feet wide.

SLIDE 7 ON. "CAMP ROBERTS"

This slide shows the fixed SATCOM ground station at Camp Roberts, California, which is the electronic twin of our Fort Dix station. It first began transmitting to the Kingsport via SYNOOM II on August 8th, when the satellite first came in "sight". The next day in a public demonstration, messages were transmitted which set a

JOHNSTON

record for the greatest distance (about 7700 miles) ever spanned between two points on the earth's surface via an active communications satellite.

SLIDE 8 ON. "USNS KINGSFORT"

This slide shows the shipboard terminal located initially in the harbor of Lagos, Nigeria. This location was selected because it is near the area where SYNOOM first entered its synchronous orbit over east Africa, thereby providing a facility to control the satellite and to communicate with it. The Kingsport terminal features a 30-foot diameter antenna housed beneath the radome, and has a tracking accuracy of plus or minus 0.05 degrees. Some five and one-half hours after launch, the satellite was given the final command to place itself in a near-synchronous orbit; and you can imagine our relief when this command was executed properly and we were able to send our first message through the satellite.

SLIDE 9 ON. "LAKEHURST TERMINAL"

This slide shows one of our transportable terminals, the AN/MSC-44, in its location at Lakehurst Naval Air Station, N.J. It was designed to work with the SYNOOM satellite, and to provide the main terminal in the United States to work with the Kingsport. This terminal is capable of being moved aboard aircraft or ships or on the road, and is completely housed in eleven vans. At present, we have an identical terminal installed in the Philippines in preparation for the launch of SYNOOM III. The antenna reflector is disassembled and hauled on a separate trailer. The antenna weighs approximately 25 tons, and when erected, is about 40 feet high. Its dish diameter is 30 feet, and has the same accuracy as that aboard the Kingsport.

SLIDE 10 ON. "SYNOOM SPACECRAFT ORBIT CONFIGURATION"

I would now like to illustrate by this slide the journey of the SYNOOM spacecraft since its launching on 26 July 1963. The launch was from Cape Kennedy, then Canaveral (Point 1) and was boosted by a Thor-Delta rocket, which carried the satellite as shown on this sequence diagram (Point 2) where the final burn of the Delta vehicle occurred and it was separated from the spacecraft. The satellite then coasted for approximately five and one-half hours until it reached the 22,300 mile altitude (Point 3), where a solid fuel rocket aboard the spacecraft was fired to place it in a near synchronous orbit. From that point, the satellite began to drift to the east, but following a series of commands sent to the satellite by a NASA control group aboard the Kingsport at Lagos, Nigeria, the direction of drift was changed to the west, and on 15 August SYNOOM II arrived at its final position of 55 degrees west longitude, centered over Brazil. After final orbital velocity corrections, it remained at this nominal longitude until 17 March 1964 (Point 5). During the drift period the attitude of the spacecraft was changed on command to optimize the antenna pattern with

JOHNSTON

respect to the earth.

Because of booster limitations, it was not possible to place the satellite in a completely equatorial synchronous orbit, and as a result the plane of the satellite's orbit is inclined to the equator. Therefore, the satellite during each 24 hours, moves back and forth across the equator and, with respect to a point on the earth, traces a figure eight (Point 4). It always crosses the equator at approximately the same point each day; for example, it drifted only one quarter of one degree in the 30 days between 15 August and 14 September. I should point out that sufficient fuel is aboard the spacecraft to perform "station keeping", or the maintenance of a nearly constant longitudinal position for a period of at least three years.

We leave the radio beacon in the satellite on 24 hours a day, but only turn the telemetry on periodically to determine the condition of the spacecraft. All of the stations described except Camp Roberts could see our relay station in space for 24 hours each day. Camp Roberts could see it for 16 hours each day.

SLIDE 10 OFF

On August 23rd, President Kennedy formally inaugurated the SYNCOM-SATCOM network in a conversation with the Nigerian Prime Minister, employing our Lakehurst transportable terminal and our ship station in Lagos, Nigeria.

Last September we participated in a SYNCOM-RELAY test circuit which provided voice communications from Africa to South America via North America. This was the first time two satellites were linked in the same circuit. The voice circuit began at our Kingsport station in Lagos, Nigeria, and passed through SYNCOM II to our Lakehurst station. It then went by overland wires to the International Telephone and Telegraph Corporation's space station at Nutley, N.J., which passed the circuit through RELAY to another ITT station on the outskirts of Rio de Janeiro.

SLIDE 11 ON. "MARK IV(X) TRANSPORTABLE TERMINAL"

This slide shows the highly transportable terminal acquired by the SATCOM Agency just last December. It is a "breadboard", using mostly off-the-shelf components. It has a 15 foot segmented antenna on a searchlight mount. Its overall weight is under nine tons. Together with its crew, it can be transported in one C-130 aircraft and be set up for satellite communications in a few hours. In spite of its size and simplicity, it performs remarkably well with SYNCOM. It has been demonstrated to representatives of a great many military agencies, including the STRIKE Command at Tampa, Florida, and the Army Commanders Conference in Washington, D.C. Its obvious tactical potentialities have created a great deal of interest.

JOHNSTON

SLIDE 11 OFF

Spectacular demonstrations of the SYNCOM Satellite Communication network, demonstrating its space communications capability for transmitting voice, teletype, and photographs, have overshadowed the very extensive and thorough experimental engineering test program which has been in progress since the successful launch of the SYNCOM II satellite. Over 2,250 hours of communications tests have been conducted. These include fifteen separate types of tests for evaluating the feasibility of global communications using a satellite such as SYNCOM II as a 24-hour relay for voice, teletype, and photo facsimile traffic between U.S. Army SATCOM Agency ground terminals. These tests, when completely evaluated, will define the technical capabilities of the SYNCOM-SATCOM type network and derive appropriate information as a basis for future systems. This test program is an extremely comprehensive one which will provide a concept of the testing which will be required before a space communication network can become truly operational.

The communication test objectives for the SYNCOM program are to demonstrate that a synchronous satellite relay system can provide a satisfactory global communication link and to determine the overall performance of these links in a number of communication modes. The philosophy upon which these system tests were based is to maintain simplicity of operation while still deriving a maximum of information. Within the framework of this philosophy, it was decided to record, on magnetic tape, all test signals that would lend themselves to this technique. In addition, three special test tapes were devised which contained highly concentrated variations of all of the test tapes. These special tapes were used in the early part of the program to obtain the most comprehensive data as soon as possible in the event of an early failure of the satellite repeater. Each test tape included its own calibration tones and test signals as well as a record of the operating mode applying to that particular test.

To understand the variety of testing required, it must be realized that the SYNCOM ground terminals have seven operating modes. These include receiver band widths, which can be selected in discrete steps from 10 kilocycles to 100 kilocycles; transmitter deviation, which can be adjusted in discrete steps from 4 kilocycles to 50 kilocycles; and a capability for accommodating information base bands at either 4 kilocycles, 20 kilocycles, or 50 kilocycles. In addition, the receiver is capable of operating as a frequency modulation detector or in a frequency modulation with feed back mode.

Because of the large number of variables affecting system performance, such as environmental changes, variations in equipment characteristics and operating conditions, the evaluation of the data derived was made on a statistical basis. The tests and evaluation of the results were therefore planned to provide a maximum of statistical information.

JOHNSTON

The communication experiments performed were classified into three major groups: namely, demonstration, technical performance, and technical characteristics tests. The demonstration tests, performed primarily with live inputs, were designed to demonstrate qualitatively the system performance and to evaluate those system characteristics, such as the effects of echo and delay, which are best measured through user appraisal. The technical performance tests, performed primarily with input signals derived from the prerecorded magnetic tapes were intended to measure quantitatively the system communication performance as a communication link and provide the basis for statistical analysis. The technical characteristics tests, using both live and taped signal inputs, were intended to measure quantitatively the link characteristics.

While the SYNCOM II satellite communication links could support a greater channel capacity, primary emphasis was placed on evaluating the system performance based on 4 kilocycle channels. Thus, the results which I will briefly review are those obtained for a single nominal 4 kilocycle voice channel or its equivalent.

SLIDE 12 ON. "RESULTS OF COMMUNICATIONS TESTS"

This slide gives a summary of the test results for a typical 24-hour period. The transmitter power radiated from the ground station during this test was 10 kilowatts, which saturated the limiter of the spacecraft transponder, thereby removing any level variations associated with the up link.

The first entry on this slide gives the maximum and minimum signal levels over a 24-hour period and those which have been calculated from the system parameters. Agreement between the measured values and the calculated values is excellent as shown in the last column.

The second entry of the chart gives the maximum and minimum communication signal levels obtained at two of the terminals participating in this experiment. These show the normal variations of signal level with range as the SYNCOM satellite drifted from its maximum northern excursion to maximum southern excursion, about its synchronous longitude.

In the last entry we see the signal plus noise to noise ratio for the 4 kilocycle baseband mode which offered the best performance, ranged from a mean value of 31 db to a maximum of 40 db. The mean average of 31 db provides a communication link which is considered to be of sufficient quality to consistently satisfy an average telephone subscriber. It should be noted that an equatorial synchronous satellite would not be subject to these variations in signal level, since it would remain fixed both in latitude and longitude.

JOHNSTON

SLIDE 12 OFF

In addition to the technical tests, various demonstration and special tests were conducted through SYNCOM. Among the most significant of these were in the study of echo suppression and delay effects. To date, round trip time delays of about 0.6 seconds, typical of the circuit lengths associated with a synchronous altitude satellite repeater, have been theorized by many authorities to constitute an unacceptable by-product of synchronous satellite communications. Two effects can be attributed to this delay. The first is that a result of an imperfect termination of the hybrid associated with a two-wire telephone circuit is the generation of an echo. When this echo is delayed by an appreciable period of time, due to the extremely long circuit length, the result is an almost intolerable communications circuit. Echo can be virtually eliminated as I will describe in a moment. The second effect is that of the delay alone, even in the absence of echo. This is an unalterable factor dictated by the circuit length and its presence may so disturb the psyche of the average talker as to render the circuit unacceptable.

In all operations of the SYNCOM communications satellite system, involving transitions from two-wire to four-wire circuits, echo suppressors were employed to eliminate the echo which is attendant on such transitions. The major problem encountered with two-wire circuits employing echo suppressors has been directly associated with low levels occasionally encountered on telephone lines. Operation of the echo suppressors at the higher signal levels encountered in most of the demonstrations, however, was completely satisfactory.

The psychological effect on an average talker of the 0.6 second round trip delay with the echo suppressed has been the subject of much discussion. Out of more than 1,000 participants in these demonstrations, only two cases have been observed where talkers experienced difficulty in attempting to speak over the spacecraft circuit because of the psychological effects of delay. These are considered rare and extreme cases. What has been indicated is that if an adjustment is required, the average talker makes this adjustment unconsciously in a matter of a few seconds. These reactions have been observed without any prior briefing of the talker.

I had hoped to tell you about SYNCOM III and the results of the tests which we will conduct, utilizing it as a communication relay. It was scheduled for an early May launch; however, the launch has been delayed until a later date. I will, therefore, only be able to describe, briefly, its scheduled performance.

It will be launched to the 22,300-mile altitude by a Thrust-Augmented Delta launch vehicle, which will be performing its first NASA mission. Because of the increased performance of this booster,

JOHNSTON

it will be possible to put SYMCOM III into a near equatorial synchronous orbit. It is expected that both SYMCOM II, which has been drifting westward at a very slow rate, and SYMCOM III will be positioned in the Pacific area and after a series of engineering tests similar to those that I have described for SYMCOM II, both SYMCOM II and SYMCOM III will be utilized in an R&D operational exercise to support trans-Pacific communication circuits.

What are the implications for the Army of the successful launch of the SYMCOM satellite? This is difficult to answer since it is the Department of Defense, not the Army, which manages our communications satellite programs. However, by utilizing SYMCOM, the Army is gaining valuable experience in the conduct of communications employing satellites. Furthermore, with SATCOM fixed stations on both the east and west coasts, with several transportable terminals, and a mobile shipborne terminal that can be deployed anywhere in the world, we have in being now a small strategic communications capability within the R&D system configuration we are testing.

We have been fortunate that we have had working with us in this project the Army Electronics Laboratories at Fort Monmouth, which developed the transportable and fixed ground terminals. We were also fortunate to have the Navy Bureau of Ships develop the floating terminal aboard the Kingsport.

CONCLUSION

This entire project has been a demonstration of cooperation, understanding, and a whole-hearted desire to make the project succeed. We have demonstrated technical achievements, effective coordination between departments of the government, and complete service cooperation. I should also like to credit the very important role played by industry in helping NASA and the DOD achieve this important advance in space communications.

With such technical capability, cooperation and respect among the participants, I am confident that this country will soon possess the capability of having a world-wide operating communication satellite system.

 SATELLITE COMMUNICATION CAPABILITY

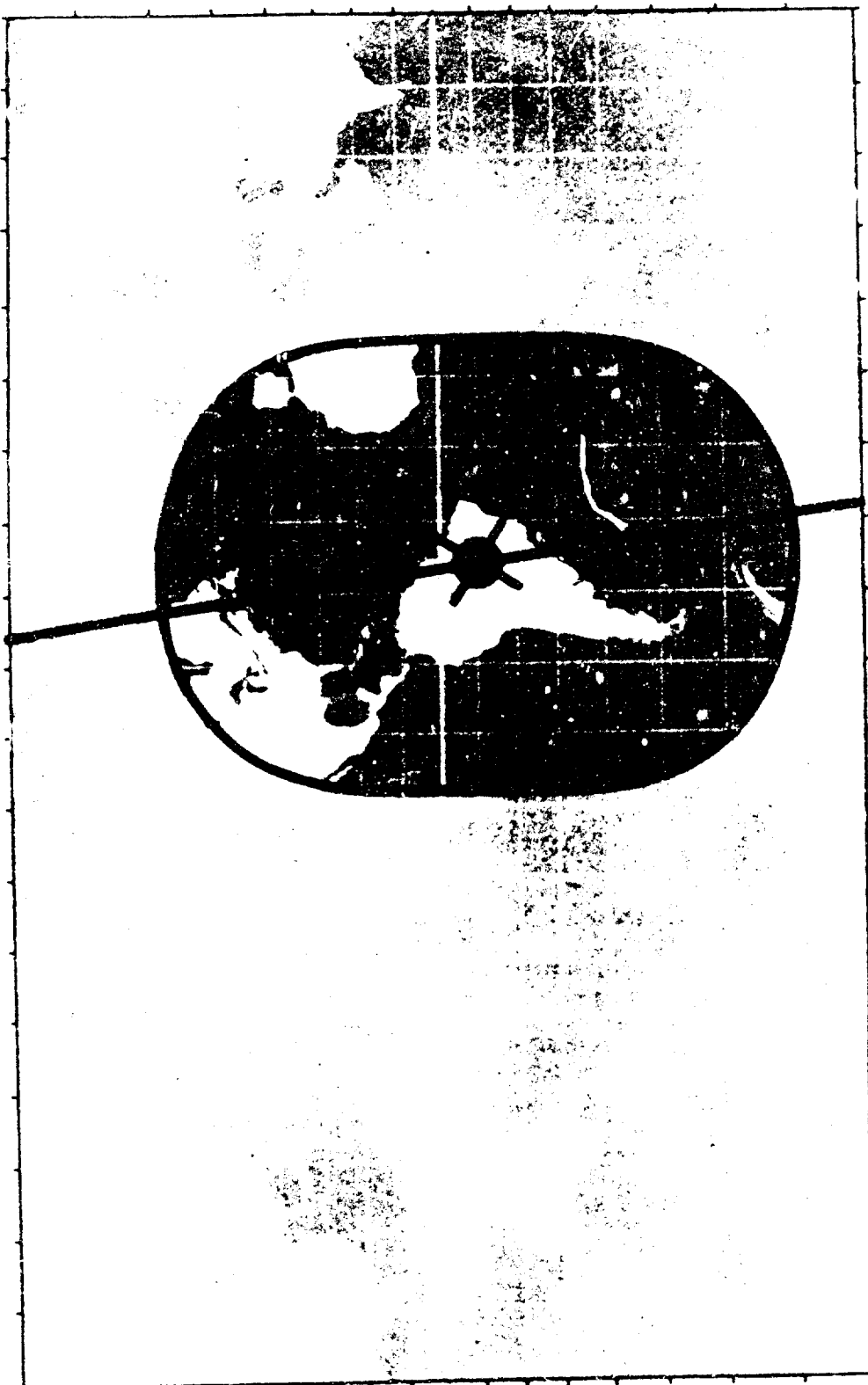


SLIDE 1



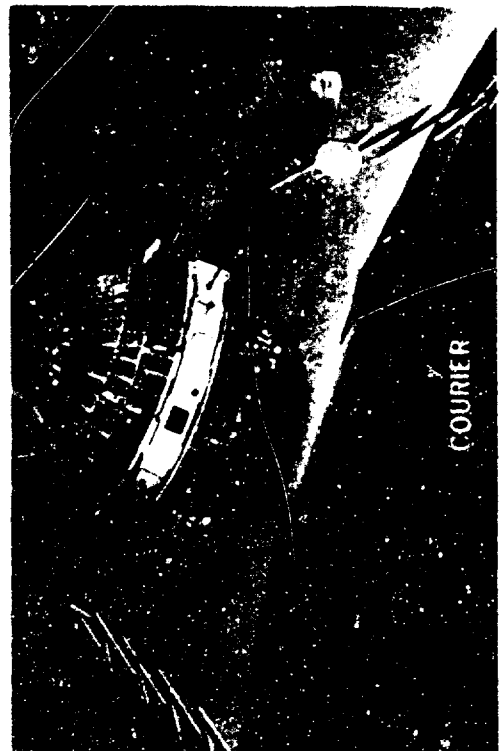
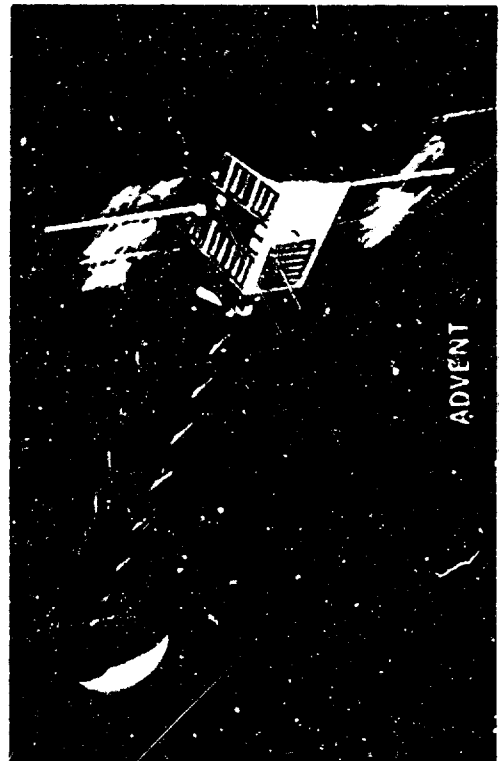
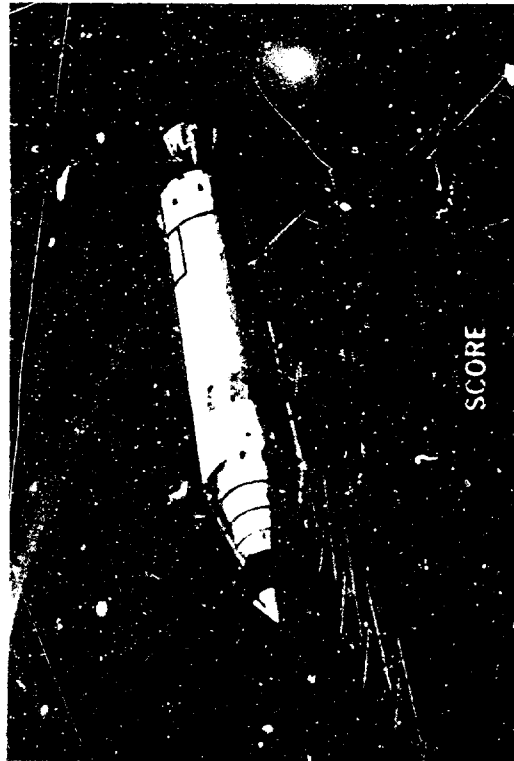
SLIDE 2

MEDIUM ALTITUDE ORBIT EARTH COVERAGE



SLIDE 3

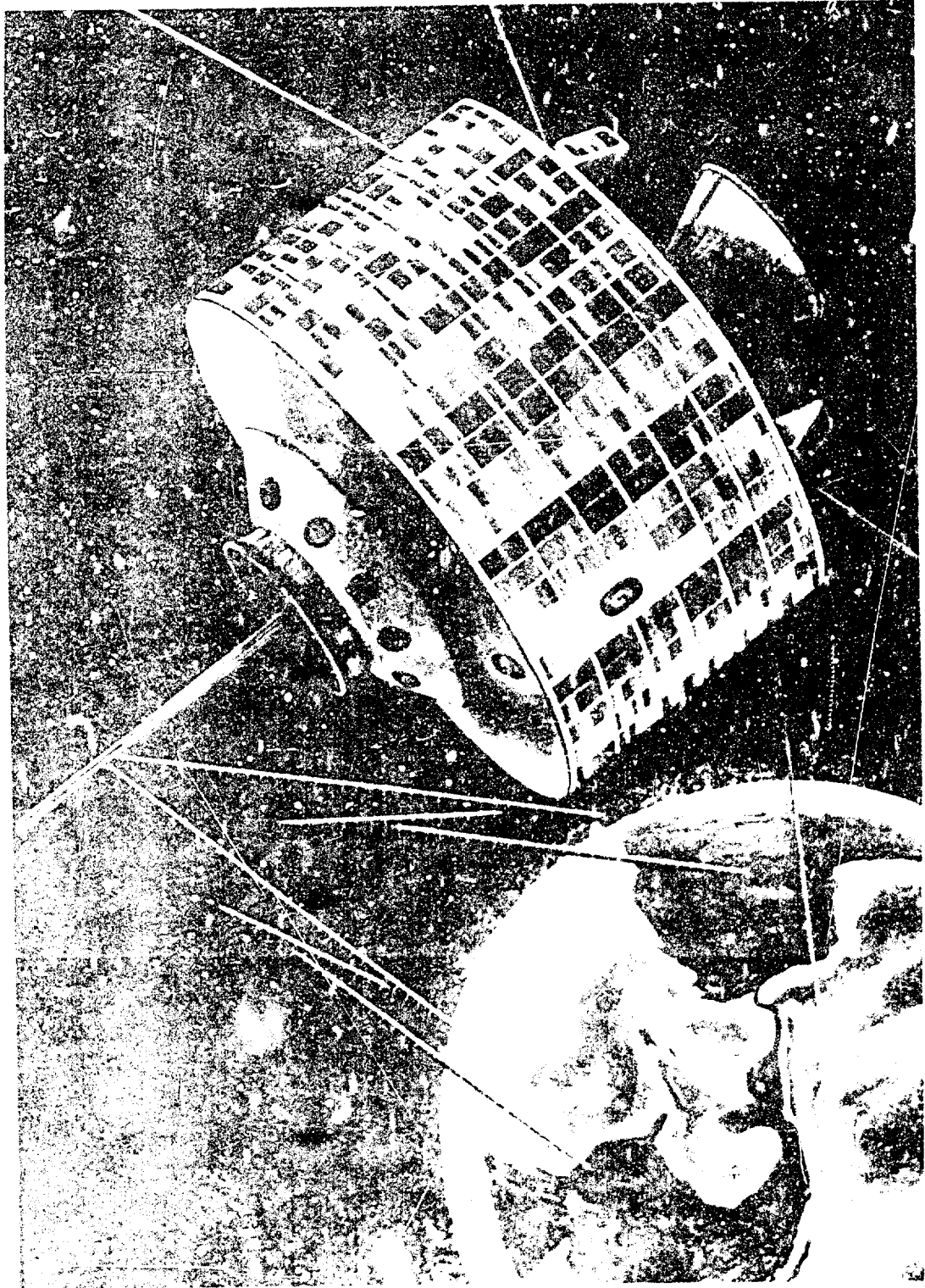
SATELLITE BACKGROUND



U.S. AIR FORCE PHOTOGRAPHIC CENTER (AFPC) - 84-100-04

SLIDE 4

PROJECT SYNCOM



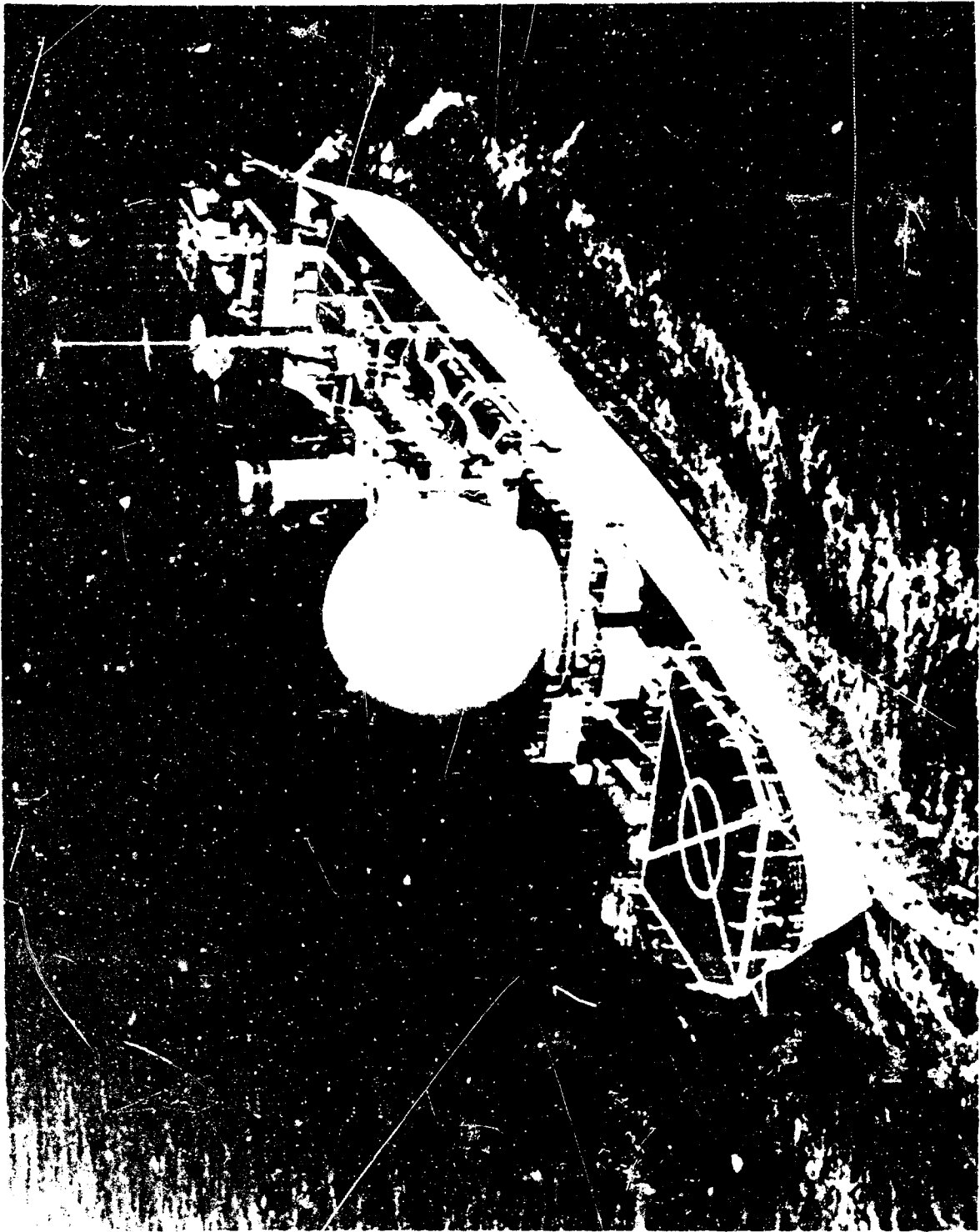
SLIDE 5



SLIDE 6

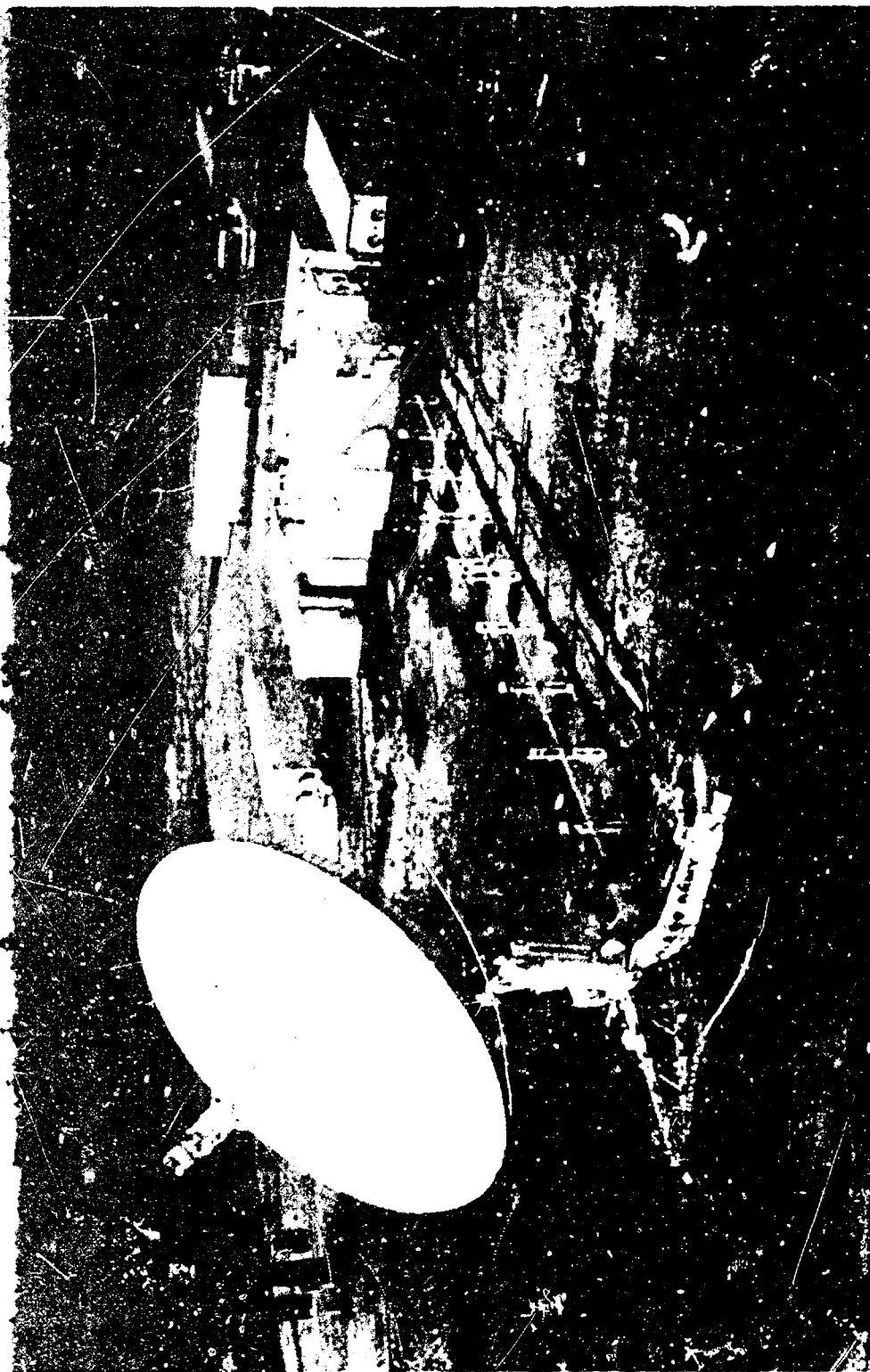


SLIDE 7



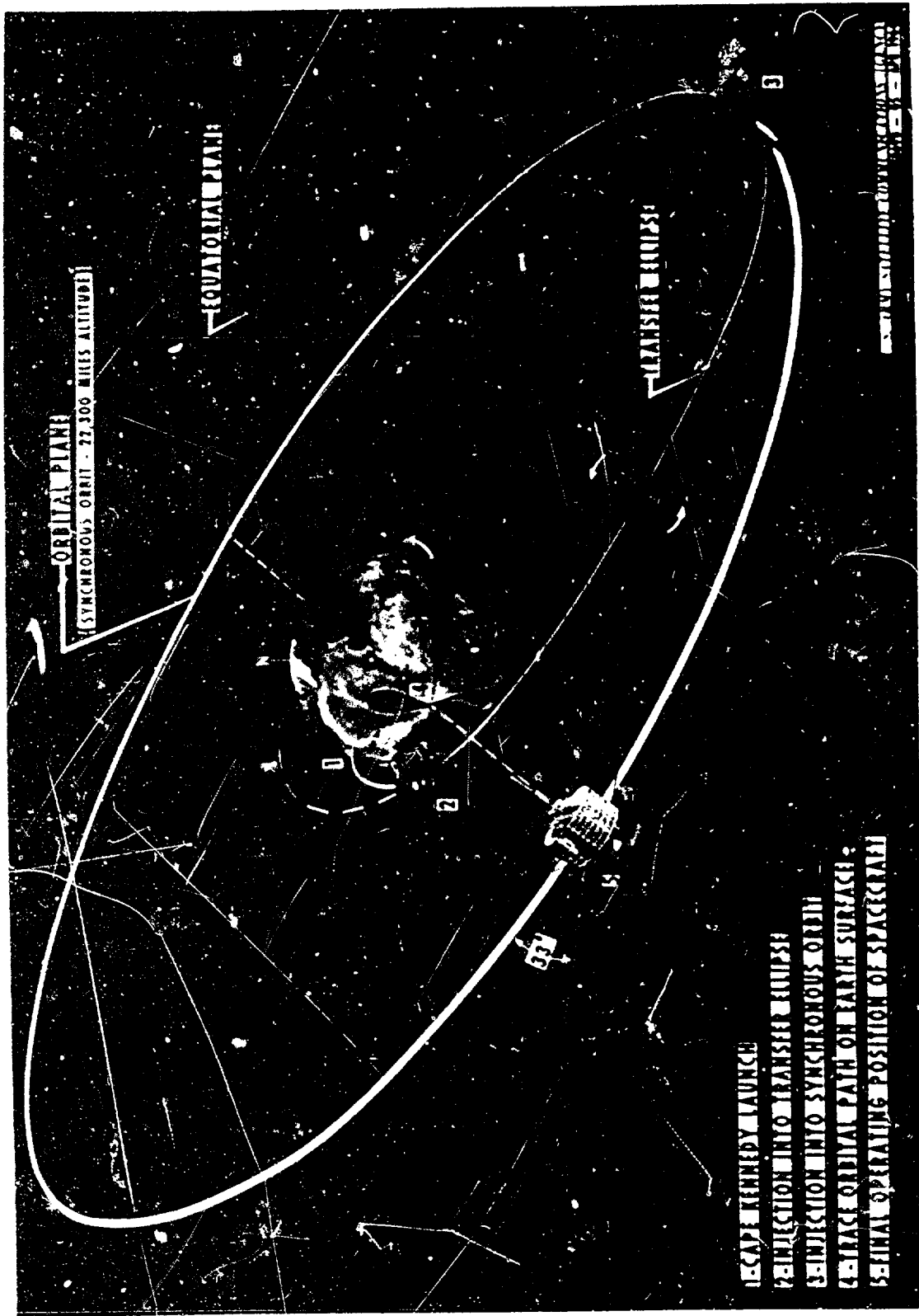
SLIDE 8

JOHNSTON

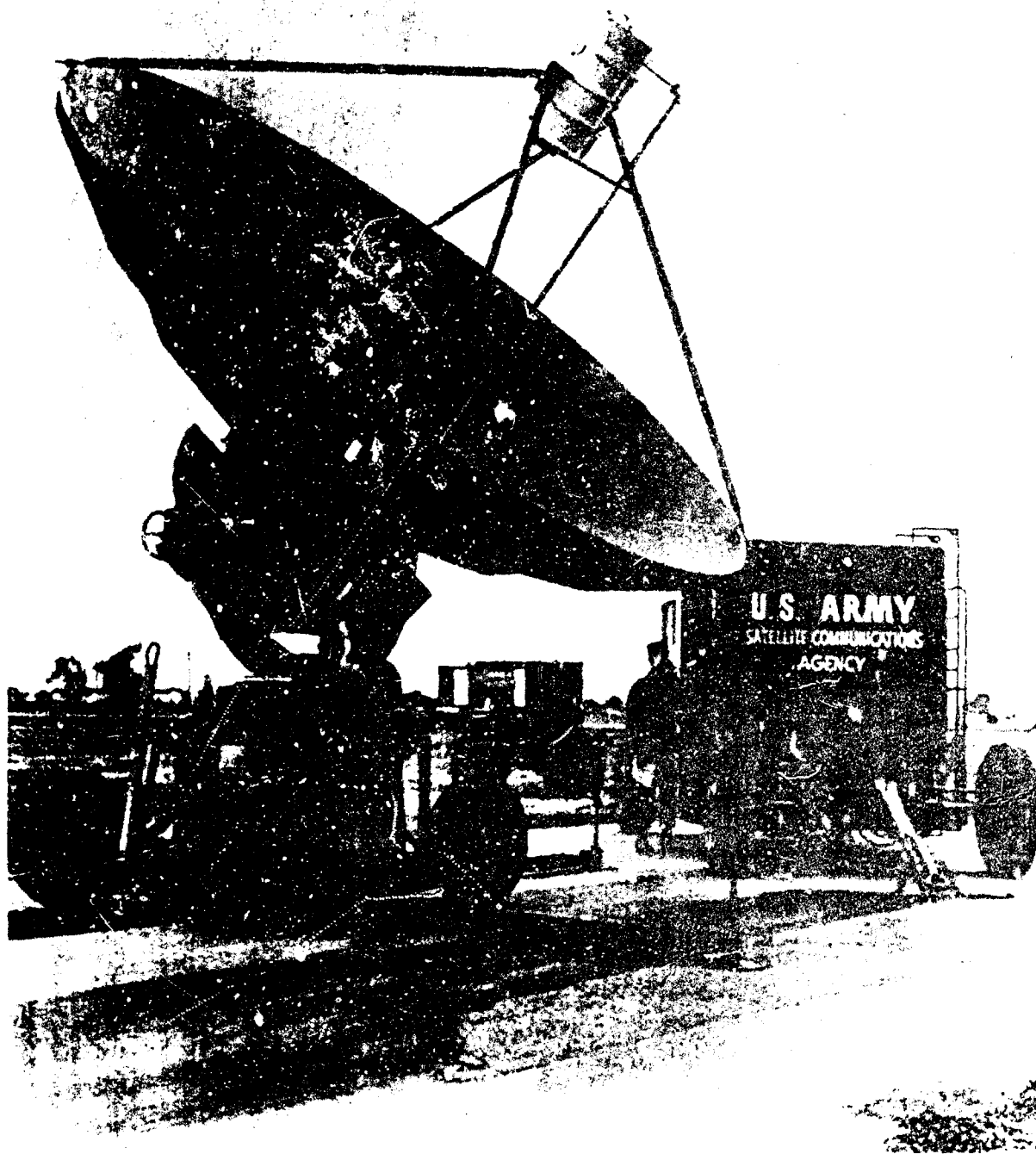


SLIDE 9

SYNCOM SPACECRAFT ORBIT CONFIGURATION



SLIDE 10



SLIDE 11

RESULTS OF COMMUNICATION TESTS

1. Communication Carrier Level (24 hr Variation)	Observed Value		Predicted Value	
	Maximum	Minimum	Maximum	Minimum
	-108 dbm	-111.4 dbm	-108.5 dbm	-115 dbm

2. Communication Carrier Levels Fort Dix USMS Kingsport	Maximum Levels		Minimum Levels	
		-108 dbm	-116 dbm	-114 dbm

3. Signal Plus Noise to Noise Ratio	Maximum Value	Mean Value
		40 db

EFFECTS OF ALLOY SEGREGATION ON SPALLING
OF METALLIC ARMOR MATERIALSK. H. ABBOTT
U. S. ARMY MATERIALS RESEARCH AGENCY
WATERTOWN, MASSACHUSETTSINTRODUCTION

Metallic armor alloys display a number of undesirable behavioral characteristics when attacked by armor-defeating ammunition. Among these is the spallation phenomena which is a process by which disks of armor are ejected from the rear surface regardless of whether the armor is otherwise defeated. These disks, or spalls, are of much larger diameter than the hole which would normally have been formed in the armor by the projectile and vary in thickness from about 20% to 50% of the armor thickness. They may be ejected with a sufficiently high velocity to effect considerable damage to personnel and equipment behind the armor. Obviously, from an armor-protection viewpoint, spallation is undesirable.

The spallation phenomenon in metallic armor materials is well recognized and is utilized as a mechanism for the defeat of armor by HEP shell. With HEP shell, spallation is induced by tensile stresses resulting from the rear surface reflection of a compressive shock wave caused by the HEP shell deformation. Spallation also occurs in armor materials when impacted by kinetic energy armor-piercing projectiles or fragments from high explosive shell. A typical example of spalling in titanium alloy armor under attack by kinetic energy projectiles is contained in Figure 1. In these cases, shock waves are usually not involved but elastic wave propagation in the armor usually provides the stress field in which spallation develops.

Regardless of the terminal ballistic situation in which it occurs, spallation is a fracture process. All fractures are characterized by a fracture origin and some particular fracture path which is dependent on the stress field and inhomogeneities which usually exist in metals. The fracture origins are located at discontinuities in the metal which cause stress intensification. In alloys, these discontinuities may exist in the form of inclusions, small voids,

grain boundary defects, or other metallurgical discontinuities where high residual stress gradients exist. A typical tensile fracture origin located at a nonmetallic inclusion is contained in Figure 2.(1) Once a fracture originates it usually propagates along a preferred path requiring low energy absorption as long as the direction of propagation is compatible with the stress field which is providing the fracture driving force. Examples of low energy preferred fracture paths are metallurgical grain boundaries, dendrites in castings, and alloy segregation bands in rolled plate such as those illustrated in Figure 3. Since spallation begins under tensile stress perpendicular to the armor plate surface, fractures propagate parallel to the armor plate surface. As the crack propagates, however, bending moments and shear stresses develop which turn the fracture direction normal to the plate surface to complete the spallation process.

One way to minimize spalling in armor alloys is to make them softer. This permits energy absorption by plastic deformation of the metal even if fractures develop. However, reducing the strength of armor metal also reduces its resistance to penetration by armor-piercing projectiles (2) and is undesirable. Another way to minimize spalling is to reduce the number of fracture origin sites and low energy fracture paths which exist in the metal. This paper is concerned with the second of these two alternatives and is restricted to spallation under kinetic energy projectile attack.

SOME ALLOY SOLIDIFICATION PHENOMENA

When steel or titanium alloy ingots of non-eutectic composition solidify, dendrites are formed during the freezing process. These dendrites have been characterized as "pine-tree like" in shape as is evident from the photograph of a steel armor casting dendritic structure in Figure 4. For information on the morphology of dendrites, the reader is referred to another paper given at this conference by Ahearn and Quigley. (3) A few general comments about dendritic structures will be made, however, because of their relevancy to fracture in general and the spallation phenomenon in particular.

During the dendritic freezing process, the dendrites freeze prior to the interdendritic volumes and with a lower concentration of solute atoms. This concentration difference can be as great as a factor of 20 or more, (4) and it is not at all uncommon in steel for the concentration difference to be as great as a factor of 4. (5) Hence, the dendrites are of significantly leaner alloy content than the interdendritic volumes. When dendritic growth is essentially complete, the remaining metal in the interdendritic volumes solidifies. When this occurs, liquid metal feeding to these volumes is inhibited by the dendrite arms and microporosity forms. Examples of these small voids which are potential fracture origin sites are contained in Figure 5. Also evident in the figure are dendrites (light areas) and interdendritic volumes (dark structure). Another type of potential fracture

origin site, inclusions, generally float in the liquid metal until they are frozen in place and, as a result, tend to occur predominantly in the interdendritic volumes. Figure 6 displays several small inclusions (white) in an interdendritic volume (black area). Microporosity and inclusions which become distributed in an ingot during solidification always remains in spite of forging, rolling, or other deformation processes. (4) Inclusions tend to become elongated during working of an ingot into plate form, and microporosity becomes very difficult to locate as a result of the working deformation. Microporosity and inclusions then can be considered as incipient fracture origin sites. Alloy segregation, on the other hand, theoretically can be removed from an ingot through diffusion.

In steel castings and ingots, homogenization treatments have been employed historically to permit alloy diffusion into the dendrite. Such treatments have been singularly unsuccessful. (6) In large steel castings and ingots, the dendrite arm spacing (the distance through which diffusion must take place) is usually between 300 and 1000 microns, except for surfaces which have smaller dendrite spacings because of solidification under high thermal gradients. The coefficients for diffusion of nickel, manganese, and chromium in iron are such that for complete diffusion to occur over these distances, several hundred to several thousand hours at temperatures of 2100 F to 2300 F would be required. (4) Such a time-temperature situation is economically impractical and, as a result, steel and titanium armor ingots contain alloy segregation as well as microporosity and inclusions as a result of ingot solidification.

SPALLATION IN CAST STEEL ARMOR

Large steel armor castings such as M-60 tank turret or hull castings receive no mechanical working but are given a liquid quench and temper heat treatment. Lean alloy steels are employed for cast steel armor as a result of the World War II alloy conservation program. These castings have a very coarse dendritic structure with the spacing between dendrite arms of about 500 to 800 microns with a large amount of microporosity present. (Micropores are smaller than the resolution limit of X rays used to inspect the castings and are not detected by this technique.) Because of the coarse dendritic structure, alloy segregation associated with dendritic growth is very evident and is not removed by commercial homogenization treatments. The alloy content difference between dendrites and interdendritic volumes results in a hardenability difference during the quench and temper heat treatment. Evidence of this hardenability difference is contained in Figure 7, where the higher alloy interdendritic area (dark) has a much higher hardness, because of its higher alloy content, than the surrounding dendrites. The casting area contained in Figure 7 is about 2 inches below the surface of an M-60 turret casting where the section thickness is about 5 inches. In Figure 8 are views of a dendrite in the same casting location and the microstructures of both the dendrite and adjacent interdendritic volume. Notice that the

ABBOTT

interdendritic volumes with their high alloy content are quenched to a martensitic microstructure which is desired in steel armor. However, the dendrite with its lower alloy content contains significant amounts of bainitic microstructure.

A martensitic microstructure is desired in steel armor since, at low ambient temperatures, martensite provides much more resistance to crack initiation and propagation than does bainite.⁽⁷⁾ In heavy steel armor castings, however, at depths greater than about 1-1/2 inch below any surface, martensite is present in interdendritic volumes with the dendrites containing bainite. Hence, when ballistic impacts are made on these castings at sub-zero temperatures, the dendrites serve as low energy absorption fracture paths. Since, in these casting areas, the dendrites are oriented heterogeneously, cracks can propagate in almost any direction. That they can propagate in a favorable direction for spallation is evident from the many spalls which have been observed. These crack surfaces have a very rough texture caused by the crack propagating along one dendrite arm and then jumping to an adjacent dendrite. When viewed at higher magnification (Figure 9), the rough texture is evident, as is the dendritic nature of the spallation crack process.

Spall crack initiation sites in cast armor are extremely difficult to detect, but these sites in general are discontinuities which are responsible for stress intensification. Microporosity and inclusions are present in these castings and are most probably the spall crack origins. In any event, solidification phenomena such as alloy segregation, microporosity, and inclusions provide both the fracture origin sites and low energy fracture paths which result in spallation of the castings under ballistic test by kinetic energy projectiles at sub-zero temperatures.

WROUGHT STEEL ARMOR

When a steel ingot is processed to billets and finally to plate form (1/2 to 1 inch thick), it undergoes extensive deformation at moderately elevated temperatures. Such severe working breaks up the dendritic structure and rearranges the alloy segregation such that it becomes manifest as a banded structure similar to that contained in Figure 3. One might assume that such working would effectively reduce the distance through which segregated alloy elements would diffuse and, hence, homogenization treatments could eliminate the alloy segregation. However, recent work⁽⁵⁾ has shown that even in steels reduced by working by a ratio of about 3000 to 1 the severity of alloy segregation is not reduced significantly from that of the original cast structure. Wrought steel armor plates, then, contain alloy segregation bands. Although they are not easily visible, the casting micropores are also probably present as very thin crack-like defects. With deformed inclusions and micropores present in the wrought plates and banded structures containing alloy segregation, both potential spall crack origins and low energy absorption planes parallel to the plate surface are present.

An example of spallation associated with banding in air-melted hot-rolled steel armor is contained in Figure 10. Since this alloy is strengthened by solid solution hardening, microhardness variations from one band to the next are very evident. Notice that the hardness difference between light and dark bands is about a factor of 2. It is also interesting to note that the spacing between adjacent band centers is about 600 to 800 microns which is equivalent to the dendrite arm spacing usually found in large low alloy steel ingots.

New techniques for obtaining and controlling higher purity steel melts, such as vacuum melting, have little effect on alloy segregation since the segregation arises during freezing. An example of spallation in vacuum-melted steel armor is contained in Figure 11. This maraging steel was double consumable electrode vacuum melted, yet both spallation and alloy segregation banding are evident in the figure. Since this steel is strengthened by precipitation hardening rather than solid solution hardening, no hardness variation between adjacent bands is evident. There is probably a difference in precipitate concentration between adjacent bands due to alloy segregation, but the precipitates are so small that they are difficult to detect even with electron microscopy.

Both examples (Figures 10 and 11) of spallation in rolled steel armor occurred during ballistic tests at room temperature and would probably be much more severe at sub-zero temperatures due to the ductile-brittle fracture transition in martensite.

WROUGHT TITANIUM ARMOR

That spallation is also a problem in wrought titanium armor is evident from Figure 1.⁽⁸⁾ All titanium alloys are double consumable electrode vacuum melted. In addition, since the interstitial elements, C, H, O, N, cause severe embrittlement in titanium alloys, they are held to very low concentration. With low interstitial contents, inclusions are also minimized. Alloy segregation, however, is still evident through dendritic growth and exists as a banded structure after hot working to plate form.

A typical banded structure in titanium alloy armor is displayed in Figure 12.⁽⁸⁾ The 6Al-4V alloy is strengthened by solid solution hardening; and, as a result, rather large hardness differences exist between bands. The light bands in the figure are harder than the dark bands with the hardness difference between adjacent bands as large as 13 points Rockwell C. Generally, the harder bands have less resistance to crack propagation than the soft bands, and spall cracks would be expected to propagate in the white bands. Evidence of such crack propagation is contained in Figure 13⁽⁸⁾ where a spall crack is shown propagating in a white band.

Although this paper is concerned primarily with spallation and its relation to alloy segregation, inclusions are significant

ABBOTT

contributors to the spallation phenomena in titanium alloy armor. Even with the carbon content held to a maximum of 0.02 weight percent, titanium carbide inclusions form and appear to concentrate in the white banded layers. In Figure 14 (8) is a photograph of a titanium carbide inclusion aligned with a spall crack in 6Al-4V titanium alloy armor. The presence of elongated hard particles such as these contribute to the propagation of cracks during the spallation phenomena.

SUMMARY

Several illustrations of the spall crack propagation in cast and wrought metallic armor alloys have been presented. Dendritic or banded structures resulting from alloy segregation during solidification of the metals have been shown to be related to spall crack propagation. Neither working nor short-time homogenization treatments significantly effect the alloy segregation. Hence, minimization of the banding (alloy segregation) appears to be dependent on a better understanding of, as well as control of, the solidification process. The Army Steel Castings Research Program has been initiated by AMRA to study solidification phenomenon in steel castings. Results from that program will be of importance to both cast and wrought armor alloys where minimization of alloy segregation will reduce spallation and lead to the development of alloys with increased resistance to penetration.

REFERENCES

1. LARSON, F.R., and CARR, F.L., "Tensile Fracture Surface configurations of a Heat-Treated Steel as Affected by Temperature." Transactions Quarterly, ASM, v. 55, no. 3 (September 1962) pp. 599-612
2. AHEARN, P.J., and QUIGLEY, F.C. "Morphology of Dendritic Structures in High Strength Steels." Army Science Conference, June 1964.
3. MASSACHUSETTS INSTITUTE OF TECHNOLOGY. "Investigation of Solidification of High Strength Steel Castings." Contract DA-19-020-ORD-5443, Watertown Arsenal Laboratories WAL TR 311/49, February 1962.
4. FLEMINGS, M.C. "Potential Improvements in Armor Alloys by Controlling Solidification Variables," Meeting of Metallurgical Advisory Committee on Rolled Armor, U.S. Army Materials Research Agency, Watertown, Mass. 22 January 1964.
5. BRIGGS, C.W., "The Effect of Heat Treating Variables on Toughness of Cast Steels and Cast Armor." Steel Founders' Society of America, Contract DA-33-019-ORD-1510, Watertown Arsenal Laboratories WAL TR 716.3/1, June 1958.
6. RIFFIN, P.V., and HURLICH, A. "Comparative Evaluation of the Toughness of 1 "Rolled Homogeneous Armor by Ballistic and Non-Ballistic Tests." Watertown Arsenal Laboratories WAL TR 710/904, August 1949.
7. SLINEY, J.L., JR. "Status and Potential of Titanium Alloy Armor." Meeting of Metallurgical Advisory Committee on Rolled Armor, U.S. Army Materials Research Agency, Watertown, Mass. 22 January 1964.



Figure 1. TITANIUM ALLOY (6Al-6V-2.5Sn ANNEALED) AFTER BALLISTIC ATTACK WITH CAL. .30 AP M2 PROJECTILE AT 0° OBLIQUITY

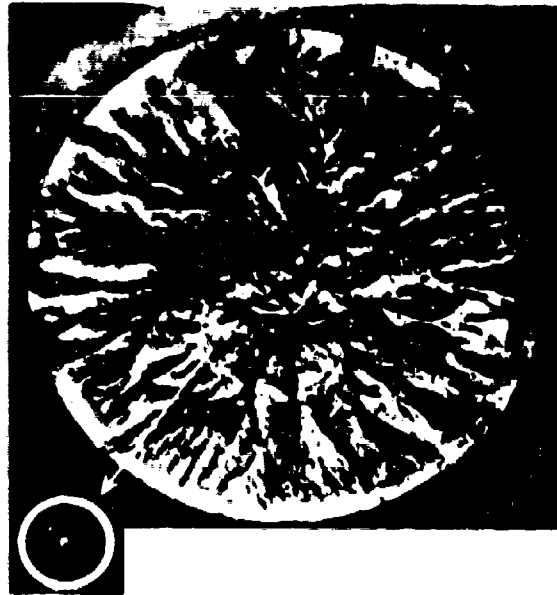


Figure 2. INCLUSION AND CONE ON FRACTURE SURFACE OF TENSILE SPECIMEN

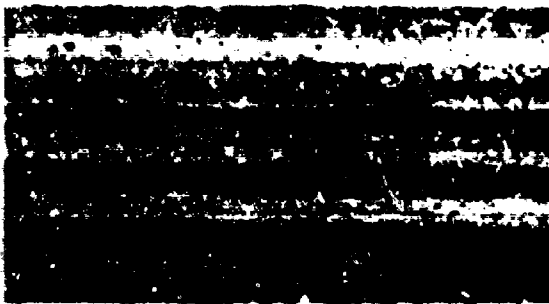


Figure 3. TYPICAL BANDED STRUCTURE IN ROLLED STEEL ARMOR PLATE

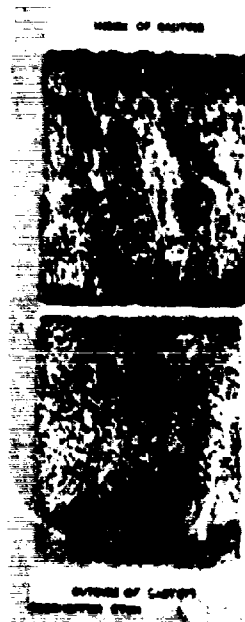


Figure 4. TYPICAL SOLIDIFICATION PATTERN IN FIVE-INCH-THICK CAST ARMOR

U. S. ARMY MATERIALS RESEARCH AGENCY

19-066-274/AMC-64



Figure 5. MICROPOROSITY IN CAST STEEL ARMOR

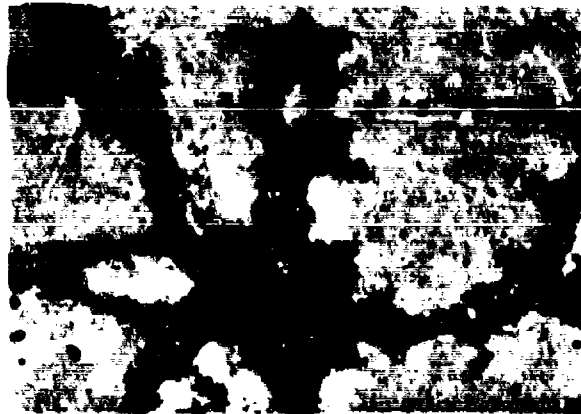


Figure 6. INCLUSIONS AND ALLOY SEGREGATION IN CAST TURRET TEST BLOCK

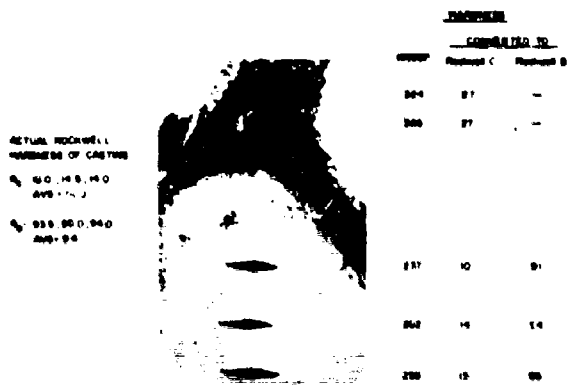


Figure 7. HARDNESS VARIATION IN CAST TURRET TEST BLOCK

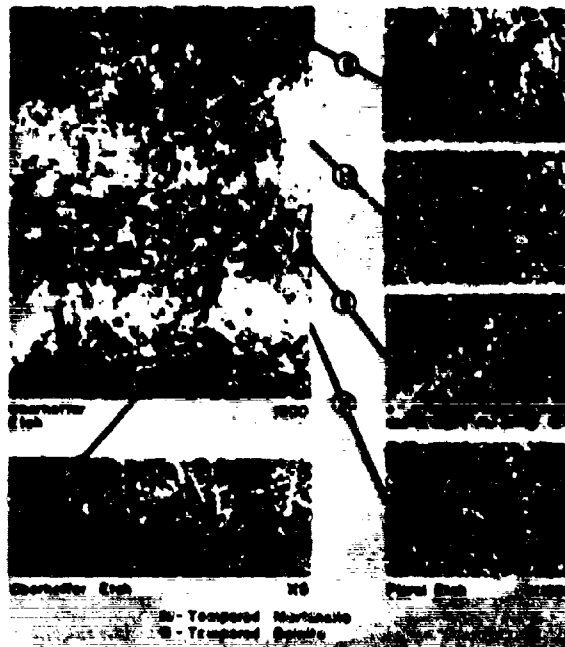
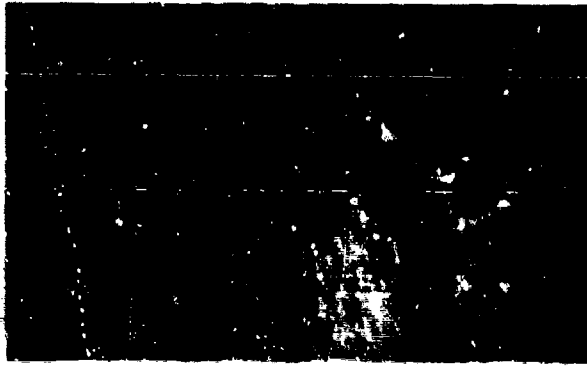


Figure 8. MICROSTRUCTURES OF DENDRITIC AND INTERDENDRITIC VOLUMES OF HEAT-TREATED CAST STEEL ARMOR

U. S. ARMY MATERIALS RESEARCH AGENCY



OBERMOPFER ETCH

X 80

Figure 9. FRACTURE THRU DENDRITE NEAR CENTER OF 6-INCH SECTION



Figure 10. SPALL FRACTURES AND BANDED STRUCTURE IN AIR-MELTED ROLLED STEEL ARMOR

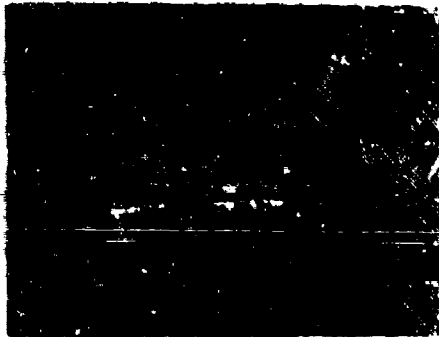


Figure 11. SPALL FRACTURES AND BANDED STRUCTURE IN VACUUM-MELTED STEEL ARMOR

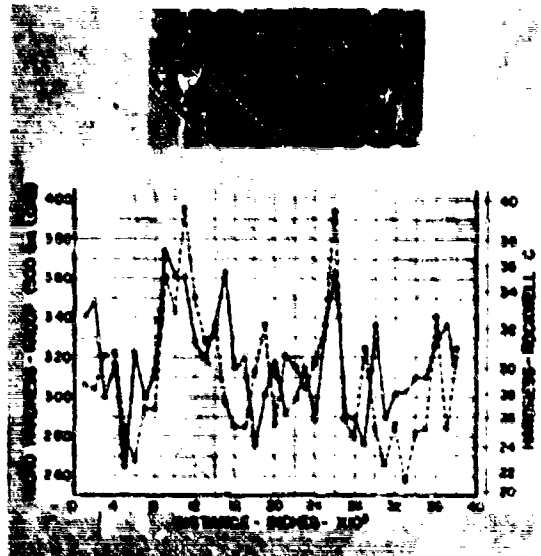


Figure 12. MICRONARDNESS SURVEY ACROSS HEAVILY BANDED 2-INCH Ti-6Al-4V ARMOR PLATE

U. S. ARMY MATERIALS RESEARCH AGENCY

19-066-271/AMC-64



Figure 13. BALLISTIC TEST PLATE
TJ-6A1-4V ALLOY (R-ETCH)

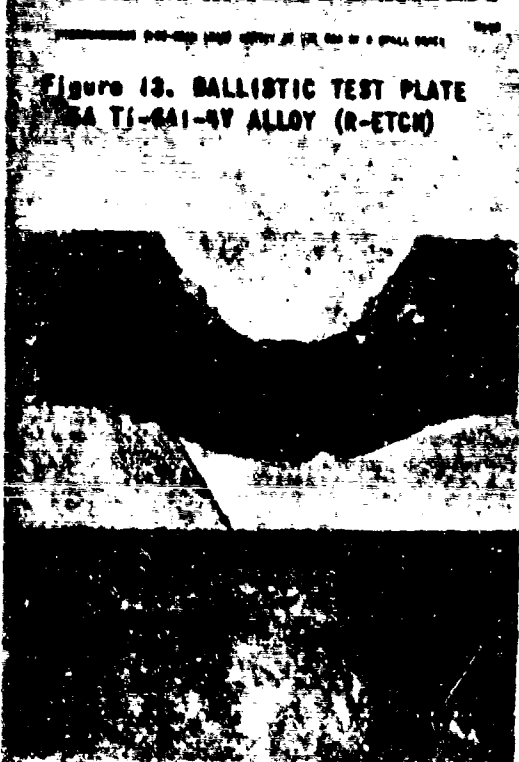


Figure 14. CARBIDES LINED UP
WITH DELAMINATION IN
TJ-6A1-4V ARMOR

U. S. ARMY MATERIALS RESEARCH AGENCY

BLANK PAGE

TE₀₁⁰ CROSS-GUIDE TRANSDUCER AND ISOLATOR

JOHN P. AGRIOS, C. D. NEUDORFER, R. A. STERN
 U. S. ARMY ELECTRONICS RESEARCH & DEVELOPMENT LABORATORIES
 FORT MONMOUTH, NEW JERSEY

In the comparatively short time between World War II and the present, the requirements of military communication and surveillance systems have not only become more complex but also have been increasing at a fantastic rate. As a result, we require our transmitting systems to deliver larger amounts of power and our receiving systems to be extremely sensitive. To date, rectangular waveguide has been utilized primarily from 1000 mc and up, in view of its low-loss characteristics and high power capabilities. However, when the transmission line distance over which microwave energy must be transmitted from one point to another becomes appreciable, rectangular waveguide ceases to be efficient. For example, if one used rectangular waveguide to transmit a signal over a distance of 100 feet at 9000 mc, half of the power would be lost in the waveguide, which is very costly when one is considering high power. The converse is also true. When a low level signal has to be transmitted from the antenna to the receiver complex, minimum attenuation is not only required from the loss point of view but also from a consideration of the overall noise figure of the receiving system. Hence, another type of transmission line is required to satisfy the requirements of these systems.

Investigators have been seeking other efficient types of transmission lines for years. Among the many varieties of such lines, the circular waveguide offers some exceptionally attractive features. The TE₀₁ mode of the circular waveguide not only offers the minimum loss characteristics of the majority of transmission lines, but its high power capabilities are superior to others. For example, at 9000 mc this waveguide offers a loss of .45 db per 100 feet or 15% of the loss of rectangular waveguide. Its high power capability is approximately seven times that of rectangular waveguide. Unfortunately, this mode is not the dominant mode and other modes can propagate in the line; such mode conversion will be reflected as an additional loss. However, if extreme care is taken in the design and fabrication of circular waveguide components, these modes will

not be excited and the advantage of minimum loss will be maintained.

We at USAELRRL have recognized the advantages of circular waveguide components for quite some time. We have conducted several investigations, externally and internally, to design appropriate components. This paper discusses two particular circular waveguide components, which we at USAELRRL have developed internally. The TE_{10} mode rectangular waveguide to TE_{01} circular waveguide transducer, and the TE_{01} ferrite isolator.

THE TRANSDUCER

The great majority of present day microwave power sources are only available in rectangular waveguide outputs, operating in the dominant TE_{10} mode (TE_{10}). Hence, transducers are required to transform the microwave energy from this mode in rectangular waveguide to the TE_{01} mode in circular waveguide (TE_{01}). Furthermore, since linearly polarized waves are required at the antenna for radiation of the desired signal, transducers are also required at the antenna to convert the TE_{01} mode back to the TE_{10} mode. These transducers must be not only compact, but efficient. The presence of other modes must be at a minimum, or the transducer must have a high mode purity. We at USAELRRL require that all circular waveguide components exhibit a mode purity of at least 20 db. That is, the power level of all the spurious modes found in the system must be at least 20 db below the level of the TE_{01} mode. Our particular requirement was such that we found it necessary to develop a broadband transducer which was compact, operated over a 12% bandwidth centered at 7.5 mc with a VSWR of 1.2:1 maximum and with a mode purity of 20 db minimum.

Various approaches were studied and considered to determine the best approach in the development of the required TE_{10} to TE_{01} transducer. The design which employs a gradual transition from rectangular to circular waveguide through sector sections was reviewed (6) This design offered an extremely low VSWR over a wide band of frequencies; however, the transducer not only was long but its modal content was such that the mode purity of the device was in the order of 15 db. The TE_{11} mode prevailed in this design. A mode filter could have been inserted in the line to improve the mode purity; however, this would not only represent a loss but resonances may also occur in the line which are not desirable. Another design which was studied in detail was the technique of launching the desired TE_{01} mode by means of four rectangular waveguides ending in the circular waveguide. Several models were designed and fabricated; however, even though the mode purity results looked encouraging the VSWR of the unit was prohibitive. After several studies we decided to broadband a transducer developed several years ago for USAELRRL. (1) This unit had a minimum mode purity of 20 db; however, its bandwidth was only 6% based on a VSWR of 1.2:1 maximum. The design consisted of a

section of circular waveguide with its longitudinal axis perpendicular to the "b" dimension of two parallel rectangular waveguides as shown in Figure 1a. The two waveguides are coupled together by an array of four resonant slots as shown in Figure 1b.

The orientation of the irises of the above described transducer allows two polarizations of magnetic field to couple to the circular waveguide. One polarization is a desired magnetic field for TE_{01} propagation, while the other sets up a mode in the circular waveguide other than the TE_{01} . Therefore, the coupling of this unwanted magnetic field into the circular waveguide has to be inhibited to achieve a pure TE_{01} mode. In the above mentioned model, this was done by using a thick iris plate; however, this decreased the magnetic coupling which resulted in an increase in insertion loss.

In order to eliminate the unwanted magnetic field, we felt that a new slot orientation should be employed. Hence the geometry of the transducer was changed to that of a "cross-guide" to yield only one axis of polarization, Figure 2. The top of the cross-guide block was designed to accept an iris plate and the output circular waveguide whose recommended diameter for this frequency range is 2.625 inches in diameter. However, this size of waveguide is able to propagate the TE_{11} mode and the slot configuration of the irises is such that the TE_{11} may be developed. Hence, a smaller size of circular waveguide had to be used at the output so that the TE_{11} mode will be below cut-off. A short taper was then utilized to transform the diameter of the circular waveguide to the desired 2.625". Since the unit was to operate from 7 to 8 kmc, the cross-guide block was fabricated using the basic RG-51/U waveguide dimensions. It may be observed from Figure 2, that the unit is fed by two rectangular waveguides with the opposite ends being terminated in short circuits. A septum is so arranged as to maintain the identity of the two transmission lines, each one having two irises spaced half-a-guide-wavelength apart from the center of each iris. Energy is transferred from the rectangular waveguides through these irises to the circular waveguide by magnetic coupling. An electric field vector appears across each iris such that the TE_{01} mode is developed.

If the two plane waves which emerge into the cross-guide block are in exact phase and magnitude, and no discontinuities or obstructions occur from the block to the circular waveguide, the electric field vectors across the irises will be equal in magnitude and hence a pure TE_{01} mode will be excited. If on the other hand all of the above conditions are not met, the electric field vectors will not be equal in magnitude and other modes will be excited, depending upon the electric field vector distribution. The theoretical analysis for the prediction of the exact modal content in this transducer is not only very complex but not available. In view of this, we felt the only way to develop the required transducer was by empirical means. Since both transducers, the cross-guide and that of Figure 1, utilize irises in the same manner to excite the TE_{01} mode,

it was felt that the iris plate should not be changed. Hence, the dimensions of the irises were scaled from those used in the transducer of Figure 1.

To determine the effect of the bend section within the cross-guide block, VSWR tests were conducted on each side. The resultant VSWR's were not acceptable. It was apparent that the corners within the block were seriously contributing to the VSWR. Hence, we proceeded to miter these corners in discrete steps. Optimum results were obtained when the "b" dimension of the rectangular waveguide was maintained through the bend to the short circuit. The results of the tests are given in Table I.

TABLE I
RESULTS OF MITERING CORNERS OF CROSS-GUIDE BLOCK

<u>Frequency</u>	<u>VSWR Side 1</u>	<u>VSWR Side 2</u>
7.0	1.06	1.06
7.25	1.07	1.08
7.5	1.08	1.15
7.75	1.04	1.12
8.0	1.04	1.05
8.25	1.07	1.04
8.5	1.11	1.09

Originally, the required power division to properly drive the cross-guide block was achieved through the use of a Magic-T with waveguide elbows symmetrically arranged for proper physical orientation. The VSWR results obtained with this arrangement were quite unacceptable. The accuracy of the power split and the equality of the phase at the output of each arm were questionable. These factors could cause serious problems with respect to the mode purity of the device. Therefore, a new power divider was designed which consisted of a transformer from single height to double height waveguide and an H-plane bifurcation to achieve the required power split, Figure 3. The transformers were tested back-to-back to obtain their VSWR characteristics. The results are given in Table II.

TABLE II
SINGLE HEIGHT TO DOUBLE HEIGHT TRANSFORMERS BACK TO BACK

<u>Frequency</u>	<u>VSWR</u>	<u>Frequency</u>	<u>VSWR</u>
7.0	1.035	7.7	1.015
7.1	1.04	7.8	1.02
7.2	1.035	7.9	1.035
7.3	1.03	8.0	1.03
7.4	1.035	8.1	1.015
7.5	1.025	8.2	1.025
7.6	1.02	8.3	1.04

The power divider was then tested to determine the quality of the phase and power split. The results are given in Table III.

TABLE III

BG-51 POWER DIVIDER

Freq	Po Arm 1(db)	Po Arm 2(db)	Min Pos on Sl. Arm 1 Shorted	Min Pos on Sl. Arm 2 Shorted	VSWR
7.0	-3.0	-3.0	10.39	10.38	1.06
7.1	-3.1	-3.1	11.80	11.80	1.02
7.2	-3.0	-3.0	10.14	10.15	1.05
7.3	-3.0	-3.0	11.65	11.68	1.06
7.4	-3.2	-3.15	10.16	10.19	1.12
7.5	-3.0	-3.1	10.46	10.45	1.07
7.6	-3.0	-3.1	12.50	12.50	1.04
7.7	-3.1	-3.1	10.90	10.91	1.01
7.8	-3.0	-3.0	9.26	9.25	1.03
7.9	-3.0	-3.0	10.33	10.33	1.03
8.0	-3.1	-3.1	11.37	11.38	1.06

The power split was found to be acceptable and superior to that obtained with the previous power divider. The phases of the output arms were also satisfactory. This was determined by observing the minimum position of the standing wave on a slotted line while alternately shorting arms 1 and 2 as indicated in Table III. The complete transducer was then assembled and tested. The VSWR of the unit was 3.0:1. Impedance calculations were made on the rectangular waveguide of the cross-guide block and the circular waveguide in an attempt to determine the source of the mismatch. The calculations showed that the impedance of the rectangular waveguide was approximately one-tenth that of the circular waveguide. The "b" dimension of the rectangular waveguide was increased in steps, to 0.685". The latter dimension yielded a VSWR of 2.4. Step transformers were designed to match the .685" dimension to the .497" dimension of the power splitter Figure 4. In order to match the impedance of the two lines, a cruciform section was designed to reduce the impedance of the circular waveguide at the irises. By trial and error, several cruciform designs were fabricated and tested. The dimensions of the optimized cruciform section which yielded the desired results were 0.530" in the "b" dimension, 2.328" in the "a" dimension with an overall length of 8.675". Figure 5 illustrates the complete cruciform section. The mode purity and VSWR characteristics of this transducer are given in Figure 6. The results indicate that the mode purity of the unit is 25 db minimum across the band with a maximum VSWR of 1.19.

The mode purity technique utilized during this period for developing the transducer was of the rotating probe type. The accuracy of this technique however was found to be questionable since it does not take into consideration spatial beating. In order to overcome this shortcoming, a new test procedure was used. (7) This approach

utilizes a variable short technique in which, while the frequency is held constant, the short position is varied and the resulting VSWR's recorded versus short position. This method eliminates the problem of spatial beating and yields accurate results with regard to mode purity. Using this technique mode purity measurements were again obtained and the results are given in Figure 7. Even though these mode purity results are not as high as Figure 6, they met our original goals. Figure 8 is a photograph of the final model. The insertion loss of the device measured from 7 to 8 kmc did not exceed 0.25 db.

Since our intent was mainly to develop a transducer with specific requirements, our empirical development ceased as soon as these objectives were met. Hence, it is not the intent of this paper to infer that the optimum in transducer design has been achieved. For instance, the dimensions of the irises, the height of the rectangular portion of the block ("b" dimension) and the cruciform have not been optimized with respect to mode purity and VSWR. With regard to the physical dimensions of the transducer, it is felt that although the present model is relatively small, much can be done to achieve an even more compact device.

THE ISOLATOR

Among the great variety of microwave ferrite components available to date, the circular waveguide TE_{01} mode ferrite devices are virtually non-existent. Since microwave and millimeter systems generally require ferrite isolators, a program was initiated at URAKIRIL to develop an isolator for use in TE_{01} mode systems.

There are two distinct types of structures suitable for the development of such an isolator; the circular waveguide supporting the TE_{01} mode and the cross-guide or cruciform type. The field configurations in these structures together with the ferrite geometries required, determined the type of design that was used. The magnetic field configuration in circular waveguide is such that a ferrite ring positioned concentrically within the waveguide must be circumferentially magnetized in order to achieve nonreciprocity at resonance. Research efforts in this area have shown that this is a very difficult approach. (2) The cross-guide structure on the other hand, is capable of supporting an electric field distribution in the transverse cross section of the guide as shown in Figure 9. This distribution is analogous to the TE_{20} mode in rectangular waveguide. Furthermore, the field distribution in each of the four channels appears to be similar to that of the TE_{10} mode. Thus, if a ferrite slab is positioned in a region of circular polarization in each channel, nonreciprocity will occur at resonance.

As the operational frequency of the device is increased, the size of the permanent magnetic structures required also increases in direct proportion to the frequency. Consequently, large heavy magnetic structures are required for this device in the millimeter region. Hexagonal ferrite materials however, exhibit such inherent magnetic properties that the requirement for external magnets is almost eliminated. In view of this it was decided to investigate the cross-guide structure, since with these materials available (8) a compact device could be designed.

Since no specific system requirement existed, we set our design goals as follows: Isolation 20 db minimum; VSWR 1.2:1 maximum; mode purity 20 db minimum; insertion loss 2.0 db maximum; and bandwidth approximately 10%, at a frequency of 35 Gc.

Just as in the case of the TE_{10} mode in rectangular waveguide, two regions of circular polarization also exist in the cross-guide. The region of circular polarization used for this device was the one located near the narrow wall of the channel. This location also offers a convenient place for the biasing magnets. These magnets are required to provide a small additional field to neutralize the detrimental demagnetizing fields introduced by the geometry of the ferrites, and also to protect the device against stray external magnetic fields. Since the theoretical knowledge of the propagation characteristics of the cross-guide was limited, the development proceeded on an empirical basis.

The magnitude of the anisotropy field required for resonance at the desired frequency is calculated using Kittel's Equation given below:

$$f_r = \frac{\gamma}{2\pi} \left[\left(H_{an} + H_{ap} + (N_x - N_z) 4\pi M_s \right) \left(H_{an} + H_{ap} + (N_y - N_z) 4\pi M_s \right) \right]^{\frac{1}{2}}$$

- where
- H_{ap} = applied magnetic field
 - f_r = resonant frequency
 - γ = gyromagnetic ratio
 - H_{an} = anisotropy field
 - N_x, N_y, N_z = demagnetizing factors
 - $4\pi M_s$ = saturation magnetization

Using the analogy of the TE_{10}^{\square} mode isolator, the ferrite slabs were mounted on dielectric substrates (Wesgo AL-995) with DUCO cement. Figure 10 illustrates the physical arrangement of the components of the cross-guide isolator. The dimensions of the dielectric substrate and ferrite slab are given in Figure 11. The position and thickness of the dielectric slabs and ferrites were optimized empirically to yield maximum isolation and mode purity. The parameters of the hexagonal ferrite materials used were $H_{an} = 11,000$ oe; $4\pi M_s = 3,500$ oe. and $\gamma/2\pi = 2.68$ mc/oe. The resonant frequency of the ferrite was 35.1 Gc. It should be noted that the location of the substrates with the ferrites in the channel was not found to be

critical; however, overall symmetry of the substrates in the channels proved to be extremely important for maximum mode purity. The demagnetizing field intensity given by $H = H (4\pi M_s)$, where H for this geometry is .03, is approximately 105 oe. Hence, with the use of small biasing magnets of approximately 500 oe., the demagnetizing fields are fully compensated and the device is not sensitive to any stray fields.

To transform the cross-guide into the desired circular waveguide, a transition was used in which the cross-guide is tapered to the circular waveguide.⁽¹⁾ The length of the taper was approximately 2.0". The complete isolator (including transitions) was fabricated in two halves from solid brass stock and then silver plated. Figure 12 is a photograph of the two halves. Figure 13 illustrates the isolation and forward loss characteristics of the device. It should be noted that the device exhibits an isolation of at least 20 db over an 8% bandwidth (33.85 - 36.70 Gc), with the insertion loss having a maximum value of 1.9 db at the upper end of the frequency band. The VSWR of the device would not be measured since a slotted line in the TE_{01} mode was not available.

Since our objective was that the isolator have a mode purity of at least 20 db, a transducer was required that also exhibited a mode purity of at least 20 db. Unfortunately, such transducers were not available at our frequency of interest. Hence, a commercial transducer was used of the type of Figure 1. Based on the tests conducted on the transducer of Figure 2, it was evident that the variable short technique for measuring mode purity was preferable. The mode purity characteristics of the transducer, and the cross-guide without any dielectric substrates and ferrite slabs (empty cruciform) are shown in Figure 14. These data were obtained using the variable short technique. However, this technique requires that the effective waveguide wavelength of each detectable mode present in the line be known very accurately. When the dielectric substrates were mounted in the cross-guide the wavelengths could no longer be determined accurately. Hence, the mode purity measurements of the dielectrically loaded cross-guide (dielectrically loaded cruciform) and the isolator, were conducted using the rotating probe technique. The results are given in Figure 14.

Even though the results obtained with the rotating probe technique may not be considered as accurate as the results using the variable short technique, they are reliable provided the necessary steps are taken to account for spatial beating. Based on the mode purity results obtained during the development of the transducer, the results indicated that the rotating probe technique yielded results which were slightly optimistic. However, when different lengths of line were inserted to determine the effect of spatial beating, the minimum mode purity obtained with the rotating probe technique approaches that obtained with the variable short technique. Hence, the results given in Figure 14 are considered reliable when we say that the final unit exhibited a mode purity of at least 15 db.

In examining the results of Figure 14, it is observed that at some frequencies the device under test yielded a mode purity which is higher than that of the transducer. It is believed that this is caused due to the interaction of the spurious modes between the transducer and the unit under test. Such interaction is possibly causing cancellation of certain spurious modes which result in an "observed" higher mode purity for the unit under test.

The results obtained during this investigation have been very encouraging. The isolation characteristics of the isolator are highly acceptable. The insertion loss of the device is slightly high; however, further work can be conducted in this area to reduce these losses. The use of the hexagonal ferrite materials results in a unit which is quite compact and does not require large and heavy external magnets. Even though the mode purity results indicate a purity of 15 db minimum it is felt that the isolator does not degrade mode purity much below the mode purity of the transducer. Therefore, the use of a better transducer would be reflected in better mode purity of the isolator. Also, the final model could be improved by better tapers to the cross-guide, and slight changes in the dimensions of the channels.

In conclusion two designs have been developed, one for a TE_{10} mode to TE_{01} mode transducer and the other for a cross-guide isolator for use in a TE_{01} mode system, with results which have not been attained heretofore. The transducer not only represents a device which has immediate application but also provides the research engineer an indispensable tool for the development of TE_{01} components. We are presently attempting a scaled version of this transducer for use at 35 Gc in order to better evaluate the mode purity characteristics of the cross-guide isolator. The design of this isolator with hexagonal ferrite materials offers a method for designing ferrite isolators for use in the TE_{01} mode systems operating in the millimeter region, which are compact and have good electrical characteristics.

Acknowledgment:

The authors wish to express their appreciation to Messrs. N. Lipetz and E. Freibergs for their contribution and assistance in the preparation of this paper. Our appreciation is also extended to Mr. J. LoCicero for his resourcefulness in the fabrication of these devices.

References:

1. D. A. Lanciani, "Basic Design Criteria for Circular Waveguide Components," Final Report, Contract DA36-039 SC-5518, Microwave Associates, Inc.
2. V. E. Dunn, "Ferrite Devices for Receiving Systems," Final Report, Contract DA36-039 SC-88949, Melabs, Inc.

AGRIOS, NEUDORFER, and STERN

3. M. Lipetz and C. Neudorfer, " TE_{01}^0 Cross-Guide Transducer," UBAKLRDL Technical Report No. 2389, October 1963.
4. E. Freibergs and R. Stern, " TE_{01}^0 Cross-Guide Isolator," UBAKLRDL Technical Report No. 2430, February 1964.
5. A. Simmons, "Three Millimeter Waveguide Components," Final Report, Contract DA36-039 80-88779, Technical Research Group.
6. S. Iiguchi, "Mode Conversion in the Excitation of TE_{01} Waves in a TE_{01} Mode Transducer" Proceedings of the Symposium on Millimeter Waves, 1959, Polytechnic Institute of Brooklyn.
7. Y. Klinger "The Measurement of Spurious Modes in Over-Moded Waveguides," The Institute of Electrical Engineers Paper No. 3040E, September 1959.
8. D. R. Taft and G. P. Harrison "Millimeter Wave Components (Hexagonal Ferrite Isolators) Final Report, Contract DA36-039 80-89193, Sperry Microwave Electronics Corporation.

NARROW BAND END FEED TYPE TRANSDUCER

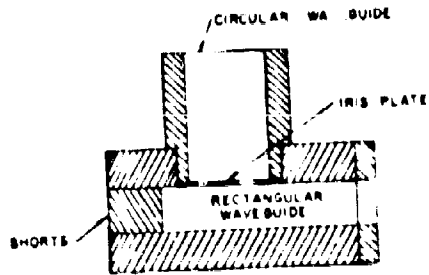


Fig. 1a

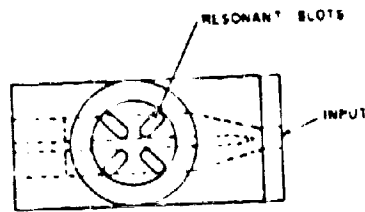
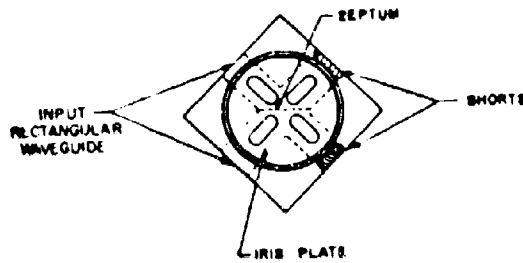
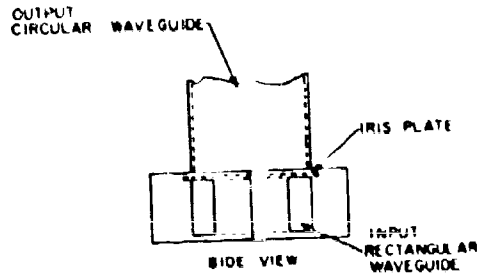


Fig. 1b

CROSS GUIDE BLOCK WITH IRIS PLATE



TOP VIEW

Fig. 2

POWER DIVIDER

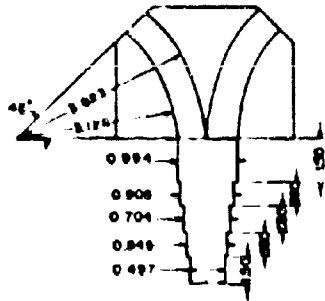


Fig. 3

FINAL MODEL OF CROSS GUIDE BLOCK AND WIRES

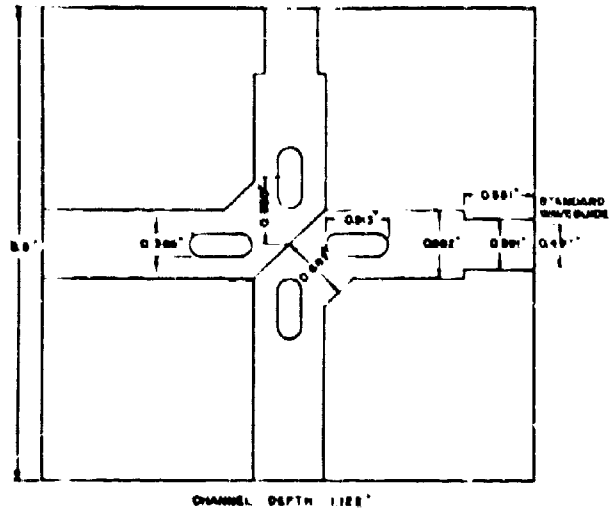


Fig. 4

CROSSIFORM SECTION

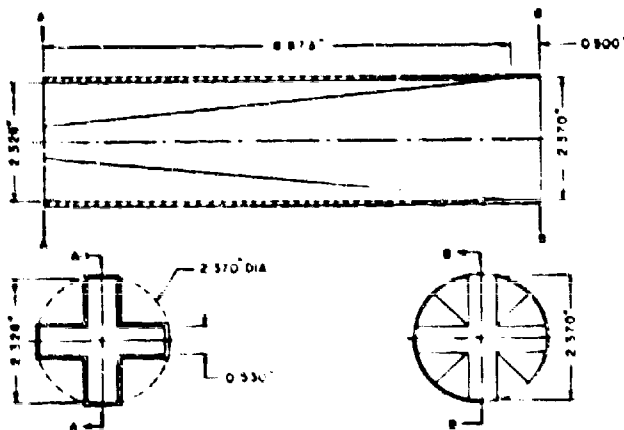


Fig. 5

CROSS GUIDE TRANSDUCER

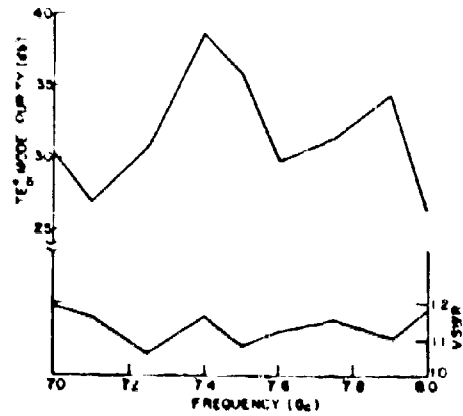


Fig. 6

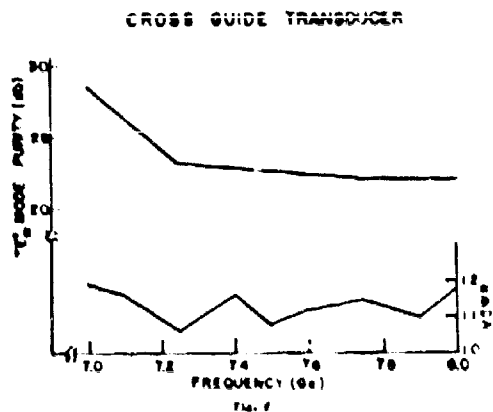


FIGURE 8 CROSS GUIDE TRANSDUCER

Fig. 8

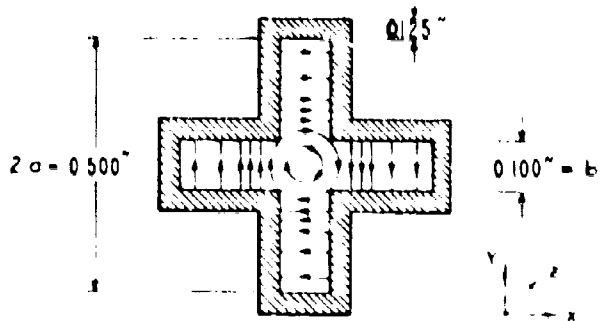


Fig. 9 ELECTRIC FIELD DISTRIBUTION IN CRUCIFORM WAVEGUIDE

Fig. 9

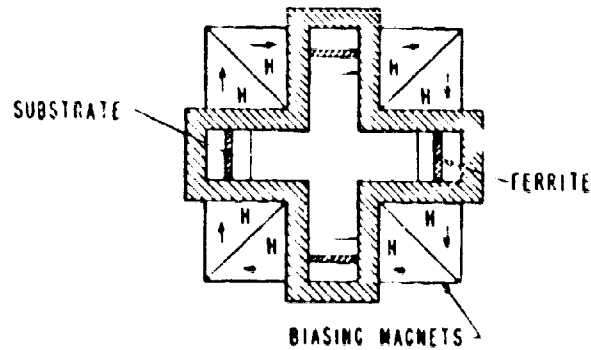


Fig. 10 CROSS-SECTIONAL VIEW OF CRUCIFORM WAVEGUIDE ISOLATOR

Fig. 10

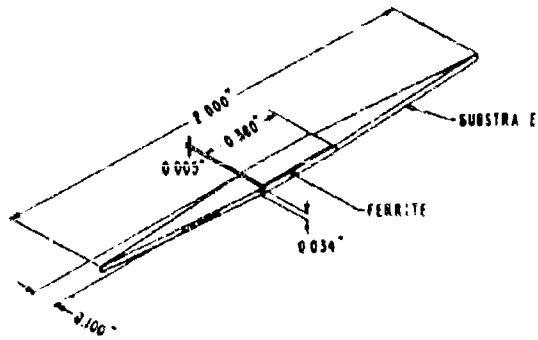


FIG. 11 FERRITE - DIELECTRIC CONFIGURATION

Fig. 11



FIGURE 12 COAXIAL ISOLATOR

Fig. 12

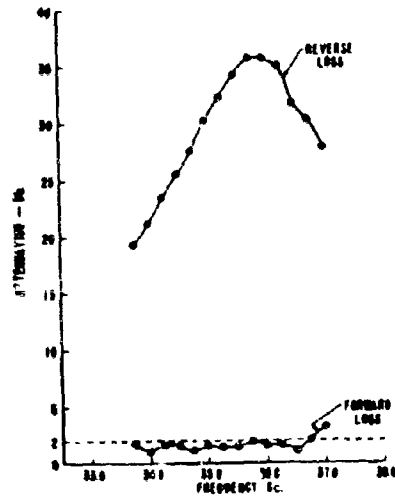


FIGURE 13 CIRCULATOR ISOLATOR CHARACTERISTICS

Fig. 13

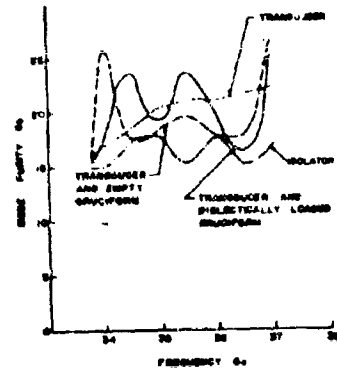


FIGURE 14 WAVEGUIDE ISOLATOR CHARACTERISTICS

Fig. 14

DENDRITIC MORPHOLOGY OF HIGH-STRENGTH STEEL CASTINGS

PAUL J. AHEARN and FRANCIS C. QUIGLEY
U.S. ARMY MATERIALS RESEARCH AGENCY
WATERTOWN, MASSACHUSETTS 02172

INTRODUCTION

The purpose of this investigation was to study a number of important parameters associated with solidification in order to establish an interrelationship between these parameters and the mechanical properties of the resulting casting. It is impractical to cause a given variation in solidification conditions and then to directly observe a change in mechanical properties due to this variation. The practical method is to consider the independent variable of casting thickness and to study other variables as functions of casting thickness or distance from the location of interest to the surface of the casting. The dependent variables were dendritic morphology, dendritic arm-spacing, dendritic arm-length or cross-arm length, microsegregation, and finally, mechanical properties before and after homogenization of the structure. All of these were studied as functions of distance from the surface of the casting.

Individual variables have been considered before (1-5) but here an attempt was made to examine many variables associated with the solidification of a single given casting in order to establish an interrelationship that possibly could be used to advantage in improving the mechanical properties of high-strength steel.

PROCEDURE

Morphological Study

The casting selected for this study was a unidirectional ingot casting 5-1/2 inches in diameter and 10 inches high. The method of manufacture consisted of preheating a hollow ceramic cylinder to 2600°F, placing the cylinder on a massive copper chill, and then pouring molten steel into the cavity formed by the ceramic cylinder chill

configuration. Upon filling the mold, a preheated ceramic slab was placed on the top surface of the mold. Unidirectional solidification progressed from the chill and resulted in the macrostructure shown in Figure 1. This type of solidification pattern not only yields outstanding mechanical properties, but also can be used as a valuable research tool.⁽⁶⁾ The unidirectional solidification produced long columnar dendrites that are lined up in such a way as to simplify study of dendritic morphology and microsegregation. It is easily seen that a normal sand casting with randomly oriented equiaxed grains would cause experimental complications.

The basic mechanical properties of this columnar casting 69 were determined by removing test bar blanks parallel to the cast surface in a progressive fashion starting from the chill and working to a depth of six inches into the casting. The blanks were heat treated in accordance with the schedule in Table I, and then were tested to determine the basic properties of this unidirectionally solidified steel. The results were compared with those developed by a conventional sand casting possessing the same chemistry and having undergone the same heat treatment.

TABLE I
Chemistry and Heat Treatment of Castings 48 and 69

Section A - Chemistry									
	Element (weight percent)								
	C	Mn	Si	P	S	Ni	Cr	Mo	Al
Casting 48:	0.31	0.73	0.29	0.015	0.027	2.78	0.74	0.53	0.08
Casting 69:	0.31	0.84	0.35	0.009	0.023	2.79	0.79	0.53	0.08
Section B - Normal Heat Treatment (following Homogenization)									
1750°F for 1 hr, air cooled 1650°F for 1/2 hr, furnace dropped to 1350°F for 15 min, water quenched 400°F for 4 hr, water quenched 400°F for 4 hr									

The macrostructure appearing in Figure 1 created an overall impression of a directionally solidified casting but it does not provide detailed information on the exact dendritic structure. Accordingly, samples were removed from the casting for detailed macrostructural examination. Samples horizontal and vertical to the chill surface were studied. The actual procedure consisted of polishing the samples, macroetching with Stead's reagent, and then photographing the structures at a magnification of 6X. Interesting dendritic parameters, such as the length of dendritic cross-arms and the interdendritic arm-spacing were measured and plotted versus distance from the casting surface.

Although exact information can be obtained on the size and the shape of dendrites, the macroetch procedure does not reveal quantitative information on the amount of microsegregation occurring in various portions of the dendritic structure. Quantitative microsegregation

gation information can be obtained through the use of the electron beam microprobe.⁽⁷⁾ Thus samples were removed from the casting for microprobe analysis. Such samples were always in the plane of the chill casting interface surface and were taken at increasing distances from this interface.

The steps involved in preparing samples for microprobe analysis consisted of polishing and etching the specimens, selecting various locations on the dendritic surface for study, and marking these locations with a hardness tester. The appropriate locations were either a scan along the full length of a dendritic arm or a scan of equal length oriented 45° from the dendritic arm. After the scans had been selected and fixed by the hardness indenter, the etched surface was removed by polishing. At this stage the samples were ready for microprobe analysis. The actual microprobe scans yielded detailed information on microsegregation but for purposes here, emphasis was placed only upon the maximum and minimum chemistries observed in the microprobe study. These maxima and minima were recorded and plotted versus distance into the casting. Microprobe scans were made to determine variations in alloying elements, Mn, Ni, Cr, and Mo.

The results of the microprobe study combined with the macrostructure observations of dendritic arm-spacing and cross-arm length made it possible to theoretically formulate homogenization conditions that would eliminate or cut down on the microsegregation. Such calculations lead to homogenization experiments at extreme temperature and for extended periods of time.

Homogenization Experiments

Two castings were involved in homogenization experiments. The first consists of columnar casting 69 employed in the morphological study; and the second was a zonal casting, 48, as demonstrated in Figure 2. The dimensions of the zonal casting and the chemical analysis are similar to the columnar casting (see Table I), but the method of manufacture was slightly different in that an exothermic compound was added to the top of the mold twelve minutes after filling the mold. This slight modification in procedure produced the four solidification zones shown in the figure. The first zone was columnar, the second was equiaxed, the third was columnar, and the fourth was equiaxed. The change from the first equiaxed to the second columnar zone was undoubtedly caused by an increase in the temperature gradient at the liquid solid interface, which was in turn caused by the addition of exothermic material. The major purpose in changing the mode of solidification was to make possible direct mechanical property comparisons of columnar and equiaxed material wherein the rate of solidification of the material was greater than that of the columnar. Basic mechanical properties of this zonal casting 48 were determined after normal heat treatment. These mechanical properties plus the normal mechanical properties of the columnar casting 69 were then compared with companion results from test blanks that underwent an extensive homogenization

treatment. Blanks from the columnar casting 69 were removed at various distances from the chill surface and then were subjected to a 2500°F homogenization treatment for 64 hours. In similar fashion, bars were removed from the first three zones of the zoned casting 48 and these were homogenized at 2500°F for either 20, 64, or 186 hours. In all cases the high temperature homogenization treatment was followed by normal heat treatment procedure.

The extremely high temperature homogenization (2500°F) and the extended time at temperature presented experimental difficulties and subsequent precautions were taken to avoid these difficulties. A cross-sectional sketch of the furnace is shown in Figure 3. The furnace consisted of a globar arrangement with a heating zone of 1-by 2-foot dimensions. For these experiments the door of the furnace was raised and the opening left by the door was bricked up except for a ceramic tube that was inserted into the heating zone through the brick work. The thermocouple for the furnace was mounted in line with the ceramic tube so that the temperature reading of the thermocouple coincided with the temperature of the ceramic tube. The arrangement in Figure 3 has the practical advantage that temperature fluctuations within the ceramic tube were minimized by the large hearth volume outside of the tube. Actual temperature control was maintained through a proportioning controller rather than a power on-power off arrangement. Test bars were placed inside the tube and a protective atmosphere was maintained in the area of the test bars by means of the arrangement indicated in Figure 3. Although the argon was 99.96 percent pure, it was felt that precautions should be taken not only against leaks within the ceramic tube, but also against the possibility that a small amount of impurities within the argon could damage the test bars. Such thinking is justified in view of the extreme high temperature and long times involved. Two basic precautions were taken: first, the test bar blanks were coated with alumina; and second, steel wool was placed just beyond the argon inlet. The steel wool acted as a getter and did in fact show substantial oxidation damage at the end of each experimental run. The alumina coating, on the other hand, was abandoned as an unnecessary precaution half way through the program.

RESULTS

Morphological Study

The mechanical properties of the unidirectional casting 69 are compiled in Table II. Yield strength, tensile strength, and true breaking strength are recorded, as well as ductility in terms of reduction of area and elongation. Interpretation of this data is facilitated by a plot (Figure 4) of the mechanical property information versus distance from the surface of the casting. Mechanical properties are also presented for a conventional sand casting (8) of the same analysis that received the same heat treatment. The properties of unidirectionally solidified steel are indicated by the dashed curves while the properties of the conventional sand casting are shown by

TABLE II
Mechanical Properties of Casting 69

Section A - Homogenized 2200°F, 4 hours, followed by Normal Heat Treatment					
Test Location Distance from Chill (Inches)	Yield Strength at 0.2% Offset (psi)	Tensile Strength (psi)	True Fracture Stress (psi)	Elongation (%)	RA (%)
0.544	197,400	253,000	327,700	11.5	31.4
1.100	190,000	254,500	316,400	9.4	25.0
2.003	186,000	253,000	320,000	11.0	29.0
2.800	200,100	253,000	322,000	9.0	28.4
3.430	194,400	255,000	320,200	10.5	28.0
4.100	196,000	255,000	320,500	10.0	28.0
5.000	204,400	260,000	313,200	10.0	24.0
5.900	182,000	236,500	301,000	9.0	20.0
Section B - Homogenized 2500°F, 64 hours, followed by Normal Heat Treatment					
0.30	183,000	237,000	-	15.0	51.5
1.55	194,000	238,000	-	13.5	44.6
2.72	193,000	220,000	-	11.5	41.0
4.10	193,000	240,000	-	12.0	39.0
5.48	193,000	235,200	-	10.0	37.2
6.73	196,000	236,400	-	9.0	17.0

solid lines. At the strength level of 255,000 psi, the unidirectional solidified casting exhibits mechanical property superiority over the conventional casting. The conventional casting undergoes a rapid decrease in both tensile strength and ductility with distance from the surface of the casting, while the unidirectional casting displays uniform properties of 255,000 psi tensile strength and 30 percent reduction of area, up to depths of three inches from the casting surface. Thereafter, mechanical property deterioration occurs, but in a gradual fashion. The deterioration in mechanical properties is called the mass effect and it is readily seen that the effect in unidirectionally solidified material is far less pronounced than in material that has been solidified in the conventional manner. The most sensitive measure of mass effect is shown by a decrease in ductility as measured by percent reduction of area versus distance from the casting surface. The same trend is noted when the true breaking strength versus distance from the casting surface is observed. The two types of castings exhibit different mechanical properties associated with different rates of solidification.

The cause of the mass effect was considered next by examining the macrostructure and microsegregation. Figures 5 and 6 present typical macrostructural results. Figure 5 is for samples parallel to the direction of solidification, and Figure 6 is for samples perpendicular to the direction of solidification. In both cases, the structure was brought out by etching with Stead's reagent and then photographing the surface at magnification of 6X. Magnified photographs yield far more information than a typical photomicrograph as in Figure 1. Figure 5 clearly shows that the columnar dendritic structure extends from the

bottom to the top of the casting, and that the structure is finest near the chill and undergoes a gradual coarsening as distance from the chill increases. The photograph at the left of Figure 5 was taken from material close to the chill while that at the right represents material at a considerable distance from the chill.

Figure 6 yields similar information to that of Figure 5, although now the dendrites are viewed from a different orientation. The tops of the dendrites are seen as cruciforms that increase in size as distance from the casting increases. By combining the photographic information in Figures 5 and 6, one can construct a model of columnar dendrites. Individual dendrites consist of a continuous spine with side branches forming at 90° in a cruciform configuration. Such a model does not take into account further branching off the cruciforms. Dendritic arm spacing and the length of dendritic cross-arms were measured with the aid of Figures 5 and 6. Cross-arm lengths were determined by two methods: the first consisted of counting the number of cruciforms per unit area in a series of photomicrographs, determining the cross-sectional area per cruciform and then finding the square root of this area to obtain a representative value of the cross-arm length. The second method simply involved measuring the length of a series of cross-arms in various photomicrographs and then plotting these lengths versus the distance from the casting surface. Such results are shown in Figure 7 where cross-arm length is plotted versus distance from the casting surface. The curve obtained by simply recording the length of cross-arms does not have as great a slope as the curve derived by the alternate method of counting the number of cruciforms per unit area of macrosurface. Regardless of the counting method employed, there appears to be a linear relationship between the length of cross-arms and distance from the casting surface. This is not surprising in view of more traditional results presented in Figure 8 wherein interdendritic arm-spacing is plotted versus distance from the casting surface. Once more a linear relationship is obtained. The linear relationship between arm-spacing and distance into the casting has previously been reported. (9)

Photomicrographs of typical samples that were employed in the microsegregation studies are shown in Figure 9. The photomicrograph on the left (9-a) is from material 0.125 inch from the chill while the photomicrograph on the right (9-b) represents material approximately 3.44 inches from the chill. In all cases the boundaries of the scans were marked with a Knoop indenter. The macrostructures were polished away and finally the samples were subjected to microprobe analysis. Two types of scans were involved as indicated in Figure 9. The first was a scan along a dendritic arm through a dendritic spine, while the second type was always oriented 45° from the first. Presumably the 45° scan started in a highly segregated area some distance from the cruciform and progressed in a straight line through the center of the spine and then out the other side into a high alloy area. Microprobe analyses were taken for Mn, Ni, Cr, and Mo content at distances 0.125, 0.8, 1.9, 2.3, and 3.44 inches from the chill surface.

The minimum chemistry occurred at the center of the cruciform, that is, at the spine of the dendrite, while the maximum concentration of alloying elements generally was observed at a location oriented 45° from the cruciform at a considerable distance from the cruciform. These maximum and minimum concentrations are plotted versus distance from the surface of the casting. Such results appear in Figure 10. The dashed lines indicate the wet chemical analysis while the maximum and minimum concentration at the location in question are shown by the ends of the vertical lines in the plot. This plot establishes the result that microsegregation is severe and can involve a very large fluctuation in alloy content across small distances of the casting. Yet this variation from maximum to minimum seems to remain fairly constant; that is, the amount of microsegregation does not change appreciable in moving into the casting. Microsegregation is as severe at the surface as at the center, while the basic dimensions of the dendritic structure, dendritic arm length, etc., do increase with distance into the casting. Any attempt to even-out the chemistry by means of a high temperature diffusion treatment should take into account the increasing diffusion distance arising because of the increase in basic dendritic dimensions. These two observations are important. The extent of microsegregation is fairly constant while dendritic dimensions increase with distance into the casting. With the aid of such information it is theoretically possible to prescribe homogenization treatments that should even-out the chemistry on a microscale. Calculations indicated that a 2500°F treatment would be desirable and that the required time would vary depending upon the size of the dendrites involved.

The actual mathematical treatment involved the use of an equation first employed in connection with a study of banding.⁽¹⁰⁾ If the microsegregation is described by a sinusoidal two-dimensional model, then homogenization will proceed in accordance with the following equation:

$$\frac{C}{C_m} = e^{-\frac{\pi^2 Dt}{l^2}}$$

where

C = maximum variation from average chemistry after homogenization

C_m = maximum variation from the average chemistry before homogenization

$\frac{C}{C_m}$ = degree of homogenization

D = Diffusion coefficient of alloying elements

t = time (sec)

l = diffusion distance (cm)

A representative value of $3.75 \times 10^{-10} \text{cm}^2$ was selected for the diffusion coefficient of alloying elements at 2500°F. The diffusion distance was considered to be one-half the value of cross-arm length as recorded in Figure 6. For example, a 64-hour homogenization time would result in $\frac{C}{C_m} \sim 0.8$, and a diffusion distance of 0.06 cm which corresponds

to a location of 2.5 inches from the surface of the casting. Recently a more realistic model has been formulated specifically for cast structures. (11)

Homogenization

The actual mechanical properties developed by both castings after conventional heat treatment and extensive homogenizations are presented in Tables II and III. Interpretation of the data is facilitated by the plot in Figure 11 of mechanical properties versus distance from the chill surface of the castings. Mechanical properties

TABLE III
Mechanical Properties of Casting 48

Section A - Homogenized 2200°F, 1 hour, followed by Normal Heat Treatment				
Test Location Distance from Chill (inches)	Yield Strength at 0.2% Offset (psi)	Tensile Strength (psi)	Elongation (%)	RA (%)
columnar zone 5.75	183,700	236,000	5.0	8.3
columnar zone 3.25	190,700	237,200	8.0	11.3
equiaxed zone 4.00	181,500	237,200	5.0	9.1
equiaxed zone 3.375	192,500	238,400	6.0	10.4
columnar zone 1.68	184,000	242,800	10.0	30.2
columnar zone 1.00	186,000	242,800	9.5	25.4
Section B - Homogenized 2500°F, 186 hours, followed by Normal Heat Treatment				
columnar zone 5.75	182,000	238,000	14.0	39.8
equiaxed zone 4.00	186,000	230,000	13.5	44.6
columnar zone 1.68	188,000	237,600	14.0	45.2
Section C - Homogenized 2500°F, 64 hours, followed by Normal Heat Treatment				
columnar zone 5.25	188,000	248,000	10.0	32.8
equiaxed zone 3.375	176,000	229,600	7.5	18.4
columnar zone 1.00	188,500	248,600	13.0	45.8
Section D - Homogenized at 2435°F, 20 hours, followed by Normal Heat Treatment				
columnar zone 5.75	188,500	248,400	7.0	15.6
equiaxed zone 4.00	187,000	238,200	6.5	27.4
columnar zone 1.68	188,000	246,800	9.5	30.2

are presented for both castings in the conventional heat-treated condition, and for columnar casting 48 under conditions of 20, 64, and 186 hours of 2500°F homogenization, and columnar casting 69 after 64 hours of 2500°F homogenization. These last four results are for material that has been homogenized and then heat treated by a conventional quench and temper. For all practical purposes the mechanical properties as represented by yield strength and tensile strength remain reasonably constant regardless of the treatment employed or the location of the test. The significant change brought about by extensive homogenization can be seen in the plot of percent reduction of area versus distance into the castings. The solid curves represent ductility results associated with high temperature homogenization while the dashed curves are associated with conventional heat treatment. In both castings there is a decrease in ductility as distance from the casting surface increases. This is the mass effect and has already been discussed.

The significance of the homogenization experiments is seen when one compares the heat-treated properties of these two castings with the 2500°F homogenized and heat-treated properties. In all cases where bars received high temperature homogenization, the ductility was improved and the improvement in ductility is directly related to homogenization time. In spite of the improved ductility with homogenization, the mass effect is still observed; that is, ductility decreases as distance from the casting surface increases. It is also observed that up through the extended time of 186 hours, the mechanical properties near the chill in zonal casting 48 continued to improve; therefore, it is suggested that maximum possible ductility was not obtained.

It is interesting to compare the results demonstrated in Figure 11 with mechanical properties of forged material. For example, AISI 4340 1/2-inch diameter forged bars, oil quenched and tempered to a strength level of 250,000 psi, exhibit a reduction of area of 48 percent.⁽¹²⁾ These results are naturally longitudinal properties and can be compared directly with the results appearing in Figure 11 and Tables II and III. In some cases the ductility of test bars located close to the chill surface slightly exceed the forged bar properties just mentioned; however, it is significant that ductilities of 45 percent reduction of area are maintained at depths of two or three inches into the castings.

Finally, zonal casting 48 results indicate that there is no significant difference between the ductility of equiaxed material and columnar material *parisse*. Other factors such as microporosity and microsegregation must be considered.

The improvement in the ductility of the casting must be assigned to the success of the 2500°F homogenization treatment. The homogenization treatment can be considered from a theoretical point of view by plotting ductility versus homogenization parameter for columnar casting 69 as shown in Figure 12. Here ductility is plotted against the dimensionless quantity diffusion coefficient of alloying element times the time of homogenization, divided by the square of one-half the dendritic cross-arm length taken from Figure 7. A continuous curve is obtained. Since the parameter is derived from consideration of first principles, it is suggested that the improved ductility is associated with the decrease in microsegregation that occurred during the high temperature homogenization.

The ductility data from both castings can be combined on a single graph by assuming that interdendritic dimensions increase in a linear fashion (9) with distance from the chill. Such reasoning leads to the simple substitution of distance from the surface of the chill for diffusion distance. The ductility data for both castings is placed on these plots and in both cases a reasonably smooth curve is obtained. In Figure 13 the diffusion coefficient employed had a value of $3.75 \times 10^{-10} \text{cm}^2$ except in the case of three bars from casting 48 that were homogenized for 20 hours. It was noted that the actual temperature of

AHEARN and QUIGLEY

homogenisation here was 2435°F. The appropriate diffusion coefficient under such circumstances would be $2.34 \times 10^{-10} \text{cm}^2$. Such a substitution yielded a slight smoothing of the curve in Figure 13.

The changes that occur during the 2500°F homogenization can easily be visualized by study of the fractured tensile bars from sonal casting 48 in Figure 14. Bars that received the standard treatment and bars that underwent 2500°F homogenization were macroetched and then set side by side for comparison. Macroetching clearly shows that the macrostructure tends to be eliminated by the 2500°F treatment. This is especially true at great distances from the casting surface. The fracture characteristics also change. Interdendritic fracture can be seen in bar 1-N while its companion 1-2500°F fractured in a cup-and-cone fashion. Finally, the increase in ductility resulting from the 2500°F treatment is clearly indicated by the necking of the 2500°F bars.

CONCLUSIONS

A number of practical and theoretical conclusions emanate from this work. It has been shown that microsegregation and dendritic size play a major role in determining mechanical properties of high-strength steel castings, and that microsegregation can be eliminated by means of 2500°F homogenization treatment. The resulting cast material exhibits mechanical properties identical to longitudinal properties of forged bar stock at the same strength level. Transverse forged properties are always inferior to longitudinal properties, thus the mechanical property superiority of this cast steel is established.

REFERENCES

1. ALEXANDER, B.H., and RHINES, F.H. Dendritic Crystallization of Alloys. Trans. AIME, v. 188, 1950, p. 1267.
2. WAGNER, C. Theoretical Analysis of Diffusion of Solids during the Solidification of Alloys. Trans. AIME, v. 200, 1954.
3. WEINBERG, F. Solute Segregation during Dendritic Growth. Trans. AIME, v. 221, 1961, p.845.
4. COLLING, D.A., et al. Cast Steel Solidification Variables Effect on Mechanical Properties. Trans. AFS, v 70, 1962, p.1088-1094.
5. MASSACHUSETTS INSTITUTE OF TECHNOLOGY. Investigation of Solidification of High Strength Steel Castings. Contract DA-19-020-ORD-5443, Interim Report, 1962.
6. BARONE, R.V., et al. Solidification of Steel Castings and Ingots. Trans. AFS, v. 69, 1961, p.422-435.
7. KOHIV, H., and PHILIBERT, J. The Solidification of Alloys. Metal Treatment and Drop Forging, v 27, September 1960, p. 351-355.
8. AHEARN, P.J., et al. High Strength Steel Castings Mass Effect on Tensile Properties. Trans. AFS, v. 70, 1962, p. 1154-1160.
9. BROWN, P.E., and ADAMS, C.M., Jr. Structures Produced in Rapidly Solidified Alloys. Trans. AFS, v. 69, 1961.
10. LAVENDER, J.D., and JONES, F.W. An Investigation of Banding. Journal of the Iron and Steel Institute, v. 163, 1949, p. 14-17.
11. MASSACHUSETTS INSTITUTE OF TECHNOLOGY. Investigation of Solidification of High Strength Steel Castings. Contract DA-19-020-ORD-5443, Interim Report, 1963.
12. BETHLEHEM STEEL CO. Modern Steels. Fifth edition.



Figure 1. CASTING 69 - UNIDIRECTIONAL SOLIDIFIED COLUMNAR GRAIN HIGH STRENGTH STEEL CASTING



Figure 2. CASTING 48 - FOUR-STAGE UNIDIRECTIONAL SOLIDIFIED HIGH STRENGTH STEEL CASTING

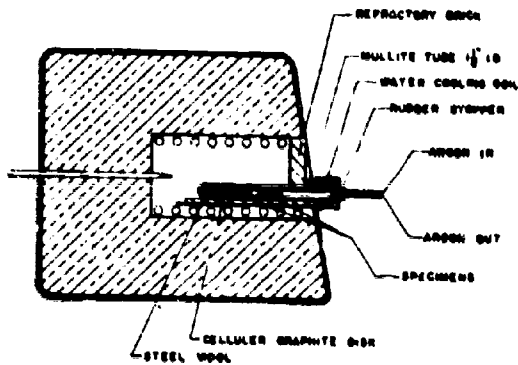


Figure 3. HIGH TEMPERATURE HOMOGENIZATION APPARATUS

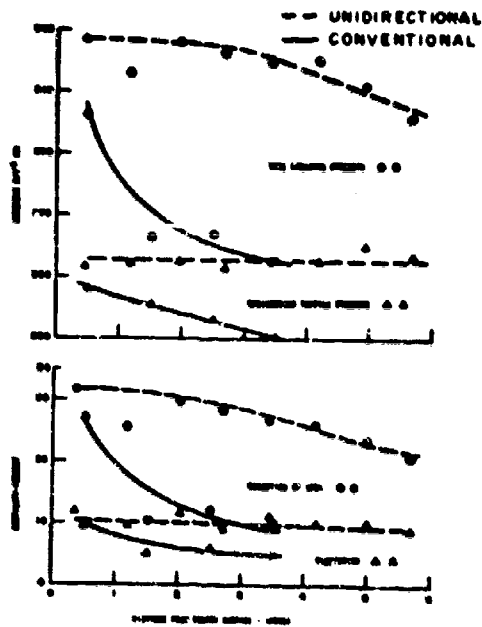


Figure 4. MECHANICAL PROPERTIES OF CONVENTIONAL AND UNIDIRECTIONAL HIGH-STRENGTH STEEL CASTINGS

AHEARN and QUIGLEY

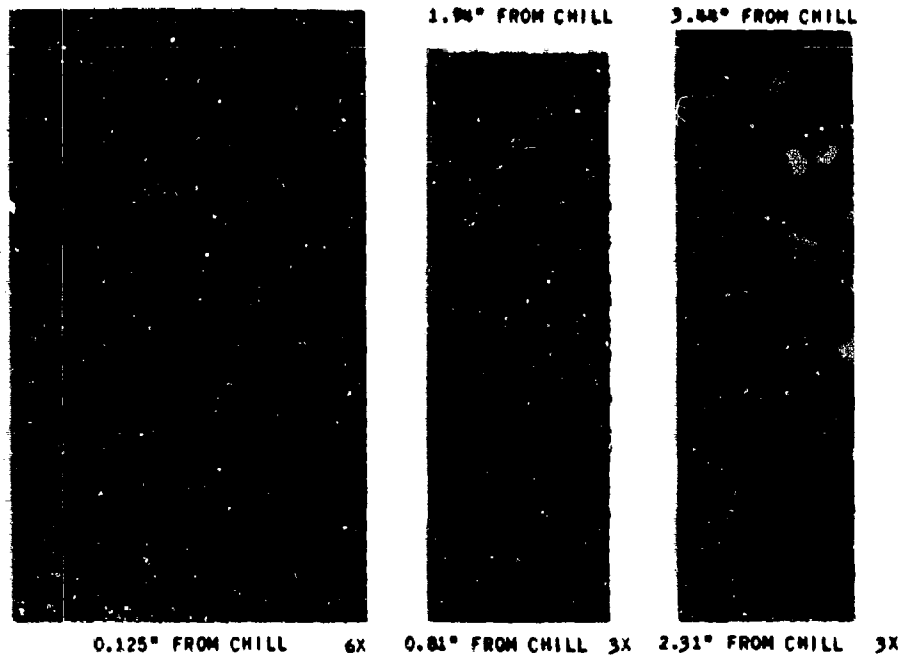


Figure 5. MACROSTRUCTURE PARALLEL TO THE DIRECTION OF SOLIDIFICATION (STEAD'S REAGENT)

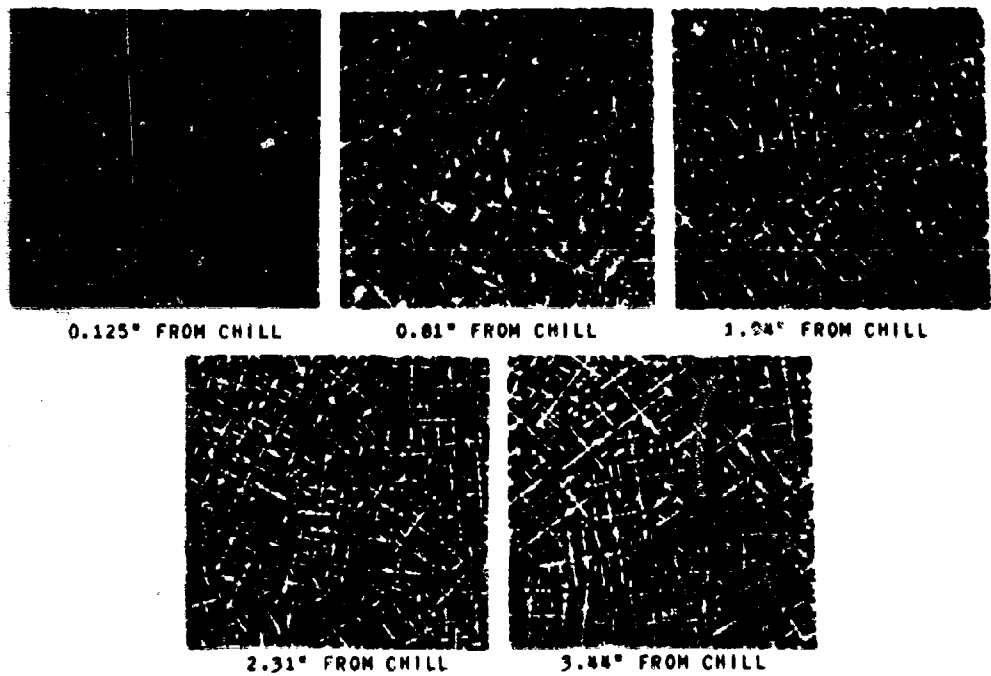


Figure 6. MACROSTRUCTURE PERPENDICULAR TO THE DIRECTION OF SOLIDIFICATION (STEAD'S REAGENT). 6X.

U. S. ARMY MATERIALS RESEARCH AGENCY

19-066-298/AMC-64

AHRAFI and QUIGLEY

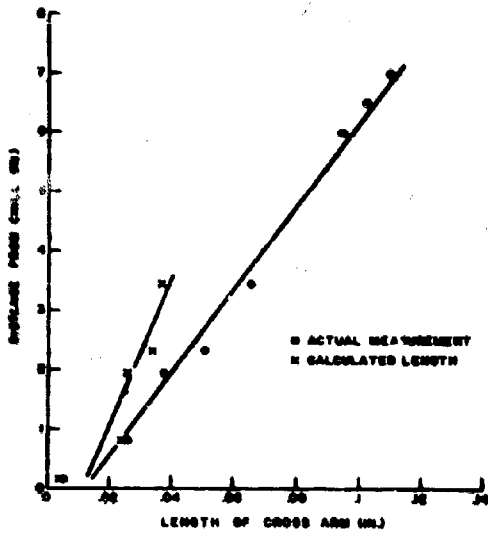


Figure 7. DISTANCE FROM CHILL VERSUS LENGTH OF CROSS ARM

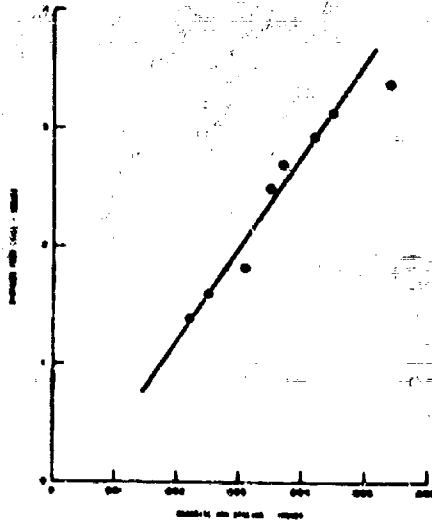


Figure 8. DENDRITE ARM SPACING PARALLEL WITH THE DIRECTION OF SOLIDIFICATION IN UNIDIRECTIONAL HIGH-STRENGTH STEEL CASTINGS

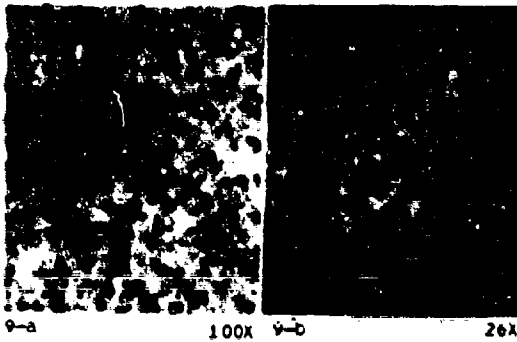


Figure 9. LOCATION OF MICROPROBE SCANS AT 0.125" FROM CHILL (100X) AND 3.44" FROM CHILL (26X)

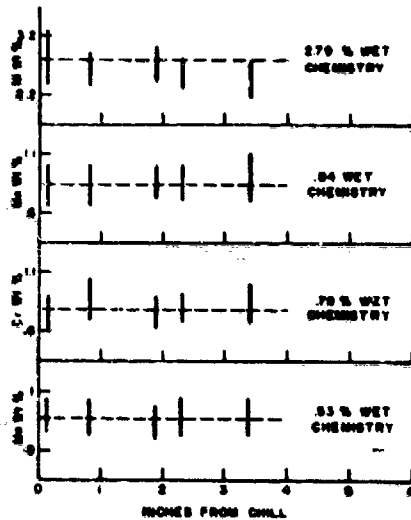


Figure 10. MAXIMUM AND MINIMUM WEIGHT PERCENT OF ALLOYING ELEMENTS VERSUS DISTANCE FROM THE CHILL (INCHES)

U. S. ARMY MATERIALS RESEARCH AGENCY

19-366-300 AMC-64

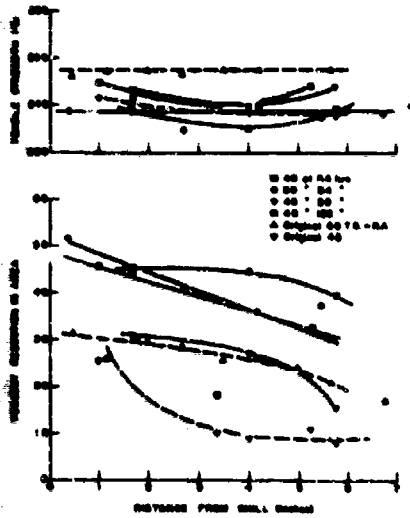


Figure 11. MECHANICAL PROPERTIES OF CASTINGS 48 AND 89 BEFORE AND AFTER HIGH TEMPERATURE HOMOGENIZATION

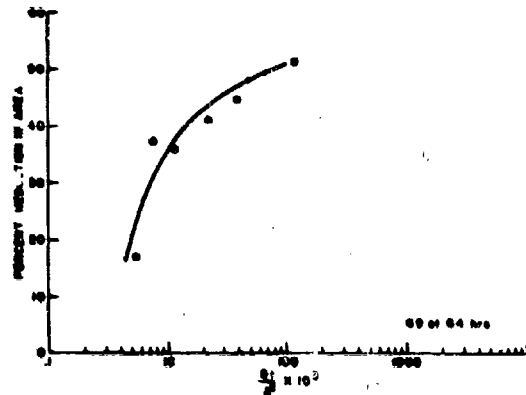


Figure 12. PERCENT REDUCTION OF AREA VS DIFFUSION COEFFICIENT X TIME $\left(\frac{\text{cm}^2 \cdot \text{SEC}}{\text{IN}^2}\right)$ LENGTH OF CROSS ARM (one side) (cm^2)

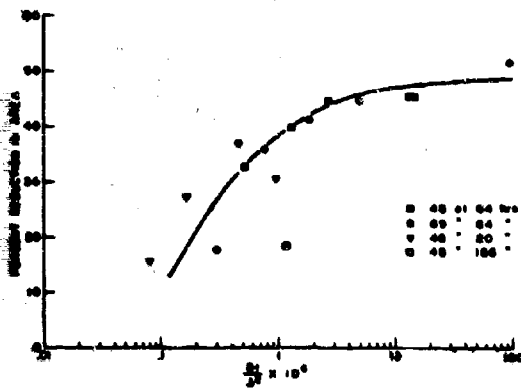


Figure 13. PERCENT REDUCTION OF AREA VS DIFFUSION COEFFICIENT X TIME $\left(\frac{\text{cm}^2 \cdot \text{SEC}}{\text{IN}^2}\right)$ DISTANCE FROM CHILL (cm^2)

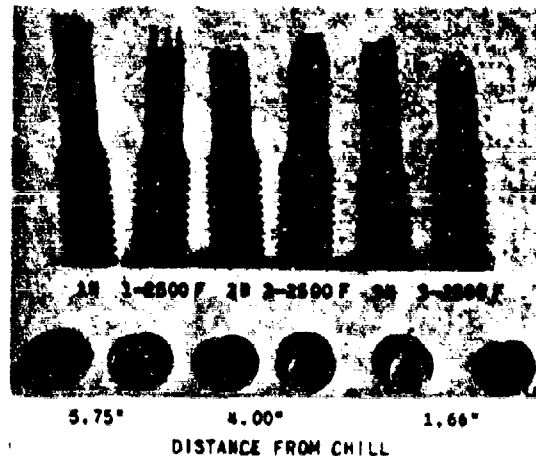


Figure 14. CASTING 48 - TENSILE BREAKS FOR NORMAL HEAT TREATMENT AND 2500 F HOMOGENIZATION TREATMENT FOR 168 HOURS

NUMERICAL SOLUTION
OF THE DISTRIBUTION
OF WIND AND TURBULENCE
IN THE PLANETARY BOUNDARY LAYER

J. F. APPLEBY, and W. D. OHMSTEDE
U. S. ARMY ELECTRONICS RESEARCH & DEVELOPMENT ACTIVITY
FORT HUACHUCA, ARIZONA

1. INTRODUCTION

Meteorology is subdivided according to the scales of micrometeorology, mesometeorology and macrometeorology. This division appears real since research interests and successes have been widely separated. If meso-scale phenomena are regarded as perturbations within the macro-scale, two elements can be selected to characterize the divisions. These would be the logarithmic wind profile to characterize micrometeorology and the geostrophic wind to characterize the others. Both of these relationships are rather successful formulas. The geostrophic wind equation has been used to evaluate the horizontal field of wind in the free atmosphere from the pressure distribution. The logarithmic wind profile equation has been used to relate the vertical distribution of wind speed immediately above the earth's surface to the momentum transport and to the roughness characteristics of the surface. On the face of it, there does not appear to be any direct connection between these two relationships as they deal with different regions of the atmosphere.

The purpose of this report is to present a hypothesis which unifies these two concepts. The model resulting from the hypothesis is solved and the results are presented in a form usable for many situations where it is necessary to predict the vertical distribution of the wind and turbulence characteristics within the planetary boundary layer. The treatment will be limited to the barotropic, adiabatic, steady state case.*

2. DEVELOPMENT OF THE MODEL EQUATIONS

* A non-steady state case is required to incorporate the meso-scale.

APPLEBY and OHMSTEDE

The model equations are developed from the equation of motion:

$$d\vec{V}/dt = \partial\vec{V}/\partial t + \vec{V} \cdot \nabla \vec{V} = - 2\Omega \times \vec{V} + \vec{g} - 1/\rho \nabla p + 1/\rho \nabla \cdot \vec{\tau} \quad (1)$$

where \vec{V} is the wind velocity vector relative to a point on the earth's surface, Ω is the angular velocity of the earth's rotation, \vec{g} is the acceleration of gravity, and $\vec{\tau}$ is the stress dyadic. This equation is difficult to solve without simplification.

For this study we assume the wind is steady, uniform in the horizontal and that the horizontal pressure gradient force is uniform and constant. As a result equation (1) can be reduced to

$$\begin{aligned} -fu &= 1/\rho \partial p/\partial y - 1/\rho \partial \tau_y/\partial z \\ fv &= 1/\rho \partial p/\partial x - 1/\rho \partial \tau_x/\partial z \end{aligned} \quad (2)$$

where $\vec{V} = u\vec{i} + v\vec{j}$, $f = 2\Omega \sin \theta$ where θ is the latitude, p is the pressure, and τ_x and τ_y are stresses in the \vec{i} and \vec{j} directions, respectively.

At this point we introduce the geostrophic wind (G) which is defined as $G = G_x\vec{i} + G_y\vec{j} = 1/\rho f (\vec{i} \partial p/\partial y - \vec{j} \partial p/\partial x)$, substituting this in equation (2) we obtain

$$\begin{aligned} f(G_x - u) + 1/\rho \partial \tau_y/\partial z &= 0 \\ -f(G_y - v) + 1/\rho \partial \tau_x/\partial z &= 0 \end{aligned} \quad (3)$$

Equations (3) are then the fundamental equations of the model. The solution of equation (3) requires a knowledge of stresses throughout the layer. Since the atmospheric boundary layer is in most cases fully turbulent, the stresses in equations (3) are the Reynolds stresses—statistical quantities that represent the covariances of the turbulent fluctuations. Our knowledge is still inadequate to derive transport relationships from the primitive equation, so we must rely on semi-empirical or phenomenological theories to obtain a solution. Prandtl's mixing-length hypothesis has been used in turbulent transfer models for three decades. We have used this concept in the solution of equation (3) knowing the results will not be precise but believing they will be useful.

APPLEBY and OHMSTEDE

The mixing-length relationship can be expressed by

$$\tau/\rho = U_*^2 = (\lambda dU/dz)^2 \quad (4)$$

where $U_* = (\tau/\rho)^{1/2}$ is the friction velocity, λ is the mixing-length and U is the wind speed. This expression is more common in the form:

$$\tau/\rho = U_* \lambda dU/dz = K dU/dz \quad (5)$$

where K is the eddy viscosity. There have been conflicting views on the factors which control the mixing-length. von Karman believed that the magnitude of the mixing-length is determined by local flow conditions, while Prandtl relates the magnitude to the gross features of the flow. von Kármán's hypothesis has certain desirable features but has not been generally accepted (see Lettau [1]). In free turbulence Prandtl views the mixing-length as constant in a cross-section of the mixing zone and its magnitude as proportional to the width. In a wall turbulence regime, it is assumed that the magnitude of the mixing-length is proportional to the distance from the wall and can be expressed by:

$$\lambda = kz \quad (6)$$

where k is von Kármán's constant. Combining equations (4) and (6) and integrating gives the logarithmic profile of the surface boundary layer

$$U = U_*/k \ln(z/z_0) \quad (7)$$

where z_0 is a length characterizing the roughness of the surface.*

Neglecting the variation of the air density with height, we can incorporate equation (5) into equation (3) with the result:

$$\begin{aligned} f(G_x - u) + \partial/\partial z (U_* \lambda \partial v/\partial z) &= 0 \\ -f(G_y - v) + \partial/\partial z (U_* \lambda \partial u/\partial z) &= 0 \end{aligned} \quad (8)$$

Actually, equation (7) is inconsistent with equation (8). The logarithmic wind profile assumes that the shearing stress ($U_* \lambda dU/dz$) is constant with height, equation (8)

* A more general discussion of the roughness length can be found in ERDAA-MET-7-63 report, "A Model for Wind Flow In An Idealized Vegetative Canopy".

APPLEBY and OHMSTEDE

shows that the gradient of the stress is greatest near the surface. This can be reconciled by recognizing that the percentual rate of change near the surface is so small as to make the wind profile logarithmic. At great height the stress will vanish and by equation (8) the wind becomes geostrophic; thus the solution should have the desired features of our objective.

To solve the model equations the specification of the vertical distribution of the mixing-length is required. According to equation (6), the mixing-length is independent of kinematic characteristics and solely a function of proximity to the boundary. It seems unreasonable that equation (6) should be valid ad infinitum. The outer portion of the boundary layer behaves more like a region of free turbulence than of wall turbulence, which suggests that the mixing-length should approach a limiting value far from the surface.

It is proposed that for a steady-state, neutral atmospheric boundary layer, the mixing-length increases with height at the rate given by von Kármán's constant, but there is a linear feedback which prevents unbounded growth. This can be expressed by the following differential equation:

$$d\ell/dz = k - \ell/L \quad (9)$$

where L is a constant. The solution of equation (9) which satisfies the boundary conditions is

$$\ell = \ell_0(1 - e^{-z/L}) \quad (10)$$

where $\ell_0/L = k$. This equation has the characteristics that ℓ increases linearly with height near the boundary, but approaches the limiting value ℓ_0 at great height. Figure 1 compares equation (10) with the mixing-length distribution computed from Lettau's^[2] reanalysis of the Leipzig wind profile. The agreement is reasonably good except in the upper portion of the profile. In this region the evaluation of the mixing-length is less reliable because of the small values of the velocity derivative.

3. NUMERICAL SOLUTIONS OF DIFFERENTIAL EQUATIONS

Equation (10) has been adopted as the general expression for the mixing-length and now it is possible to solve for the velocity profile in the planetary boundary layer. For computing purposes, the equations are expressed in non-dimensional form by defining new variables:

APPLEBY and OHMSTED

$$U = (G_x - u)/G, \quad V = (G_y - v)/G, \quad s = \tau_x / \rho G^2, \quad t = \tau_y / \rho G^2$$

$$T = U_s / G = (s^2 + t^2)^{1/4} \quad R_0 = G / f z_0 \quad x = z / z_0$$

where R_0 is termed the surface Rossby number. Using these new variables, equations (3) and (5) are expressed as:

$$V = R_0 \partial s / \partial x, \quad U = - R_0 \partial t / \partial x \quad (11)$$

and

$$s = - T(z_0/z_0) \gamma \partial U / \partial x, \quad t = - T(z_0/z_0) \gamma \partial V / \partial x \quad (12)$$

where $\gamma = 1 - \exp(-z/L)$. Differentiating equations (11) and substituting the results in equation (12) we obtain

$$s = A_0 T \gamma \partial^2 t / \partial x^2, \quad t = - A_0 T \gamma \partial^2 s / \partial x^2 \quad (13)$$

where $A_0 = R_0 z_0 / z_0$.

The boundary conditions require that $s=t=0$ at great height and we can choose axes of orientation so that at the lower boundary $t=0$ and $s=s_0$. For each value of the parameter A_0 , there must be a corresponding value of s_0 . The actual relationship of s_0 and A_0 cannot be determined until the solutions are known. It is necessary to rewrite equations (13) in a form containing only a single parameter. This is done by defining new variables:

$$P = A_0 T \quad Q = A_0^2 s \quad R = A_0^2 t$$

Using these terms, equations (13) become

$$Q = P \gamma \partial^2 R / \partial x^2, \quad R = - P \gamma \partial^2 Q / \partial x^2 \quad (14)$$

with boundary conditions of

$$x = 1 : Q = Q_0, \quad R = 0$$

$$x \rightarrow \infty : Q = R = P = 0$$

From equation (14) and the boundary conditions, a particular solution is completely specified by the quantity Q_0 , provided γ is solely a function of the distance from the boundary.

Difficulties were encountered in solving equations (14). It was necessary to assume that Q and R became negligible at a finite but large height rather than at infinity. Several mathematical methods were

APPLEBY and OHMSTEDE

tried to solve the equations. The one adopted was the finite difference method of Crank and Nicholson [3]. Briefly, the solution consisted of transforming the independent variable to $y = \sqrt{2.0} x$ and establishing a grid of 100 equally spaced points in the y domain between $x = 1$ and $x = 11$. The derivatives at the grid points were approximated from second-order Lagrangian polynomials utilizing the values of the variables at adjacent grid points. On the basis of ordinary differential equations, a form of parabolic partial differential equation was assumed. The final solution was derived by integrating the partial differential equation until the solution decayed to the steady state ordinary solution.

Choosing $z_0/l_0 = 3.125 \times 10^{-5}$, solutions of equation (14) were computed for boundary values (Q_0) ranging over three orders of magnitude. The solutions can be applied to any arbitrary surface roughness length by accepting as boundary conditions the values of Q and R at the corresponding height. The roughness length plays a significant role in translating the solutions of Q and R into terms of physical variables. By definition, at height z_0 the velocity is zero. In terms of nondimensional variables, this requires that $U^2 + V^2 = 1$. From equation (11) $R_0^2 [(\partial s/\partial x)^2 + (\partial t/\partial x)^2] = 1$. The surface Rossby number (R_0) can then be determined for a given solution at any roughness length by the equation:

$$R_0 = (z_0/l_0)^2 [(\partial Q/\partial x)^2 + (\partial R/\partial x)^2]^{1/2} \quad (15)$$

or the quantity

$$G/fl_0 = (z_0/l_0)^3 [(\partial Q/\partial x)^2 + (\partial R/\partial x)^2]^{1/2} \quad (16)$$

We found that the derivatives of Q and R at the lower boundary could not be adequately determined from finite difference formulae, so an approximate analytical solution was developed. Since Q and P are nearly constant with height at the lower boundary we can express Q and P in terms of the average values in the first meter above the surface. We know from Equation (6) that $\gamma = (z_0/L)x$ near the boundary. The first equation of equations (13) can be expressed as:

$$\partial^2 R/\partial x^2 = L\bar{Q} 1/x\bar{P}z_0 \quad (17)$$

where the bar denotes average values. Integrating twice, we derive an expression for the derivative at the lower boundary. A similar approach can be applied to the second equation of equations (14). Numerical values of the derivatives can then be found for any arbitrary height (x)

from the resulting expressions. Then the surface Rossby number and G/fz_0 can be determined for any roughness lengths from equations (15) and (16).

The surface Rossby number (G/fz_0) is a useful parameter for characterizing the planetary boundary layer just as the Reynolds number characterizes turbulent flow through ducts. Two additional parameters which are useful are the angle between the surface wind and the geostrophic wind (α) and the geostrophic drag coefficient (C). The surface wind angle can be determined by the following equation:

$$\alpha = \arctan(\partial Q/\partial x/\partial R/\partial x) - \arctan(R/Q) \quad (18)$$

evaluated at the roughness height z_0 . The geostrophic drag coefficient is given by $C = U_s/G$ where U_s is evaluated at the roughness height z_0 . This leads to the following formula for evaluating the surface stress in terms of the geostrophic wind:

$$\tau_0 = \rho C^2 G^2 \quad (19)$$

The solutions of equations (13) give the vertical distribution of the stress vector. According to equations (11) the nondimensional velocity is proportional to the derivative of the shearing stress. Near the surface these derivatives were determined from the approximate analytic solutions. For the remainder of the profile the derivatives were approximated from second-order Lagrangian polynomials using values at the adjacent grid points. Multiplying the derivatives by the Rossby number gives the velocity components.

4. BOUNDARY LAYER MODEL

The development to this point has been formal in that we have stated the differential equations, postulated a functional form for the mixing-length, and found a number of solutions of the equations using the assumed mixing-length relationship. The important question is— which of these solutions are relevant to the real planetary boundary layer? To answer this question we shall invoke the idea of similarity. We assume that the Reynolds number of the planetary boundary layer is so large that a universal equilibrium structure for turbulent eddies exists. The scale of the turbulent eddies, as measured by the mixing-length, is universally related to the gross scale (depth) of the boundary layer for any given surface Rossby number. Equation (10) shows that either z_0 or L is adequate to characterize the eddy struc-

APPLEBY and OHMSTEDE

ture. The gradient wind level (that level where the wind velocity first reaches the direction of the geostrophic wind) has been used widely in meteorology and is a direct scale length analogous to the boundary layer depth of Fluid Mechanics. Consequently, we have chosen z_0 to characterize the eddy structure and the gradient wind level (H) to characterize the gross scale of the planetary boundary layer.

According to our notion of similarity, the ratio of the gross length scale of the boundary layer and the length scale of the eddy structure (H/z_0) is a universal constant. To determine its magnitude we again resorted to the observations of the Leipzig wind profile. These data yield an $H/z_0 = 33.44$. This constant defines the surface wind angle (α) and the geostrophic drag coefficient as unique functions of the surface Rossby number. With this constant we are able to select the solution which is proposed as the universal solution of the steady state, barotropic, neutral planetary boundary layer. Table I presents this solution in terms of the nondimensional variables.

We can now demonstrate how this solution is used to determine the distribution of wind and turbulence in the planetary boundary layer. We can also illustrate some of the more salient features of the model. To use the model, it is necessary to specify the three parameters which determine the surface Rossby number, namely, the geostrophic wind, the latitude, and the surface roughness. The geostrophic wind can be readily evaluated from synoptic (or prognostic) weather maps while several sources [4,5] can be consulted to evaluate the roughness at the site.

The relationship of Hf/G and z_0f/G to the surface Rossby number as derived by the model is illustrated in figure 2. The magnitude of z_0 and H can be determined from figure 2, consequently, the nondimensional height (z/z_0) of Table I can be converted to true height z . We can also determine the relationship of the angle of the surface wind (α) and the geostrophic drag coefficient (C) to the surface Rossby number. These relationships are illustrated in figures 3a and 3b. Small angles of the surface wind are associated with strong geostrophic wind, low latitude, and small roughness. This dependence on roughness can be seen on any surface synoptic weather chart. The cross-isobaric flow is considerably less over the sea than over the land. The trend of the geostrophic drag coefficient is similar to that of the surface wind angle. If one compares figure 2 with

APPLEBY and OHMSTEDE

figures 3a and 3b, it is apparent that the surface wind angle (α) and the geostrophic drag coefficient (C) are less sensitive to changes in G/f than are H and z_0 .

Unfortunately, very little data exists which can be used to verify the model. With the exception of the Leipzig wind profile, the available data do not satisfy the conditions of the model (barotropic, neutral, steady state) and/or the method of computation used to analyze the data is of doubtful validity. Nevertheless, all known observations have been plotted on figures 3a and 3b to permit comparison with the results of the model. Although there is qualitative agreement between the model and the observations, the scatter is too great and the number of observations too few to permit any formal statistical analysis of the results. A graphic presentation of the salient feature of the model is present in figures 4-5. Figure 4a is the hodograph of the wind vector in arbitrary units. A line representing the geostrophic wind vector is drawn from the origin to the desired value of z_0/l_0 . The velocity at any height z is found by drawing a line from the value of z_0/l_0 to the z/l_0 corresponding to z . In a similar manner, the hodograph of the shearing stress vector is shown in figure 4b. The magnitude of the surface stress (τ_0) can be determined by finding the drag coefficient from figure 3b and solving equation (19).

Figure 5 shows the vertical distribution of the relative eddy viscosity ($K/l_0 G$). From previous discussion we concluded that for a fixed roughness and latitude, l_0 increases with increasing geostrophic wind. Therefore, the eddy viscosity is proportional to the geostrophic wind to a power greater than one. A simple computation shows that the maximum of the eddy viscosity occurs at a height approximately 1/5 of the gradient wind level.

5. DISCUSSION

The planetary boundary layer model presented here contains the features sought in the objective. In the free atmosphere the wind approaches the geostrophic wind while in the surface boundary layer the wind profile is logarithmic. It would be misleading to imply from this that the idealized planetary boundary layer could be as successful as the two independent concepts. The fact that the wind appears to respond quickly to pressure gradient changes, resulting in a quasi-geostrophic wind in the free atmosphere even when conditions are not ideal, account for the success of the geostrophic wind equation. Similarly, the surface wind profile responds quickly to stress and roughness changes. Batchelor [6, p 255] emphasizes that

APPLEBY and OHMSTEDE

the stress at the ground is likely to respond slowly to pressure gradient changes which indicates that for the real atmosphere a true steady state condition rarely exists. Lettau^[6, p 257] in reply stated that in the lower portion of the boundary layer the profile will respond quickly to pressure gradient changes, but more prolonged inertial oscillations will complicate the upper portion of the boundary layer.

In solving equations (14) the authors encountered computational instability in the upper portion of the boundary layer solution. This instability led to inertial oscillations of the type anticipated by Lettau. In spite of these oscillations, the lower portion of the profile was stable and the distribution of stress was in consonance with the steady state solution. This is encouraging for it implies that at the boundary such features as the stress and angle of the wind are in balance with the geostrophic wind even if the upper portion of the boundary is in oscillation.

The restriction of the solution to barotropic conditions was not essential, but simplified the problem in terms of the number of solutions required. A simple reformulation of the problem representing the geostrophic wind as a function of height will allow the thermal wind to be taken into account and the same numerical technique is applicable.

The restriction of the model to adiabatic conditions limits its applicability to conditions over the sea or land surfaces with overcast skies. However, the model can be extended to cover most conditions by including a suitable expression for the effect of stability on the mixing-length.

6. CONCLUSIONS

We believe that the development and solution of the planetary boundary layer model represents a significant step forward, since it furnishes a connecting link between micrometeorology and macrometeorology. We do not contend that the model presented is the final answer as it is based on several hypotheses which require experimental verification. The model is restricted to a class of atmospheric conditions. However, it represents a foundation which can be refined as more experimental data become available, and extended to cover a broader class of meteorological conditions.

The present model has strategic and design applications which require a knowledge of the distribution of wind or turbulence characteristics for differing situations. For example, it is believed that the model can be

APPLEBY and OHMSTEDE

used to evaluate the effect of roughness on the diffusion of CBR agents. Such a study is now under way. In like manner, the model can be extended to a broader class of micrometeorological problems such as determining climatic and synoptic estimates of evapotranspiration from large-scale watersheds, turbulent transfer of heat, water vapor, and carbon dioxide from the earth's surface.

From a long range standpoint, one of the great challenges facing meteorology is the development of a truly diabatic model for the prediction of the hemispheric circulation. An important feature of such a model would be the energy dissipation associated with the planetary boundary layer. The model furnishes estimates of the dissipation and surface shearing stress derived from knowledge of only the geostrophic wind, latitude and roughness.

It is concluded, therefore, that the planetary boundary layer model presented has useful applications and is worthy of extension to cover a broader class of situations.

BIBLIOGRAPHY

1. Lettau, H. H., "A Generalized Mathematical Model of the Mean-Velocity Distribution in Fully Turbulent Duct Flow", "Studies of the Three Dimensional Structure of the Planetary Boundary Layer," Annual Report, DA Contract 36-039-56-80282, August 1961.
2. Lettau, H. H., "A Re-examination of the Leipzig Wind Profile Considering Some Relations Between Wind and Turbulence in the Friction Layer", Tellus, 2, 125-129, 1950.
3. R. D. Richtmyer, "Difference Methods for Initial-Value Problem", Interscience Publishers, Inc., 1957.
4. U.S. Weather Bureau, "Meteorology and Atomic Energy", prepared for the U. S. Atomic Energy Commission U.S. Government Printing Office, July 1955.
5. Kung, E.G. and Lettau, H.H., "Regional and Meridional Distribution of Continental Vegetational Cover and Aerodynamic Roughness Parameters", "Studies of the Three Dimensional Structure of the Planetary Boundary Layer", Annual Report, DA Contract 36-039-56-80282, August 1961.
6. Lettau, H. H., "Wind Profile, Surface Stress, Etc., Atmospheric Diffusion and Air Pollution", Symp. Proc. Oxford 24-29 August 1958, 241-257, Advances in Geophysics, Vol 6, Academic Press 1959.

Table I
Basic Model Solution as a Function of Relative Height

z/z_0	$(U_G - u)/G$	$(V_G - v)/G$	$\tau_x/\rho G^2 \times 10^3$	$\tau_y/\rho G^2 \times 10^3$	U_s/G	$K/z_0 G$
53.1	.0000	.0000	.0000	.0000	.0000	.0000
50.0	.0021	-.0009	.0007	.0005	.0009	.0010
45.0	.0043	.0052	.0005	.0035	.0019	.0019
40.0	-.0030	.0147	-.0080	.0048	.0031	.0031
35.0	-.0112	.0174	-.0151	.0020	.0039	.0047
30.0	-.0465	+.0030	-.0316	-.0319	.0067	.0067
25.0	-.0653	-.0306	-.0218	-.0772	.0089	.0090
20.0	-.0676	-.0809	.0213	-.1315	.0115	.0115
15.0	-.0429	-.1411	.1096	-.1781	.0144	.0144
10.0	.0099	-.2011	.2472	-.1942	.0177	.0174
6.4	.0711	-.2381	.3741	-.1719	.0203	.0187
5.0	.1010	-.2503	.4285	-.1527	.0213	.0184
2.5	.1724	-.2690	.5327	-.0988	.0233	.0147
1.0	.2460	-.2787	.5978	-.0503	.0245	.0081
.5	.2956	-.2819	.6200	-.0283	.0249	.0045
0.25	.3420	-.2835	.6317	-.0162	.0251	.0024
0.10	.3999	-.2845	.6383	-.0028	.0252	.0010
3.125×10^{-2} *	.4785	-.2852	.6413	.0000	.0455	.0000
3.125×10^{-3} *	.6665	-.2855	.6426	.0000	.0350	.0000
3.125×10^{-4} *	.8124	-.2855	.6427	.0000	.0294	.0000
3.125×10^{-5} *	.9584	-.2855	.6428	.0000	.0254	.0000
3.125×10^{-6} *	1.1043	-.2855	.6428	.0000	.0222	.0000

*Applicable to z_0 heights only.

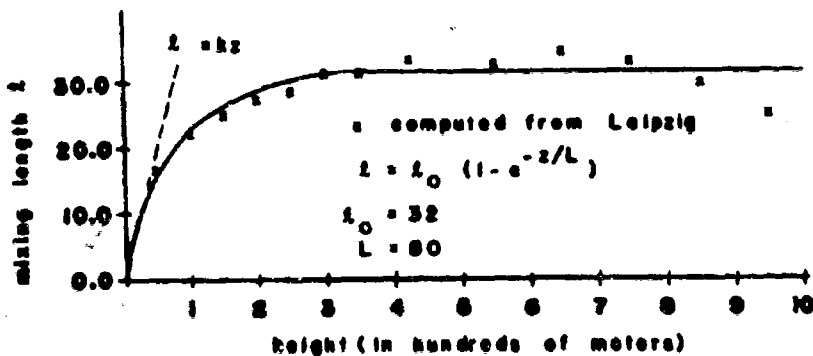


FIG. 1 Mixing Length Profile

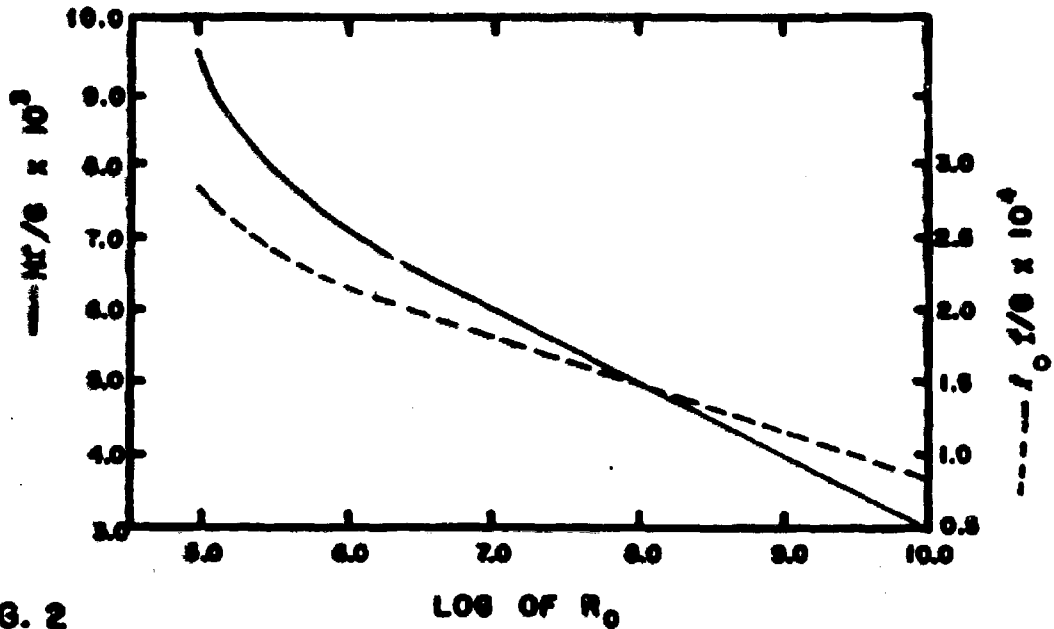


FIG. 2
 H_1/G AND L_0/G AS FUNCTIONS OF THE SURFACE ROSSBY NO.

———— H_1/G L_0/G - - - - -

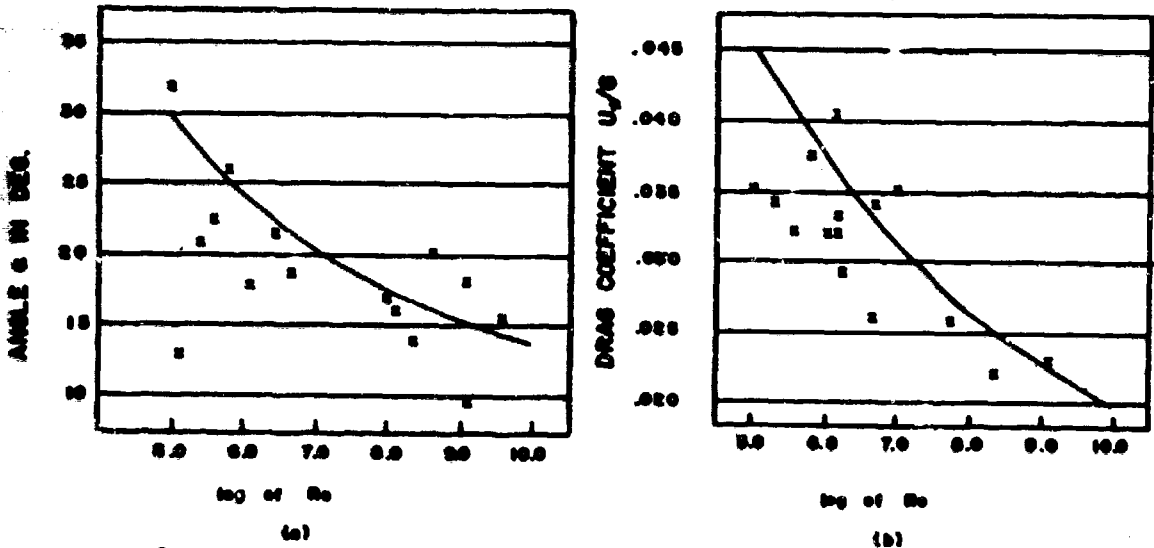


Fig 3
 MODEL RELATIONSHIPS VS OBSERVATIONS
 ——— MODEL x OBSERVATIONS

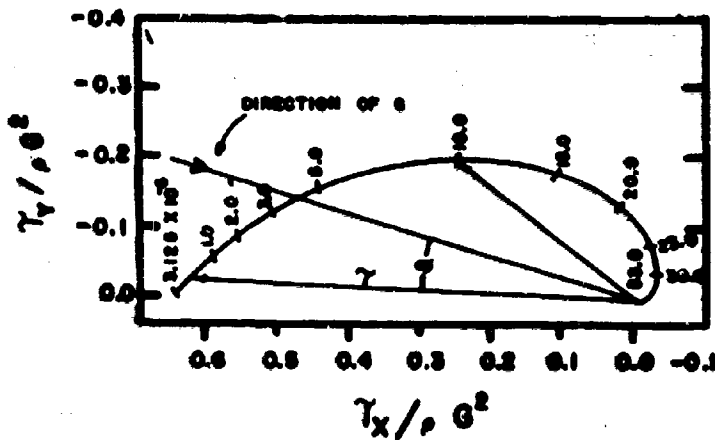
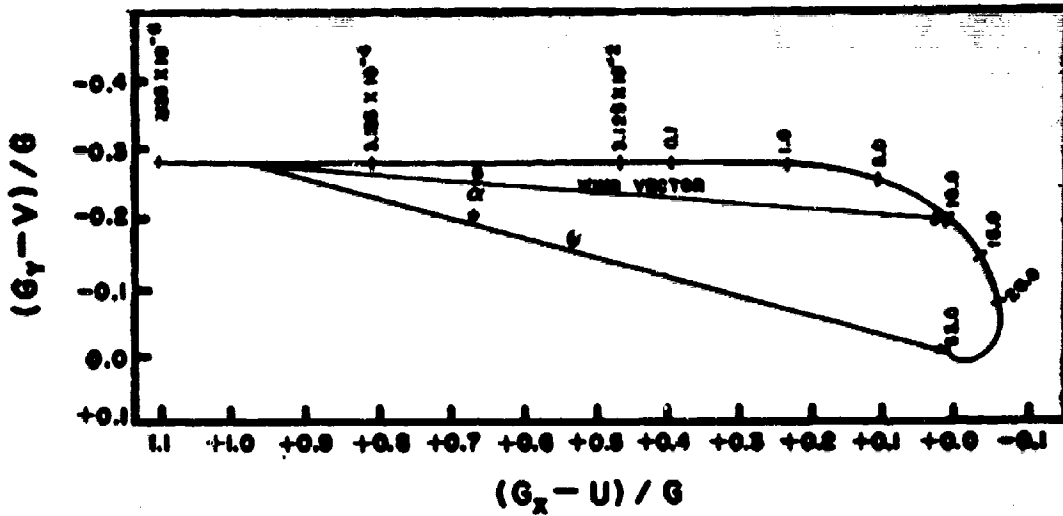
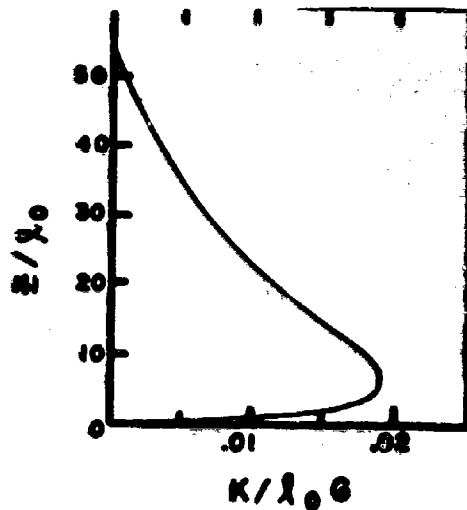


FIG. 4b SHEARING STRESS HODOGRAPH



DETERMINATION OF THE IONOSPHERIC ELECTRON
CONTENT UTILIZING SATELLITE SIGNALS

PAUL R. ARENDT, ARISTOLÉ PAPAYOANOU, and HAIM SOICHER
U.S. ARMY ELECTRONICS RESEARCH AND DEVELOPMENT LABORATORIES
FORT MONMOUTH, NEW JERSEY

INTRODUCTION

Ionospheric electron content may be determined from satellite signals by both active and passive methods: the former involve measurements by the satellite itself, the latter involve measurements by ground stations. This paper touches on both methods, limiting the discussion to topside ionospheric sounding, for the active case, and to c-w transmission from the satellite and ground reception, for the passive case. Active methods, which yield complete information about the altitude profile of the electron density, depend on specific information regarding the satellite's position and the geomagnetic field strength along the radio path. Under restricted circumstances, passive methods are independent of such specific information.

PASSIVE METHODS

The passive method has been used since the launching of the first artificial satellite. Application of this method is based on stable c-w satellite transmissions, the stability of which should be high compared with the measurement tolerances (1). When two or more frequencies are required, the frequencies should be phase-coherent. These conditions have been or will be met by various satellites: Sputnik III, the Navigational series, and the forthcoming S-66 are examples.

Passive methods involve determination of the Doppler frequency, the Faraday phase rotation, or both. The equations governing these phenomena are well known (2)(3)(4); however, it is useful to review their inter-relation under various conditions (Tables I and II). A comparison (Table II) indicates that none of these equations can be used alone for determination of the electron content. We must introduce either simultaneous data from different frequencies, data taken at different times, or hybrid methods using both effects. However, all these methods require additional information regarding the orbital

ARENDE, PAPAYOANOU, and
SOICHER

parameters and the geomagnetic field.

The various conditions for simplification of the general equations are indicated in Table II. Equations 1a and 1b describe the observable effects due to the motion of the source and to the nature and variation of the medium along the signal path. For ordinary ionospheric conditions, the medium may be assumed to be static during the time of observations. This assumption is permissible provided that ionospheric frequency scintillations (5)(6) in the range of lower frequencies (20 to 50 Mc) remain small. Under the static condition, we have the Doppler Equation 2a (Table II). An additional simplification is obtained for the Faraday equation in assuming that the geomagnetic component ($B \cos \Theta$) is constant over the entire ionospheric section of each particular ray path. This constant is equivalent to a value that in principle is different at each time moment, and is placed before the integral in Equation 3b and designated as "effective" value $\overline{B \cos \Theta}$. Note that both Equations 2a and 3b contain the local density N_p multiplied by the velocity component dS/dt and the integral

$\int \partial N / \partial t ds$. However, in the equation for the Faraday effect, the factor $B \cos \Theta$ is multiplied by N . Since Equations 2a and 3b cannot be solved for the electron content without special information about the inherent parameters, further simplifications have been introduced with the assumption of a straightline path of propagation in a stratified ionosphere. The corresponding coordinate transformations are:

for a circular orbit (7)(8)

$$\int_0^{S(p)} [\partial N / \partial t] ds = (N - N_p) dS/dt \quad (8)$$

for an elliptical orbit (9)

$$\int_0^{S(p)} [\partial N / \partial t] ds = (N - N_p) \{dS/dt - (S/h_0) dh_0/dt\} \quad (9)$$

These transformations result in Equations 4a, 4b, 5a, 5b, respectively (Table II). Assuming a flight vertically upward $dS/dt = dh_0/dt$, we arrive at Equations 6a and 6b which for a vertical flight in empty outer space ($N_p = 0$) reduce to Equations 7a and 7b. Further simplifications involve a restriction to small zenith angles in nearly overhead passages. If these conditions are assumed, and the observation point is located at medium geomagnetic latitude, we can neglect in Table II, during the center part of an overhead passage, all terms that are related to a variation of $\overline{B \cos \Theta}$. This is a sharp geometric restriction, however, as shown below, it has to be made independently and rigorously for the applicability of the differential Doppler methods. Thus, with $\partial(B \cos \Theta) / \partial t = 0$, we arrive at a simplified version (10) of Equation 4b:

$$d(\Delta a)/dt = (k/f^2) \overline{B \cos \Theta} dS/dt \quad (4b)$$

1. Assuming the first order of high-frequency approximation for the refractive index for frequencies larger than the plasma frequency

ARENDT, PAPAYOANOU, and
SOICHER

which together with Equation 4a results in

$$\bar{N} = f^3 / (c \Delta f K \overline{B \cos \Theta} [d(\Delta a)/dt]^{-1} \pm kf) \quad (10)$$

when a reasonable approximation for $\overline{B \cos \Theta}$ can be found. As a first approximation, the term kf can be neglected and Equation 10 becomes:

$$\bar{N} = f^3 d(\Delta a)/dt / c \Delta f K \overline{B \cos \Theta} \quad (10a)$$

i.e. the integrated electron density is proportional to the ratio of Faraday fading rate and Doppler shift. This method (Equation 10 or 10a) has the advantage that it involves only one frequency. On the other hand, if \bar{N} is known from different information, the above measurement gives an estimate for $\overline{B \cos \Theta}$ that may be used further.

Having made the assumption that $\partial(B \cos \Theta)/\partial t = 0$, we also transform Equation 1b into a simplified form:

$$d(\Delta a)/dt = (K/f^2) \overline{B \cos \Theta} d/dt \int_0^{S(p)} N ds \quad (1b)$$

where the coordinate transformation inherent in Equation 4b has not been applied. Combining Equation 2a and the above form of 1b, and eliminating $d/dt \int N ds$, we obtain (11):

$$d(\Delta a)/dt = (K/k) \overline{B \cos \Theta} [dS/dt - c \Delta f/f] \quad (11)$$

Thus, the Faraday fading rate is, other things being equal, proportional to the difference between the change of the geometric path length dS/dt and the change of the optical path length $c \Delta f/f$ (11).

Considering Equation 2a, we can solve for dS/dt when two observations at different frequencies (f_1 and f_2) are performed simultaneously:

$$dS/dt = c(f_2 \Delta f_2 - f_1 \Delta f_1) / (f_2^2 - f_1^2) \quad (12)$$

With Equations 11 and 12, we arrive at another expression for $\overline{B \cos \Theta}$.

$$\overline{B \cos \Theta} = [d(\Delta a)/dt] (k/cK) (f_2 - f_1 f_1/f_2) / (\Delta f_2 - \Delta f_1 f_2/f_1) \quad (13)$$

Considering two simultaneous Doppler measurements, we can also apply Equation 4a and solve for \bar{N} (7):

$$\bar{N} = (f^2/k) (\Delta f_2 - \Delta f_1 f_2/f_1) / (\Delta f_2 - \Delta f_1 f_1/f_2) \quad (14)$$

To obtain added confidence for the result, we differentiate Equation 4a:

$$d(\Delta f)/dt = f/c (1 - k\bar{N}/f^2) d^2S/dt^2 - (k/cf) dS/dt d\bar{N}/dt \quad (15)$$

The second term vanishes for the assumed small zenith angles when $dN/dt = 0$, or at Doppler inflection point where $dS/dt = 0$. Using the maximum slope σ at inflection moment (which data can be based on many single measurements), we have (Fig 1):

$$\bar{N} = (f^2/k) (\sigma_2 - \sigma_1 f_2/f_1) / (\sigma_2 - \sigma_1 f_1/f_2) \quad (16)$$

When the differential Doppler methods are applied, special precautions are necessary; the applicability depends on the equality of the pertinent geometry for the propagation of the two frequencies. Equation 2a (Table II) indicates that the velocity component dS/dt , and the shape and length of the line integral, must be identical. Experimental experience (12) has proved that these identities are not warranted. Even if it is assumed that the curvature of the line integral is close to a straight line for both frequencies, we notice (Equation 2a) the effect from dS/dt , which is very sensitive with respect to small changes of the direction of the ray directly at the satellite. Changes in N_p (horizontal gradients) may produce such

directional defects that affect the basic term defining the momentaneous Doppler frequency shift. All other types of ionospheric irregularities can produce similar defects. Obviously, the time function of the differential Doppler shift $\Delta f_2 - \Delta f_1$, f_2/f_1 , when applied to identical geometry, should have its inflection point at the same time at which inflection occurs for Δf_1 and Δf_2 ; otherwise, the assumption of a straightline path would not be justified. Thus, an important test is the precise calculation of the time moment of Doppler inflection for each single frequency from a large number of subsequent Doppler measurements. When these time moments for the two frequencies differ by an interval larger than the precision of the calculation (about 30 msec), we must assume divergent geometry and cannot apply Equations 13, 14, 16, or their equivalents. Doppler inflection measurements show that a large number of differential Doppler data do not correspond to such restrictions. This basic requirement is realized only for close-to-overhead passages during stable ionospheric conditions (small scintillations) for radio path penetration nearly vertical to stratified layers. Thus, conclusions from differential Doppler data that have not been checked for equal geometry may lead to substantial errors¹ (compare Figs 2 and 3).

The restriction for the selection of passages to which differential Doppler methods can be applied leads directly to the assumption applied above for the simplification of the Faraday Equations 1b and 4b. Within the boundaries of this restriction, we can combine Equation 13 with the following integrated version of 1b:

-
1. For the same reason, the phase-defect method (4) should be restricted to nearly overhead passages and small zenith angles.

BLANK PAGE

$$\Delta a = (\Delta a)_{\text{observed}} + 2\pi(m+\epsilon) = (K/f^2) \overline{B \cos \Phi} \int_0^{S(p)} N ds \quad (17)$$

and obtain (11) the electron content integrated along the radio path as follows:

$$\int_0^{S(p)} N ds = (c/k) \Delta a_1 [d(\Delta a)/dt]^{-1} f_1^2 (\Delta f_2 - \Delta f_1 f_2/f_1) / (\Delta f_2 - \Delta f_1 f_1/f_2) \quad (18)$$

The uncertainty in the total number of phase rotations, Equation 17, $[2\pi(m+\epsilon)]$ at any moment is resolved (13) by observing the change of the phase between that moment and the moment when the Faraday effect is zero. The same result is obtained by forming the difference $\Delta a_1 - \Delta a_2$ from simultaneous phase measurements at two different but adjacent frequencies. Insofar as the difference remains greater than .2 radians, we can neglect the experimental error ϵ , which is in the order of .1 radians if recent phase-following systems (14) are applied.

In case the moment for $\Delta a = 0$ cannot be observed, we use $\Delta a_1 - \Delta a_2$ from two closely spaced frequencies, and arrive at:

$$\int_0^{S(p)} N ds = (c/k) (\Delta a_1 - \Delta a_2) [d(\Delta a)/dt]^{-1} [f_1^2 f_2^2 / (f_2^2 - f_1^2)] (\Delta f_2 - \Delta f_1 f_2/f_1) / (\Delta f_1 - \Delta f_2 f_1/f_2) \quad (19)$$

Note, three frequencies (f_1, f_2, f_3) are used, and the assumption $m_1 = m_2$ has been made, f_1 and f_3 being sufficiently close.

The apparent symmetry of the Doppler curves with respect to the inflection point has led to a method that is based on the statement that the absolute value of Δf is always equal at corresponding times before and after inflection moment. E. Golton and B. Burgess (15) concluded that at those moments \bar{N} is the same. However, Equation 19, indicates that this fact is justified only when dS/dt is also equal. This assumption holds only for a polar orbit in overhead observation when we can neglect the Doppler component due to the relative motion of the earth. Then, the difference of corresponding Faraday phase shifts at time t_1 and t_2 becomes:

$$(\Delta a)_{t_1} - (\Delta a)_{t_2} = (\overline{B \cos \Phi}_{t_1} - \overline{B \cos \Phi}_{t_2}) K/f^2 \int_0^{S(p)} N ds \quad (20)$$

The application of Equation 13 yields the integrated density:

$$\int_0^{S(p)} N ds = c/k [(\Delta a)_{t_1} - (\Delta a)_{t_2}] f^2 [\Delta f_2 - \Delta f_1 f_2/f_1] / [d(\Delta a)_{t_1}/dt - d(\Delta a)_{t_2}/dt] (\Delta f_2 - \Delta f_1 f_1/f_2) \quad (21)$$

From the assumed straight-line propagation it follows that both Δf_1 and Δf_2 must be equal at times t_1 and t_2 . A check of this equality is mandatory for the application of this method.

1. A method (2) replacing $(\Delta a_1 - \Delta a_2)$ by $d(\Delta a)/df$ cannot achieve a different result.

ARENDE, PAPAYOANOU, and
SOICHER

Equation 18, 19, and 21 are independent of data acquisition pertaining to the orbital parameters and the effective component of the geomagnetic field. They are bound to the prerequisites of geometric identity and of constant field component (i.e. of small zenith angles).

In addition, other experimental conditions must be considered. For example, when the phase measurements are performed by counting the rate of amplitude fadings, the measurement of $d(\Delta a)/dt$ becomes an average over a time interval that is different in real time for the two frequencies. Thus, continuous but moment-by-moment measurements are preferable. The new phase-follower system (14) yields Δa and $d(\Delta a)/dt$ every second from phase comparisons that are integrated over only a small fraction of a second (Fig 4). For very close frequencies, the differential Doppler shift becomes so small that it is submerged in the uncertainties produced by ionospheric frequency scintillations and by experimental errors.

The question arises whether some advantage could be gained by the introduction of a differential phase-rotation method, similar to the two differential Doppler methods mentioned. Unfortunately, such efforts are handicapped by the fact that all simultaneous $d(\Delta a)/dt$ are proportional to f^2 . Thus simultaneous rate data always end in the trivial equation:

$$f_1^2 d(\Delta a_1)/dt = f_2^2 d(\Delta a_2)/dt \quad (22)$$

One can use Equation 22 only to check the prerequisite of equality of path geometry. It has been found, for example, that with 20 and 40 Mc (Sputnik III), Equation 22 is not strictly satisfied for most observations (Fig 5). This fact¹ confirms strong refraction and bending effects, and indicates the care one must exercise in the data interpretation.

The various passive methods are revised in Table III; a comparison of actual data is given in Table IV. The application of Equations 18, 19, and 21 must be postponed until data becomes available from Satellite S-66. Equation 21 deserves restrictive consideration because the S-66 orbit will not be polar.

ACTIVE METHODS

Experience with active methods is limited thus far to observations of the US-Canadian Topside Sounder Alouette (S-27), which

1. The effects of so-called second-order rate of rotation (16) might be responsible at least for a part of those deviations; however, there is no experimental method to differentiate between second-order rate effects and path inequalities; further, different paths for the ordinary or the extraordinary mode have to be considered.

ARENOT, PAPAYOANOU, and
SOICHER

will be followed by the US Topside Sounder S-48 and by various other vehicles. Active methods are based on the following procedure. The satellite transmits a pulse at a sounding frequency (in the range of .5 to 11.5 Mc) that is considered to be vertically incident to the ionosphere, and receives the echo. The frequency is varied either continuously (17) (18) (19) S-27, or in steps (20) S-48. The sounding information (ionograms) is transmitted to the ground via a telemetry channel. The appearance of the ionograms (Fig 6) suggests a nearly exponential decay (21) of the height distribution of electron density. In order to convert the measured delay times (or virtual paths) into electron densities at true height, the virtual path must be expressed in terms of the group refractive index. Since the sounding frequencies are close to the gyrofrequency and the plasma frequencies, the complete expression, and not an approximation, for the refractive index that accounts for the geomagnetic field and its variation must be applied. A value for the gyrofrequency and the plasma frequency at the satellite itself must be obtained from the spikes (22) visible in the ionograms. Finally, a numerical integration over several zones of different exponential decay allows one to arrive at the electron-density profile (see appendix).

The electron dependence near the F₂ layer of the ionosphere can be approximated by a parabolic function. If this section of the profile is evaluated together with the exponential regions, one obtains a detailed picture of the electron-density distribution at and above the F₂ maximum. If bottomside ionograms obtained simultaneously with topside ionograms are reduced to true-height profiles, then the total electron-density distribution from the earth to the satellite is known. Fig 7 shows examples of such curves revealing the total structure of the electron content below the satellite. It is obvious that a numerical integration of the entire profile will result in a figure for the integrated electron content that can be compared with corresponding figures from passive methods. Table V shows typical values for day-time integration that fit the data from the differential Doppler slope method shown in Fig 1. It will be interesting to make direct comparisons of the active and passive methods when the S-27 and S-66 Satellites can be observed simultaneously.

APPENDIX

The assumption that the ionosphere above the F₂ layer is describable by an exponential function leads to the following equation for the electron density distribution:

$$N(z) = N_{m1} e^{-\alpha(z-h)} \quad 0 \leq z \leq h \quad (1.1)$$

-
1. Different from the usage of the phase-velocity refractive index for the passive experiments.

ARENDR, PAPAYOANOL, and
SOICHER

where z is measured vertically downward from the satellite. The basic equation relating the virtual path h' , ($h' = 1/2 \sigma_T$) to the true height of reflection h_T , is given by:

$$h' = \int_0^{h_T} n'(z) dz \quad (1.2)$$

If one introduces $X = k'N(z)/f^2$, then for an exponential function $N(z)$, Equation 1.2 becomes:

$$h' = (1/\alpha) \int_{X_0}^{X_1} [n'(X)/X] dX \quad (1.3)$$

where X_0 is just

$$k'N_{m1} e^{-\alpha h_0}/f^2$$

at the satellite, and X_1 is obtained from the reflection condition. The extraordinary ray has been used to perform the calculations and so

$$X_1 = 1 - f_H/f \quad (1.4)$$

Thus from Equation 1.3, one can calculate α . The spikes on the ionogram allow a calculation of X_0 and consequently of h_0 . These spikes correspond to plasma resonances (22).

The problem becomes more complicated when a single α does not suffice to describe the electron-density distribution. In this case, one assumes that the topside ionosphere is composed of various regions each of which has a characteristic α . A computer program has been devised to find the various alphas and the depth of the region over which each is applicable (21). If one has the appropriate α 's and h 's for the ionosphere between $f F_2$ and the satellite, one may obtain the integrated electron content^o for this region from:

$$\int Ndh = \sum_{\lambda=1}^m [N_{m\lambda} / \alpha^{(\lambda)}] [1 - e^{-\alpha^{(\lambda)} (h_\lambda - h_{\lambda-1})}] \quad (1.5)$$

where $\alpha^{(\lambda)}$ is the suitable α for the λ th region, h_λ is the depth of this λ th region measured from the satellite, $h_{\lambda-1}$ is the depth of the $(\lambda-1)$ region measured from the satellite, and $N_{m\lambda}$ is the maximum electron density of the λ th region.

REFERENCES

1. P.R. Arendt, "Actual Measurement of the Aging of Frequency Oscillators in Outer Space"; Shock, Vibration and Associated Environments Bulletin, No. 30, March 1962, p 55 (Classified).
2. G.W. Swenson, Jr., "The Utilization of Ionosphere Beacon Satellites, Goddard Space Flight Center, X-250-62-32, presented Cospar Meeting, 1962.

ARENDR, PAPAIOANOU, and
SOICHER

3. B. Burgess, "Techniques for Analysis of Faraday Data," NATO Advanced Study Institute, Radio Astronomical and Satellite Studies of the Atmosphere, 1963 (North-Holland Publ., editor J. Aarons); p 313.
4. O.K. Garriot and R.N. Bracewell, "Satellite Studies of the Ionization in Space by Radio," Advances in Geophysics, New York Academic Press, 1961, Vol 8, p 85.
5. H.P. Hutchinson and P.R. Arendt, "Ionospheric Scintillations of Satellite Signals, Proc. 8th Intern. Astronautical Congress, 1959, p 402.
6. P.R. Arendt, "Frequency Scintillation of Satellite Signals before and after the Argus Experiments," IRE Trans. Ant. and Prop., AP-8, No. 1, p 73, 1960.
7. Ya. L. Al'Pert, "On the Method of Investigating the Ionosphere by Using an Artificial Satellite of the Earth," Usp. Fiz. Nauk, 64, p 3, 1958.
8. S.A. Bowhill, "The Faraday-Rotation Rate of a Satellite Radio Signal," Journ. Atmos. Terrest. Phys., 13, 1/2, p 175, 1958.
9. F.B. Daniels and S.J. Bauer, "The Ionospheric Faraday Effect and its Applications," Journ. Frankl. Inst., 267, No. 3, p 187, March 1959.
10. P.R. Arendt, "Measurement of Ionospheric Electron Content with a Single Satellite Frequency," Proc. IEEE, 51, No. 2, p 363, Feb. 1963.
11. P.R. Arendt, "Measurement of Ionospheric Electron Content Regardless of Approximations, Nature, 197, No. 4867, p 579, Feb 1963.
12. P.R. Arendt and H. Soicher, "Diurnal Variation of Ionospheric Electron Content Obtained from Doppler Slopes of Simultaneous Satellite Signals," 1964 IEEE Convention, Session 62.
13. O.K. Garriott, "The Determination of Ionospheric Electron Content and Distribution from Satellite Observations," Journ. Geophys., 65, No. 4, p 1139, 1960.
14. G. Vogt, "An Analog Polarization Follower for Measuring the Faraday Rotation of Satellite Signals," The British Institute of Radio Engineers, Joint Symposium on Signal Processing in Radar and Sonar Directional Systems, to be presented in July 1964.
15. B. Burgess, "Ionospheric Studies Using Satellite Radio Transmissions," Electron Density Profiles, 1962 (Mc Millan Co., editor B. Maehlum), p 224.

ARENDE, PAPAYOANOU, and
SOICHER

16. K.C. Yeh, "Second Order Faraday Rotation Formulas," Journ. Geophys. Res., 65, No. 8, p 2548, Aug 1960.
17. J.H. Chapman, "Alouette the Topside Sounder Satellite," URSI, Ottawa, Oct 1963.
18. J.E. Jackson, "The Alouette Satellite," Goddard Space Flight Center Scientific Symposium, Mar 1963.
19. E.S. Warren, "Some Preliminary Results of Sounding of the Topside of the Ionosphere by Radio Pulses from a Satellite, Nature 197, No. 4868, p 636, Feb 1963.
20. "Fixed Frequency Topside Sounder S-48," NASA, N-90, 014, Nov 1963
21. P.R. Arendt and A. Papayouanou, "Conversion of Topside Ionograms to Exponential Electron-Density Profiles," USAELRDL Technical Report 2424, Feb 1964, presented URSI Spring Meeting 1963.
22. W. Calvert and G.B. Goe, "Plasma Resonances in the Upper Ionosphere," Journ. Geophys. Res. 68, No. 22, p 6113, Nov 1963.

TABLE I - List of Symbols

Δ	phase shift between ordinary and extraordinary mode	n	refractive index (phase velocity)
B	geomagnetic field strength	n'	group refractive index
c	velocity of light, free space	N	electron number density
d_s	line element of path S	N_p	electron density at satellite
f	emitted frequency	\bar{N}	average integrated electron density
f_h	gyrofrequency	C	observer's position
Δf	Doppler shift	p	satellite's position
h_o	altitude of satellite, measured upward	$S, S(p)$	geometric length of radio path
h^i	virtual path, measured downward	t	time
h_T	true light	z	path of sounder pulse, measured downward
dh	line element of altitude	α	exponent of profile function
$k = 4.03 \times 10^7$	cgs units	Θ	angle between B and propagation
$k' = 8.065 \times 10^7$	cgs units	ϵ	experimental error
$K = 2.36 \times 10^4$	cgs units	σ	maximum slope
m	integer		
τ	delay time		

	DOPLER	FARADAY
GENERAL EQUATION	$\Delta f = \frac{1}{2} \frac{d}{dt} \int_0^{2\pi} \omega dt$	$\frac{d(\Delta\omega)}{dt} = \frac{1}{2} \frac{d}{dt} \int_0^{2\pi} \omega \cos \phi dt$
STATIONARY MEDIUM, NO OTHER RESTRICTIONS $f = f_{\text{PLASMA}}$	$\Delta f = \frac{1}{2} \left(\frac{d}{dt} \int_0^{2\pi} \omega dt - \frac{1}{c} \int_0^{2\pi} \omega v dt \right)$	
AS ABOVE EFFECTIVE GEOMAGNETIC FIELD CORRECT FOR EACH PATH		$\frac{d(\Delta\omega)}{dt} = \frac{1}{2} \left(\frac{d}{dt} \int_0^{2\pi} \omega dt - \frac{1}{c} \int_0^{2\pi} \omega v dt \right)$
AS ABOVE STRAIGHT LINE PROPAGATION CIRCULAR ORBIT	$\Delta f = \frac{1}{2} \left(\frac{d}{dt} \int_0^{2\pi} \omega dt \right)$	$\frac{d(\Delta\omega)}{dt} = \frac{1}{2} \left(\frac{d}{dt} \int_0^{2\pi} \omega dt \right) - \frac{1}{c} \int_0^{2\pi} \omega v dt$
STRAIGHT LINE PROPAGATION ELLIPTICAL ORBIT	$\Delta f = \frac{1}{2} \left(\frac{d}{dt} \int_0^{2\pi} \omega dt \right) + \frac{1}{c} \int_0^{2\pi} \omega v dt$	$\frac{d(\Delta\omega)}{dt} = \frac{1}{2} \left(\frac{d}{dt} \int_0^{2\pi} \omega dt \right) - \frac{1}{c} \int_0^{2\pi} \omega v dt$
AS ABOVE ROCKET IN VERTICAL FLIGHT $\omega = \omega_0 \left(1 - \frac{2h}{R} \right)$	$\Delta f = \frac{1}{2} \left(\frac{d}{dt} \int_0^{2\pi} \omega dt \right)$	$\frac{d(\Delta\omega)}{dt} = \frac{1}{2} \left(\frac{d}{dt} \int_0^{2\pi} \omega dt \right) - \frac{1}{c} \int_0^{2\pi} \omega v dt$
AS ABOVE ROCKET IN FREE SPACE $\omega_0 = 0$	$\Delta f = \frac{1}{2} \frac{d\omega}{dt}$	$\frac{d(\Delta\omega)}{dt} = 0$

TABLE II

EFFECTS FROM MOTION OF SATELLITE TRANSMITTER

EQUATION	FOR SMALL ZENITH ANGLES ONLY	ADDITIONALLY REQUESTED
10	$\int N dh = \frac{N_0 r^2}{c \Delta t \cos \phi \left(\frac{d(\Delta\omega)}{dt} \right)^{-1} \kappa + \Delta t}$	$N_0 \cos \phi$
14	$\int N dh = \frac{N_0 r^2}{h} \frac{\Delta t_2 - \Delta t_1 \frac{v_2}{v_1}}{\Delta t_2 - \Delta t_1 \frac{v_1}{v_2}}$	N_0
16	$\int N dh = \frac{N_0 r^2}{h} \frac{\sigma_2 - \sigma_1 \frac{v_2}{v_1}}{\sigma_2 - \sigma_1 \frac{v_1}{v_2}}$	N_0
18	$\int N dh = \frac{c r^2}{h} \Delta \omega_1 \left(\frac{d(\Delta\omega_1)}{dt} \right)^{-1} \frac{\Delta t_2 - \Delta t_1 \frac{v_2}{v_1}}{\Delta t_2 - \Delta t_1 \frac{v_1}{v_2}}$	—
19	$\int N dh = \frac{c}{h} \frac{v_1^2 v_2^2}{v_2^2 v_1^2} (\Delta \omega_1 - \Delta \omega_2) \left(\frac{d(\Delta\omega)}{dt} \right)^{-1} \frac{\Delta t_2 - \Delta t_1 \frac{v_2}{v_1}}{\Delta t_2 - \Delta t_1 \frac{v_1}{v_2}}$	—
22	$\int N dh = \frac{c r^2}{h} (\Delta \omega_{11} - \Delta \omega_{12}) \left(\frac{d(\Delta\omega_{11})}{dt} - \frac{d(\Delta\omega_{12})}{dt} \right)^{-1} \frac{\Delta t_2 - \Delta t_1 \frac{v_2}{v_1}}{\Delta t_2 - \Delta t_1 \frac{v_1}{v_2}}$	POLAR ORBIT

TABLE III

METHODS FOR THE DETERMINATION THE IONOSPHERIC ELECTRON CONTENT

ARENDE, PAPAYOANOU, and
SOICHER

CALCULATION OF INTEGRATED ELECTRON CONTENT

SATELLITE 1960 EYA (TRANSIT), ORBIT 2370, DEC. 6, 1960

MEASURED DATA $\Delta f_{94} = 763$ cy/sec, $\Delta f_{102} = 2273$ cy/sec } AT
 $\frac{d(\Delta f)}{dt} = \frac{\pi}{4.0}$ sec⁻¹ } 08:17:26
 EAT
 $\nu_{94} = 12.092$ cy/sec² } AT
 $\nu_{102} = 36.807$ cy/sec² } 08:16.08.6
 EAT
 OTHER DATA $h_p = 680$ km (FROM EPHEMERIS)
 $B \cos \phi = .454$ GAUSS (PER EQUATION 13)

TABLE IV

INTEGRATED CONTENT ELECTRONS/CM ²	METHOD	EQUATION
3.27×10^{15}	TWO DOPPLER SHIFTS	(14)
3.34×10^{15}	TWO DOPPLER SLOPES	(18)
3.32×10^{15}	ONE DOPPLER SHIFT, ONE FARADAY FADING RATE	(10)

ELECTRON CONTENT (10^{13} e/cm²)

ACTIVE METHOD		PASSIVE METHOD	
TIME EST	CONTENT	TIME EST	CONTENT
AUGUST 1963 { 15:30 17:00	2.10 2.75	JULY-SEPT 1961 { 15:00 17:00	2.1 1.9
NOV. 1963 16:00	2.0*	NOV. 1961 16:00	2.1

TABLE V

* CORRECTED FOR VARIATION OF SUNSPOT NUMBERS

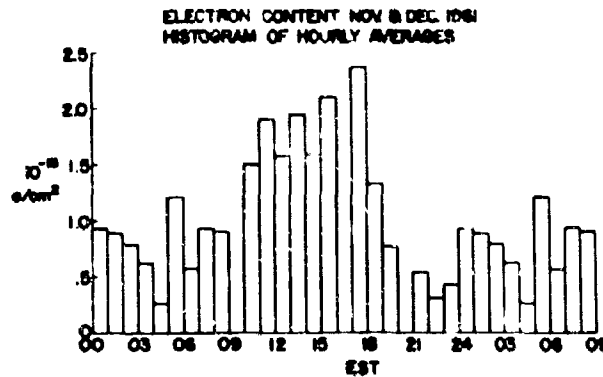


FIGURE 1. Electron Content Produced from Differential Doppler Slopes

ARENDET, PAPAYOANOU, and
SOICHER

1961 OMICRON ORBIT NO. 192

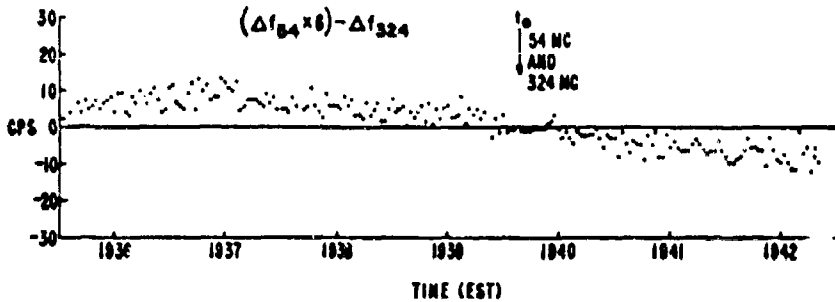


FIGURE 2. Differential Doppler Shift in Identical Geometry; note the simultaneity of the inflection moments

61 OMICRON ORBIT 205

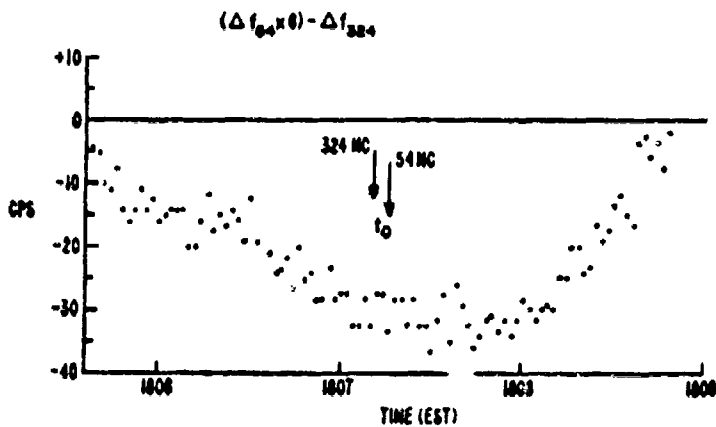


FIGURE 3. Differential Doppler Shift in Non-identical Geometry; the two inflection moments are 5 seconds apart.

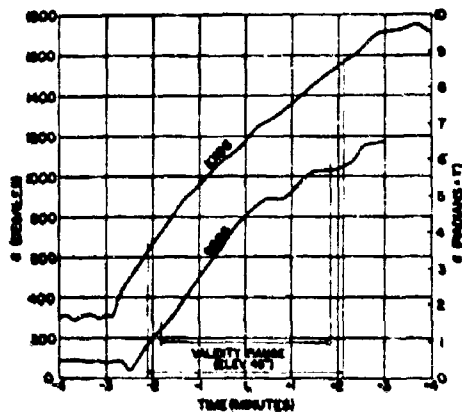


FIGURE 4. Δa Measured with Phase Follower System; after Vogt

ARENDT, PAPAIOANOU, and
SOICHER

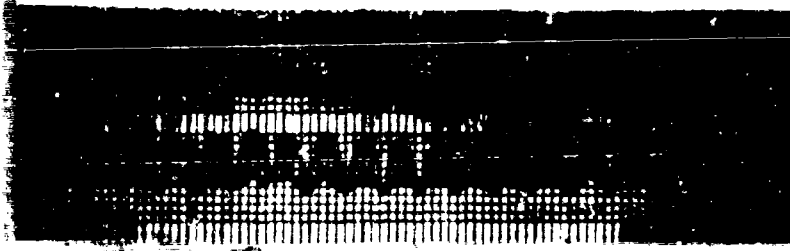


FIGURE 5. Faraday Fading on 20 and 40 Mc from Sputnik III



FIGURE 6. Topside Ionogram from Alouette (S-27) obtained at Fort Monmouth

ELECTRON PROFILE FROM S-27 AND BOTTOM SOUNDER

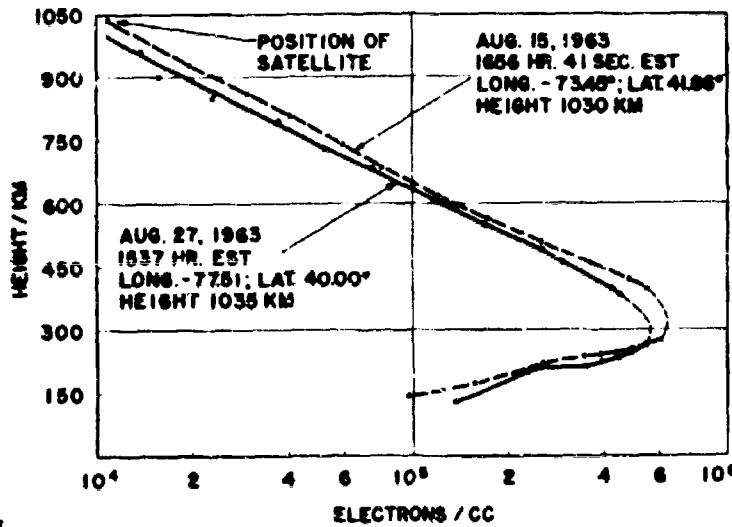


FIGURE 7. Complete Electron Profiles above Fort Monmouth, N.J. from Topside and Bottomside Ionograms

POTENTIAL RADIATION-PROTECTIVE COMPOUNDS.
 SYNTHESIS OF THE THREE ISOMERIC THREE-CARBON
 AMINOHYDROXY BUNTE SALTS AND RELATED COMPOUNDS ¹

2

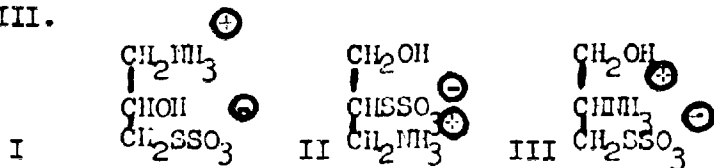
D.H.BALL, J.M. WILLIAMS, and L.LONG, Jr.
 PIONEERING RESEARCH DIVISION
 U.S.ARMY NATICK LABORATORIES
 NATICK, MASSACHUSETTS

3

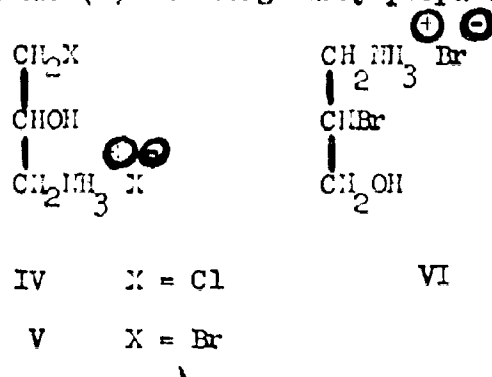
In recent years, Bunte salts, S-alkyl thiosulfates, have been the subject of investigation in many laboratories. A Bunte salt related to glutathione, "S-sulfo γ -glutathione", has been isolated from calf lens extracts.⁴ Sulfite has been used to obtain soluble protein or peptide fractions from wool,⁵ flour,⁶ ribonuclease and insulin,⁷ and from trypsinogen and α -chymotrypsinogen;⁸ the cystine disulfide bonds are cleaved under mild conditions to form "S-sulfocysteinyl residues". The properties and potential uses of the "S-sulfocysteinyl residues" have been discussed by Swan.⁹

Several examples of Bunte salts containing amino or alkyl-amino groups were prepared by Bretschneider.¹⁰ These, like the amino-acids, are internal salts and the simplest compound of this type, S-2-aminoethyl thiosulfuric acid, was found to have significant radiation-protective activity in mice. It was also 2.4 times less toxic than 2-aminoethanethiol (cysteamine) hydrochloride.¹¹ Other aminoalkyl thiosulfates have been prepared by Rosenthal and Citarel¹² who found that they were stable compounds which possessed significant anti-radiation activity. The low activity of 3-amino-1-propanethiol compared to the activity of S-3-aminopropyl thiosulfuric acid¹³ suggests that the protective activity of the Bunte salt is not, at least in this case, due to the formation of the thiol.

Aminoalkyl thiosulfuric acids are, in general, stable, odorless, crystalline, water-soluble substances and are thus attractive potential anti-radiation drugs. This paper describes the synthesis of the three isomeric internal Bunte salts I, II and III.



Bunte salts are conveniently prepared by the reaction of alkyl halides with thallos thiosulfate,¹⁴ the insoluble thallos halide formed being easily removed from the reaction mixture. An aqueous solution of 1-amino-3-chloro-2-propanol hydrochloride (IV),^{15,16} unlike 2-chloroethylamine hydrochloride,¹⁴ did not react readily with thallos thiosulfate at room temperature. Formation of the internal Bunte salt took place at higher temperatures but extensive decomposition also occurred. The corresponding bromo hydrobromide (V) was originally prepared by condensation of

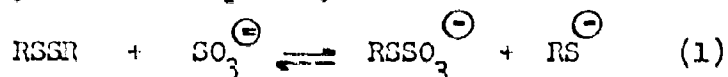


potassium phthalimide with boiling epibromohydrin, followed by hydrolysis of the product with hydrobromic acid.¹⁷ When N,N -dimethylformamide was added in the condensation step,¹⁸ the reaction proceeded at 35-40° and the yield was improved slightly. Under these conditions, the epoxide ring is slowly opened by the potassium phthalimide. This was demonstrated by the formation of a small amount of a diamino-propanol (presumably 1,3-diamino-2-propanol) after hydrolysis of the condensation product. The bromo hydrobromide reacted more readily than IV with thallos thiosulfate and, after 2 days at 50°, the Bunte salt I, S-3-amino-2-hydroxypropyl thiosulfuric acid, was isolated in 67% yield.

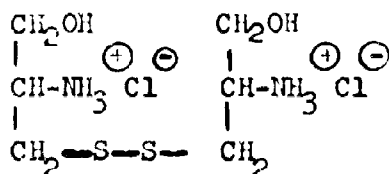
No suitable precursor for the Bunte salt II was known and attempts to prepare 3-amino-2-bromo-1-propanol hydrobromide (VI) from 2,3-dibromo-1-propanol and from 2,3-dibromopropylamine and their derivatives, by selective displacements of the primary bromine atoms, were unsuccessful. The hydrobromide (VI) was eventually obtained from the reaction of the allylammonium ion with hypobromous acid. The latter was conveniently generated *in situ* by the method of Leibman and Fellner¹⁹ whereby a mixture of bromine vapor and air is passed into the reaction solution which contains one equivalent of silver nitrate. Both isomeric aminobromo alcohols were isolated, the desired 3-amino-2-bromo isomer VI being the major product. The hydrobromide VI did not react with thallos thiosulfate at 40° but, at 75-80°, conversion to the Bunte salt II was essentially complete after 12 hours. Amine impurities, which hindered crystallization, were removed by chromatography on silica gel. Crystallization of the purified material from methanol-ether gave the pure Bunte salt II, S-2-amino-1-(hydroxymethyl)ethyl

thiosulfuric acid in 51% yield.

Bunte salts have also been prepared by the action of sulfites on disulfides in the presence of an oxidizing agent. Air or oxygen are often sufficient but cupric ions,²⁰ iodosobenzoate and tetrathionate⁷ have also been used. The oxidant converts the thiol formed in the reversible reaction (1) to disulfide and eventually this is completely converted to the Bunte salt.



Cystinol dihydrochloride (VII), 3,3'-dithiobis [2-amino-1-propanol] dihydrochloride, is therefore a suitable precursor for the internal Bunte salt, III. It was prepared from 2-phenyl-2-thiazoline-4-methanol by Crawhall *et al.*²¹ These authors designated their product as the DL-form but the method of preparation would give

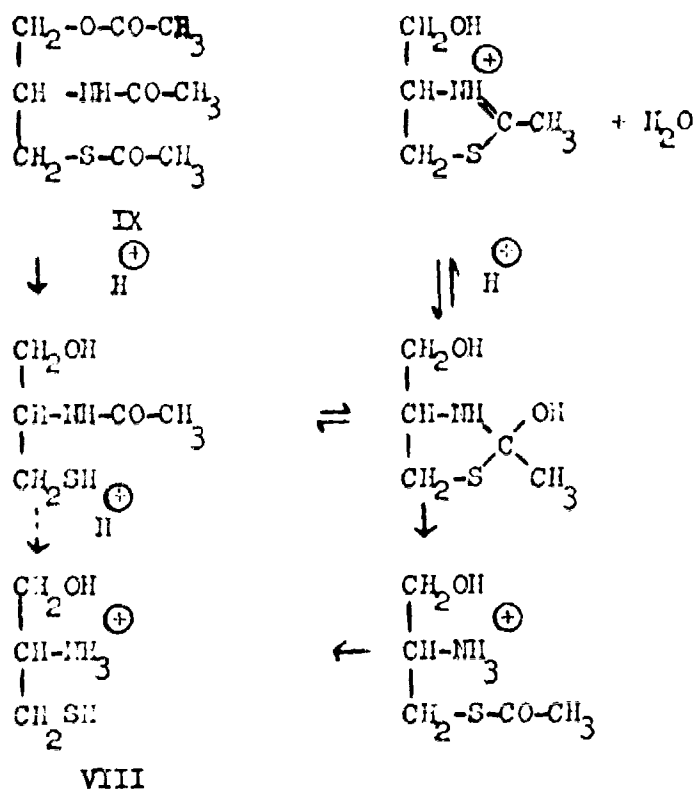


VII

a mixture of DL- and meso-isomers. We used the L-isomer of the thiazoline in an attempt to obtain the pure L-isomer of VII but hydrolysis of the thiazoline ring required much more vigorous conditions than those reported²¹ and racemization occurred at this stage. The resultant DL-thiol ("cystinol") was oxidized without isolation and a crystalline, optically inactive mixture of, presumably, the DL- and meso-forms of VII was obtained. No attempt was made to separate these isomers and the mixture was treated at room temperature with ammonium sulfite solution (pH 7) and oxygen. The reaction was followed by paper electrophoresis and was complete after 6-7 hours. The product was freed from inorganic salts by fractionation on a column of a cation exchange resin in the lithium salt form²² and subsequent recrystallization from methanol-ether gave S-2-amino-3-hydroxypropyl thiosulfuric acid (III) in 65% yield.

L-Cystinol, L-2-amino-3-mercapto-1-propanol (VIII), has recently been prepared by Enz and Cecchinato²³ by a method which should not cause appreciable racemization. L-Cysteine ethyl ester was reduced with lithium aluminum hydride and the L-cystinol formed was isolated as the H,O,S-triacetate (IX) in 29% yield. These authors hydrolyzed the triacetate with dilute hydrochloric acid and obtained L-cystinol as the crystalline hydrochloride, although they

did not record the specific rotation. Thin layer chromatography indicated that during hydrolysis of the triacetate, an intermediate was rapidly formed and that this was slowly converted to cysteinol. In dilute hydrochloric acid at 90°, the absorption at 231.5 mμ (thiolacetate) decreased rapidly. A second absorption at 261 mμ (thiazolinium cation) appeared, increased to a maximum after ca 1½ hours and then decreased slowly (FIG.1). It has been shown recently that thiazoline formation occurs in acid solutions of N-acetyl cysteine²⁴ and N-2-mercaptoethyl-acetamide²⁵ and hydrolysis of cysteinol triacetate (IX) probably proceeds via the same mechanism.



L-Cystinol dihydrochloride was prepared from L-cysteinol triacetate by acid hydrolysis followed by oxidation of the resultant thiol which was not isolated. The L-form of VII had a specific rotation of -108° and a much lower melting point than the values obtained for the mixtures of DL- and meso- forms.

The internal Dunte salt (III) prepared from this disulfide had a higher melting-point than the DL-form and a specific rotation of -31° in water.

The three isomeric internal Dunte salts are stable crystalline solids, very soluble in water and slightly soluble in methanol. They decompose slowly in boiling water. All three compounds showed slight activity when tested as radiation-protective

agents by the Walter Reed Army Institute of Research.

Footnotes

- (1) This work was supported by a grant from the Surgeon General's Office, Medical Research and Development Command, U.S. Army.
- (2) Present address: Department of Chemistry, Bristol University, Bristol, England.
- (3) H. Bunte, Ber., 7, 646 (1874).
- (4) S. G. Waley, Biochem. J., 71, 132 (1959).
- (5) J. M. Swan, Nature, 180, 643 (1957); Australian J. Chem., 14, 69 (1961).
- (6) E. E. McDermott and J. Pace, Nature, 184, 546 (1959).
- (7) J. L. Bailey and R. D. Cole, J. Biol. Chem., 234, 1733 (1959).
- (8) J. F. Pechère, G. H. Dixon, R. H. Maybury, and R. Neurath, J. Biol. Chem., 233, 1364 (1958).
- (9) J. M. Swan, "Sulfur in Proteins," R. Benesch et al., ed., Academic Press, New York, 1959, p. 3.
- (10) H. Bretschneider, Monatsh. Chem., 81, 372 (1950).
- (11) B. Holmberg and B. Sorbo, Nature, 183, 832 (1959).
- (12) H. A. Rosenthal and L. Citarel, 141st Meeting of the American Chemical Society, Washington, D.C., 1962; Abstracts of Papers, 29H.
- (13) A. Maluszynier, P. Czerniak, and E. D. Bergmann, Radiation Research 14, 23 (1961).
- (14) H. Z. Lecher and E. M. Hardy, J. Org. Chem., 20, 475 (1955).
- (15) S. Gabriel and H. Ohle, Ber., 50, 819 (1917).
- (16) Unless otherwise indicated, all compounds containing an asymmetric carbon atom are racemic mixtures.
- (17) M. Weizmann and S. Malkowa, Bull. soc. chim. France, 47, 356 (1930).
- (18) J. C. Sheehan and W. A. Bolhoffer, J. Am. Chem. Soc., 72, 2786 (1950).
- (19) K. C. Leibman and S. K. Fellner, J. Org. Chem., 27, 436 (1962).
- (20) I. M. Kolthoff and W. Stricks, J. Am. Chem. Soc., 73, 1728 (1951).
- (21) J. C. Crawhall, D. F. Elliott, and K. C. Hooper, J. Chem. Soc., 4066 (1956).
- (22) J. K. N. Jones, R. A. Wall, and A. O. Pittet, Can. J. Chem., 38, 2285 (1960).
- (23) W. Enz and M. Cecchinato, Helv. Chim. Acta, 44, 706 (1961).
- (24) H. A. Smith and G. Gorin, J. Org. Chem., 26, 810 (1961).
- (25) R. B. Martin, S. Lowey, E. L. Elson, and J. T. Edsall, J. Am. Chem. Soc., 81, 5089 (1959).

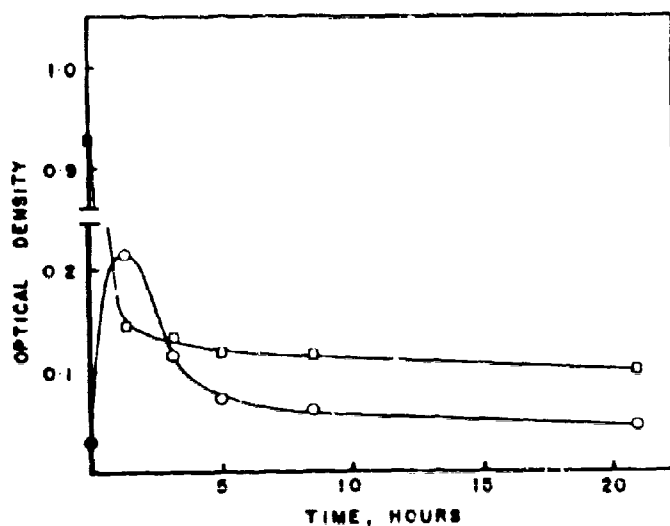


Fig. 1. -- Hydrolysis of L-cysteinol triacetate (2.08×10^{-4} M) in hydrochloric acid (1.7 N) at 90° . \square - optical density at 231.5 mμ. \circ - optical density at 261 mμ.

**BUSER, KADZ
and SULLIVAN**

**THE USE OF INTENSE PINCH
DISCHARGES FOR LASER
ILLUMINATION**

**R. G. BUSER, J. J. KADZ, and J. J. SULLIVAN
U. S. ARMY ELECTRONICS RESEARCH AND DEVELOPMENT LABORATORIES
PORT MONMOUTH, NEW JERSEY**

1. INTRODUCTION

The standard light sources currently in use for optical pumping of laser materials in military and commercial applications as well as in basic research are high pressure noble gas discharges in a linear or helical geometry, Fig. 1a. They all exhibit certain limitations:

a) Due to their mechanical construction they only can handle limited energies in a given time interval (in the order of 10^6 watts). The high inductivity of the geometry prevents high currents of μ sec duration and therefore limits peak light output and pump repetition rate.

b) Focusing cavities necessary in high power application are bulky and difficult to manufacture.

c) Their (except for a few superimposed lines) continuous spectrum often provides poor matching with the absorption bands of the laser material and therefore gives low pump efficiency.

d) In connection with c) a large amount of the available energy is transformed to heat, thus drastically reducing the possible repetition rate of the pump pulses.

Under these circumstances it is very tempting to investigate two discharge types which have been extensively used in connection with high temperature plasma research, as to their application as a light source: the theta pinch (electrodeless) and the linear pinch (with electrodes) Figs. 1b and 1c. We will show that pinch discharges can indeed be used very effectively as light pumps eliminating most of the problems mentioned before. In both discharge types ultra high currents are possible because

BUSER, KAINZ
and SULLIVAN

inductivities of the circuit can be made extremely small. Current conducting elements can be used as reflectors and the laser material can be placed in the center of the discharge. Especially in the blue and ultraviolet region of the (discrete) spectrum often an improved optical match between light source and laser seems possible by proper choice of gases, pressures and discharge conditions. The spectral energy distribution of the pinch discharge also leads to a greatly reduced rate of heating of the laser material which makes high pump repetition rates feasible. Furthermore, large amounts of light are available in short times which is particularly attractive for certain laser materials and switching techniques.

There are a great number of parameters and possibilities. In the following we discuss the main principles and considerations and present some of the more important results obtained so far.

2. THE MECHANISM OF PINCH DISCHARGES

2.1 Theta pinch.

The basic principle of the so-called theta-pinch is the following (Fig. 1b): A condenser is discharged over a single loop coil; A rapidly changing longitudinal magnetic field H will be induced which produces in the gas cylinder (located inside of the coil) a high electric field which, after breakdown, leads to an intense current I_0 in form of a sheath near the wall. Interaction forces between H and I_0 ($F = l \cdot I_0 \times H$) push the plasma sheath to the inner cylinder. As the plasma is compressed, inward moving radial shockwaves are set up which heat and ionize the gas further. During this pinching of the gas large amounts of radiation are emitted. The plasma motion is stopped if internal and external pressure is equal or if the plasma hits the inner cylinder of the discharge tube.

High current, short duration discharges have been studied for some time and a considerable amount of information is available. However, most experiments have been done in hydrogen or deuterium and only few with other gases or gas mixtures. Several theoretical models have been developed to explain the observations in detail. Here we are only interested in the "light source" aspect of the pinch discharge. There are two modes of operation: If the ringing period τ of the circuit is larger than the characteristic pinch time τ_0 (in this case the time between the pinch starting at the outer wall and maximum compression) the plasma hits the wall producing a very intense light emission until the driving force is released. If the circuit is oscillatory a series of compressions may take place. This is shown in Fig. 2a(01) in a streak camera picture.

BUSER, KAIKZ
and SULLIVAN

If $\tau < \tau_0$, the driving force is turned off in time and the sheath does not reach the inner cylinder. Again at maximum compression an intense light flash is produced.

The pinch discharge with a single driving coil (coil length \ll laser length which is approximately $2'' - 3''$ in all experiments reported here) naturally is a very inhomogeneous light source. Most of the light is produced in the drive coil plane as is to be expected from the H-field structure of a single loop. This does not present a problem for the 4 level laser, but is of great disadvantage for the 3 level laser, as will be discussed in Sec. 6. There is also a time delay between the intense illumination of the laser in the center and the regions further out. It takes some time for the axial shockwaves to move away from the center plane (Fig. 2b(01)). This additional illuminating effect produced by the shock waves can be used only in materials with sufficiently long relaxation time of the metastable states. In this connection proper choice of the length of the discharge tube may also be important.

At this point there are several alternatives: The use of multiple loops in parallel or a single loop of sufficient length will greatly improve the homogeneity of the discharge but equally cut down the induced current densities and therefore the light emission unless several independent circuits are used. For example, the replacement of the single loop by a double loop did not produce a significant change of the overall light emission.

Experiments also indicate that preionization of the gas by means of a d.c., a.c., or high frequency current will in many cases increase the light output as there are differences to be expected between cold and preionized shocks. Furthermore it is possible to improve the laser illumination using reflecting coatings around and on the ends of the pinch tube, but no quantitative measurements of this effect have been made so far.

2.2 r- Pinch

In contrast to the θ pinch where actually the imbalance of the magnetic driving field and the field of the induced current provides the compression, the current of the r-pinch produces its own compression, (Fig. 1c). Again the interaction force F is $\sim \epsilon I \times H$ and acts like a cylindrically compressing piston. However there are some basic differences: Electrodes must be used which will produce sputtering effects and therefore shorten the lifetime of the tube. The kevar sealing glass which is necessary in connection with the kevar-stainless steel electrodes is by far inferior in its optical and thermal qualities to the quartz glass used in θ -geometry. Of course there are better technological solutions, but they have not yet been explored. On the other hand the energy coupling is improved and the light source is obviously very homogeneous. As has been pointed out before, preionization

BOGER, KAINZ
and SULLIVAN

improves the light efficiency of the discharge mechanism.

3. ELECTRIC CIRCUITRY AND ELECTRIC MATCHING

The circuit used in all θ and z - experiments consists of a trigger switch, a condenser (energy storage) with a proper power supply and an inductive load (in case of the θ pinch a single loop coil, in case of the z pinch the discharge tube and its connections), Figs. 1a, 1b, 1c. We have the following circuit equations:

$$\begin{aligned} \text{Current} \quad I &= \frac{U}{\omega L} e^{-\frac{Rt}{2L}} \sin \omega t \\ \text{Peak current} \quad I_p &\approx \omega UC \text{ if } \frac{R^2}{4L^2} \ll \frac{1}{LC} \\ \text{Current derivative} \quad \frac{dI}{dt} (t=0) &= \frac{U}{L} \\ \text{with} \quad U &= \text{voltage} \\ t_0 &= \frac{2L}{R} \\ \omega &= \text{frequency} = \sqrt{\frac{1}{LC} - \frac{R^2}{4L^2}} \approx 2\pi\nu \\ L &= \text{inductivity} \\ C &= \text{capacity} \\ R &= \text{resistance} = \frac{4\mu}{\pi} \ln(A_n/A_{n+2}) \\ A_i &= \text{ith current amplitude} \end{aligned}$$

With these expressions the dynamics of the pinch (aside from certain instabilities) is determined if gas type and pressure is given. For example in the θ pinch case the induced electric field

$$E_\theta = \frac{1}{2} r E_0 \sqrt{\frac{R^2}{4L^2} + \omega^2} e^{-\frac{Rt}{L}} \sin(\omega t + \varphi)$$

After breakdown we have a current density $I = \sigma E_\theta$, σ being the electrical conductivity. For reasons apparent in Sec. 4 it appears most advantageous to have extremely high current pulses for very short times. This essentially means one should use a high voltage, low inductance circuit with an overall resistance near critical damping ($R = R_C \approx 2 \sqrt{\frac{L}{C}}$), if possible. From a practical point of view there are several limitations, the most important one being the inductivity of the condenser. Another problem: to avoid unnecessary inductances with connecting lines and the switch. In the following table we present some data on combinations of circuit elements which have been used or are currently being tried in this Laboratory for θ and z pinch geometry.

Condenser Mfg.	μf	U kv	E kjoule	$\left(\frac{dI}{dt}\right)_{\text{avg}}$ amp/sec	I_p amps	V ke	t_0 μsec	E/t_0 joule/sec	$L_{\text{CLR}}/L_{\text{Lead}}$	Symbol
TOMI DEUTSCHMANN	12.0	20	2.4	2.2×10^{11}	2.1×10^3	140	13.8	1.7×10^9	2.9	01, z1
AXEL	0.85	50	10.6	2.7×10^{11}	1.1×10^3	400	6.5	1.6×10^9	4.9	02, z2
G.E. 14F734	1.0	40	0.8	5.3×10^{11}	1.4×10^3	980	2.6	3.1×10^9	2.0	03
T.D. BSC261	0.2	20	0.04	1.0×10^{12}	0.6×10^3	2500	0.8	0.5×10^9	0.4	z3
CONVENTIONAL	10^4	1	5.0	3.1×10^9	0.6×10^3	1	2000	2.5×10^6	---	C1

Symbol	Remarks
01, z1	No laser action positively identified in θ or z geometry. Inner cylinders show destructive wear after a few firings in the case of z pinch geometry.
02, z2	Laser action observed with Nd in θ geometry.
03	Experiments not yet completed. However laser action has been reported for Nd and ruby in θ -pinch geometry. ^{2,4}
z3	Experiments not yet completed.
C1	Conventional methods used for comparison purposes.

**BUSER, KAINZ
and SULLIVAN**

Symbol: identify the experimental conditions at which the measurements shown in Figs. 2 and 4 were made and some results are given. In Fig. 3 an optimal design for the x-pinch discharge is shown which could be used in connection with a portable unit (Z3).

4. RADIATION OF FAST PINCH DISCHARGES - OPTICAL MATCHING

Considering the optical properties of light sources there are two extreme cases: If we wish to obtain a more or less continuous spectrum, it is standard practice to pass high currents through a high pressure gas. We approach here the properties of a black body radiator and this means that the radiated energy is spread over a considerable range of frequencies. If such a light source is used for pumping only the radiation matching the absorption bands of the laser material is useful, this being the condition which prevails in standard light sources. Large quantities of energy contribute only to unwanted heating of the laser rod. This heating effect makes high repetition rates virtually impossible and leads sometimes to the destruction of the crystal. A low efficiency in useful energy conversion is to be expected; a representative value is $.5 \cdot 10^{-3}$ joule output/joule input.

However, by proper design of circuit elements and light source we can produce an extremely fast discharge so that no equilibrium of the plasma constituents is reached. The spectrum in this case will essentially be a line spectrum of the gas atoms and ions present and will be similar to the one of high energy sparks of short duration. Only little heat production is expected and therefore high pump repetition rates should be possible. Furthermore, by using the right gases, gas mixtures and gas-vapor mixtures, and pressures, often an improved spectral matching might be obtained. Also, the system itself will be simplified, because no Q - spilling techniques are necessary: If the discharge current is short enough, the light source itself acts as a "Q-switch" and only one laser spike will be produced.

What guide lines do we have if we want to use fast pinch discharges for a particular laser material? The theoretical outlook is difficult: Let us assume that we have electrons with a Maxwellian velocity distribution and furthermore an energy source which heats the electrons. As the electrons increase their energy they start to lose some of it by inelastic collisions. If the energy source is sufficiently strong very soon a high degree of ionization is reached. Since all collision cross sections are functions of the (changing) electron temperature and many times not even known quantities, it is difficult to calculate the population of all the states involved. Spontaneous radiative transitions compete with the electron collisions as de-excitation processes. Their intensity as a function of time is actually the quantity which determines the characteristic emission of the light source. Many of the processes are coupled, and in a quantitative analysis a number of other processes

BUSER, KAINZ
and SULLIVAN

would have to be included also. It is fair to say that the theoretical study of nonequilibrium plasma radiation is only at its beginning.

Under these complex circumstances it seems reasonable to apply at first an experimental approach. From a study of their known discharge characteristics at standard conditions (using the information contained in the MIT wavelength tables) we selected the elements neon, argon, krypton, xenon, mercury, iron and silicon for a first closer investigation. The important point here is the number of lines in the wavelength region of interest, their intensity, the availability and technical applicability of the particular element. While the noble gases and mercury are generally used in flash tubes, iron may be present in connection with the stainless steel electrodes in the z-pinch configuration. When the compressed plasma hits the glass wall with high speed usually Si lines are excited which might be useful in particular applications. The radiation from other impurities often present or produced in strong gas discharges might also be of interest.

5. EXPERIMENTAL SETUP AND GENERAL RESULTS

The electric circuit has been discussed in Sec. 3 with the exception of the triggered airgap switch the characteristics of which was the same in all experiments (switching times in the order of 100 nanosec.) For the vacuum system we used a vacuum pump, scription pump and stainless steel tubing and obtained a base pressure below 10^{-7} Torr without bakeout. Only high purity gases (supplied by Linde and Airco) have been used to be able to control the impurity level sufficiently. In the routine measurements of the spectrum a Zeiss prism instrument in connection with a STL image converter camera has been used which provides time integrated and time resolved spectra. For precision measurements a monochromator (Czerny-Turner; Spex Industries, Inc.) has to be employed. The laser output and the total light output of the pinch tube is measured in standard fashion with suitable photomultipliers and filters in connection with oscillographic recording.

In view of the many parameters (geometry, discharge conditions, pressure, gas type and gas mixtures, impurity content) only a limited number of results have been obtained so far (See also Sec. 6).

However, in general we may state the following: The time integrated and the time resolved spectra are mainly unbroadered, unreversed intense line spectra. Fig. 4 (02,z2) shows as an example some spectra in the range between 4000 - 7000Å.

The character of the line spectrum is determined mainly by the peak current and its time derivative at a given pressure. The high induction current seems to produce most of the electron heating and there is little shock wave heating. If the current densities are

BUSER, KAINZ
and SULLIVAN

sufficient a high degree of ionization will be reached and large amounts of radiation should be expected in the ultraviolet.

Only a small amount of continuum radiation is found. The continuum is more pronounced in the z-pinch than in the θ pinch as can be seen from a comparison of the spectra in Fig. 4.

The total light output as a function of time depends also mainly on peak discharge current and its derivative. Peak amplitudes are pressure sensitive. For each given discharge current there seems to be an optimum pressure beyond which the light output decreases. The time duration of the light pulse also depends on pressure. It is interesting to note that under equal geometrical and pressure conditions theta and z-pinch give very similar records.

Spectral matching is possible: If two gases are mixed, for example, one finds in general an addition of their spectra with a considerable shift of the original intensity. This is shown in Fig. 4 for argon, krypton and an Ar-Kr mixture. Using effects of this kind one might get an improved match between laser and pump. Adjustment of the pump spectrum to fit the absorption curves of the laser material will be possible particularly in the UV region of the spectrum, since it is expected that with increased discharge currents more and more lines of single and multiple ionized atoms will be excited. Extensive measurements are under way to obtain a more detailed picture.

6. APPLICATION OF PINCH DISCHARGES FOR THE RUBY, NEODYMIUM AND TERBIUM LASER

6.1 The ruby laser

In Fig. 5a we present the absorption A ($A = \ln \frac{1}{T}$, T = transmission) as a function of wavelength for ruby (1). There are two large absorption regions corresponding to the absorption band indicated on the schematic diagram of the three level laser in the same figure. Curve (2) shows the emission of a standard Xe flash tube in the same wavelength region. Obviously only little of the emitted overall energy gives useful pumping light for the laser.

We also show the line spectrum of ArI and ArII (3) in the range 3400Å-6000Å indicating Argon as a possible choice for the gas in the pinch tube. (The "length" of the lines represents qualitatively their intensity in a logarithmic scale). In an ideal case as many lines as possible should be situated in the absorption region, so that in a time $t < t_0$ (t_0 = lifetime of spontaneous emission of the metastable level), the emitted power integrated over the absorption band is sufficient to produce laser action.

At this point let us make a rough estimate which shows the path to follow in a more accurate calculation.

BUSER, KAINZ
and SULLIVAN

If we assume that 1/10th of the gas atoms at 10 mm Hg radiate (in a steady-state condition) at an "average wavelength" $\lambda = 4000\text{\AA}$ we find for the emitted power per cm^3 gas

$$P_e = \int_{\text{Abs. band}} h \nu_1 N_1 A_1 = h \bar{\nu} \cdot \bar{N} \cdot \bar{A} \approx 10^3 \frac{\text{Watt}}{\text{cm}^3}$$

where h Planck's constant

N_1 Concentration of emitting atoms or ions

ν_1 Pump frequency

A_1 Transition probability

This value has to be multiplied by a factor < 1 depending on geometry and reflectors used. On the other hand the total power absorbed per unit volume in a three level laser is

$$P_a = \frac{h\nu_1}{Q} \frac{5A \left[\frac{(1-R)}{l} + \alpha_s \right] \Delta\nu}{\lambda^2} + \frac{h\nu_1 N_0}{Q 2\tau}$$

(In a four level laser the second term is replaced by $\frac{h\nu_1 N_0 e^{-\frac{\Delta E}{KT}}}{Q\tau}$).

where R coefficient of reflection = .9

l Laser length = 10 cm

α_s Scattering loss/unit length ≈ 0

$\Delta\nu$ Linewidth of laser output = 6\AA

λ Wavelength of laser emission = 6943\AA

N_0 Concentration of active ions = $1.6 \cdot 10^{19}$ ions/ cm^3

Q Total quantum efficiency = .6

τ Decay time of radiative spontaneous emission = $4 \cdot 10^{-3}$ sec

Putting in numbers in the case of ruby we obtain for $P_a \approx 1400 \frac{\text{Watt}}{\text{cm}^3}$. Since for $P_a \geq P_e$ the laser starts to oscillate under the conditions assumed here laser action should indeed occur.

Considering the gases mentioned in Sec. 4 in the case of ruby only a rather limited "spectral match" seems to be possible since pinched light sources have most of their intensity below 5000\AA . Of course, if by some means the ultraviolet light could be transformed to frequencies appropriate to the absorption bands of the ruby in the blue and green region of the spectrum, the efficiency would be improved. This could be done by a coolant liquid, by changing the composition of the ruby or by an appropriate fluorescent mirror on the outside of the discharge tube, and several experiments in this direction are under way.

BUSER, KAINZ
and SULLIVAN

The pumping of the three level ruby laser in θ -pinch geometry is particularly sensitive to the inhomogeneity of the pump source. Since the main light emission occurs in the coil plane very soon the material becomes saturated there and cannot absorb any more energy. In contrast, the less illuminated areas of the laser absorb strongly the light coming from its central regions and considerable time will pass until laser action occurs if at all. In spite of these difficulties it has been possible to pump successfully a ruby at 77°K in an θ pinch geometry with Ar as fill gas.²

6.2 The neodymium laser

In Fig. 5b we present the absorption of a Nd-laser as a function of wavelength in (1) and indicate in (2) the line spectrum of NdI and NdII. Under the conditions of interest here the strong absorption band around 3500Å is important: We have tried argon in θ - pinch geometry (02) and indeed obtained laser action. Using a z-pinch geometry laser action has been found also.³ Again a very much reduced heat transfer to the laser rod is observed. Consequently high repetition rates in laser pumping should be possible and have been reported in the meantime.⁴ The four level Nd laser is also insensitive to light source inhomogeneities.

6.3 The terbium laser

Terbium starts to absorb strongly below 2300Å. Pumping with conventional light sources at these low wavelengths is extremely difficult because their emission is very low. Here pinch discharges should be exceptionally useful. Experiments with terbium doped glasses are under way and showed an intense fluorescence around 5500Å with a long time constant.

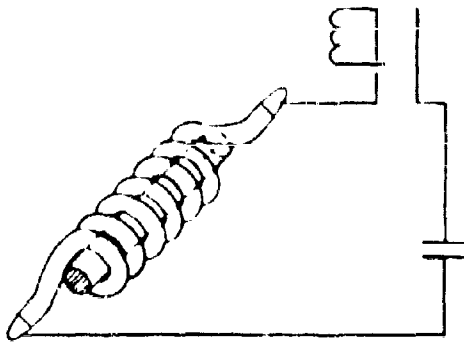
7. CONCLUSION

Results so far indicate a) Pinch discharge light sources provide excellent means for laser pumping with increased efficiency in comparison to conventional light sources in particular in the spectral range below 5000Å. b) For the first time high repetition rates are feasible which are necessary for certain military applications. The fast pinch discharge mechanism also may eliminate the need for mechanical or electronic Q - spilling and simplifies the system. c) For each laser material the light source should be tailored for optimum performance. It might also be possible to develop new laser materials matching the particular properties of pinch discharges. For example, intense radiation in the ultraviolet is expected at high degrees of ionization. Therefore new laser materials with absorption bands in the UV become promising.

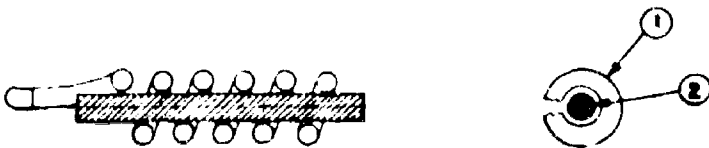
REFERENCES: ¹R. D. Haun, Electro-Technology, 63, Sept 63; ²R. A. Brandow, et al, J. Appl. Phys. 34, 3415, 1923; ³C. H. Church, et al, J. Opt. Soc. Am., 52, 514, 1933; ⁴Philco ATL Tech Note 63-14.

CONVENTIONAL LIGHT SOURCE

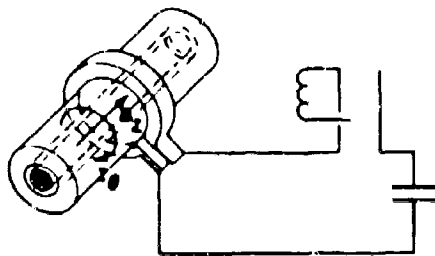
FIG. 1a



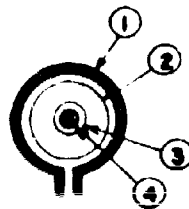
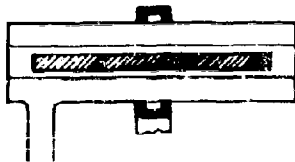
- 1. HELICAL FLASH TUBE
- 2. RUBY



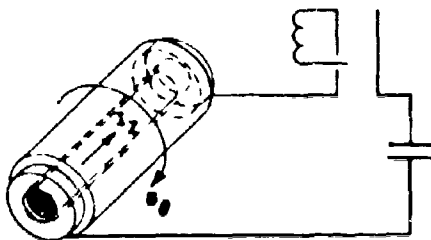
Ø PINCH LIGHT SOURCE FIG. 1b



- 1. SINGLE LOOP COIL
- 2. OUTER GLASS CYLINDER
- 3. INNER GLASS CYLINDER
- 4. RUBY



Z PINCH LIGHT SOURCE FIG. 1c



- 1. OUTER GLASS CYLINDER
- 2. RING ELECTRODE
- 3. INNER CYLINDER
- 4. RUBY

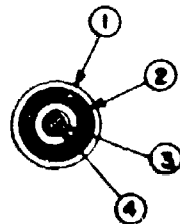
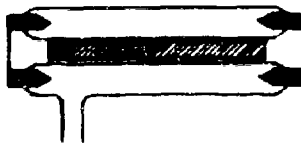


FIG. 2a Radial compression of θ -pinch

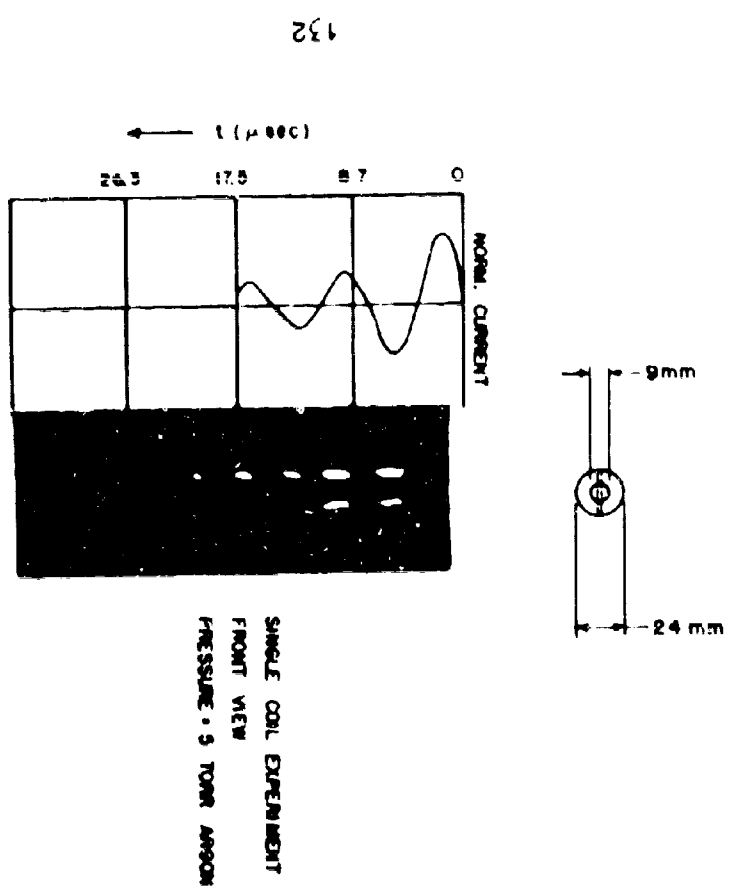
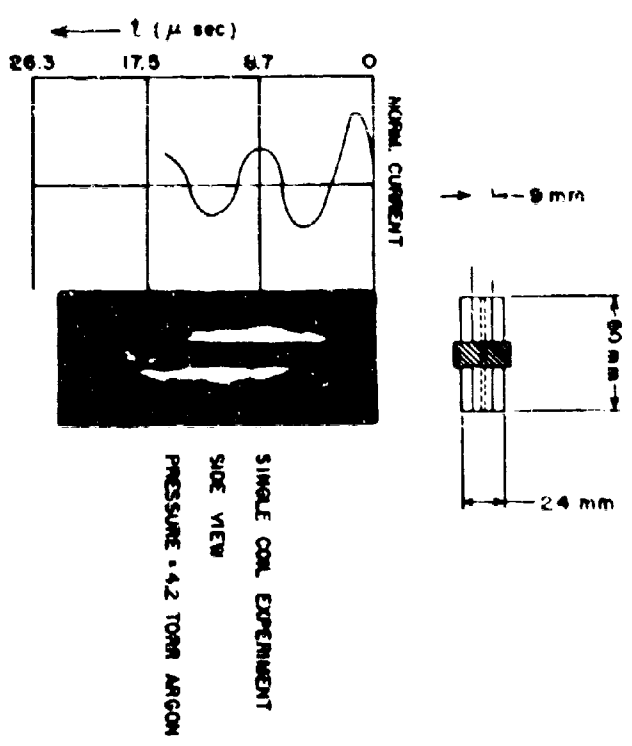


FIG. 2b Axial shocks in θ -pinch geometry



BUSER, KAINZ
and SULLIVAN

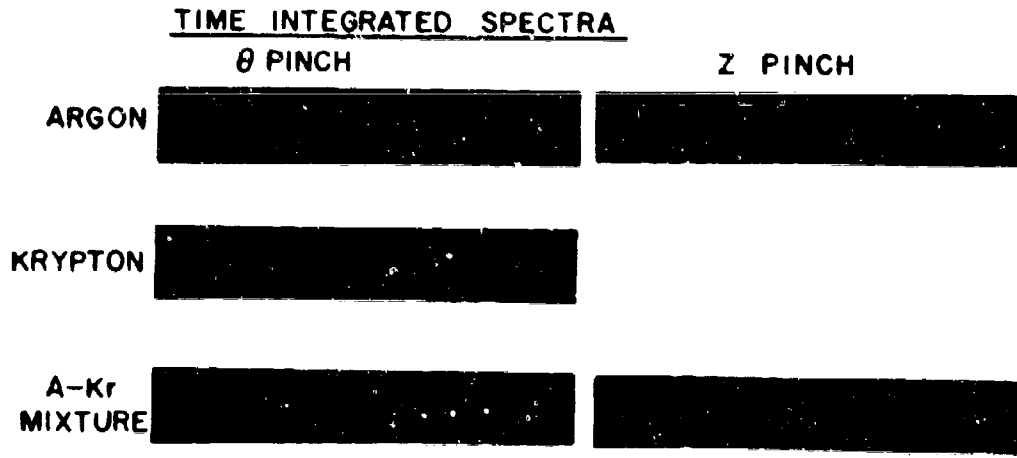
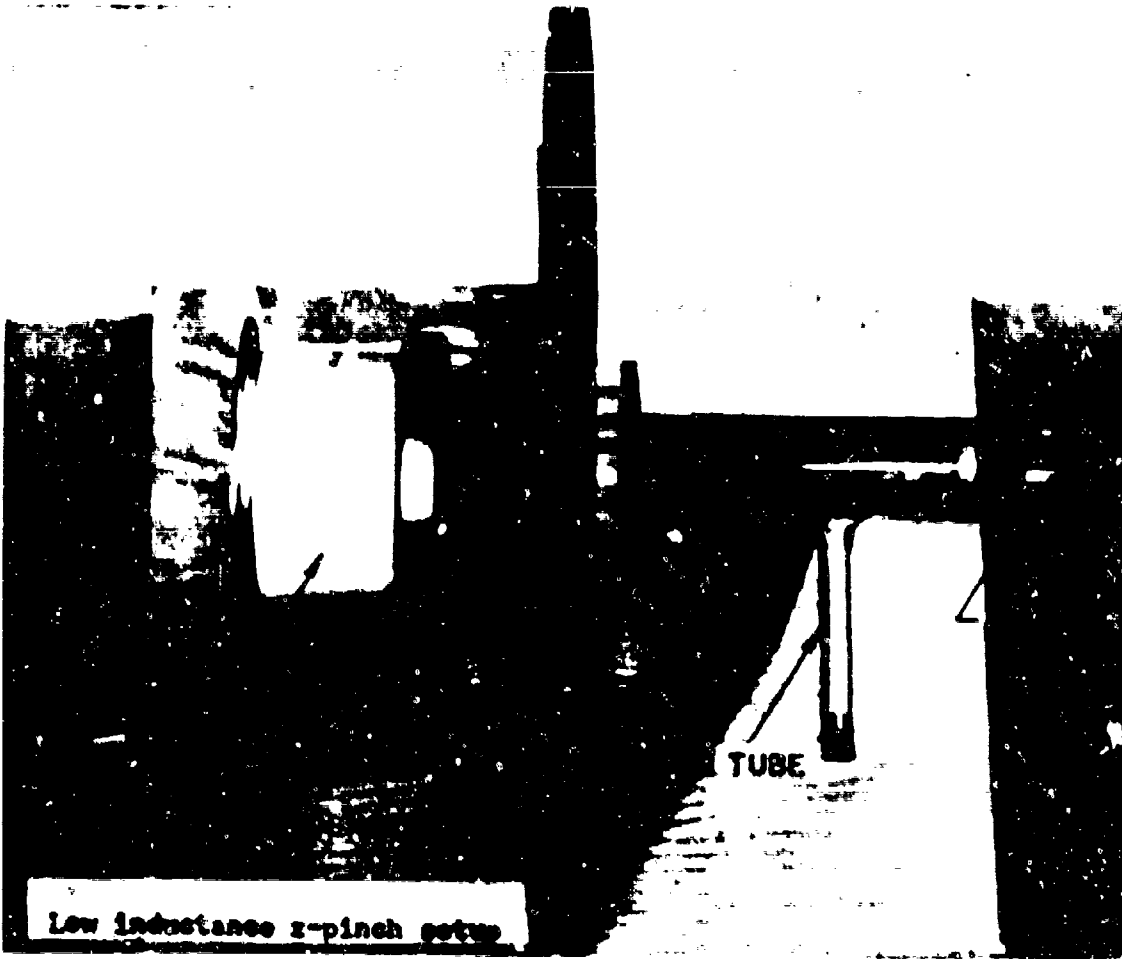


FIG. 4 Time integrated spectra of θ and z-pinch_s in krypton, argon and Kr-Ar-mixtures

BUSER, KAINZ
and SULLIVAN

Fig. 5a. Absorption: Ruby

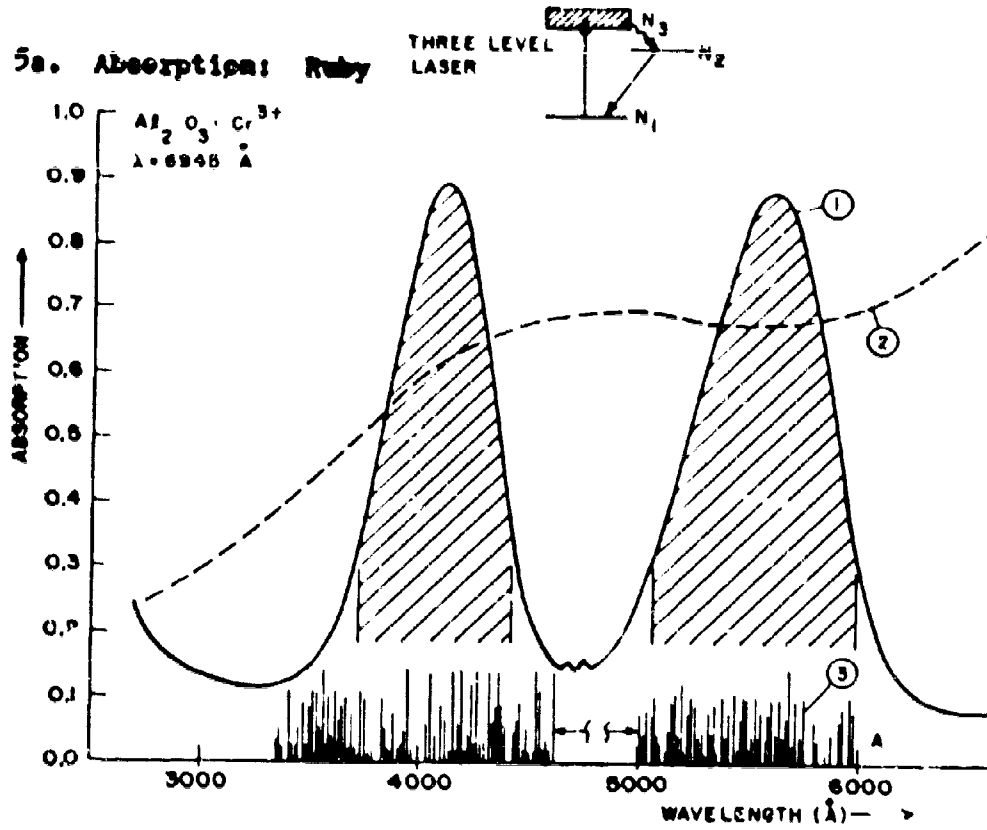
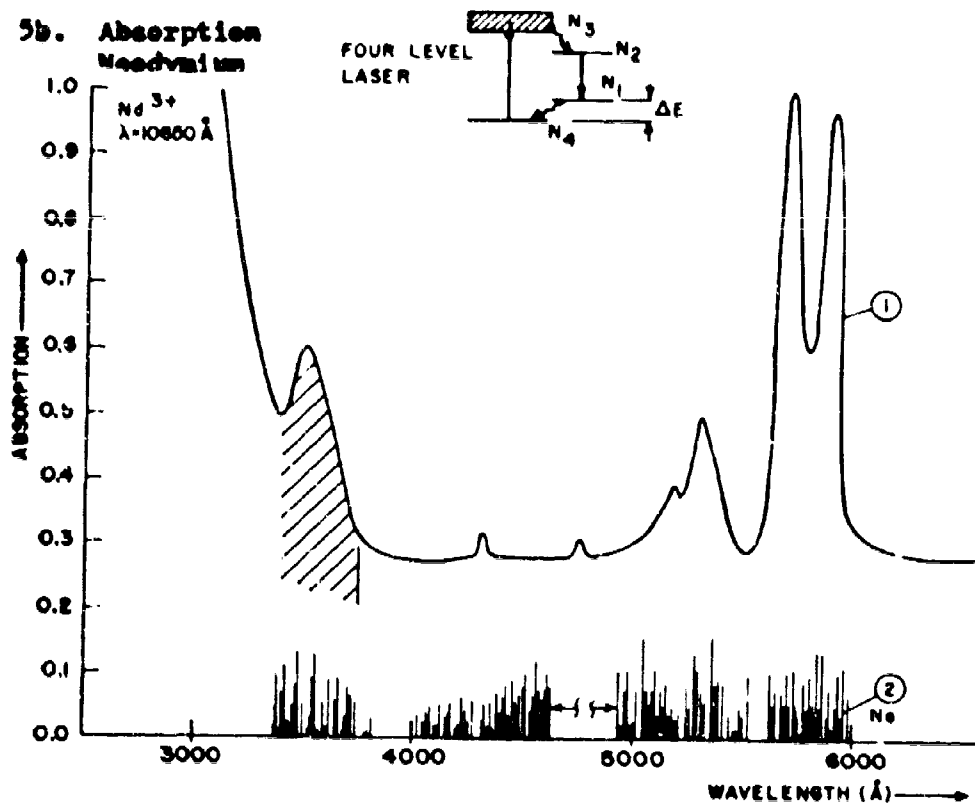


Fig. 5b. Absorption Neodymium



CARLSON

A Paramagnetic Resonance Study of Nitrogen Atoms
Trapped in X-Irradiated Alkali Azides

FREDERICK F. CARLSON
U. S. ARMY ENGINEER RESEARCH AND DEVELOPMENT LABORATORIES
FORT BELVOIR, VIRGINIA

INTRODUCTION: A radical is a chemical building block which is extremely reactive. It may be a molecule or an atom. Radicals which have not yet reacted with anything are called free radicals. OH, NO₂, atomic hydrogen and atomic nitrogen are examples of free radicals. Reactions involving free radicals are usually very energetic. Thus, when nitrogen atoms combine to form molecules of diatomic nitrogen gas, each reaction gives off about 10 electron volts of energy. It is apparent that if a large number of free radicals were to react simultaneously an explosion would occur. If the reaction were moderated in some way, a more controlled release of energy could be obtained.

In the late 1950's there was a considerable effort to obtain high concentrations of trapped free radicals stabilized in some matrix. The National Bureau of Standards was the center of much of this work. Several reviews of this work have been published in book form.¹

The classical way of investigating stabilized free radicals arising from a gas (N atoms, H atoms, etc.) is to pass the diatomic gas through a radio frequency discharge thus forming a small percent of gas atoms among the molecules. The gas is then condensed as an amorphous solid on a cold finger at 4°K (-269°C), and the free radicals are trapped among the condensed gas molecules. The concentration of free radicals stabilized by these techniques is usually 0.001% to 0.05%. Fontana² reported that in his experiments with N atoms he observed small explosions. He estimated the concentration to be about 0.3%. This of course did not represent a successful stabilization of a high concentration of free radicals since the mixture detonated even at 4°K.

CARLSON

We at Fort Belvoir felt that the decomposition of various solid azides such as NaN_3 , KN_3 , PbN_6 , etc., might be a means for obtaining useful concentrations of trapped N atoms. Furthermore, a study of their formation could provide clues about the decomposition mechanism of these important metastable materials. Since nitrogen atoms are paramagnetic they may be observed by electron paramagnetic resonance spectroscopy. This is a most fortunate method of study since it allows us to look into the crystal to observe the nitrogen atoms without disturbing either the nitrogen atoms or the crystal. In addition, if there are electric or magnetic interactions between the nitrogen atoms and the surrounding crystal, we should see them as additional complications in the spectrum. These additional complications in the spectrum will allow us to identify the location of the trapped atoms and tell us something about the strength of the trap. In the next sections of this paper we shall discuss the experimental methods and the results. A brief discussion of the current theory will be given. It will be noted that nitrogen atoms in the alkali azides provide an excellent system in which to study the spectroscopy of S state atoms, a field in which there is much discussion concerning the inadequacy of the present theories. At the end we shall present some conclusions concerning the concentration of N atoms and their usefulness.

EXPERIMENTAL PROCEDURE: In these experiments we used crystals grown from aqueous solutions of the salts. Tetragonal KN_3 and RbN_3 crystallized in small dipyramids (3 mm x 3 mm x 2 mm). These were used directly. Tetragonal CsN_3 crystallized in plates which were cleaved (3 mm x 3 mm x 0.5 mm) and were comparable to the dipyramids but with an undeveloped c axis. Rhombohedral NaN_3 also crystallized in plates. Samples, 2 mm to 3 mm square, were cleaved from these plates. All samples were oriented optically by observing their natural faces. A sample was glued with G. E. 7031 varnish, in a desired orientation to a glass rod which, when suspended into the microwave cavity, served as the crystal rotator. The samples were irradiated in liquid nitrogen with a Machlett OEG 60 x-ray tube operated at 40 kV and 40 mA. The glass rod was protected from the x rays so that signals from irradiated glass would not interfere with the work. Other samples irradiated in evacuated quartz tubes showed identical spectra. The irradiated sample was placed with as little warming as possible into the precooled variable temperature Dewar insert of the Varian spectrometer. The other end of the glass rod was fastened to a rotatable chuck. A pointer on the rod measured the angle of rotation against a protractor fixed at the axis of the chuck. Spectra were made at 9.1 kMc using a rectangular H_{012} cavity with 100 kc field modulation and phase-sensitive detection. This leads to the familiar 1st derivative display of the absorption spectrum. Low temperatures were obtained by using the Varian variable

CARLSON

temperature apparatus. The temperature of the sample was measured by a chromel alumel thermocouple inserted in the gas stream just beneath the cavity.

THE SPECTRA: The electron paramagnetic resonance spectrum of $^4S_{3/2}$ atomic nitrogen in various matrices is well known.³⁻⁸ For atomic nitrogen trapped in a crystal lattice, one would expect the crystalline electric field to shift the four magnetic states so that the three transitions between them would occur at different magnetic fields. This is known as the fine structure. (See the article by Bleaney and Stevens⁹ for an excellent discussion with diagrams.) In addition, the magnetic field of the nitrogen nucleus will split each fine-structure component into $2I + 1 = 3$ hyperfine lines. Thus, the spectrum should consist of three groups of three lines. The intensity ratio of the outer triplets to the center one should be three to four. The splitting between the fine structure components will depend upon the intensity of the crystal field at the trapped nitrogen atom. It will also be a function of the angle between the crystal field and dc magnetic field.

The spectrum of N in the azides we are discussing appeared more complex than this nine-line model. However, it was possible to show that the complexities were caused by the presence of several groups of nine line spectra. This would result if there were more than one site for the trapped nitrogen in each crystal. Further study showed that each spectrum corresponds to a nitrogen atom in one of several similar sites with noncoincident magnetic axes. The center triplets of each spectrum coincide except for small second-order effects. Thus the center triplet appears stronger than the 3:4 ratio predicts. In NaN_3 there are three such similar sites, each one with its magnetic axis parallel to one of the three rhombohedral axes of the crystal. In the tetragonal azides there are two sites,¹⁰ one with its magnetic axis along the 110 direction and the other with its magnetic axis along the $\bar{1}\bar{1}0$ direction. Fig. 1 shows typical spectra of trapped nitrogen atoms. Figs. 2 and 3 show the location and symmetry of trapped nitrogen atoms for NaN_3 and for the tetragonal azides respectively.

In paramagnetic resonance spectroscopy, we describe our results in terms of a spin Hamiltonian. The spin Hamiltonian for these experiments is $\mathcal{H} = g\beta H \cdot S + D[S_z^2 - \frac{1}{3}S(S+1)] + E(S_x^2 - S_y^2) + A I \cdot S$

Each term of the Hamiltonian represents the energy of a particular interaction. The first term represents the Zeeman splitting. Here g is the Landé factor, β is the Bohr magneton, and H is the dc magnetic field. The second and third terms represent the perturbation of Zeeman splitting by the interaction of axial and rhombic crystal fields (the fine structure). D and E are constants. The fourth term represents the additional splitting of the energy levels by magnetic

CARLSON

interaction of the nitrogen nucleus (the hyperfine structure). A is the isotropic hyperfine constant and I is the magnetic moment of the nitrogen nucleus. The values of the various constants obtained in these experiments are shown in the table. It will be noted that the values of the various constants given in the table appear to be orderly functions of the metal ion. D & E increase as one proceeds down the first column of the periodic chart, and the g value decreases. We shall not discuss the g shift in this paper, but confine our attention to the fine structure parameters.

As was mentioned previously the fine structure splitting is a function of the angle between the magnetic axis of crystal field surrounding the trapped nitrogen and the dc magnetic field which provides the Zeeman splitting. Maximum splitting occurs when the magnetic axis of the trap is parallel to the dc magnetic field. For a $3/2$ state this corresponds to a splitting of $4D$. In our experiments this is observed to occur when the line from one nearest neighbor metal ion through the nitrogen atom to the other nearest neighbor metal ion is parallel to the dc magnetic field. This is taken as the z axis, as is shown in Figs. 2 and 3. We may thus relate D to the axial field of the metal ions and E to the rhombic asymmetry caused by the nearest neighbor azide ions. In order to relate the spin Hamiltonian to the experiment, it is convenient to transform it from wave numbers to magnetic energy in Gauss and to coordinates involving angles, which may be related to the rotation of the crystal in the dc magnetic field. The coordinates relating the Hamiltonian to the crystal are shown in Figs. 2 and 3. This transformed Hamiltonian is¹¹

$$\begin{aligned} \mathcal{H} = & H_0 + (M - \frac{1}{2}) [D(3\cos^2\theta - 1) - 3E(\cos^2\theta - 1)\cos^2\varphi] \\ & + \frac{1}{2H_0} [4S(S + 1) - 24M(M - 1) - 9] \{D - E \cos 2\varphi\}^2 (\sin\theta \cos\theta)^2 \\ & + (E \sin 2\varphi)^2 (\sin\theta)^2 \} \end{aligned}$$

where H_0 is the magnetic field (in Gauss) at which the fine structure coalesces θ is the angle between the z axis and the dc magnetic field and φ is the angle in the xy plane. $\varphi = 0$ is the x axis. It is the fit of the experimental data to this spin Hamiltonian which provides us with the information that the resonance is due to a spin state of $3/2$ and indeed belongs to ground state atomic nitrogen. Fig. 4 shows a plot of this theoretical Hamiltonian for N in NaN_3 , only the $M = \pm 1/2 \leftrightarrow \pm 3/2$ transitions are shown. The angular variation of all three sites is included. The points are the experimental data. It is seen that the agreement is quite good. Fig. 5 shows a similar plot for N in RbN_3 ; however, the plot is made for only one site.

THE THEORY OF THE FINE STRUCTURE: The central theme of this paper is the discussion of the fine structure and what we may learn from it about the trap which holds the N atoms. Thus, for the remainder of this

CARLSON

paper I shall focus my attention on the fine structure parameters D & E in an attempt to understand how these zero field splittings arise in an S state atom. The origin of these splittings is not yet fully understood, since, for spherically symmetrical, orbitally nondegenerate states, i. e. S states, the usual mechanisms of splitting are not applicable. The problem is an old one having first been considered in the early 30's by Van Vleck and Penney.¹² Their theory, based on spin orbit interaction and crystal field in a high order perturbation, has been rejected since the splittings it predicts are much smaller than the observed ones. Pryce,¹³ in 1950, suggested an alternative mechanism for the splitting in $6s_{5/2} \text{Mn}^{++}$. He proposed that the electric field gradient

and spin-spin interactions combine to cause the observed splitting. He made a rough calculation which predicted a D of the right order of magnitude. Briefly, what these theories say is that one can explain how a spherically symmetrical state is split, by recognizing that a non-spherical field gradient may induce a non-sphericity in the electron charge configuration. Then the anisotropic dipole-dipole interaction will cause a splitting of the spin levels. Recently, Chakravarty and Bersohn have presented detailed calculations of the zero field splitting of atomic nitrogen based on the field-gradient-spin-spin mechanism. I shall summarize their results and then discuss my model which indicates the importance of nearest neighbor interactions. Chakravarty and Bersohn write expressions for the distorted states based on a Hamiltonian representing the perturbation by the crystalline field. Then they calculate the expectation value of the dipole-dipole interaction operator between these states. After a suitable transformation, they compare the result with the original spin Hamiltonian for the crystal field, and they obtain

$$D = -\frac{3}{2}g^2\beta^2q \left(\frac{a_0}{e}\right)^3 \langle DD \rangle \text{ and } E = -\frac{1}{2}g^2\beta^2nq \left(\frac{a_0}{e}\right)^3 \langle EE \rangle$$

where a_0 is the Bohr radius, e is electronic charge and q is the field gradient in the z direction nq is the field gradient in the xy plane. They evaluate the integrals $\langle DD \rangle$ and $\langle EE \rangle$ by using a Slater type function and obtain appropriate numbers. Thus we see that D & E are linear functions of the field gradients. If one could measure the field gradients by a different experiment, say quadrupole resonance, one would have a check of the theory. Unfortunately no measurements are available or likely, since NMR studies of nitrogen involve great signal to noise difficulties. One could also check the theory by calculating the field gradients from a lattice sum.¹⁵ Thus

$$q = (1-\gamma_\infty) \sum_j e_j \sum_i \frac{3z_{ij}^2 - r_{ij}^2}{r_{ij}^5} \quad nq = (1-\gamma_\infty) \sum_j e_j \sum_i \frac{3(x_{ij}^2 - y_{ij}^2)}{r_{ij}^5}$$

where γ_∞ is a Sternheimer anti-shielding factor.^{16, 17}

CARLSON

Such a calculation ignores the size of the nearest neighbors and may or may not be meaningful. Chakravarty and Bersohn have made such a calculation for KN_3 and report no agreement at all with the values taken from the EPR data and their theory. Mr. McMillan of our group has made some point charge lattice sums for the tetragonal azides in order to see in which direction the field gradient changed as one progressed down the first column of the periodic chart. The result was that q decreased as one progressed from KN_3 to CsN_3 . According to the theory, q should increase since the measured D & E increase.

The most likely conclusion from these discrepancies is not that theory is incorrect but that one must add to the point charge model by taking into account the special effects of the nearest neighbor ions. We have begun such a study.

Consider the following qualitative picture. The nitrogen atom is trapped between two large metal ions which are very near (particularly so for the tetragonal azides). We can perhaps get a picture of how important it is to include the effects of these ions by plotting their radial electron densities on a scale next to the nitrogen atom radial electron density at the proper inter-ionic distance. Fig. 6 shows such a plot for nitrogen trapped in sodium azide. In making these plots the following assumptions were made.

1. We used Hartree radial wave functions.¹⁸ The metal ion functions do not include exchange since exchange calculations were available only for Na^+ and K^+ . Exchange interactions tend to shrink electron cloud and would thus lessen the overlap.

2. Available lattice constants were used. This means that no corrections were made for lattice expansion because of the interstitial nitrogen atom, and that no corrections were made to account for the thermal contraction of the lattice at the low temperature of the measurements.

Since we are looking for relationships between similar materials, it is hoped that these approximations are valid. Fig. 7 shows a similar plot for the alkali metal azides. If one uses the area of common electron density as a measure of how strongly the metal ion disturbs the spherical charge distribution of the nitrogen atom and plots D as a function of this disturbance, one obtains the exponential graph shown in Fig. 8. It is clear that there is a functional relationship between the amount of asymmetry induced in the nitrogen free atom wave functions by the nearest neighbor metal ions and the fine structure parameter, D . The exponential form of this graph is characteristic of interactions involving exchange as a function of r . This is what one might expect since interactions between the electron clouds would involve the Pauli exclusion principle and Coulomb effects. The interpenetration of the electron densities implies a mixing of the metal ion wave functions with the nitrogen

CARLSON

atom wave functions. This mixing grafts some of the aspherical characteristics of the metal ion wave function onto the spherical nitrogen atom. We can describe this by a superposition of excited states. The result is a splitting of the spin levels, as shown by the fine structure. This more or less empirical model suggests several things concerning the fine structure parameters D&E:

That overlap charge effects are important in the calculation of fine structure splittings of S state atoms, and that further study of the overlap of the wave functions involved should be made. Furthermore, in sodium azide the effect of the sodium ion is small and hence one might hope that a calculation of the field gradient for nitrogen trapped in sodium azide might give reasonable agreement with the resultant experimental value of D, thus supporting the basic ideas of the theory given by Chakravarty and Bersohn.

DISCUSSION: The theory discussed in the previous section gives us a picture of the importance of the size of the nearest neighbor ions and the size of the trap. It may be noted that where the overlap of the wave function is large the formation of a chemical bond may occur. Thus we might expect that in the tetragonal azides, KN_3 , RbN_3 , and CsN_3 , the N atoms would be more tightly bonded to the surrounding ions than in NaN_3 . That this is the case is borne out by experiments in which the change in intensity of the resonance is observed over a period of time as a function of temperature. These thermal bleaching experiments show that the trapped nitrogen is more stable in KN_3 than in NaN_3 . In NaN_3 the nitrogen resonance bleaches instantly at 170°K . In KN_3 it bleaches around 190°K .¹⁰ Preliminary measurements on nitrogen in RbN_3 indicate that this trend continues.

From the preceding discussion it would appear that cesium azide would be the best of these materials in which to attempt to stabilize large quantities of nitrogen atoms. However, other factors are involved. The production of the radicals requires that the x rays penetrate through the crystal. The heavier metal azides, RbN_3 and CsN_3 , are so dense that the intensity of x rays reaching the interior of the crystal is very low. Thus a long time is required to produce even a low concentration of radicals in the interior of the crystal. More energetic x rays or gamma rays might improve the situation.

So far we have obtained our highest concentrations of N atoms in the lightest of these materials, NaN_3 . By comparison with other resonances of known concentration, we have estimated this concentration to be less than 0.01%. This is about the same as others have previously observed. It does not represent a satisfactory concentration for use as an energy source. However, even though we have been unsuccessful in obtaining a useful concentration of free radicals, we have gained considerably in terms of stability. Trapped nitrogen atoms in azide crystals are relatively long-lived even at 150°K . This represents a great improvement in stability over nitrogen atoms in a frozen gas

CARLSON

matrix where the nitrogen atoms are not stable above 20°K or 30°K. We hope to exploit this improved stability by investigating trapping phenomena at liquid helium temperatures. At 4°K it may be possible to increase the concentration of trapped nitrogen atoms many fold. However, it seems improbable that the improvement would be so spectacular that the result could be used as an energy source.

I wish to thank Mr. G. J. King for providing many of the crystals which made this study possible. I also wish to thank Mr. King, Mr. R. C. McMillan, and Mr. B. S. Miller for their assistance and for many helpful discussions. The interest of Dr. Z. V. Harvalik is gratefully acknowledged.

REFERENCES

1. Arnold M. Bass and H. P. Broida ed. Formation and Trapping of Free Radicals (Academic Press Inc. New York 1960)
2. B. J. Fontana, J. Chem Phys. 31, 148 (1959)
3. Frederick F. Carlson, J. Chem Phys. 39, 1206 (1963)
4. T. Cole, J. T. Harding, J. R. Pellam, and D. M. Jost, J. Chem Phys. 27, 593 (1957)
5. T. Cole and H. M. McConnell, J. Chem Phys. 29, 451 (1958)
6. C. K. Jen, S. N. Foner, E. L. Cochran, and V. A. Bowers, Phys Rev. 112, 1769 (1958)
7. S. N. Foner, C. K. Jen, E. L. Cochran, and V. A. Bowers, J. Chem Phys. 28, 351 (1958)
8. G. J. King, F. F. Carlson, B. S. Miller, and R. C. McMillan, J. Chem Phys 33, 1499 (1961)
9. E. Bleaney and K. W. H. Stevens, Repts. Progr. Phys 16, 108 (1953)
10. D. W. Wylie, A. J. Shuskus, C. G. Young, O. R. Gilliam, and P. W. Levy, Phys Rev. 125, 451 (1962)
11. K. Morigaki, M. Fujimoto, and V. Itoh, J. Phys Soc (Japan) 13, 1174 (1958)
12. J. H. Van Vleck and W. G. Penney, Phil. Mag. 17, 961 (1934)
13. M. H. L. Pryce, Phys. Rev. 80, 1107 (1950)
14. A. S. Chakravarty and R. Bersohn, Proceedings of the First International Conference on Paramagnetic Resonance in Israel, Academic Press Inc., New York (1963)
15. R. Bersohn, J. Chem Phys. 29, 326 (1958)
16. R. Sternheimer, Phys Rev. 84, 244 (1951)
17. H. M. Foley, R. M. Sternheimer, and D. Tycko, Phys Rev. 93, 734 (1954).
18. D. R. Hartree and W. Hartree, Proc Royal Soc. (London) A193, 299 (1948) (nitrogen); D. R. Hartree, Proc Camb. Phil Soc. 24, 111 (1928) (sodium⁺, rubidium⁺); D. R. Hartree, Proc Royal Soc. (London) A143, 506 (potassium⁺, cesium⁺); D. R. Hartree, Proc Royal Soc. (London) A141, 282 (1933) rubidium⁺)

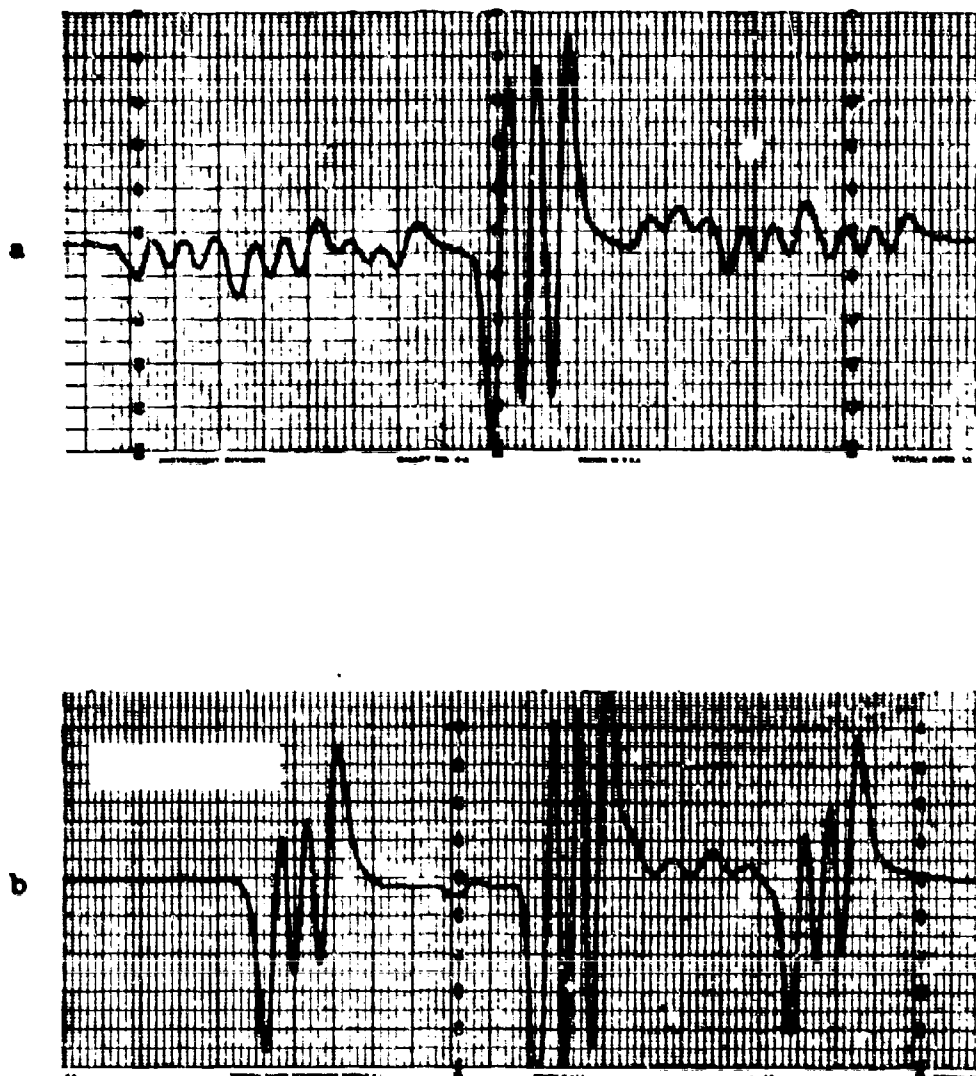


Fig. 1: The first derivative of the electron resonance absorption of nitrogen atoms in sodium azide. (a) With the crystal in an arbitrary orientation showing all the observed nitrogen lines. (b) With lll parallel to the dc magnetic field.

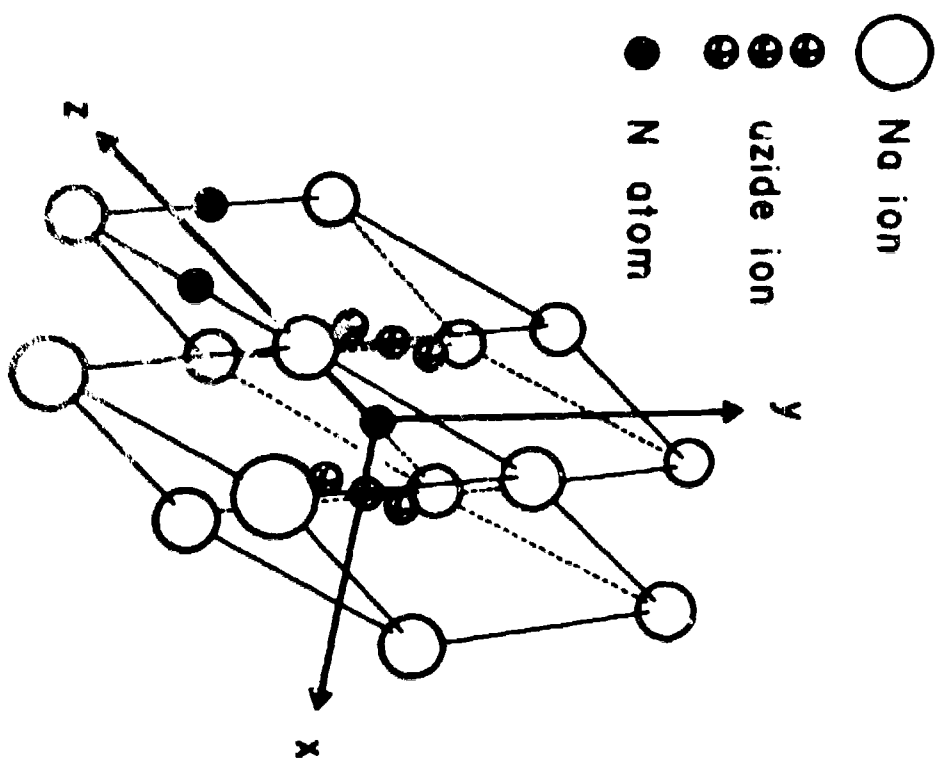


Fig. 2: Two rhombohedral unit cells of NaN₃. The

nitrogen atoms are trapped in the interstitial sites as shown by the solid circles. The coordinates for the spin Hamiltonian are shown for one of the sites.

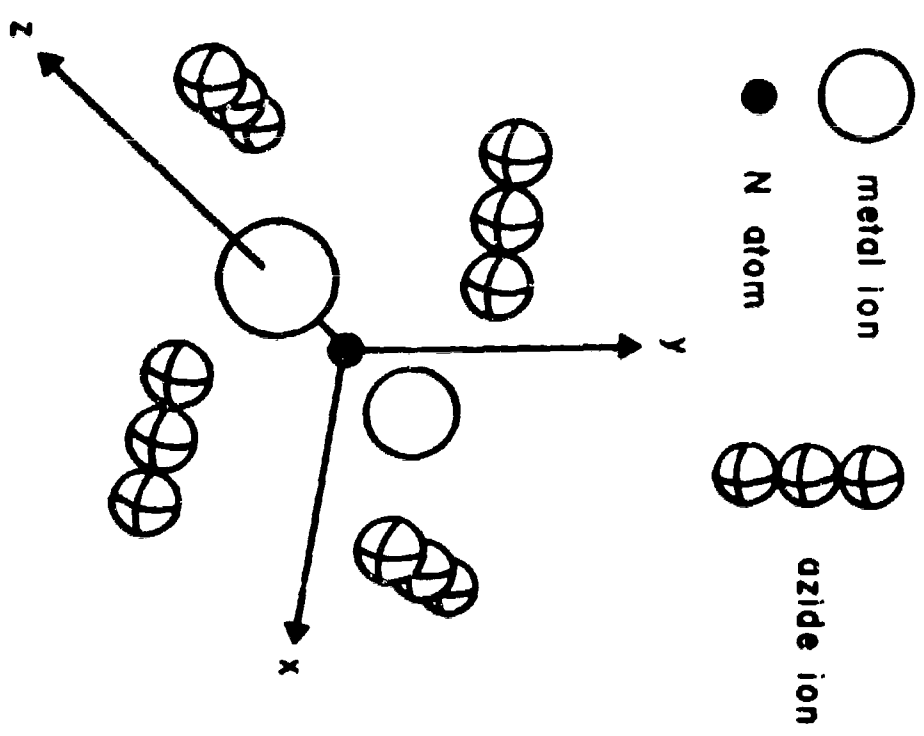


Fig. 3: The local symmetry about the trapped nitrogen in the tetragonal azides.

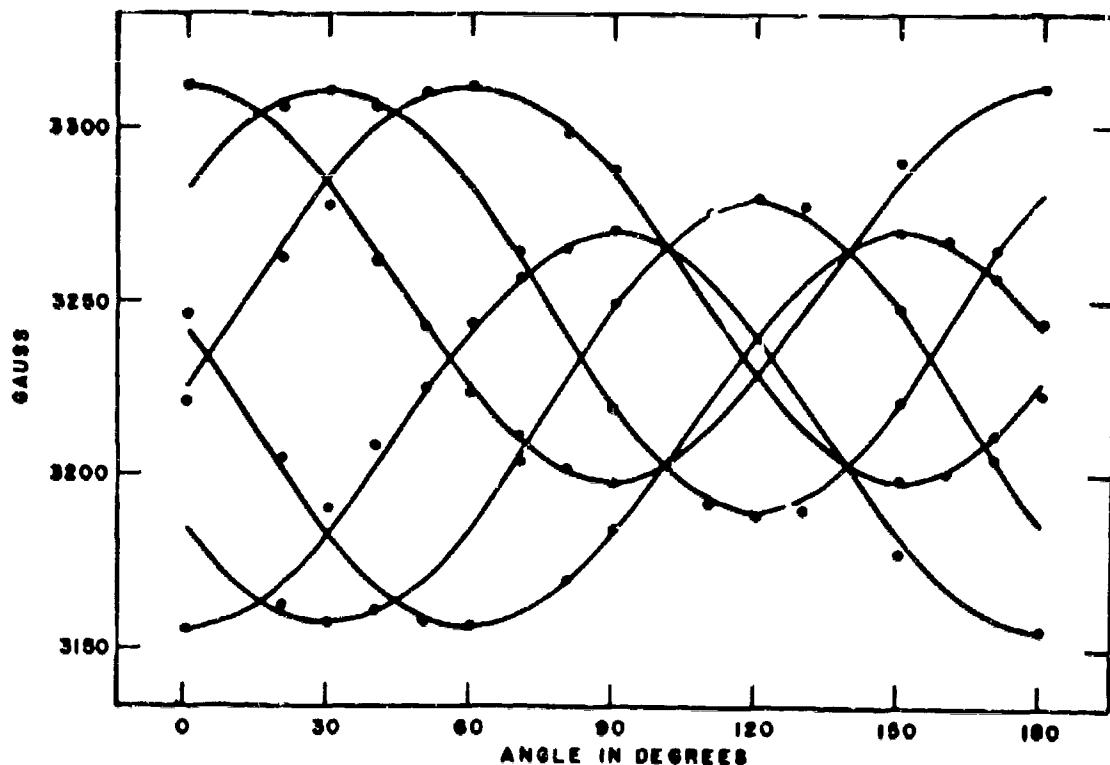
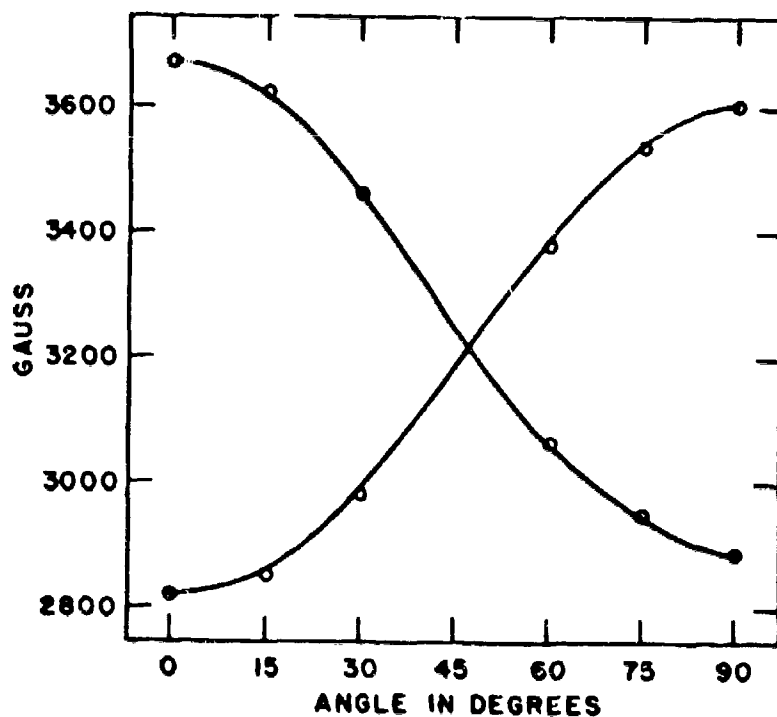


Fig. 4: The angular variation of the fine structure as the crystal of Na^{3+} is rotated in the dc magnetic field. The drawing shows the three similar patterns corresponding to the three orientations of the defect in the crystal.

Fig. 5: The angular variation of the fine structure as the crystal of tetragonal RbN_3 is rotated in the dc magnetic field. Only one of the two sites is plotted.



CARLSON

Table of the experimentally determined values of the spin Hamiltonian parameters for nitrogen atoms in the alkali azides.

	NaN ₃	KN ₃ [*]	RbN ₃	CsN ₃
D(cm ⁻¹)	0.00407 ± 0.0001	0.0143 ± 0.0001	0.0199 ± 0.0001	0.0341 ± 0.0002
E(cm ⁻¹)	- 0.00026 ± 0.00005	- 0.00199 ± 0.00002	- 0.00463 ± 0.00005	- 0.0103 ± 0.0003
g	2.002 ± .0002	2.001 ± .0001	1.997 ± .002	1.992 ± .002
A(cm ⁻¹)	0.00053 ± 0.00001	0.00051 ± 0.00002	0.00053 ± 0.00003	0.00054† ± 0.00005

*Data from reference 10

† Estimated.

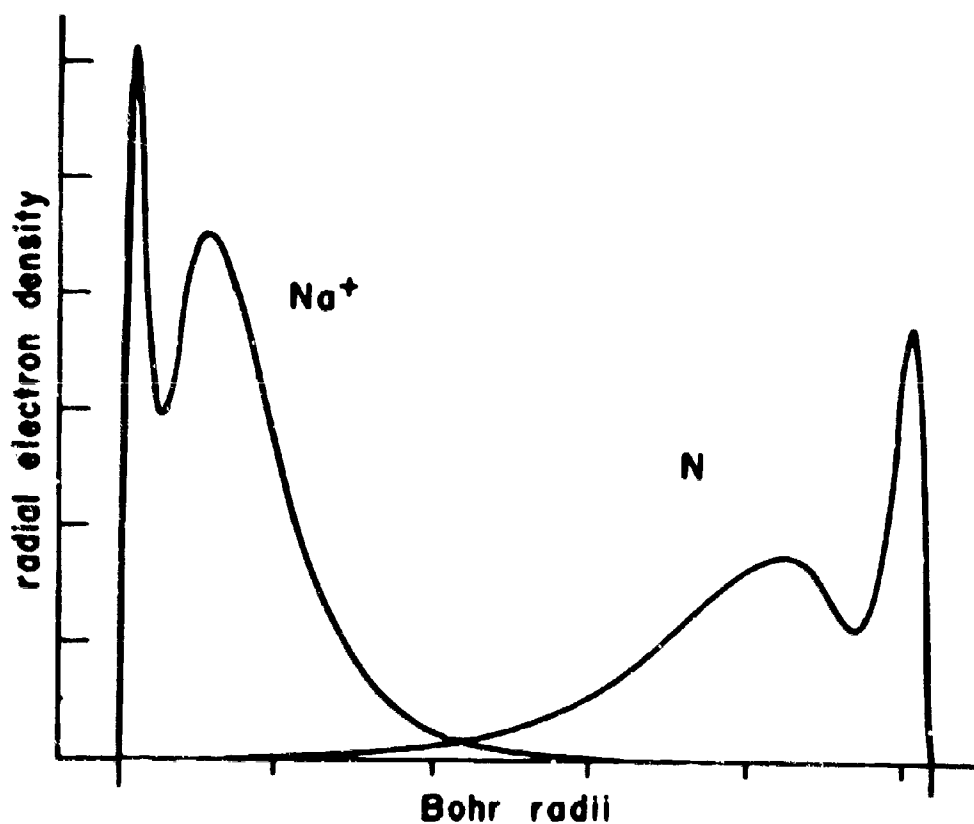


Fig. 6: A plot of the Hartree radical electron densities of atomic nitrogen and ionic sodium. The plots are separated by one half the sodium-sodium distance in sodium azide. Note the small overlapping areas.

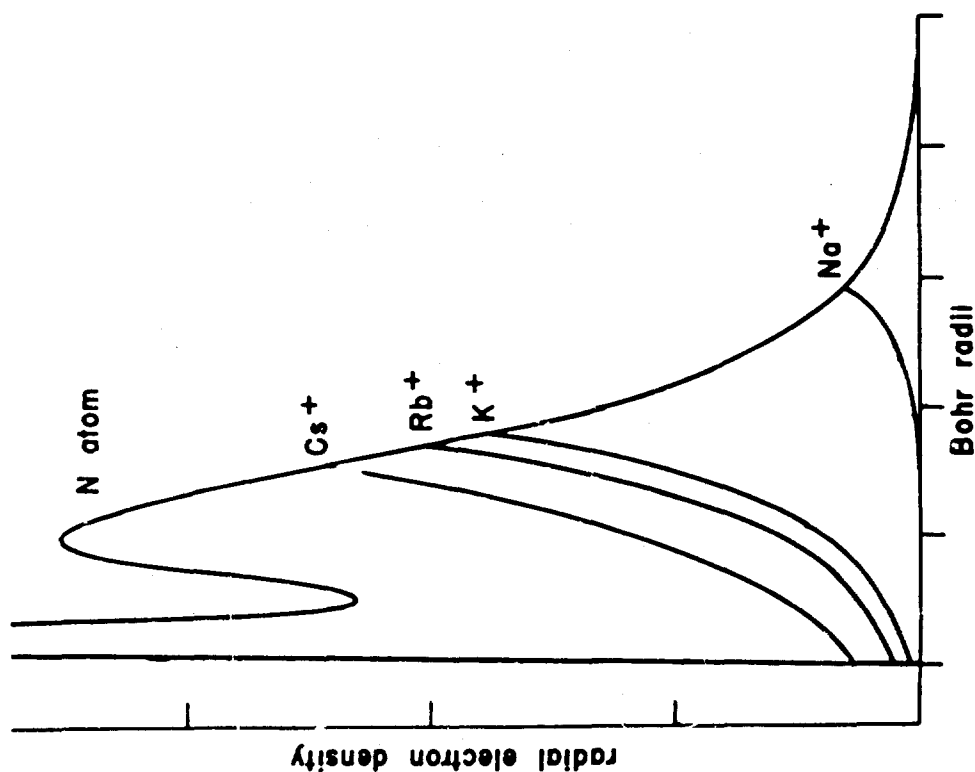


Fig. 7. A graph, similar to figure 6 showing the overlapping areas of electron density for nitrogen in the alkali azides.

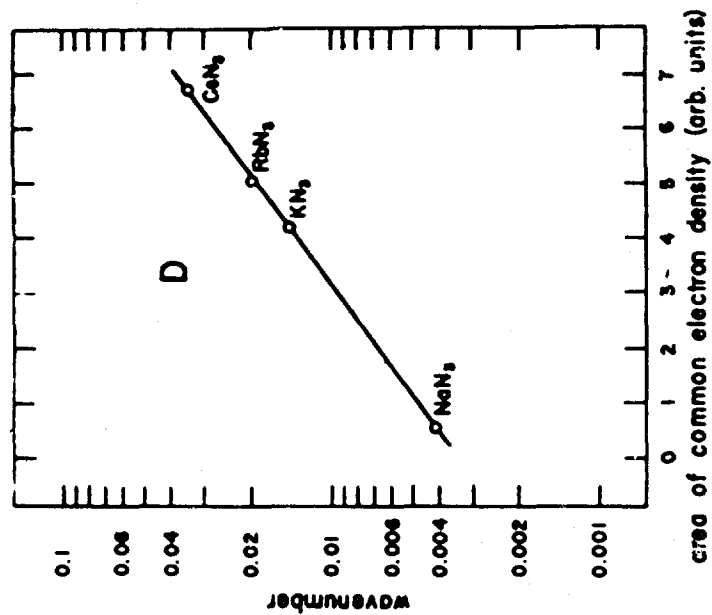


Fig. 8: A semilog plot of D as a function of the area of overlapping electron density.

GAS VELOCITY PROBE FOR MOVING IGNIZED GASES

CHARLES CASON
U. S. ARMY MISSILE COMMAND
REDSTONE ARSENAL, ALABAMA

1. INTRODUCTION

Methods for measuring gas velocity in a flowing plasma may be categorized as average type measurements and local measurements. Betchov and Fuhs (Ref. 1) measured the gas velocity of a plasma jet with a pair of pickup coils, one located upstream from the other one. "Noise" signals from each coil were displayed on a double gun oscilloscope and the record photographed. Gas velocity was estimated by the time delay in the signal between the upstream and the downstream coils. Gourdine (Ref. 2) also used coils but he developed an RF method in which a moving plasma distorted the magnetic field within the coils and this variation was related to the gas velocity. Beasley and co-workers (Ref. 3) used photomultiplier tubes with lenses as the sensors. This method was essentially the same as Betchov's and Fuhs'. Methods which use detectors to observe disturbances at different axial positions to relate time delays to velocity will give, at best, average velocities for wide separations, or inaccurate velocities for very close spacings of detectors. This property is comparable in a way to an "uncertainty principle".

Disadvantages of the above methods are that the data must be photographed point by point and analysed at a later time and also that good spatial resolution is not achieved. Streak camera photography using a focal plane shutter camera as suggested by Pasley (Ref. 4) and Freeman (Ref. 12) will produce the same results. The chief advantage of these "average" measurements is that the plasma flow is undisturbed by the measuring equipment.

Probes immersed in a plasma have been successfully used to measure many plasma properties. Jahn and Grosse (Ref. 5) used paired electrostatic probes to measure characteristics of shock velocities. As before, the arrangement of the probes was axial or time separated. Baker and Hammel (Ref. 6) suggested a new way of measuring plasma

CASON

gas velocities. They investigated the properties of a plasma streaming transverse to a magnetic field. In general the behavior of a moving plasma in the presence of a magnetic field rarely follows the simple classical theory due to a multitude of plasma processes. However, Baker and Hammel demonstrated that when a plasma streams through a transverse magnetic field B , with a velocity v , an orthogonal electric field E , is present. According to simple classical theory a polarization electric field, $\bar{v} \times \bar{B}$, would be generated in the plasma to allow it to pass through the magnetic field. The value of E/B which they observed was in agreement with the plasma velocity as determined by magnetic probes.

Clayden and Coleman (Ref. 7) applied the above method to an arc-heated low density wind tunnel. They inserted a pair of symmetric probes mounted 1 cm apart along the flow radius normal both to the applied D.C. magnetic field and the gas flow to detect the electric polarization field present. A recording galvanometer was used to simultaneously measure the voltage across the probes and the current to the coil. Their analysis indicated a linear variation in polarization voltage with applied D.C. magnetic field. Their study did not estimate the sensitivity or linearity of the equipment and they did not report any special experimental difficulties in applying this apparatus to wind tunnel research.

2. THEORY

Important parameters in the design of velocity probe experiments are: 1) the minimum charge density for application, 2) thermionic emission effects, and 3) minimum detectable gas velocity. A review of standard electrostatic probe theory with modifications to allow for polarization voltage, measurement currents and probe thermionic emission will yield this information.

The theory of electrostatic probes was devised by Langmuir and Mott-Smith (Ref. 8) for a single probe floating in a plasma. It was modified by Johnson and Malter (Ref. 9) to apply to two equal area probes. Figure 1a. depicts an idealized potential diagram across a plasma with a $v \times B$ potential added. Subscripts 1 and 2 are probe index numbers and ϕ denotes the surface work function. The potentials of the two probes with respect to the immediate vicinity plasma are V_1 and V_2 and the electric polarization of the plasma due to a magnetic field is V_B . V_0 denotes the output voltage as would be measured by a voltmeter. The output voltage then has components due to plasma properties, electrode work functions, and applied external magnetic fields.

When an electron is thermionically emitted it has overcome a barrier of ϕ above the Fermi level and leaves with a kinetic energy proportional to the surface temperature. Energy is added as the electron falls through the probe sheath. For probes perpendicular to B field and to the stream velocity, the electron then loses kinetic

CASON

energy proportional to the polarization voltage of the plasma relative to electrons back at its source probe. It next loses kinetic energy proportional to the other probe's potential when collected by it. Then relative to its source electrons, its energy is then further reduced proportional to the work function. Arrows in the figure indicate the path this electron would take as its potential is changed. Collection and emission rates of electrons are assumed not to disturb the plasma. Charges then must be generated or absorbed by the plasma as rapidly as they are drained away or emitted by the electrodes.

Figure 1b. shows an idealized plasma and instrument system. In this figure, i_i and i_e refer to random ion and electron current densities near electrodes of areas A_1 and A_2 . The primed terms correspond to those plasma charges actually collected. Langmuir (Ref. 10) has shown that ion currents may be produced at a hot electrode when it is in an alkali vapor. This method for ion production will be neglected, even for hot probes, since the application is for air, nitrogen, and inert gas plasmas. Thermionic electron current density is represented by I_{th} while conventional circuit current is labeled I . The measurement impedance is called R and the observed voltage is then V_0 .

Complete current expressions are obtained by applying Kirchhoff's law at each of the two electrodes.

$$I = i'_{i1} A_1 - i'_{e1} A_1 + I_{th1} A_1 \quad (1)$$

$$I = -i'_{i2} A_2 + i'_{e2} A_2 - I_{th2} A_2 \quad (2)$$

The primed terms are the electrons and ions collected by each probe. Langmuir (Ref. 8) used the Boltzmann distribution function for the x component of velocity to determine collection currents. When this is done for equations (1) and (2) the potential on each probe is calculated to be,

$$V_1 = -T_{e1} \ln \left(i'_{i1} / i_{e1} \right) - T_{e1} \ln \left(1 - \frac{I - I_{th1} A_1}{i_{i1} A_1} \right) \quad (3)$$

$$V_2 = -T_{e2} \ln \left(i'_{i2} / i_{e2} \right) - T_{e2} \ln \left(1 + \frac{I + I_{th2} A_2}{i_{i2} A_2} \right) \quad (4)$$

when T_e is in electron volts. The random electron current density terms are a convenient grouping of constants from the integration while i_i is the saturated ion current density to each negative (respect to the local plasma) probe. In an idealized measurement situation I is exactly zero because R is ∞ . For this situation the

CASON

floating potential V_f , on each electrode is obtained. Equation (5), written without subscripts, is valid at either electrode.

$$V_f = -T_e \ln(i_f/i_e) - T_e \ln(1 + J_{th}/i_f) \quad (5)$$

Equation (5) has one additional term added to the usual electrostatic probe theory for the probe floating or wall potential. It may be seen from this equation that if J_{th} is sufficiently large V_f may even be positive. This condition is

$$J_{th} > i_{e1} - i_{i1}. \quad (6)$$

Generally V_f is negative except for rarefied plasma flows and hot electrodes.

The potential measured with an infinite impedance meter can be deduced from Fig. 1a.

$$V_0 = \phi_1 - \phi_2 + V_2 - V_1 + vB \quad (7)$$

Effects of absorbed atoms which would modify the work function are neglected in this study since each electrode is assumed to be hot. When equations (3) and (4) are substituted into equation (7) one obtains

$$\begin{aligned} V_0 = & \phi_1 - \phi_2 + vB + T_{e1} \ln(i_{i1}/i_{e1}) - T_{e2} \ln(i_{i2}/i_{e2}) \\ & + T_{e1} \ln\left(1 - \frac{I - J_{th1} A_1}{i_{i1} A_1}\right) - T_{e2} \ln\left(1 + \frac{I + J_{th2} A_2}{i_{i2} A_2}\right). \end{aligned} \quad (8)$$

Equation (8) shows separately the effects on the measured voltage of 1) thermionic emission, 2) thermal gradients, 3) fluctuations of velocity, charge density and temperature, and 4) magnetic fields. Problem areas may be divided into those associated with D.C. and A.C. measurement. As an example the A.C. case will be considered. When a magnetic field of $B_0 \sin \omega t$ is applied, the measurement voltage may be tuned to the angular frequency, ω , which is used. From equation (8) the D.C. terms not affected by drawing current may be dropped thereby leaving the following:

$$\begin{aligned} V_0 = & vB \sin \omega t + T_{e1} \ln\left(1 - \frac{I - J_{th1} A_1}{i_{i1} A_1}\right) \\ & - T_{e2} \ln\left(1 + \frac{I + J_{th2} A_2}{i_{i2} A_2}\right) \end{aligned} \quad (9)$$

Equation (9) is an implicit expression of circuit current. The last two terms on the right of equation (9) give the error voltage due to

CASON

the plasma-probe current drain of the measuring instrument. For low density plasmas these terms need to be evaluated to test for any changes in V_0 due to circuit current effects.

3. EXPERIMENTAL PROCEDURE AND DATA

Steady State Induced Fields

In the first experiments a D.C. magnetic field was placed across a plasma jet perpendicular to the flow. Clayden and Coleman (Ref. 7) used an iron core electromagnet greater than 100 gauss but found hysteresis in their field due to the iron. Therefore, an air core Helmholtz coil was used in this study to produce a magnetic field linear with current. The magnetic induction, B , in the center of the coil system is

$$B = 8.99 \times 10^{-7} nI/a, \quad (10)$$

where B is in Weber/meter², n is the number of turns per coil, I the current in amps and a is the radius in meters of a coil.

The magnet consisted of 8 turns of copper tubing wound on a radius of 16.5 cm. The magnetic field was calculated by equation (10) to be 0.436 gauss per ampere. Measurements of the induced magnetic field was 0.443 gauss per ampere which suggested that a partial turn was generated from a return lead. Since agreement is within 2%, the calculated value was used in the analysis. Field strengths of 85 to 100 gauss were used in the experiments.

Source of the plasma gas was a small D.C. plasma generator. The cathode was water cooled tungsten and the anode nozzle was brass with a 1/8" diameter throat. Power for the plasma generator came from a bank of 10, 24 volt truck batteries in series. The arc was initiated by means of an in-line Tesla coil arrangement. The minimum power delivered to the electrodes was 325 watts while the maximum was 1,350 watts. Argon was the gas used in all cases.

In each experiment the vacuum reservoir (13,000 cubic feet volume) was pumped down to a pressure less than 1 micron. Then the test gas was injected at a constant pre-determined flow rate. Next the arc was initiated by the Tesla coil and then the magnetic field was applied. Polarization voltage, a measure of velocity, was graphed as a function of the output voltage of a thermocouple vacuum gage attached to the vacuum reservoir. It was at this point that an undesirable feature of the D.C. system became evident. The sampling probes became white hot and began to thermionically emit electrons. One probe was in the center of the flow and one was near the side thus producing important temperature differences in the probe metal.

By alternately turning the magnet on and off a "zero" base line for measured probe voltage could be drawn on the x-y plotter in

CASON

addition to the total signal. Figure 2a. is a typical data plot. The pressure is plotted on the abscissa and probe voltage is on the ordinate; in this case the gas flow was argon at 0.35 gm/sec. Figure 2b. is a plot of the gas velocity for several flow rates for a constant power setting. Uncertainty in the reading was ± 170 meters/sec due to velocity fluctuations in the flow of the plasma. Stray D.C. effects in the data, such as plasma thermal gradients, thermionic emission rate variations at each probe, etc., are shown in several terms of equation (8) and appear on the "base line" in Fig. 2a. These effects can lead to very large errors unless they are taken into account or properly eliminated. Isolation amplifiers must be used to measure the polarization voltage because of probable interaction with the arc power supply through ground loops. This limits system response time.

The magnet had a noticeable influence on the plasma. When it was turned on the plasma jet could be seen to deflect at a shallow angle; it would then return to its original direction when the magnet was turned off. This effect is not predicted according to theory dealing with first order interaction effects but it does obviously modify the gas velocity to be measured. The same experimental procedure was then followed except with a weak A.C. field in order to avoid stray D.C. effects and operate with weaker fields.

A.C. Induced Fields

A new coil was made from #18 magnet wire. Fifty-one turns were wound on a radius of 8.25 cm. Equation (10) shows the reduced magnetic induction to be 5.5 gauss per ampere. Magnet current was supplied from a commercial high fidelity amplifier accepting a selected frequency from a signal generator. Probe voltage fed an audio interstage transformer with a 60 K Ω impedance and a 3.5 K Ω D.C. resistance. This transformer was used to isolate the amplifier from the probes to avoid ground loops through the arc power supply. An amplifier tuned to 2 KC was used for the voltage measurement. It has a sensitivity of 1 microvolt and a gain of 10^4 .

An x-y plotter was used to record RMS polarization voltage versus RMS magnet current. The data plot in Fig. 3a. has curvature due to the nonlinear response of the rectifier filter in the plotter for the signal voltage used. Only the first part of the data from this run is shown. Two "error" voltages are noted. One is the smooth curve which was taken with no plasma being present. The other, a background signal, is evident from Fig. 3b. The top curve in this figure is a time sweep of output (note displaced zeros for each curve). The level of the background 2 KC signal field is in the order of 2 to 5 millivolts. The other two curves reflect an attempt to determine linearity of the response of the plasma polarization to the A.C. field. Linearity was tested by applying a 1 KC magnetic field to the plasma and measuring response at 2 KC (2d harmonic) and again at 666.6 cycles when measured at 2 KC (3rd harmonic). Changes in background level would be produced by changes in subharmonic polarization fields if the

CASON

response was non linear. No changes were present as compared to the normal background drift seen in Fig. 3b. It appears that 2 KC was not the best choice in frequency because of an apparent fluctuation present in the plasma but was required by the limitations of the amplifier.

Upon reducing the data a background "error voltage" was taken to be a constant 0.003 volts. Provision was not made to monitor its fluctuations as seen from Fig. 3b.; therefore scatter should result from this source of error. Gas velocities were calculated and plotted in Fig. 4a. Error signals proportional to the magnet current were estimated by a least square fit to the equation

$$v = m B + v_0, \quad (11)$$

where m is the proportionality constant due an error, v_0 the true gas velocity, and v the gas velocity indicated by adding in an error assumed proportional to the magnet current. The results of the data reduction showed the velocity to be 2,160 meters/sec as the gas velocity with a root mean square deviation of ± 137 meters/sec.

The average gas velocity may be estimated from mass flow rate and electric power. The enthalpy for 0.19 gms of argon per second heated at 500 watts is 2.63×10^6 calories/mole. For pressures equal to or greater than 0.01 atmospheres this enthalpy would result in a stagnation temperature of approximately 5,000°K (Ref. 11). From the relation $1/2 m \bar{v}^2 = 3/2 kT$ a velocity of 1750 meters per second is calculated.

An alternate way of making the same type of crude estimate is from the power equation for a flowing gas. Power is $1/2 \rho v^3 A$ but \dot{m} is $\rho v A$ when ρ is density, A is area and \dot{m} is flow rate. This gives

$$\text{Power} = 1/2 \dot{m} v^2. \quad (12)$$

For the example cited, equation (12) indicates a gas velocity of 2,200 meters/sec. These two results although rough, nicely bracket the estimates made from both the D.C. and A.C. data except for very low free stream pressures.

The phototube system planned to measure gas velocity for an independent comparison was found to exhibit the same difficulties as reported by Freeman, and others (Ref. 12). At best, their measurements indicate a gas velocity of one to two thousand meters per second. Freeman and co-workers also found that several disturbances such as those produced by temperature and total pressure fluctuations may propagate in a subsonic plasma jet at different velocities.

4. CONCLUSIONS

Comparison between the velocity obtained by the D.C. and A.C.

CASON

methods was quite satisfying as shown in Figs. 2b. and 4a. The plasma jet was found to exhibit rapid fluctuations in gas velocity as seen in both the D.C. and A.C. data. Also magnetic fields in the order of 100 gauss were found to slightly disturb the gas flow. For probes in a hot flow where a D.C. field is used it was found best to alternate the field between "on" and "off" to determine extraneous D.C. effects. When A.C. measurements are employed a field of the order of 5 gauss is all that is necessary since background "error voltages" can be easily eliminated at this level. It is expected that flows other than plasma jets would exhibit less fluctuations at the frequency of the tuned voltage amplifier. The lowest saturation ion current obtained to date was 1.3 ma while the minimum current required by the measurement equipment is of order of microamps. From equation (9), using 1.500° K for tungsten probes, it was found that a few microamps did not make a measurable modification of voltage present.

6. REFERENCES

1. Betchov, R., and Fuhs, A.E., "A Survey of Diagnostic Methods for Ionized Gases Using Magnetic Fields", TRD-169(3153)TR-1, Aerospace Corp., Contract AF 04(695)-169, September 1962.
2. Gourdine, Meredith C., "A Technique for Making Local Measurements of the Conductivity and Velocity of a Plasmajet", PLR-71, Plasmatyne Corp., 22 June 1960.
3. Beasley, W.D., Brooks, J.D., and Barger, R.L., "Direct Velocity Measurements in Low-Density Plasma Flows", NASA TN D-1783, May 1963.
4. Easley, Ronald L., "Diagnosis of Plasma Cylinders by Angular Scattering of Microwaves", RR-TR-63-14, U. S. Army Missile Command, 1 May 1963.
5. Jahn, R.G., and Grosse, F.A., "Response of Electrostatic Probes to Ionized Gas Flows in a Shock Tube", Phys of Fluids, Vol. 2 #4, 469-470. July-August 1959.
6. Baker, D.A., and Hammel, J.E., "Demonstration of Classical Plasma Behavior in a Transverse Magnetic Field", Phys. Rev Letter 8, 157 (15 Feb 62).
7. Clayden, W.A., and Coleman, P.L., "Distribution of Electron Density and Temperature in an Arc-Heated Low-Density Wind Tunnel", R.A.R. D.E. Memorandum (B) 57/63, Royal Armament Research and Development Establishment, Fort Halstead, Kentz, October 1963.
8. Langmuir, and Mott-Smith, (A series of Articles Concerning Electrostatic Probe Theory and Experiments). General Electric Review 27, 449, 538, 616, 762, 810 (1924).

CASON

9. Johnson, E.O., and Malter, L., "A Floating Double Probe Method for Measurements in Gas Discharges", Phys. Rev. 80, 58 (1950)..
10. Langmuir, I., and Kingdon, K., "Thermionic Phenomena Due to Alkali Vapors Part I and Part II", Phys. Rev. 21, 380 (1923).
11. Braunschwig, B., "Plasma Properties", Proceedings of the ASME Meeting on Thermal Properties, Purdue, 1959.
12. Freeman, M.P., Li, S.U., and Jaskowsky, W.V., "Velocity of Propagation and Nature of Luminosity Fluctuations in a Plasma Jet", Jour. Appl. Phys. 33, 2845, (1962)

CASON

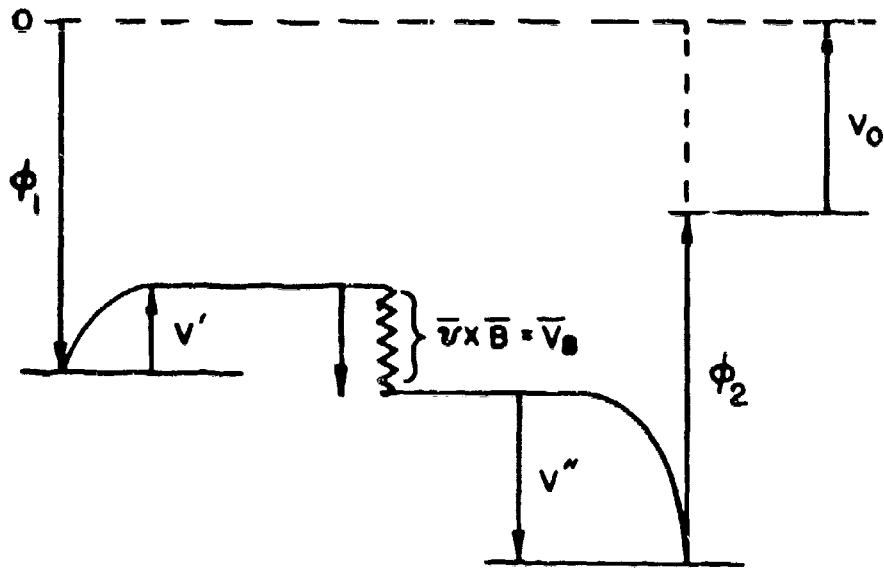


FIG. 1a. PLASMA POTENTIAL DIAGRAM

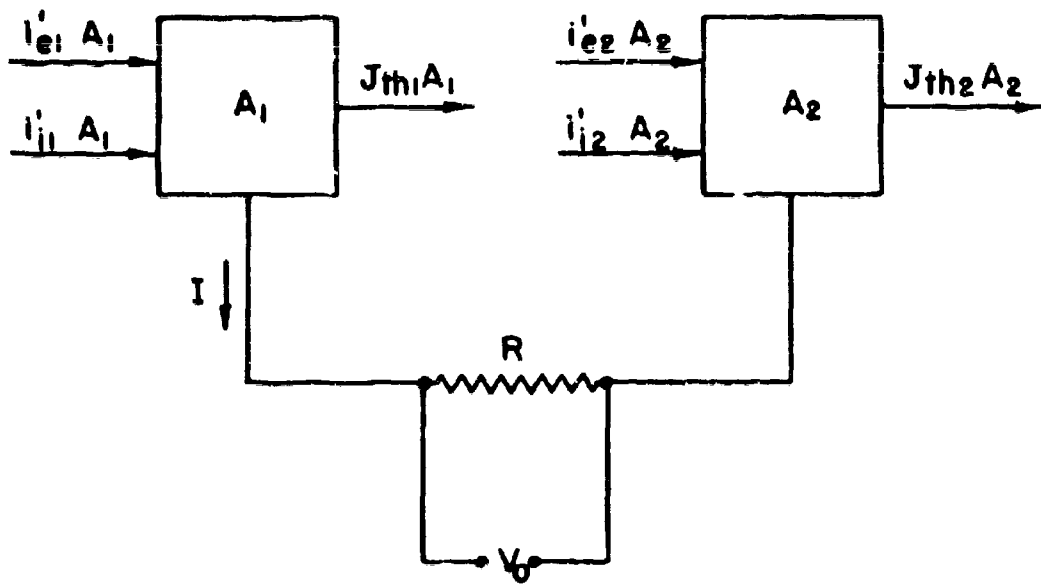


FIG. 1b. GAS VELOCITY PROBES AND VOLTAGE DETECTOR SHOWING CURRENTS

CASON

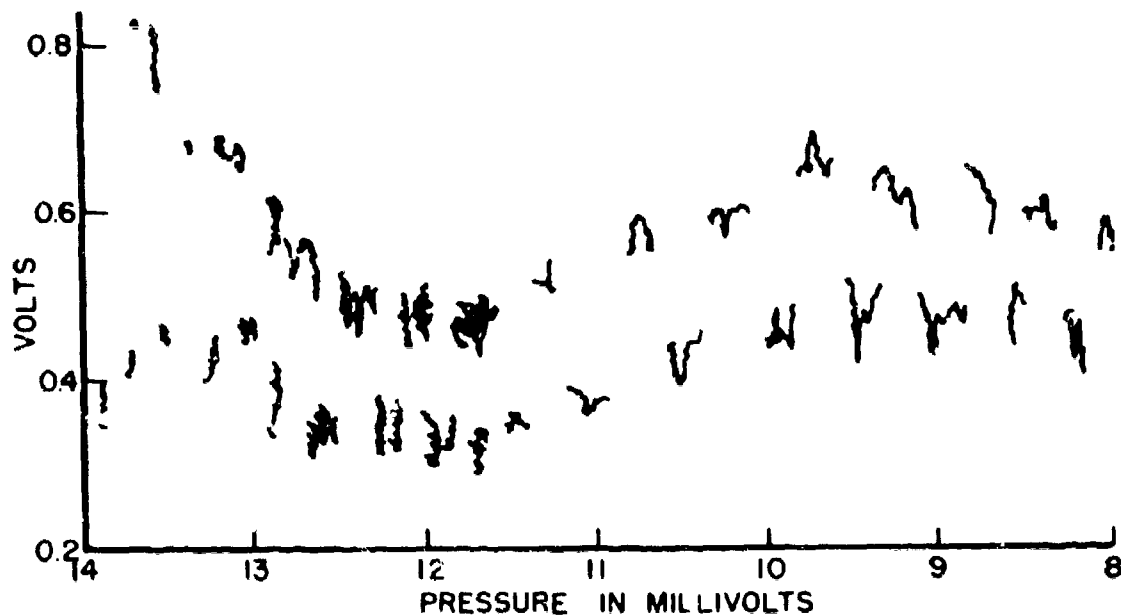


FIG. 2a. D.C. VOLTAGE VS PRESSURE INDICATION IN MILLIVOLTS
BOTTOM TRACE D.C. BACKGROUND
TOP TRACE D.C. BACKGROUND + D.C. POLARIZATION

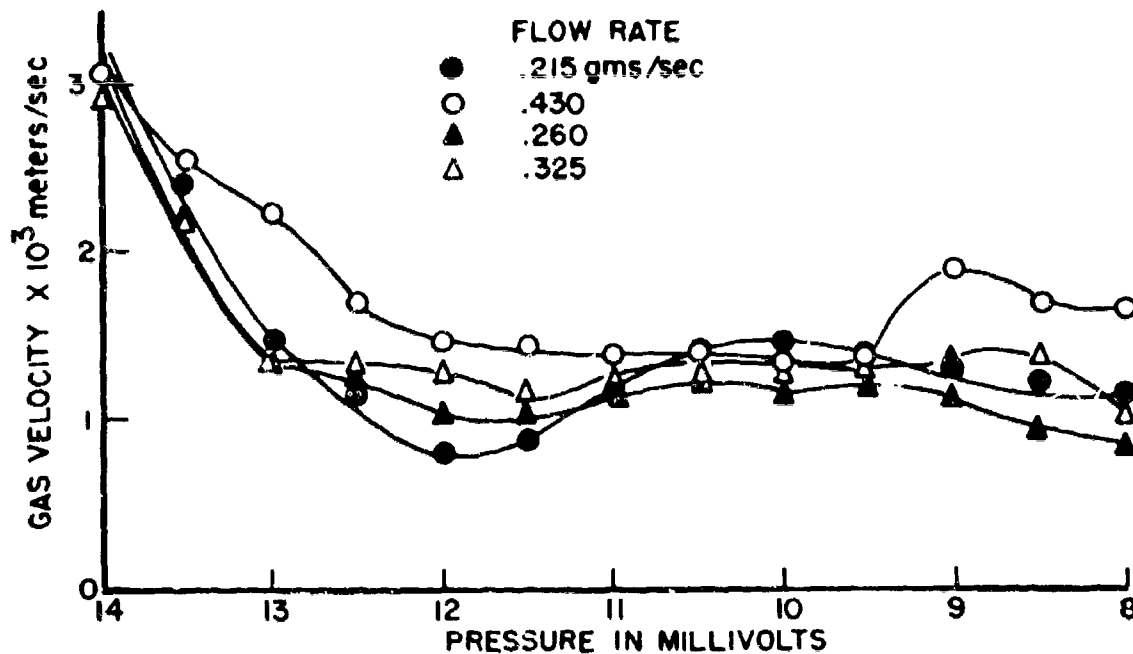


FIG. 2b. GAS VELOCITY VS. PRESSURE (REDUCED DATA FOR SEVERAL FLOW RATES)

CASON

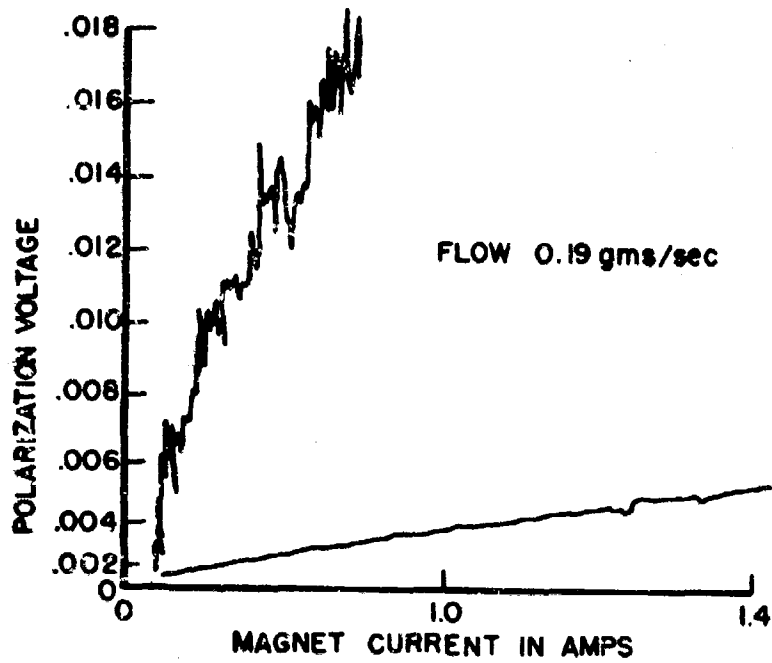


FIG. 3a. POLARIZATION VOLTAGE VS. MAGNET CURRENT

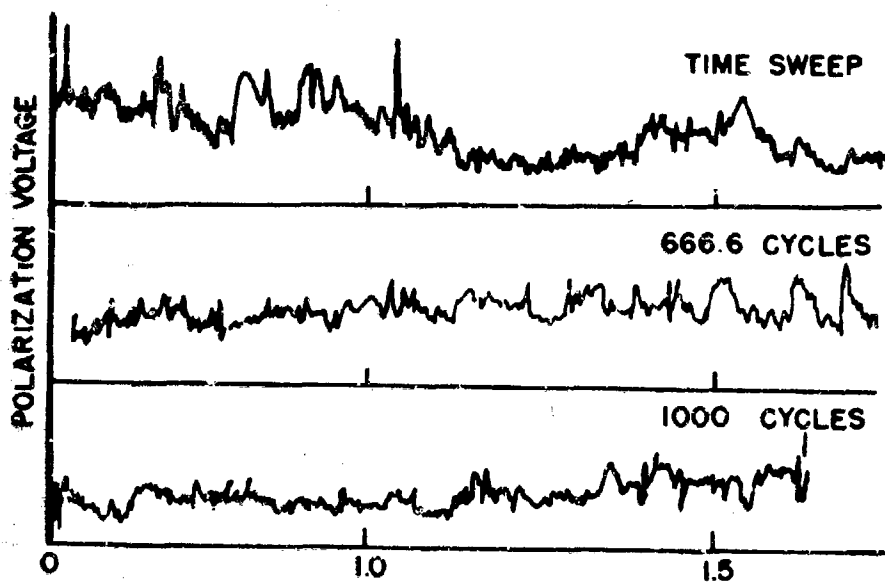


FIG. 3b. POLARIZATION VOLTAGE VS. TIME OR SUBHARMONIC MAGNET CURRENT

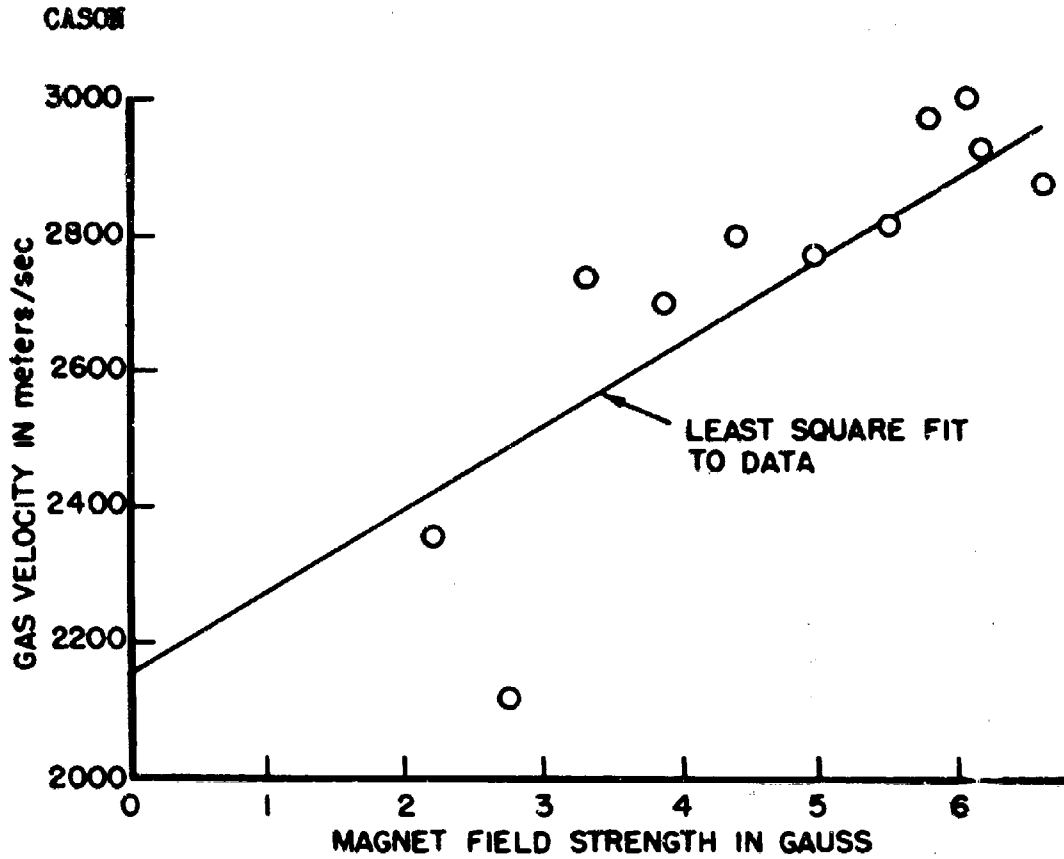


FIG. 4a. REDUCED DATA GAS VELOCITY VS. MAGNET FIELD

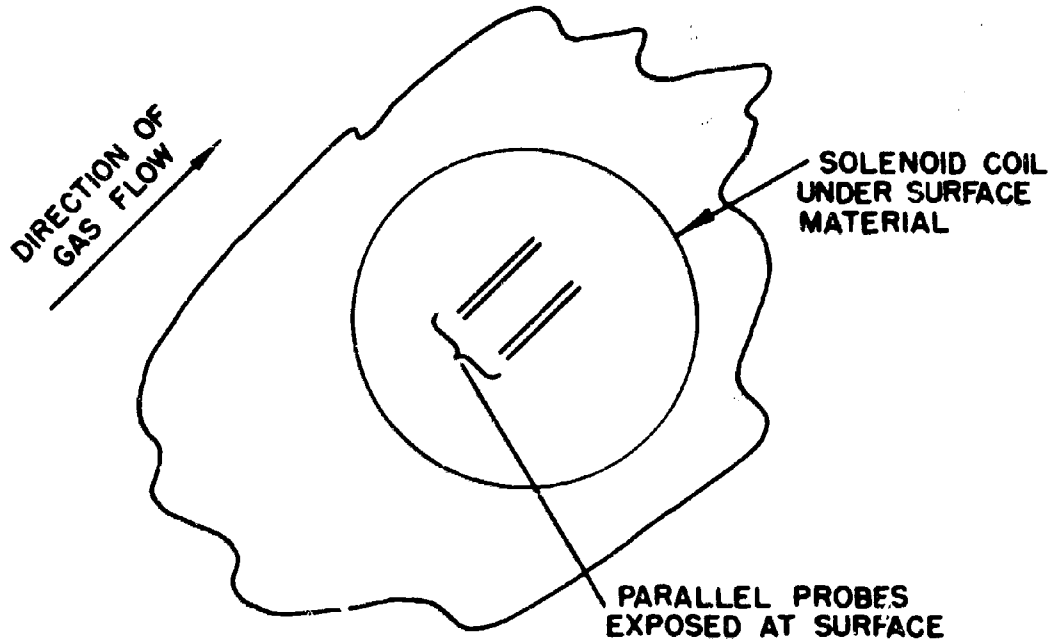


FIG. 4b. STATION FOR BOUNDARY LAYER MEASUREMENT

NITROGEN-15 TRACER STUDIES OF THE
NITROLYSIS OF HEXAMETHYLENETETRAMINE

THOMAS C. CASTORINA and JOSEPH R. AUTERA
EXPLOSIVES LABORATORY, PICATINNY ARSENAL
DOVER, NEW JERSEY

Introduction

The nitrolysis of hexamine to HMX/RDX has been studied in these laboratories (1) and using essentially the same conditions of the Holston Ordnance Works (HOW) process (2), which in turn is based on the Bachmann method (3), similar yields of products were obtained. By including a small quantity of paraformaldehyde $(\text{CH}_2\text{O})_x$ in the reaction mixture (hereafter called the PA process) the yield of mixed products was increased by approximately 10%. A study was carried out in which either hexamine or $(\text{CH}_2\text{O})_x$ was tagged with carbon-14 and the distribution of activity in the cyclic methylene nitramine products was traced during the course of nitrolysis (4). It was concluded that in the first reaction stage, methylene groups from hexamine and $(\text{CH}_2\text{O})_x$ form a common pool for the formation of DPT molecules. Similarly, in the second stage methylene groups from DPT and $(\text{CH}_2\text{O})_x$ also form a common pool for the formation of HMX and RDX molecules. Therefore, the relative ratio of HMX to RDX is not controlled by the selective cleavage of a large molecule but is more than likely influenced by the particular conditions controlling the recombination of a common methylene containing fragment. It was thus postulated that the $(\text{CH}_2\text{O})_x$ increases the concentration of methylene containing fragments which exist in precursors to HMX/RDX. The increase in yield of combined HMX/RDX products, resulting from the use of $(\text{CH}_2\text{O})_x$, can therefore be accounted for on this basis.

The methylene content is not the only criterion in controlling yield or composition of product. It is obvious that the processes involving amino groups and the formation of the nitramine structure are equally important for understanding the chemistry of hexamine nitrolysis. An indirect attempt was made to ascertain the extent to which the ammonium nitrate participates in the formation of HMX/RDX in the PA process (4). Carbon atom equilibration was found independent of the amino-nitrogen concentration. On the basis of this observation it was concluded that $(\text{CH}_2\text{O})_x$ does not react independently with ammonium nitrate to form HMX or RDX. In a more direct manner,

CASTORINA and AUTERA

by using ammonium nitrate tagged with N-15 in the amino-nitrogen position, Bachmann (5) carried out some exploratory work to determine the role played by the ammonium radical in the formation of RDX (together with HMX as the minor constituent). The results obtained were complicated by the observed isotopic exchange of amino-nitrogens in hexamine and ammonium nitrate. In spite of the complications introduced by interchange, it was concluded that the formation of RDX involves more interaction with ammonium nitrate (37%) by exchange or incorporation or both, than does the formation of HMX (15%).

These results demonstrated the possibility of also obtaining some information directly on the participation of ammonium nitrate in the PA and HOW processes of HMX formation. Experiments were therefore conducted along lines similar to those described in Reference (4) with the exception that products were assayed mass spectrometrically for N¹⁵/N¹⁴ ratios.

Results and Discussion

Since the products DPT, HMX and RDX contain two types of nitrogen $-\text{CH}_2-\overset{\text{N}}{\text{N}}-\text{NO}_2$, the method for the conversion of amino-nitrogen to N₂ for mass spectrometric analysis was checked for specificity. Standard samples of HMX and RDX were prepared from (CH₂O)_x and ammonium nitrate of known N-15 atom percent as the sole source of amino-nitrogen. A sample of the N¹⁵H₄NO₃/HNO₃ used in this study, was included for analysis to check the possibility of exchange taking place between the nitrate-nitrogen and amino-nitrogen prior to the addition of reactants in the HOW and PA processes. The N-15 atom percent in all of the compounds analyzed mass spectrometrically was found to be within $\pm 0.2\%$ of the value for the N¹⁵H₄NO₃ reference standard. These results clearly establish the lack of nitro and/or nitrate-nitrogen participation in the generation of elemental amino-nitrogen, induced either during analysis or by exchange in a nitric acid medium. Therefore, it can be concluded that any interchange of nitrogen atoms that may be observed to take place during the nitrolysis of hexamine with NH₄NO₃/HNO₃ involves amino-nitrogens only.

Bachman (5) observed exchange to take place between hexamine and ammonium nitrate in acetic acid solution and assumed that exchange took place during nitrolysis. The process of exchange in inert media need not necessarily take place in reaction mixtures. Although hexamine and formaldehyde were found to exchange in an acidic aqueous medium (6), (CH₂O)_x does not exchange with hexamine under conditions of nitrolysis (4). Therefore, to determine whether the exchange reported by Bachman (5) does indeed complicate the study of the PA and HOW processes with N-15, exchange was studied under actual conditions of nitrolysis.

The atom percentages of N-15 found in the hexamine are listed in Table I. The average percent exchange, based on 2.4 atom % N-15 calculated for the complete equilibration of hexamine and ammonium nitrate amino-nitrogens, is seen to be independent of the amount of (CH₂O)_x included in the reaction mixture. The chief

significance of these results is that the rate of exchange of amino-nitrogens is faster than the rate of reaction during the nitrolysis of hexamine in the presence or absence of $(\text{CH}_2\text{O})_x$. It may be helpful at this point to emphasize the difference between "exchange" and "reaction" processes. Although exchange does proceed by means of reactions, the structural configuration of the interchanging species remain unaltered at the steady state of dynamic equilibrium. "Reaction" on the other hand proceeds in one direction only with the initial constituents transformed to different entities at a "static" state of equilibrium.

Table I

Amino-Nitrogen Exchange ^(a) Between Hexamine and NH_4NO_3 ^(b) In a Nitrolyzing Medium				
	Expt. No.	Replicate No.	Atom % N-15 Found In Hexamine ^(c)	Average % Exchange ^(d)
Without $(\text{CH}_2\text{O})_x$	1	1	0.91	24
		2	0.84	
	2	1	0.87	
		2	0.73	
With $(\text{CH}_2\text{O})_x$	1	1	0.96	24
		2	0.92	
	2	1	0.73	
		2	0.72	

(a) All reactants added simultaneously and the mixture quenched with CCl_4 after 2 minutes.

(b) Added as the nitric acid solution and containing 7.6 atom % N-15.

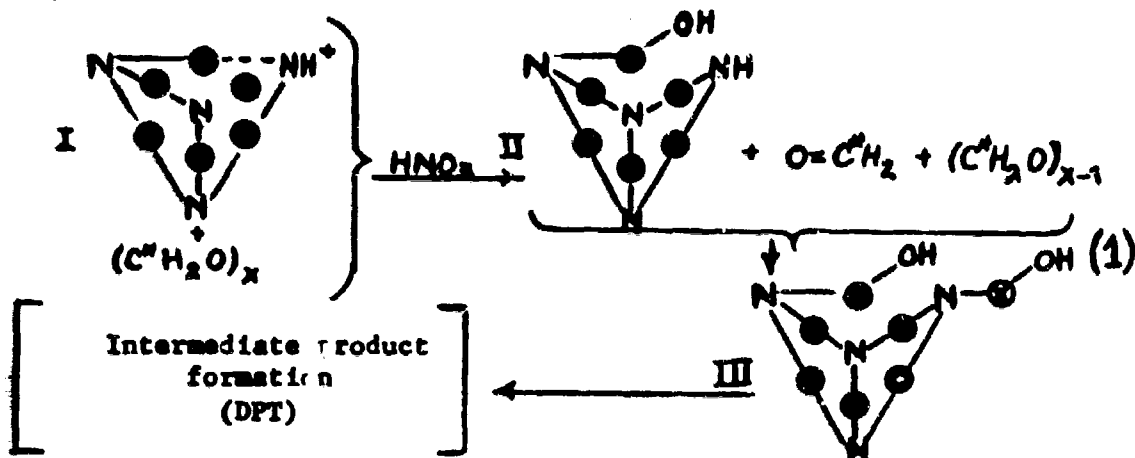
(c) Isolated as the mercuric chloride salt.

(d) Based on 2.4 atom % N-15 calculated for 100% exchange.

The question arises as to why the amino-nitrogens and not the methylene groups in the hexamine molecule exchange under conditions of nitrolysis. Exchange reactions are usually ionic and faster than metathesis reactions. Therefore, conditions conducive to exchange in competition with reactions would involve interchange between ionic species. Since $(\text{CH}_2\text{O})_x$ cannot be considered ionic in an acidic medium, even though hexamine is, the possibility of exchange is minimized and participation via reaction is favored (4). By way of inference only, the representation given below indicates the chemically equivalent nature of the methylene groups that can exist in a nitrolyzing medium. In an acidic medium hexamine exists as a protonated ion with a labile $-\text{CH}_2\text{---NH}^+$ bond. In an oxidizing medium this bond is subject to attack together with the bonds in the polymeric chain of $(\text{CH}_2\text{O})_x$ as shown in Diagram 1. In this mechanism, the implication is made that the reaction involving the fragmentation of $(\text{CH}_2\text{O})_x$ takes place irreversibly prior to equilibration and subsequent reaction leading to intermediate product formation. This is based on the experimental evidence given in Reference 4. Hexamine was isolated free of activity, indicating that III does not exist in

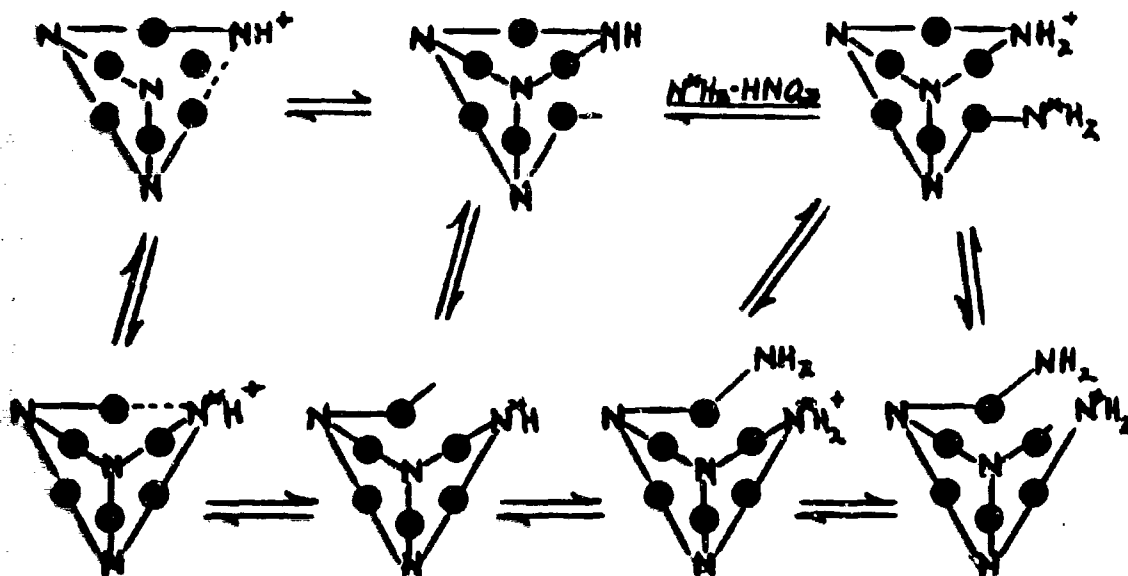
CASTORINA and AUTERA

equilibrium with I but proceeds directly to form intermediate products. The configuration III may be considered the "common pool of methylene groups" to which reference was made in the work with carbon-14.



Where $\text{---}\bullet\text{---} \equiv \text{-CH}_2\text{-}$; $\text{C}^* \equiv \text{C}^{14}$ and $\text{---}\oplus\text{---} \equiv \text{-C}^{14}\text{H}_2\text{-}$

The path of amino-nitrogen exchange between hexamine and ammonium nitrate must undoubtedly involve both species in an ionic metastable state. The charged amino group of the metastable species, I, can then interchange with its equivalent nitric acid salt of ammonia as follows:



Where $\text{---}\bullet\text{---} \equiv \text{-CH}_2\text{-}$; $\text{N}^* \equiv \text{N-15 enriched}$; and $\text{N} \equiv \text{unenriched N-15 (i.e., in natural abundance)}$.

This implies that an exchange of amino-nitrogens takes place before nitrolysis and even before $\text{R-CH}_2\text{-}$ goes to $\text{R-CH}_2\text{-OH}$ or some equivalent irreversible substitution, in which case the rate of exchange is

CASTORINA and AUTERA

necessarily faster than the rate of reaction ($R =$ residual hexamine molecule).

According to the above suggested process of amino-nitrogen exchange, any DPT isolated from the addition of N-15 ammonium nitrate during the first stage of nitrolysis should contain N-15 enrichment corresponding to 100% exchange. This, of course, is based on the fact that all of the amino-nitrogens in hexamine are chemically equivalent, having a common $(CH_2)_3N$ configuration. The experimental atom-percent N-15 values found in the DPT are listed in Table II. As expected, these values are identical to the 2.4 atom % N-15 calculated for 100% exchange of hexamine and ammonium nitrate-amino nitrogens. Here again, exchange is shown to be independent of the $(CH_2O)_x$ concentration.

Table II

NH_4NO_3 (a) Exchange In The Formation of DPT

	Expt No.	Replicate No.	Atom % N-15 Found In DPT	Average % Exchange (b)
Without $(CH_2O)_x$	1	1	2.5	102
		2	2.5	
	2	1	2.4	
		2	2.4	
With $(CH_2O)_x$	1	1	2.4	102
		2	2.4	
	2	1	2.5	
		2	2.5	

(a) Added as the nitric acid solution and containing 7.6 atom % N-15.

(b) Based on 2.4 atom % N-15 calculated for 100% exchange of hexamine and ammonium nitrate-amino nitrogens.

The incompleteness of the exchange shown in Table I can, therefore, be attributed to the mode of addition and the time of contact of the reactants. Because complete equilibration of the four amino-nitrogens occurs prior to reaction, no conclusion can be drawn concerning the extent, if any, to which the ammonium nitrate directly participates in the formation of the DPT.

To investigate the step-wise role of ammonium nitrate in the HOW process the $N^{15}NH_4NO_3/HNO_3$ was added during the second stage of nitrolysis and the HMX/RDX was analyzed for N-15 enrichment. Thus, the formation of HMX/RDX from DPT with any N-15 enrichment can be attributed solely to the participation of the ammonium nitrate added during the second stage of nitrolysis. Assuming that all of the amino-nitrogens, including the ones with nitro-substitutions, exchange during this second stage, the atom-percent N-15 calculated for HMX/RDX is 2.48%. The actual atom-percent N-15 found in HMX and RDX as listed in Table III are seen to be approximately half the calculated value, 2.48%.

The amino-nitrogens in DPT contain two types of substitutions, $(CH_2)_3N$ and $(CH_2)_2N-NO_2$ where only the former is, of course, common to

CASTORINA and AUTERA

hexamine and shown to exchange completely. The observed incomplete exchange during the second stage of nitrolysis may, therefore, be attributed to the nitrosubstituted amino-nitrogen which fails to exchange.

Table III

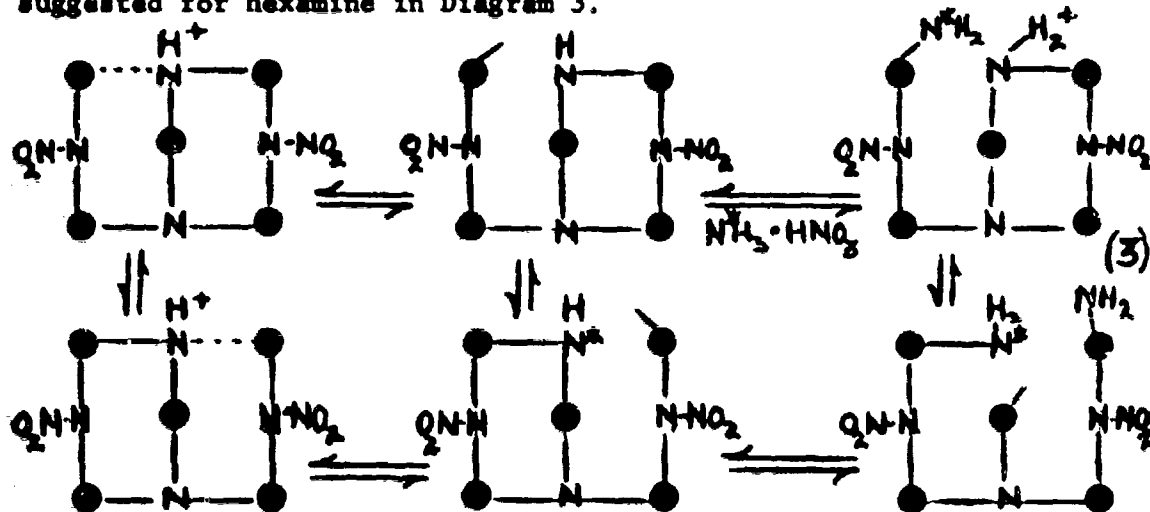
NH_4NO_3 (a) Participation In The Formation of HMX/RDX
In The Absence of $(\text{CH}_2)_3$

Expt No.	Replicate No.	Atom % N-15 Found		Avg % NH_4NO_3 Participation	
		HMX	RDX	HMX(b)	RDX(b)
1	1	1.3	1.5	0	6
	2	1.5	1.5		
2	1	1.2	1.2	0	6
	2	1.2	1.4		

(a) 7.6 atom % $\text{N}^{15}\text{H}_4\text{NO}_3$ added as the nitric acid solution during the second stage of nitrolysis.

(b) HMX and RDX as derived from NH_4NO_3 amino-nitrogens solely due to reaction.

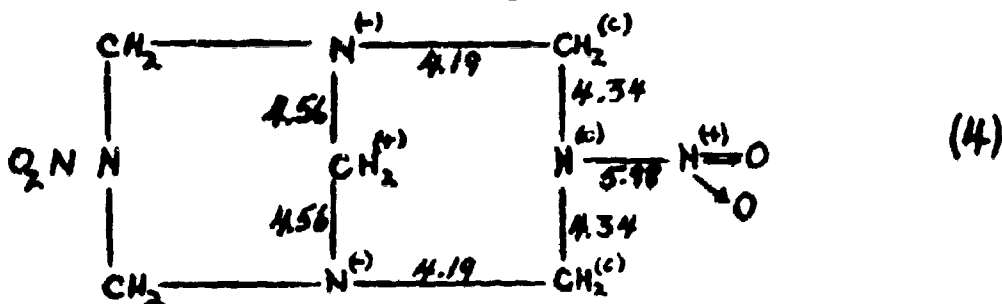
The equilibration of amino-nitrogens due to exchange being limited to the $(\text{CH}_2)_3\text{N}$ substitutions in DPT, would follow the course suggested for hexamine in Diagram 3.



Here again exchange proceeds before the methylene radical undergoes reaction to form stable products. Thus the DPT molecule contains two of its four amino-nitrogens with N-15 enrichment prior to irreversible transformation. In the case of $(\text{CH}_2)_2\text{N-NO}_2$ the amino-nitrogen does not equilibrate with the ammonium adduct based on the assumption that the N- NO_2 bond is stable under the prevailing conditions of nitrolysis.

Calculations of bond force constants based on stretching vibrations lend substantiative evidence to the above suggested mode of exchange during the second stage of nitrolysis. The method of calculation according to Gordy (7) is considered a very good approximation of relative bond strengths of the type of bonds found in DPT. The

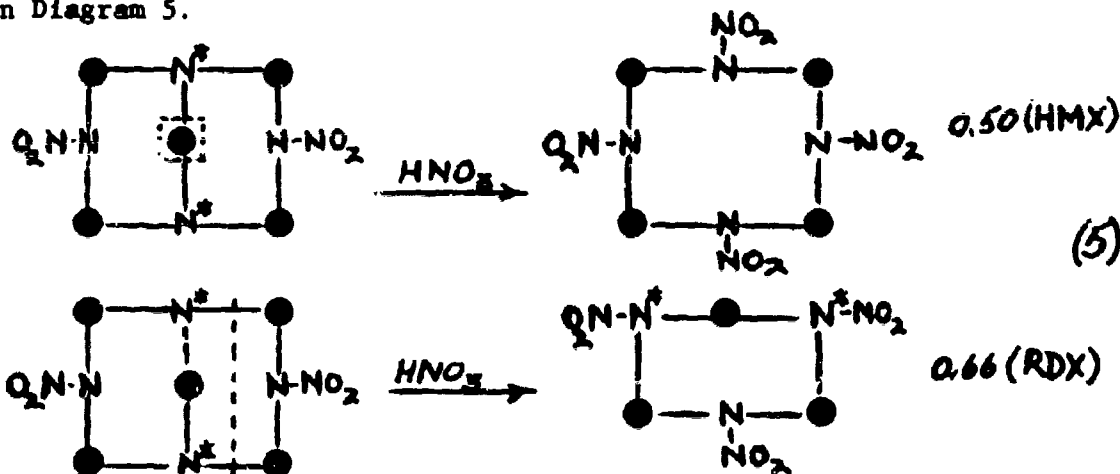
values of the bond stretching force constants ($\times 10^5$ dynes/cm) obtained for the DPT molecule are shown in Diagram 4.



The symbol in parentheses is the induced charge or relative electro-negativity assigned to the atom. It can be readily seen that the N-NO₂ bond is indeed the strongest type bond in DPT. Therefore, it is reasonable to assume that this bond is stable in a nitrolyzing medium and being strongly covalent does not undergo exchange.

NH₄NO₃ Participation in the HOW Process

When N¹⁵H₄NO₃ is added during the second stage the N-15 enrichment is, therefore, expected to be present only in the trimethylene substituted amino-nitrogen prior to reaction. Because of this selective exchange it was possible to determine the role of ammonium nitrate during the second stage as well as the mode of HMX/RDX formation from DPT. If selective cleavage of DPT were to take place in the formation of products, 0.50 and 0.66 of the amino-nitrogen in HMX and RDX, respectively, would contain N-15 enrichment, as indicated in Diagram 5.

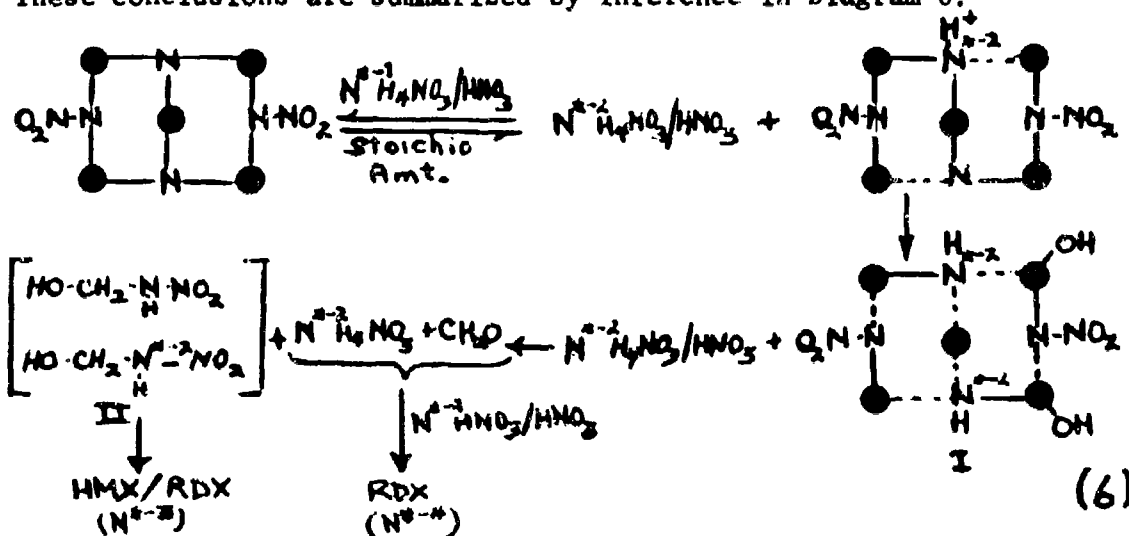


The ratio of atom % N-15 in HMX/RDX should be 0.50/0.66 or 0.75 if the above mechanism were indeed the mode of HMX and RDX formation. As can be seen in Table III the average ratio of atom % N-15 in HMX/RDX formed in the absence of (CH₂O)_x is 0.93 or essentially unity. To check these results, in one series of experiments N¹⁵H₄NO₃ was added during the first stage. In another series, N¹⁵N₄NO₃ was added during both the first and the second stages. In each case, only the HMX and RDX were analyzed for atom % N-15 content. The results are listed in

Table IV and show an over-all average atom % N-15 of 0.97 in HMX/RDX for each respective series.

The following conclusions which can be drawn from this observed N-15 ratio give a direct insight into modes of reaction that may take place during the second stage of nitrolysis in the absence of $(CH_2O)_x$:

- DPT does not cleave selectively to form HMX and RDX.
 - HMX and RDX are derived essentially from hexamine nitrogens.
 - HMX and RDX are derived from equilibrated $-CH_2-N-NO_2$ fragments, half of which are formed during the first stage and the other half during the second stage.
 - A small fraction of approximately 5% of RDX is derived from ammonium nitrate.
 - The rate of formation of HMX is greater than that of RDX.
- These conclusions are summarized by inference in Diagram 6.



Where the atom % N-15 of $N^{*-1} > N^{*-4} > N^{*-2} > N^{*-3}$ and $N =$ natural abundance. The ratio of the $HO-CH_2-NH-NO_2/HO-CH_2-NH^{*-2}-NO_2$ mixture in II is essentially equal to one, since the major product is DPT. Therefore, there is an equal distribution of N-15 and N-14 atoms in HMX and RDX derived therefrom. However, after the formation of II the $N^{*-1}H_4NO_3$ in the ammonium nitrate-nitric acid mixture is in excess of the hexamine nitrogens. This N-15 ammonium nitrate, which cannot exchange with the now completely nitro-substituted hexamine nitrogens, exchanges instead with the ammonium nitrate in situ. Any product formed from the ammonium nitrate amino-nitrogen would obviously have a higher atom % N-15 than a product formed from hexamine-nitrogen. The extent of this increase is proportional to the formation of product from ammonium nitrate. With these factors taken into consideration and correcting for the incomplete conversion of $-CH_2-N-NO_2$ to products, calculations of the atom % N-15 due to exchange alone and, therefore, implying the involvement of hexamine-nitrogens only are in excellent agreement with the values shown by analysis for HMX

in Tables III and IV. Therefore, the extent of ammonium nitrate participation in HMX formation is zero as shown in the last column of Tables III and IV. The atom % N-15 found in RDX is on the average, 4% greater than the calculated value for exchange and is attributed to ammonium nitrate participation. Calculations of the ammonium nitrate participation according to the reaction path as outlined in Diagram 6 show values corresponding to this observed atom % N-15 difference.

The condensation of **I** to products may be considered as a function of competing rates controlled by a concentration effect. It is conceivable that the rate of HO-CH₂-NH-NO₂ cyclization is faster for HMX than for RDX formation. This is based on the fact that the C-N bond angles in HMX are less strained than those in RDX. As the cyclization of the common precursor is brought to completion, the nitric acid content in the reaction mixture is reduced to negligible proportions. The ammonium nitrate and acetic anhydride concentrations are concurrently increased, approaching conditions similar to those in the Ross process (8). In that process, RDX is formed with HMX as a contaminant predominantly from the condensation of (CH₂O)_x with ammonium nitrate by the action of acetic anhydride. Therefore, because of these conditions prevailing during the latter stages in the HOW process, the residual formation of RDX is favored from ammonium nitrate and the available methylene-containing species equivalent to formaldehyde.

Table IV

Check Determination of NH₄NO₃ Participation in The Formation of HMX/RDX In The Absence of (CH₂O)_x

N ¹⁵ H ₄ NO ₃ (a) Addition	Expt. No.	Repl- cate No.	Atom % N-15 Found		Avg % NH ₄ NO ₃ Participation	
			HMX	RDX	HMX (b)	RDX (b)
1st Stage	1	1	1.8	1.9	0	0
	2	2	1.8	1.7		
	2	1	1	1.7	-	
		2	2	1.7	-	
1st & 2d Stages	1	1	2.4	2.5	0	6
	2	2	2.5	2.7		

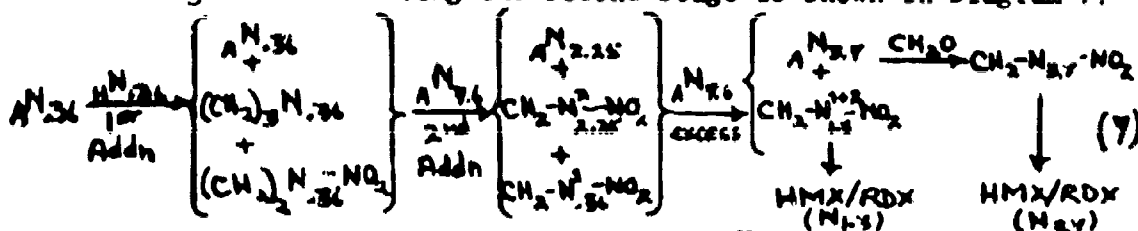
(a) Added as the nitric acid solution and containing 6.42 atom % N-15.

(b) HMX/RDX as derived from NH₄NO₃ amino-nitrogens solely due to reaction.

NH₄NO₃ Participation in the PA Process

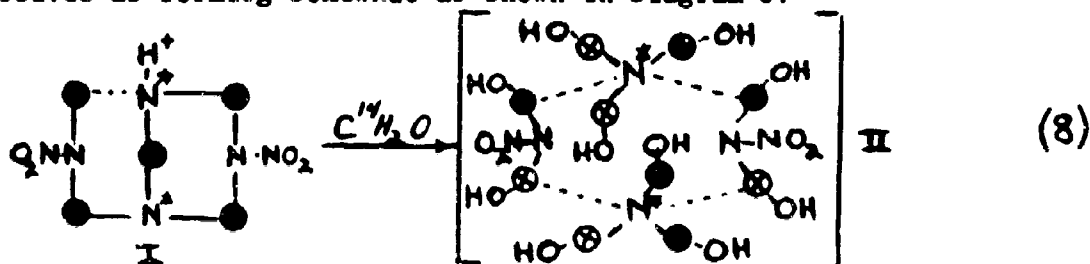
In tracing the role of ammonium nitrate in the presence of (CH₂O)_x N¹⁵H₄NO₃ was added during the second stage of nitrolysis. Check determinations were made by adding N¹⁵H₄NO₃ during the first stage and in another set during both the first and the second stages. The N-15 enrichments found in the HMX and RDX are listed in Table V. These values imply that the reaction sequences, although complicated by the inclusion of (CH₂O)_x, are nevertheless comparable to the ones outlined above for the HOW process. In major part, this mechanism

is used for the interpretation of the N-15 values obtained in the PA process. The distribution of N-15 enrichment in the reaction sequences when $N^{15}HNO_3$ is added during the second stage is shown in Diagram 7.



Where A^N = ammonium nitrate amino-nitrogen, H^N = hexamine amino-nitrogen, N^1 = nitro-substitution formed during 1st addition, N^2 = nitro-substitution formed during 2d addition. According to this mechanism the calculated over-all average percent ammonium nitrate participation shown in Table V is 8% for HMX and 40% for RDX (see experimental procedure for sample calculation).

The suggested mechanism must, of course, be compatible with the results obtained from the study with carbon-14 (4). In that study, the methylene groups from hexamine and $(CH_2O)_x$ were shown to equilibrate during the second as well as during the first stage of reaction. For such an equilibration to take place a metastable intermediate of the form $O_2N-N-(CH_2OH)_2$ must exist which then proceeds to form the common monomer precursor to HMX and RDX, $HO-CH_2-NH-NO_2$ as shown for I in Diagram 6. The common pool of methylene groups may be conceived as forming somewhat as shown in Diagram 8.

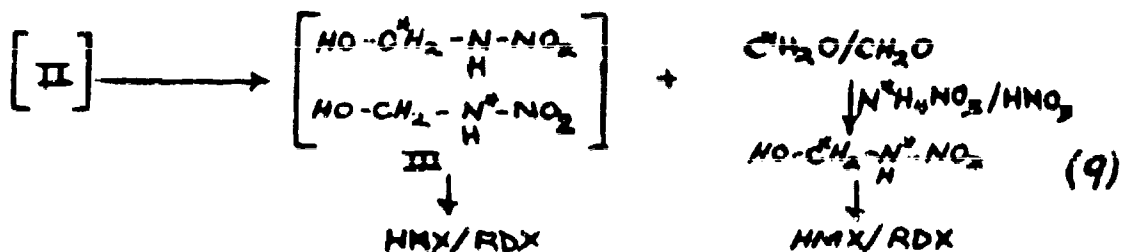


Where N^* = N-15 enriched, $\text{---}\oplus\text{---}$ = $-C^{14}H_2$, $\text{---}\bullet\text{---}$ = $-CH_2$

The DPT with N-15, I, is included to demonstrate the concurrent behavior of amino-nitrogens and methylene groups and may be looked upon as a continuation of the preceding step involving selective amino-nitrogen exchange indicated in Diagram 6. It must be emphasized again the equilibration of methylene groups following amino-nitrogen exchange is due to reaction and not exchange; and that these processes are, therefore, independent of each other.

It is not to be construed that II in Diagram 8 actually exists as such, but what is solely intended is to illustrate the probable existence of an intermediate species which can explain the experimentally observed chemical equivalence of methylene groups and amino-nitrogens. The pool of amino-nitrogens and methylene groups as is intimated by II, is then thought of as forming a common monomer precursor, III (Diagram 9), analogously to the path indicated by II

in the series of reactions shown in Diagram 6.



Here again, the participation of ammonium nitrate must take place, in this case after the release of the equilibrated hexamine and $(\text{CH}_2\text{O})_x$ methylene groups resulting from the cyclization of III to HMX and RDX. This conclusion is not only supported by the excellent agreement between the observed atom % N-15 and the calculated values based on the proposed mechanism, but also by the results obtained in the Carbon-14 tracer study. When C-14- $(\text{CH}_2\text{O})_x$ was added during the second stage an equal distribution of activity was observed in HMX and RDX, signifying methylene carbon equilibration. If any direct condensation of ammonium nitrate with $(\text{C}^*\text{H}_2\text{O})_x$ had occurred prior to equilibration, the carbon-14 content of the RDX would have been significantly greater than that of the HMX, since 40% of RDX and only 8% of HMX is derived from ammonium nitrate.

Table V

NH_4NO_3 Participation in The Formation of HMX/RDX In The Presence of $(\text{CH}_2\text{O})_x$

N ¹⁵ H ₄ NO ₃ (a) Addition	Expt No.	Replicate No.	Atom % N-15 Found		Avg % NH ₄ NO ₃ Participation	
			HMX	RDX	HMX(b)	RDX(b)
Second Stage	1	1	1.3	2.6	4	50
		2	1.6	2.3		
	2	1	1.4	2.4		
		2	1.2	2.6		
First Stage	3	1	1.7	1.6	10	40
		2	1.8	1.5		
	4	1	1.7	1.6		
		2	1.8	-		
1st and 2d Stages	5	1	2.7	3.1	9	37
		2	2.6	3.1		

(a) Added as the nitric acid solution and containing 6.42 atom % N-15 in Expts 1, 2 and 5; and 7.6 Atom % N-15 in Expts 3 and 4.

(b) HMX/RDX derived from NH₄NO₃ amino-nitrogens solely due to reaction.

The final step in the PA process involving the condensation of methylene and ammonium fragments is shown by calculations, based on the experimental data to include the formation of HMX together with RDX, whereas in the HOW process only RDX was shown to form (see last column, Table V). The formation of approximately 8% HMX as well as

CASTORINA and AUTERA

the increased production of RDX from ammonium nitrate (5% in the HOW process as compared to 40% in the PA method) may be explained on the basis of the relatively high concentration of methylene fragments released from II in Diagram 9. Only one equivalent endo-methylene group is released in the HOW process while all of the methylene groups equivalent to those added in the form of $(\text{CH}_2\text{O})_x$ and, in addition, the equivalent endomethyl group are released in the PA process. As the concentration of equilibrated methylene groups increases, their condensation with ammonium nitrate is driven to the right. The increased HMX yield as well as the increased over-all HMX/RDX yield of the PA process may be attributed in part to the final step of this postulated mechanism.

The relative rates of product formation mentioned above for the HOW process evidently prevail also in the PA process. The atom % N-15 in HMX is significantly less than that of RDX, indicating preferential HMX formation from hexamine nitrogens as compared to the more pronounced formation of RDX from ammonium nitrate. Correspondingly, the ratio of atom % N-15 in HMX/RDX is decreased from 0.96 in the HOW process to 0.56 in the PA process. Both values deviate appreciably from the 0.75 figure associated with the selective cleavage of the DPT molecule to discount the possibility of such a mechanism taking place in either process for HMX/RDX production.

Conclusions

The conclusions drawn from the data obtained in the earlier work with carbon-14 are substantiated by these tracer studies with nitrogen-15. In addition, a more complete understanding of the chemistry of amino groups and the processes involved in the formation of the nitramino structure has been obtained through the use of nitrogen-15.

With Carbon-14 indirect evidence had indicated that ammonium nitrate does not condense directly with $(\text{CH}_2\text{O})_x$ to yield HMX and RDX. With nitrogen-15 this was shown directly to be the case.

The equilibration of methylene groups was found to be independent of ammonium nitrate concentration; and with nitrogen-15 tracing the equilibration of amino-nitrogens was shown to be independent of the $(\text{CH}_2\text{O})_x$ concentration. The two processes are mutually independent of each other because their respective modes of equilibration are different. The equilibration of amino-nitrogens being ionic in character takes place by means of exchange and, therefore, precedes the equilibration of methylene groups, which is attributed solely to a reaction process.

It was concluded in the previous study with carbon-14 that hexamine and DPT degrade non-selectively to low molecular weight species which then recombine to form products. The existence of a selective cleavage mechanism has been discounted by data obtained with nitrogen-15 and with carbon-14. The complementary nature of the data obtained from the two studies makes possible the narrowing down of the common fragmentary precursor to HMX and RDX as being of the type, $\text{HOCH}_2\text{NHNO}_2$.

In the earlier work it was postulated that $(\text{CH}_2\text{O})_x$ increases the concentration of methylene groups which exist in precursors to HMX

CASTORINA and AUTERA

and RDX. It was suggested that the increase in yield of combined HMX and RDX products results from inclusion of $(\text{CH}_2\text{O})_x$ could be accounted for on this basis. With nitrogen-15 it became possible to indicate how this actually can take place, viz., by the condensation of the increased concentration of equilibrated methylene fragments with the ammonium radical. The extent of this participation with ammonium nitrate is approximately 7% for the formation of HMX and 40% for RDX, which is indeed equivalent to the 10% over-all increase in HMX/RDX yields observed when $(\text{CH}_2\text{O})_x$ is included in the reaction mixture.

The evidence given for the existence of a common precursor to HMX/RDX of the type $\text{HOCH}_2\text{NHNO}_2$ justifies any effort to prepare HMX from monomethylene containing compounds. A more direct approach to the formation of HMX free of RDX contamination in a simple reaction medium would greatly reduce the current cost of production.

Experimental Procedure

The nitrolyses reactions and the isolation and purification of products (DPT, HMX and RDX) were carried out as described in References (4) and (6), with the exception of using $\text{N}^{15}\text{H}_4\text{NO}_3$ in place of the normal NH_4NO_3 wherever indicated in this report.

Hexamine Exchange Study

The exchange between hexamine and $\text{N}^{15}\text{H}_4\text{NO}_3$ in the absence as well as in the presence of $(\text{CH}_2\text{O})_x$ was studied under actual conditions of nitrolysis in the manner described in Reference (4). During the two-minute period of intermixing of reactants the temperature was permitted to rise to 44°C .

Preparation of HMX and RDX Standards for N-15 Mass Spectrometric Analyses

To a 150 ml three-neck, round-bottom flask are added 11.43 g of ammonium nitrate (of known atom % N-15) and 6.43 g of $(\text{CH}_2\text{O})_x$ dissolved in 25 ml of glacial acetic acid. The mixture is heated on a water bath to effect dissolution. With the solution at 40°C , 50.4 ml of acetic anhydride are added during a 30-minute period. Following this addition, the mixture is aged at 45°C for one hour; the temperature is then raised to 75°C for an additional 30-minute aging. Then the mixture is refluxed with 17 ml of water for 30 minutes. Finally 150 g of ice are added and the HMX and RDX separated in the usual manner.

Mass Spectrometric Analysis of Amino-Nitrogen N-15 Enrichment

All samples were analyzed by the Isomat Corporation, Palisades Park, New Jersey, under service contract. The specificity of their method for amino-nitrogen was checked by the preparation of HMX and RDX standards of known amino-nitrogen N-15 enrichments.

Sample calculation of atom % N-15 in HMX and RDX when $\text{N}^{15}\text{H}_4\text{NO}_3$ is added during the Second Stage

The quantities of N-15 and N-14 in hexamine and ammonium nitrate due to natural occurrence, viz., 0.36%, at end of First Stage are:

CASTORINA and AUTERA

HN-15 = 0.00726 g, AN-15 = 0.00277 g, HN-14 = 2.018 g, AN-14 = 0.769g.
 The ammonium nitrate added during the Second Stage, having 7.6 atom % N-15 contains: 1.154 x .076 = 0.0877 g N-15. Correcting for the amount of ammonium nitrate equivalent to the trimethylene substituted amino-nitrogen in DPT (.634) and for the incomplete conversion to final nitramine products (.85); i.e., 0.85 x 0.634 = 0.539.

The atom % N-15 at the end of the second addition is given by:

$$0.36\% + \left[\frac{.5(.00726) + .00277 + .539(.0877)}{.5(2.018) + .769 + .539(1.154)} \right] \times 100 = 0.36\% + 2.25\% =$$

2.61% then $\frac{2.61}{2} = 1.3$ atom % N-15 in HMX/RDX

$$\text{i.e., } \left[.5(\text{CH}_2\text{-N-NO}_2) = 0.18\% \right] + \left[.5(-\text{CH}_2\text{-N-NO}_2) = 1.13\% \right] = 1.3\%$$

1st Stage Formation

2d Stage Formation

After the addition of the stoichiometric quantity, the ammonium nitrate and the equivalent hexamine nitrogen contain 2.25 atom % N-15 and the excess ammonium nitrate contains the initial 7.6 atom % N-15. The interchange of these two ammonium nitrates yields an equilibrated specie with an atom % N-15 proportional to their respective mole fractions, i.e.

$$2.25 \times \left[\frac{.055 + .539(.082)}{.055 + .082} \right] + 7.60 \times \left[\frac{.461(.082)}{.055 + .082} \right] = 3.77$$

Therefore any product formed from ammonium nitrate contains 3.7 atom % N-15 in combination with product formed from hexamine with 1.3 atom % N-15, The relative amounts are obviously governed by the atom % N-15 found in the particular product.

For the RDX formed in the absence of $(\text{CH}_2\text{O})_x$, the percentages of RDX derived from the respective amino-nitrogens are given by: $1.3x + 3.7(1-x) = 1.5$; $x = 0.92$; i.e., 8% RDX derived from NH_4NO_3 amino-nitrogen, 92% RDX derived from hexamine amino nitrogen.

The HMX formed in the absence of $(\text{CH}_2\text{O})_x$ is solely derived from hexamine nitrogen since the atom % N-15 found in HMX is equal to the calculated value, 1.3%.

The RDX's formed in the presence of $(\text{CH}_2\text{O})_x$ are: $1.3x + 3.7(1-x) = 2.5$; $x = 0.50$; i.e. 50% RDX derived from NH_4NO_3 amino-nitrogen, 50% RDX derived from hexamine amino-nitrogen.

The HMX derived from NH_4NO_3 in the presence of $(\text{CH}_2\text{O})_x$ is given by: $1.3x + 3.7(1-x) = 1.4$; $x = 0.96$; i.e., 96% HMX derived from hexamine amino-nitrogen, 4% HMX derived from NH_4NO_3 amino-nitrogen.

References

1. Ricard, J.P., et al, "An Improved Batch Process for the Preparation of HMX", P.A. Tech Rpt 2059.
2. Herring, R.B., et al, "The Laboratory and Pilot Plant Preparation of HMX and 75/25 Octol", Holston Ordnance Works, Report 20-Z-14.

References (cont)

3. Bachmann, W.E. and Sheehan, J.C., J. Am. Chem. Soc. 71, 1842 (1949).
4. Castorina, T.C., et al, J. Am. Chem. Soc. 82, 1617 (1960).
5. Bachmann, W.E., et al, J. Am. Chem. Soc. 73, 2769 (1951).
6. Castorina, T.C., et al, "The Chemistry of Hexamethylenetetramine in an Acidic-Aqueous Medium", Technical Report 2674.
7. Gordy, W., J. Chem. Phys. 14, 305-20, 1946.
8. McGill University, U.S. Patent 2,434,230, 6 Jan. 1948, Schiessler, R.W. and Ross, J. H., Progress Report CE 53 No. 1-9, July 1961.

RE-ENTRY VEHICLE FOR RADAR SELECTIVITY EVALUATIONS

R. M. COLTON and E. B. DOBBINS
U. S. ARMY MATERIALS RESEARCH AGENCY
WATERTOWN, MASSACHUSETTS 02172

INTRODUCTION

The materials and fabrication studies conducted in this program were directly concerned with the development of re-entry vehicles which would simulate objects re-entering the atmosphere. These objects could be "heat sink" or "ablative" type nose cones, spent metallic or nonmetallic missile tankage, or various types of spacecraft structural members. The identification of the re-entering object and its position in space is quite important in determining whether it is an actual warhead-bearing missile or just innocuous "space garbage."

Two types of materials were selected for the test vehicles, a refractory metal (Mo-TZM, molybdenum alloy) which produces a metallic ionization pattern upon re-entry and represents the heat sink type metallic material, and a nonmetal (phenolic-impregnated fiberglass) representing the ablative type material, which produces an ionization pattern possessing characteristics different from the type observed with metallic type materials. These two materials would thus be representative of most man-made objects which conceivably could be re-entering the atmosphere.

Several fabrication techniques were evaluated and developed in this study, including shear spinning, tape wrapping, plasma arc spraying and welding. The basic configuration of the vehicles fabricated using these techniques are shown in Figures 1 and 2. The dimensions shown in these figures correspond to the full-scale vehicles now being fabricated. The subscale vehicles evaluated in this study were approximately 1/4-scale model reproductions of these configurations.

In the re-entry vehicle material identification study, the test vehicles will be lifted to maximum altitude via a 3-stage system and upon descent, accelerated by a 4th stage motor to reach a re-entry velocity of 25,000 ft/sec. It is assumed that actual atmospheric re-entry commences at 300,000 feet and that material identification by

radar and all telemetry measurements are completed at an altitude of 100,000 feet, corresponding to an elapsed time of approximately 11 to 40 seconds, depending upon the load characteristics of the particular re-entry vehicle. During this brief period of time such measurements as vehicle pitch, yaw and deceleration, radar cross-section, and ionization pattern will be taken for subsequent analysis. In the elapsed time period of 11 to 40 seconds, the temperature of the hemispherical portion of the metallic (Mo-TZM) vehicle, as shown in Figure 1, rises from about 300 F to almost 4000 F. This 4000 F mark is well above the charring point of the ablative (phenolic-fiberglass) material, and only about 700 F below the melting point of the metallic (Mo-TZM) material. However, the mechanical properties of the Mo-TZM degenerate rapidly between 3500 F and 4000 F and its structural usefulness is practically nonexistent at the latter temperature. The ablative material, in contrast, would still be functional at 4000 F.

However, once the vehicles have passed the 100,000-foot altitude level, no further tracking or measurements are to be made and "burnup" of the vehicles is anticipated in the atmosphere, thus terminating their useful life.

GENERAL CONSIDERATIONS IN RE-ENTRY VEHICLE DESIGN

Two of the most important aspects in the design of these re-entry vehicles are the selection of materials and fabrication techniques. Both these selections are dependent upon the re-entry forces and re-entry heating conditions. Upon re-entry into the earth's atmosphere the vehicles will be traveling at velocities of 20,000 to 30,000 feet per second, and during the entire re-entry they are subject to stresses induced by drag, aerodynamic, and centrifugal forces. At altitudes below 150,000 feet the vehicles are subject to aerodynamic heating of the skin to ultimate temperatures in excess of 5000 F.

Re-Entry Forces

The actual transverse forces that will be acting on the vehicles are represented in the following expression:

$$N + \frac{WV}{g}(\dot{\theta} + \dot{\psi}) - W \cos \theta = 0 \quad (1)$$

where

- N = aerodynamic lifting forces
- W = instantaneous vehicle weight in pounds
- g = gravitational acceleration
- V = velocity - ft/sec
- ψ = angle between local horizon and horizon at launch
- θ = angle measured between vehicle axis and local horizon
- $\dot{\psi} = \frac{d\psi}{dt}$ $\dot{\theta} = \frac{d\theta}{dt}$

COLTON and DOBBINS

Simultaneously, the longitudinal forces acting on the vehicle are represented in the following expression:

$$C + \frac{W}{g} \left(\frac{dV}{dt} \right) - W \sin \theta = 0 \quad (2)$$

where

t = time in seconds

C = aerodynamic drag force in pounds

An additional analysis of the longitudinal force Equation 2 reveals the importance of another parameter, the ballistic coefficient ($W/C_D A$), which essentially is a measurement of the efficiency of the re-entry vehicle system based on its weight, shape, and drag coefficient. This parameter is derived by substituting Equation 3 for C in Equation 2, resulting in Equation 4.

$$C = C_D \frac{\rho}{2} AV^2 \quad (3)$$

and, if $C \gg W \sin \theta$ (at very high velocities and medium altitudes)

$$\frac{1}{g} \left(\frac{dV}{dt} \right) = -\frac{C}{W} = -\frac{\rho V^2 / 2}{W/C_D A} \quad (4)$$

C_D = coefficient of drag

A = aerodynamic area of vehicle

ρ = atmospheric density

The transverse and longitudinal force equations summarize the principal forces acting upon the vehicle. Some of the force factors in these equations are eliminated or modified because of various re-entry conditions and vehicle design. However, under most conditions, the vehicle is subjected to the large variety of high magnitude force vectors symbolized above.

Re-Entry Heating

The vehicle must not only serve as a structural member capable of withstanding the forces mentioned above, but must also act as a thermal barrier which insulates sensitive recording, telemetry, and measurements equipment. The equipment would be rendered inoperative if not protected from the vast heat source surrounding the vehicle. This thermal barrier is provided by means of the vehicle skin.

The skin material is composed of one of two types of heat protection systems: (1) heat capacitance cooling, and (2) ablative cooling.

The heat capacitance or "heat sink" system utilizes the properties of a material, usually a metal, which allows it to absorb the heat generated during re entry without impairing its properties below usable levels. The efficiency of such a system depends mainly upon the thermal diffusivity of the material.

$$a = K/\rho c \quad (5)$$

COLTON and DOBBINS

where a = thermal diffusivity - ft^2/hr
 K = thermal conductivity of material
 c = specific heat of the material
 ρ = density of material

Materials having low diffusivity, such as glass ($a = 0.015$) or steel ($a = 0.6$), must have sufficient mass to allow for greater heat absorption. However, a point of diminishing return is experienced in adding mass, since a condition is soon reached where additional increases in mass have little effect on reducing the skin surface temperatures. Even with materials with high diffusivity, such as aluminum ($a = 3.5$) or silver ($a = 5.8$), this point of diminishing return is reached at skin thicknesses of 1 or 2 inches. A second important factor is the melting point of the material and the temperature at which its mechanical properties are no longer suitable to the re-entry conditions of force. Consequently, the choice of heat sink materials is quite dependent upon their diffusivity and melting point.

The ablative cooling system operates on an altogether different heat dissipation phenomena than the capacitance cooling system. Capacitive cooling is usually not efficient in coping with the extremely high heat fluxes associated with re-entry. Ablative cooling can be made to operate at a relatively low surface skin temperature, thus reducing the cooling problem in the interior of the vehicle. The ablative process is not heating-rate limited and can dispose of a large amount of heat for a small amount of material loss. There are basically two types of ablative systems: (1) charring oxidation system illustrated by the phenolic-impregnated glass fiber type studied in this program, and (2) melting and vaporization systems illustrated by a material such as quartz. The former system is characterized by three physical layers: (1) top gas layer (above skin); (2) middle char layer (part of skin); and (3) base unaffected (part of skin). The heat is disposed of by radiation, convective transport, and pyrolysis.

The ideal heat-resistant material is capable of radiating at high surface temperatures and has a long characteristic conduction time, so that heat penetrates through the material very slowly. These characteristics are typified mostly in ablative-type materials.

In contrast, high temperature metallic materials (heat-sink) usually encourage conduction of heat to the cool structure and introduce an insulation problem.

MATERIALS SELECTION

Density is an important consideration in the selection of the re-entry vehicle materials, both metallic and nonmetallic, since the flying weight of the test vehicle can not exceed a specified amount. This limitation is essential to allow for the weight of the telemetry equipment and instrumentation carried within the body of the vehicle. Consequently, strength/density ratio values of the candidate materials at the re-entry temperatures are equally as important as their

ablative or heat sink characteristics. Oxidation resistance is a factor which assumes added importance as the vehicle encounters increases in atmospheric oxygen content in its descent from 300,000 to 150,000 feet.

Metallic Vehicle System

The molybdenum alloy Mo-TZM was selected for the hemispherical portion of the metallic vehicle, as shown in Figure 1, because of its superior mechanical and physical property levels at the temperatures experienced during re-entry. Table I lists its mechanical properties at various temperatures.

TABLE I
Mechanical Properties of 0.050-Inch-Thick Mo-TZM Sheet
(0.5% Ti, 0.08% Zr, 0.030% C)

Test Temperature (deg F)	Yield Strength (ksi)	Tensile Strength (ksi)	Elongation (%)
2000	60	72	18
2400	38	48	26
2800	13	18	35
3000	8	12	44
3500	2	4	66
Effect of Plastic Reduction on Room Temperature Mechanical Properties			
Reduction (%)	Yield Strength (ksi)	Tensile Strength (ksi)	Elongation (%)
0	75	80	3
20	82	105	5
40	98	110	9
60	108	122	15
80	170	138	15
100	138	150	10
Effect of Plastic Reduction on Temperature of Complete Recrystallization			
Reduction (%)	Recrystallization Temperature (deg F)		
20	3200		
40	3000		
60	2870		
80	2650		

Its relatively high diffusivity ($\alpha = 2.5$) and high melting temperature (4730 F) were also important factors in its choice for the hemispherical portion. The combinations of properties cited above produced the lightest weight vehicular structure of the heat sink type of all metallic materials considered with the possible exception of beryllium. However, beryllium was not selected at this time as it could not be fabricated into the required configuration in the specified time frame.

Inconel "X" was chosen for the frustum portion of the Mo-TZM tipped vehicle, shown in Figure 2, because of its superior performance characteristics at the temperatures and stresses encountered during re-entry, its good fabricability into the shape desired, and its accessibility. Its mechanical properties at various temperatures are shown in Table II.

TABLE II

Mechanical Properties of 0.050-Inch-Thick Inconel "X"
Sheet Material, Annealed and Aged
(73% Ni, 15% Cr, 7.7% Fe, 2.5% Ti, 0.9% Al, 1.7% Cb)

Test Temperature (deg F)	Yield Strength 0.2% Offset (ksi)	Tensile Strength (ksi)	Elongation (%)
RT	80	167	20
400	80	160	25
800	80	150	30
1200	80	130	15
1800	20	40	25

Ablative Vehicle System

A phenolic-impregnated fiberglass was selected for the skin material of the ablative vehicle shown in Figure 2, because of its acceptable performance characteristics as observed during previous re-entry and ablative charring studies, its ease of fabrication and ready accessibility. The specific phenolic-fiberglass combination chosen was the type 128, 111 finish, fiberglass cloth impregnated with phenolic resin type Conlon 506. The phenolic resin was composed of 36.4% solids and 7.6% volatiles, and exhibited 18% flow. The mechanical properties of the materials are shown in Table III.

TABLE III

Mechanical Properties of Phenolic-Fiberglass Composites

Test Temperature (deg F)	Compressive Strength (ksi)	Yield Strength (ksi)	Tensile Strength (ksi)	Flexural Modulus (psi x 10 ⁶)
RT	52	50	71	3.4
300(1/2 hr)	45	41	56	3.0
500(1/2 hr)	39	41	56	2.8

The substrate for the phenolic-fiberglass vehicle was composed of a shear spun 5052 aluminum alloy cone. This material was chosen for its low density, good fabricability, and ready accessibility. Its heat dissipation properties were secondary, since the bulk of the dissipation was accomplished by the ablating skin material. The mechanical properties of this material used are shown in Table IV.

TABLE IV

Mechanical Properties of 0.50-Inch-Thick 5052 Aluminum Alloy
Sheet, H-38 Condition
(2.5% Mg - 0.25% Cr)

Test Temperature (deg F)	Yield Strength (ksi)	Tensile Strength (ksi)
RT	37	42
300	29	34
500	7	12
700	3	5

PROCEDURE AND RESULTS

Studies were initiated to determine the proper fabrication techniques compatible with each of the vehicle designs (Figures 1 and 2) and the materials selected.

This is an important consideration since the re-entry vehicle material per se must not only be capable of withstanding the devastating force attack encountered during re-entry, but must be capable of being formed into the desired shapes without degradation of its basic mechanical and physical properties, as is often experienced during fabrication. Consequently, the material may be suitable originally, but may lose this suitability when processed into the usable configuration.

In order to eliminate such an occurrence, all the techniques listed below were evaluated.

Vehicle	Section	Fabrication Technique
Metallic Skin (See Figure 1)	Hemisphere	1. Shear Spinning of Mo-TZM 2. *Plasma Arc Spraying of free standing Molybdenum cones 3. *Welding of Mo-TZM
	Frustum	1. Rolling and Welding of Inconel "X" followed by Shear Spinning
	Fairing	1. Shear Spinning of Inconel "X"
Ablative Substrate (See Figure 2)	Hemisphere	1. Shear Spinning of 5052 Al
	Frustum	
	Fairing	
Ablative Skin (See Figure 2)	Hemisphere	1. Phenolic-Fiberglass Wrap
	Frustum	
	Fairing	

*Evaluated but not employed for fabrication

It was necessary to evaluate three different techniques to insure that at least one method would be suitable for fabricating the metallic hemisphere. The ultimate selection was the shear spinning technique. The metallic frustum involved the evaluation of two techniques, both of which were ultimately used in combination. The frustum was rolled into the required shape using sheet material, longitudinally welded, and finally shear spun to size. The fairing section required the evaluation of only one technique, shear spinning.

The evaluation of only one technique was deemed necessary for fabricating the ablative vehicle substrate and skin sections.

Shear Spinning of Mo-TZM, Inconel "X" and 5052 Al

All materials were shear formed using shear spinning equipment of 42.5" diametrical capacity.

The shear spinning of the Mo-TZM, Inconel "X" and 5052 Al required markedly different reduction ratios, number of passes to obtain the desired reductions, intermediate anneal cycles, roller settings, and angles and temperatures at which shearing took place.

However, all of the materials studied were subject to deformation in accordance with what is referred to as Sine Law of shear spinning. The basic theoretical relationship between the starting blank thickness and finished part thickness is expressed in the trigonometric formula:

$$\sin \alpha = \frac{T_f}{T_s} \quad (6)$$

where α = development angle
 T_f = finished wall thickness
 T_s = starting blank thickness

From this equation it is observed that as angle α decreases, the reduction ratio of wall thickness of blank to wall thickness of the finished part increases rapidly. A complex situation arises if the shape produced from a flat blank is hemispherical, having a variable wall thickness along its hemispherical curvature. This is the type of situation which was prevalent in the case of the Mo-TZM and 5052 Al hemisphere portions shown in Figures 1 and 2. The Inconel "X" frustum section shown in Figure 1 corresponds to a less complicated situation, having a constant development angle and constant wall thickness.

The Mo-TZM was formed into a 1/4-size scale model of the hemispherical configuration shown in Figure 1, at a temperature of 700°F. The starting blank was 0.100 inch in thickness and was reduced to thicknesses between 0.089 and 0.026 inch depending upon the location on the hemisphere. The heaviest section was located at the stagnation point shown in Figure 1. The 5052 aluminum ablative substrate hemisphere was formed into a similar shape, as shown in Figure 2, at room temperature. However, in this case a starting blank of 0.063-inch thickness was reduced to thicknesses between 0.050 and 0.036 inch as shown in Figure 2. The Inconel "X" metallic frustum was formed to 0.026-inch thickness at room temperature from a rolled and welded preform of 0.050-inch thickness with a resultant reduction in wall thickness of about 40%. The Inconel "X" metallic fairing was formed to 0.026-inch thickness at room temperature from a flat, washer-shaped disk preform of 0.050-inch thickness, also with a reduction of approximately 40%.

Tensile test specimens were cut from the rear portions of the shear spun Mo-TZM and Inconel "X" test vehicle sections and the results are shown in Table V. All tests were conducted at 1200°F.

Wrapping, Machining, and Inspection Techniques for Phenolic-Impregnated Fiberglass Composite

The 1/4-scale model of the re-entry vehicle shown in Figure 2, including hemisphere, frustum, and fairing, was wrapped using 1- and 2-inch-wide type 128 fiberglass tape having a 111 finish and impregnated with Conlon 506 phenolic resin.

TABLE V
Mechanical Properties of Shear Spun Inconel "X" and Mo-TZM Alloys

Specimen Identification	Thickness (inches)	Yield Strength at 0.2% (ksi)	Tensile Strength (ksi)
INX-T1*	0.026	167.7	155.6
INX-T2*	0.026	108.5	133.1
TZM-L1**	0.057	85.7	87.5
TZM-L2**	0.052	78.8	82.7
TZM-T1*	0.064	84.1	80.6
TZM-T2*	0.063	72.7	87.8

*Specimens transverse to shear spinning direction

**Specimens parallel with shear spinning direction

NOTE: All specimens were slightly curved due to the shape of the shear spun vehicles, consequently, strength values are lower than normally expected.

Tape-Wrapping Sequence

The tape-wrapping and subsequent machining technique can best be described by referring to Figures 3-10. Figure 3 shows the aluminum substrate mandrel prior to wrapping, with a phenolic-fiberglass plug cemented and bolted to the vehicle at the stagnation point. The aluminum extension to the plug provides for sufficient area to start the tape-wrapping sequence. Figure 4 shows the tape-winding fixture with a 2-inch-wide tape roll in place, threaded through the tension and tape-backing stripping rollers, and ready for wrapping. In this view, the aluminum mandrel is not yet in place, however, both the template for tracing the vehicle contour and the tracing indicator are shown. A conventional lathe was used for the entire operation. Winding was initiated at the rear of the vehicle as shown in Figure 5. A heating lamp raised the temperature of the tape to approximately 100 F, to facilitate wrapping. The rear section was wrapped at a rate of 60 rpm. Following wrapping of the rear section, the front section was wrapped at a rate of 17 rpm. Figures 6 and 7 show the front portion of the vehicle, shown in Figure 8, was then vacuum cured. After vacuum curing the vehicle was machined on a standard lathe. Figure 9 shows machining in process and Figure 10 shows the completely machined vehicle.

Testing and Inspection of Phenolic-Fiberglass Composite

Thickness measurements at any point on the surface of the phenolic-fiberglass composite skin, Figure 2, were obtained using ultrasonic techniques. A No. 210 RADAC ultrasonic test unit was found to provide high accuracy and versatility. Thickness measurements between 0.020 and 0.300 inch could be easily determined within an accuracy of $\frac{1}{2}$ of 1%.

Variations in density, porosity, and delaminations of the phenolic-fiberglass composite skin could be resolved using a low voltage X-ray source and high-sensitivity, fine-grained X-ray film. Densit variations of about 5%, and porosity defects and delaminations of 0.001-inch thickness could be detected. The optimum X-ray voltage was 90 kv and the film resulting in the best resolution was the single emulsion type.

CONCLUSIONS AND RECOMMENDATIONS

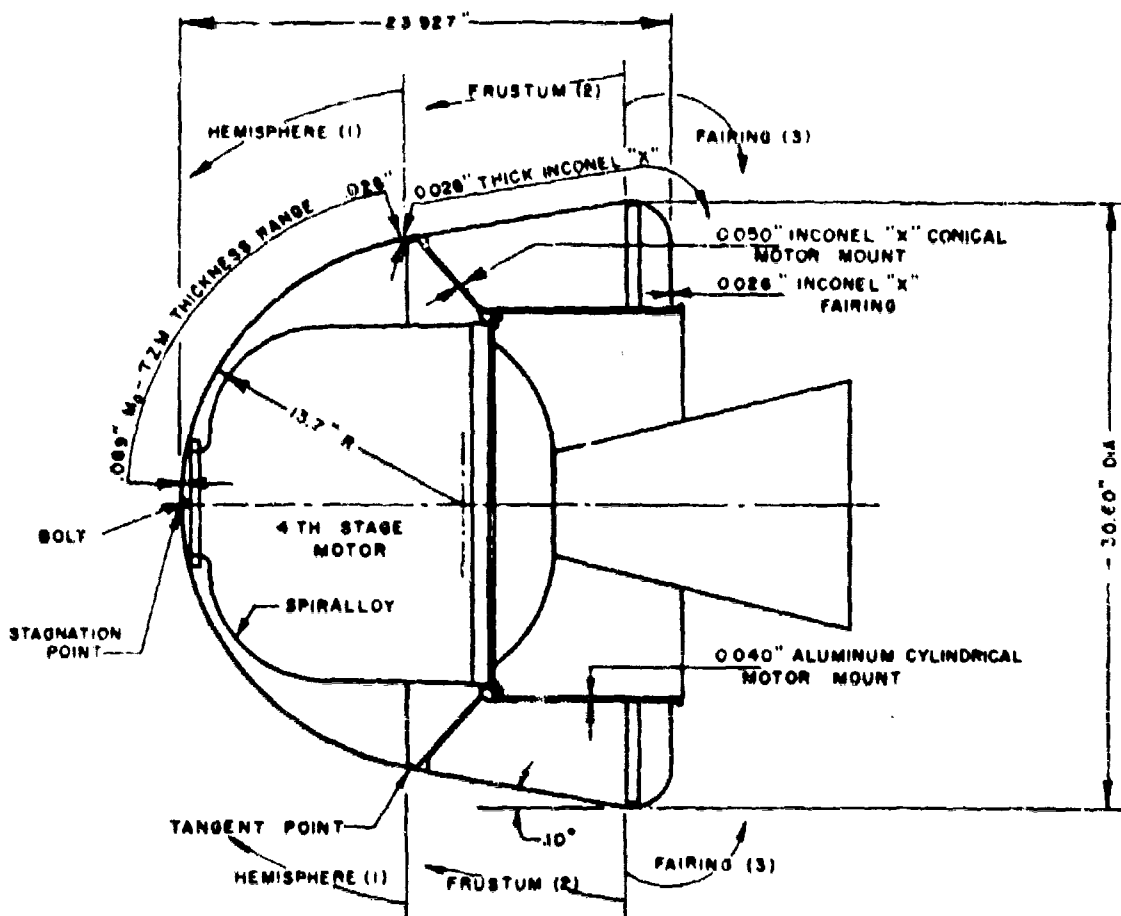
Shear spinning of Inconel "X", 5052 aluminum alloy and Mo-TZM re-entry vehicle configurations of the type evaluated can be accomplished at predetermined temperatures, using specific reduction rates and ratios. Additional studies are required to firmly establish shear spinning variables for the Mo-TZM alloy.

Ablative re-entry vehicles of the type evaluated can be successfully fabricated using a type 128-111 finish fiberglass cloth, impregnated with a phenolic type resin, which is subsequently wrapped over a type 5052 aluminum shear spun substrate. Wrapping, curing, machining, and shear spinning techniques are compatible with the vehicle design.

A more sophisticated tape-wrapping fixture has been designed to wrap the full-scale vehicle. The new fixture incorporates integral devices for varying the tension on the tape, accommodating large-diameter and variable-width rolls of tape, and heating the tape. Additional studies are required to optimize tape heating, stacking, and support techniques during wrapping.

BIBLIOGRAPHY

- ABRAHAM, LEWIS H. Structural Design of Missiles and Spacecraft, McGraw-Hill Book Co., 1962.
- GENERAL ELECTRIC CO. Re-Entry Ablation of Reinforced Plastics Descriptive Summary. Missile and Space Vehicle Department, Philadelphia, Pennsylvania, P1B-15. August 1959.
- OSOFSKY, D. B. Study of Oxidation During Re-Entry for A Molybdenum Faced AMRAD Body. Cornell Aeronautical Lab., Buffalo 21, New York, AMRAD MSMC No. 17, 3 July 1963.
- BURLESON, WILLIAM G. Final Thermal Analysis. AMRAD Target Experiment No. 4, AMXMI-RSTT-16, 23 October 1963.
- ROBERTS, LEONARD. Ablation Materials for Atmospheric Entry. NASA - University Conference on Science and Technology of Space Exploration, Chicago, Illinois, v. 2, 1-3 November 1962.

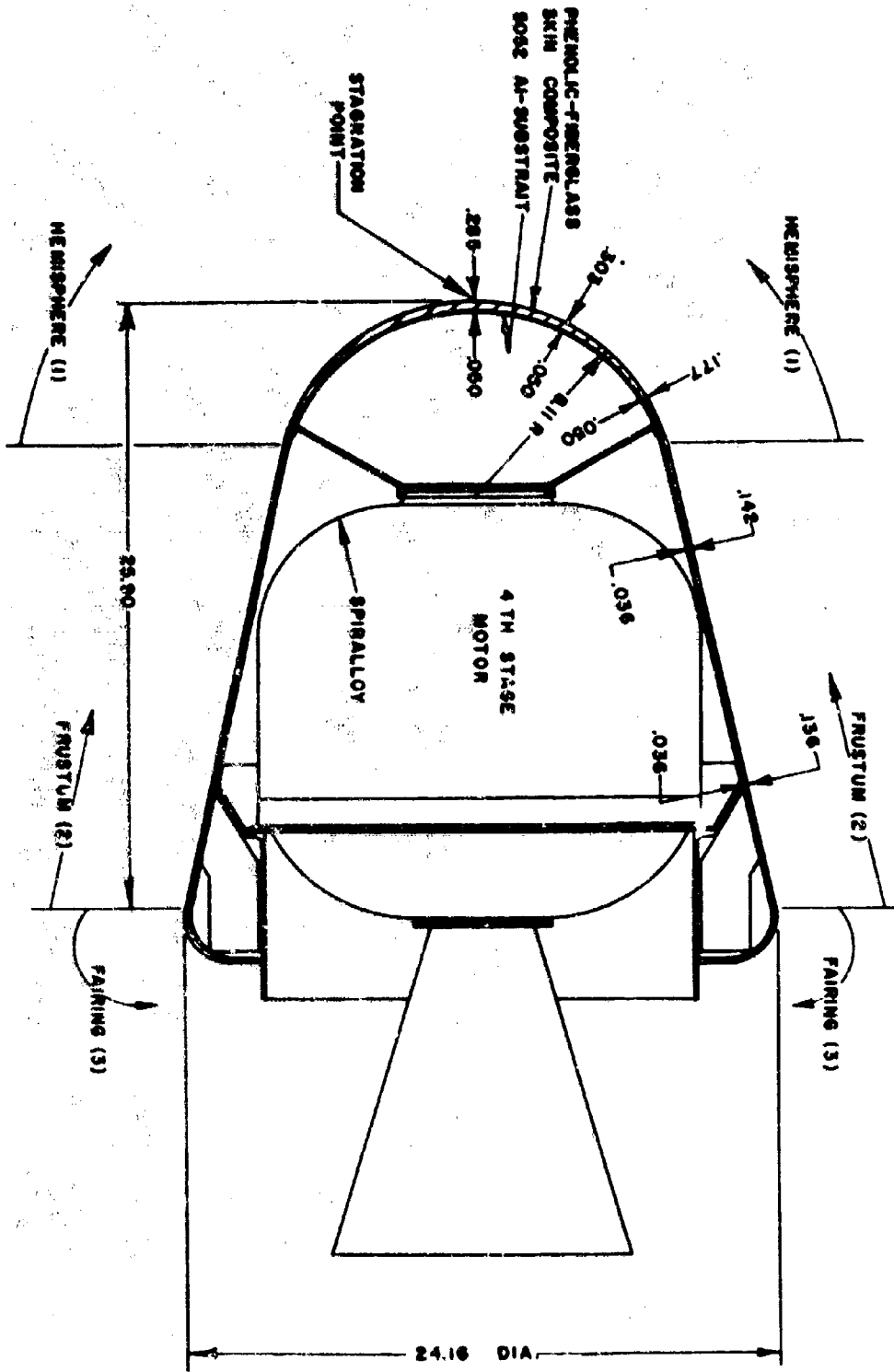


METALLIC RE-ENTRY VEHICLE

FIGURE 1

U. S. ARMY MATERIALS RESEARCH AGENCY

19-066-215/AMC-64

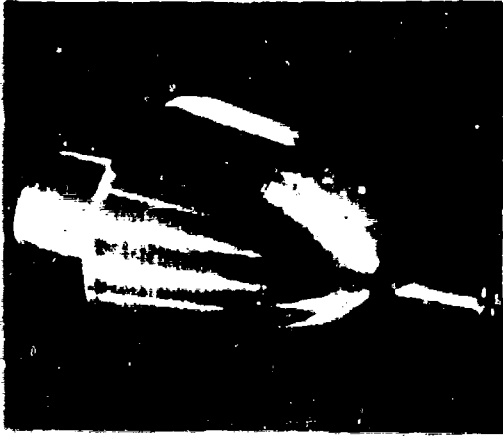


NON-METALLIC RE-ENTRY VEHICLE

FIGURE 2

U. S. ARMY MATERIALS RESEARCH AGENCY

19-066-216/ARC-44



Front View
Figure 3. ARPA RE-ENTRY VEHICLE "A"
1/4-SCALE MODEL BEFORE WRAPPING



Right Side View
Figure 4. TAPE WINDING FIXTURE



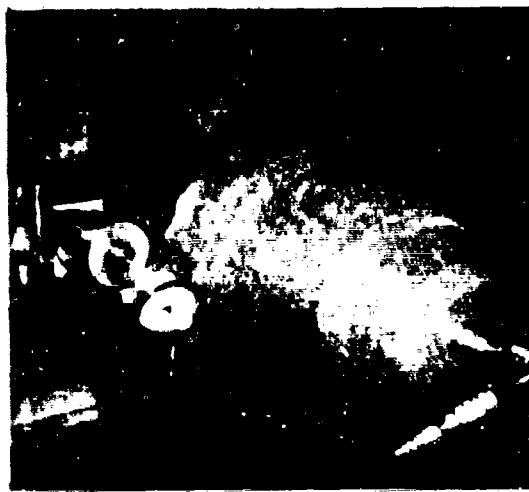
Right Side View
Figure 5. WRAPPING REAR OF ARPA
RE-ENTRY VEHICLE "A" 1/4-SCALE MODEL



Front View
Figure 6. STARTING TO WRAP FRONT OF ARPA
RE-ENTRY VEHICLE "A" 1/4-SCALE MODEL



Right Side View
Figure 7. FRONT OF ARPA RE-ENTRY
VEHICLE "A" 1/4-SCALE MODEL
(3/4 OF RADIUS WRAPPED)



Front View
Figure 8. ARPA RE-ENTRY VEHICLE "A"
1/4-SCALE MODEL AFTER WRAPPING WITH
FIBERGLASS TAPE



Figure 9. ROUGH MACHINING FIBERGLASS
TAPE-WRAPPED 1/4-SCALE MODEL
ARPA RE-ENTRY VEHICLE "A"



Closeup
Figure 10. FINISH MACHINING TAPE-WRAPPED
1/4-SCALE ARPA RE-ENTRY VEHICLE "A"

U. S. ARMY MATERIALS RESEARCH AGENCY

19-066-290/AMC-64

IMPROVED DIELECTRIC FILMS FOR OPTICAL AND SPACE APPLICATIONS

J. T. COX, G. HASS, AND J. B. RAMSEY
U.S. ARMY ENGINEER RESEARCH AND DEVELOPMENT LABORATORIES
FORT BELVOIR, VIRGINIA

I. INTRODUCTION

The designer of multilayer antireflection coatings, protected front surface mirrors, and filmed surfaces for space applications has often been seriously hampered by the fact that the number of oxides suitable for producing durable and optically transparent films with a great variety of refractive indices has been quite restricted. Many oxides, when evaporated from directly heated tungsten boats, become decomposed by reaction with the tungsten heater and condense on the substrates in the form of strongly absorbing films. Only recently, new deposition techniques and treatments, such as evaporation by electron bombardment, deposition in the presence of oxygen, and ultraviolet irradiation after the film deposition, have greatly increased the number of oxides that can now be deposited in the form of optically nonabsorbing coatings. These new techniques are especially suitable for producing durable optical coatings of Si_2O_3 , SiO_2 , Al_2O_3 , ThO_2 , and ZrO_2 .

It is the purpose of this paper to present data on the preparation and optical properties of these oxide films and to discuss some of their applications, such as the preparation of well protected front surface mirrors with high reflectance in the ultraviolet, multilayer antireflection coatings, surface films for controlling the temperature of satellites, coatings for solar energy converters, and for the study of erosion effects in outer space.

II. EXPERIMENTAL TECHNIQUES

The basic principles and techniques for producing thin films by high vacuum evaporation are well known (1), so that only a brief resume of the methods and equipment used for producing and measuring the above mentioned oxide films will be given here.

Silicon monoxide was evaporated from a directly heated tantalum container. A high-powered electron gun capable of applying up to 20 KV and 250 ma was used for evaporating SiO_2 , Al_2O_3 , ThO_2 , and ZrO_2 . The oxides to be evaporated by electron bombardment were placed into a dimple of a heavy copper block. To assure high and rather constant evaporation rates, and to avoid overheating and decomposing the oxides, an electron beam at least 0.6 inches in diameter was used. With this technique, films of SiO_2 and Al_2O_3 produced at deposition rates of about 400 A/min and at a distance 45 cm from the evaporation source showed no absorptance in the visible and ultraviolet. This is not true for films produced with a highly focussed electron gun.

The reflectance and transmittance of films deposited under various conditions onto transparent substrates and onto evaporated aluminum were measured from 2000 A to 40 μ with Perkin Elmer and Beckman instruments. The refractive indices of nonabsorbing films were computed from the measured reflectance values at the quarter-wavelength positions. Additional values for the film indices were determined from the true interferometrically determined film thicknesses and the wavelength positions of reflectance maxima and minima.

A sketch of a 72-inch vacuum chamber suitable for coating satellites uniformly with temperature controlling surface films of aluminum and silicon oxide is shown in Fig. 1. A special device was built for supporting and rotating a satellite sphere in the evaporator so as to expose all of its surface uniformly to the incident vapor of the evaporated material. The sphere is mounted in a fork so that it can be rotated simultaneously about two axes. To keep the coatings from becoming thicker at the sphere's supporting points or poles than at the equatorial region, the fork must be rotated with non-uniform angular speed having maximum velocity when the poles pass over the evaporation sources. A double universal joint is used to transmit this uneven motion to the fork from the driving shaft.

The sphere is cleaned in vacuum by a high voltage dc glow discharge using 6 KV and 400 ma. The glow discharge cleaning is applied for approximately 15 min and heats the satellite to about 1500C. Before the evaporation is started, a mask is placed between the source and the sphere to limit the angle of vapor incidence on the sphere to less than 40° in order to prevent roughening of the surface by film deposition made at high incidence angles. A glass plate fastened to the mask allows monitoring of the film thickness during the deposition by reflectance measurements with monochromatic light.

More than 30 satellites up to 36 inches in diameter including all Vanguards, Injuns, Explorers XVI and VII and all satellites

of the Naval Research Laboratory were successfully coated for temperature control in the vacuum chamber shown in Fig. 1.

III. RESULTS

1. Films Produced by Evaporation of Silicon Monoxide

Thin films produced by high vacuum evaporation of silicon monoxide are, today, the most frequently used protective layers for evaporated aluminum mirrors. However the composition and optical properties of such protective films and the reflectance characteristics of protected aluminum mirrors depend strongly upon the conditions under which SiO is evaporated (2-4). High deposition rate at low pressure results in films of true SiO which show rather strong absorptance in the ultraviolet and shorter wavelength region of the visible. Coatings prepared at a low deposition rate and rather high pressure of oxygen are strongly oxidized and show a large shift of the absorption edge toward shorter wavelengths, negligible absorptance in the visible and near ultraviolet, and lower index of refraction. Aluminum mirrors which are to be used in the visible and near ultraviolet should therefore be protected with a strongly oxidized film rather than with one of true SiO. Below $\lambda = 300 \text{ m}\mu$, strongly oxidized silicon oxide films still show rather strong absorptance which increases with decreasing wavelength. This has for a long period of time limited their usefulness as protective layers for aluminum mirrors. Recently it has been shown that exposure to the ultraviolet radiation of a quartz mercury burner completely eliminates the undesired ultraviolet absorptance of strongly oxidized silicon oxide films and thus makes it possible to produce well-protected aluminum mirrors with high reflectance throughout the far ultraviolet (5). Fig. 2 shows the effect of SiO evaporation conditions and ultraviolet irradiation on the visible and ultraviolet reflectance of silicon oxide protected aluminum mirrors. Both protective coatings are effectively $\lambda/2$ thick at $\lambda = 550 \text{ m}\mu$. The true SiO coating deposited at a very low pressure shows strong absorptance in the shorter wavelength region of the visible, which gives the mirror a yellow appearance. The strongly oxidized film material prepared by evaporation at 9×10^{-5} torr of oxygen shows almost no absorptance in the visible and near ultraviolet. It can be seen that five hours of ultraviolet irradiation eliminates the far ultraviolet absorptance of the strongly oxidized film completely but has very little effect on the true SiO coating produced without oxygen. Additional experiments made with coatings more than 1μ thick gave the same results. Ultraviolet treatment of aluminum coated with strongly oxidized silicon oxide represents, therefore, a new technique for producing well-protected aluminum mirrors with high ultraviolet reflectance.

Strongly oxidized films of silicon oxide have also been deposited on ultraviolet transparent fused quartz and sapphire and measured in transmittance and reflectance before and after ultraviolet irradiation. Their initially rather high ultraviolet absorptance could be completely eliminated and their refractive indices were greatly decreased by the ultraviolet treatment. Fig. 3 shows the refractive index of a silicon oxide film, produced by evaporating SiO slowly in the presence of oxygen, before and after ultraviolet irradiation from 0.2 μ to 1.6 μ . A film prepared under the described conditions consists predominately of Si₂O₃ (6), and shows an initial refractive index of 1.57 at $\lambda = 600$ m μ . After 5 hours of ultraviolet irradiation, its refractive index is decreased to 1.48 at the same wavelength, which brings it near to that of fused quartz. However, measurements of the infrared absorption spectra of such films have led to the conclusion that 5 hours of ultraviolet treatment does not convert Si₂O₃ into SiO₂.

2. Silicon Dioxide Films Produced by Evaporation with an Electron Gun

Undecomposed films of SiO₂ can be obtained by evaporating quartz glass with an electron gun. If properly evaporated, such films show no absorptance in the ultraviolet and visible, and have a refractive index which is identical to that of fused quartz. Fig. 4 shows the dispersion curve of SiO₂ films produced by electron bombardment in the wavelength region from 0.2 μ to 1.6 μ . The measurements were made on films of various thicknesses deposited at rates of 400 to 800 $\text{\AA}/\text{min}$ at a distance of about 18 inches. They show no interference effects when evaporated onto quartz glass since they have the same index of refraction. Such films are especially suitable for producing chemically and mechanically durable protective layers for front surface mirrors. The ultraviolet and visible reflectance of aluminum with and without a 0.78 μ thick protective film of SiO₂ is shown in Fig. 5. Even in the far ultraviolet the protected mirror exhibits, at the interference maxima, slightly higher reflectance values than the unprotected one, which proves that the SiO₂ film is free of absorptance in the wavelength region shown. Much thinner protective layers of SiO₂ should be used of course for producing the most efficient front surface mirrors for the visible and ultraviolet.

3. Aluminum Oxide Films Produced by Evaporation with an Electron Gun

Another material quite suitable for producing optical coatings is aluminum oxide (Al₂O₃) because it is both extremely hard and transparent from the far ultraviolet to the near infrared. Al₂O₃ can be evaporated from tungsten heaters but the resulting films are

slightly decomposed, due to reduction by tungsten, and are therefore absorbing in the visible. Decomposition, and therefore absorption, can be avoided by using electron bombardment for the evaporation of Al_2O_3 . With this technique extremely hard and durable Al_2O_3 films, up to several microns in thickness, can easily be prepared which show no absorptance in the visible and ultraviolet. The refractive indices of aluminum oxide films evaporated by electron bombardment onto substrates at 40°C and 300°C are shown in fig. 6 for the wavelength region from 0.2μ to 1.6μ . The refractive indices of the films produced at 300°C are slightly higher than those of the films prepared at 40°C . In the visible at $\lambda = 500 \text{ m}\mu$ an increase of the substrate temperature from 40°C to 300°C causes a rise in the refractive index from 1.60 to 1.63. These values are considerably smaller than those of crystalline $\gamma - \text{Al}_2\text{O}_3$ films produced by reactive sputtering onto glass at higher substrate temperatures (7). This is due to the fact that the evaporated films condensed on substrates of 40°C and 300°C are amorphous. Their indices agree well therefore with those of amorphous Al_2O_3 films prepared by anodic oxidation (8), and deposition from organic solutions (9).

Because of their hardness and excellent adherence, evaporated Al_2O_3 films are very suitable as protective layers for aluminum front surface mirrors. Al_2O_3 protected aluminum mirrors show very good abrasion and scratch resistance. They cannot be damaged by rubbing with rough linen and can be repeatedly cleaned with water and detergent without changing their reflectance. They are, however, more sensitive to salt spray and boiling water than silicon oxide protected mirrors.

Fig. 7 demonstrates the use of Al_2O_3 films in multilayer antireflection coatings for glass. It shows the visible reflectance of glass ($n_g = 1.51$) with and without a three-layer antireflection coating which uses Al_2O_3 as the inside layer. This coating provides a very broad region of low reflectance and is extremely durable when deposited on a glass substrate at about 300°C . The middle layer of ZrO_2 used in this film combination was also evaporated by electron bombardment and has an index of about 2.1 at $\lambda = 550 \text{ m}\mu$.

4. Thorium Dioxide Films Produced by Evaporation with an Electron Gun

Films of thorium dioxide (ThO_2) are of special interest since they are transparent throughout most of the ultraviolet region and have a rather high index of refraction. ThO_2 films have, therefore, been used in combination with low-index films of SiO_2 for producing multilayer interference filters (10) and multilayer polarizers (11) for the ultraviolet region. The films for these applications were prepared by deposition from organic solutions.

Using electron bombardment, stable and nonabsorbing coatings of ThO_2 can also be prepared by high vacuum evaporation. The refractive indices of evaporated ThO_2 are lower than those prepared from organic solutions by Schroeder (9). Their index in the ultraviolet at $\lambda = 250 \text{ m}\mu$ is about 1.95. This is high enough however to use such films in combination with evaporated SiO_2 for enhancing the reflectance of aluminum in this spectral region. The ultraviolet reflectance of aluminum coated with a reflectance-increasing film pair of $\text{SiO}_2 + \text{ThO}_2$ is shown in Fig. 8. Both dielectric films are effectively one quarter wavelength thick at $\lambda = 290 \text{ m}\mu$ and result in a maximum reflectance of more than 95%. The region of high reflectance can be shifted to shorter and longer wavelengths by changing the thicknesses of the films and its maximum can be increased by adding additional film pairs of SiO_2 and ThO_2 .

5. Coatings for Controlling the Temperature of Satellites

Evaporated films of aluminum plus silicon oxide have played an important role as surface coatings for controlling the temperature of satellites in outer space (12). The temperature control of an orbiting satellite with small internal power dissipation is established by arranging for the solar energy absorbed to be balanced by the thermal radiant energy emitted by the surface at the temperature desired for the payload. For a given orbit, the crucial parameter is the ratio of the effective solar absorptivity, a , of the surface to its hemispherical emissivity, e . For spherical satellites this ratio of a/e should be about 1.2 for maintaining the satellite at close to room temperature. Highly polished metal surfaces have a/e values of 4 to 10 which would result in very hot satellites and failure of their instruments. By using aluminum in combination with surface films that are nonabsorbing in the solar region but strongly absorbing in the far infrared, surfaces with a wide range of a/e values can easily be prepared. If evaporated aluminum is coated with surface films that are nonabsorbing in the solar region and have a thickness greater than 0.3μ , its solar absorptivity becomes essentially independent of the surface film thickness and assumes, with surface films of $n = 1.5$ to 1.6 , a value of 11.5% to 12.0%. If the surface films have strong absorption bands in the infrared, the thermal emissivity of overcoated aluminum increases steadily with increasing thickness of the surface films.

Silicon oxide films produced by evaporation of SiO in the presence of oxygen are free of solar absorptance but have strong infrared absorption bands in the 9 to 10 μ and 20 to 25 μ region. Fig. 9 shows the infrared reflectance of evaporated aluminum coated with 0.7, 1.4, and 3.6 μ thick films of strongly oxidized silicon oxide. Thin surface films of silicon oxide produce low temperature thermal radiation mainly in the 9 - 10 μ region, while for thicker

COX, HASS, and RAMSEY

films, the 20 - 25 μ region becomes an important factor in the total thermally emitted energy.

Table I gives measured values of the total solar absorptivity and thermal emissivity of aluminum coated with silicon oxide films of various thicknesses.

Table I. Measured Normal Solar Absorptivity and Thermal Emissivity of Evaporated Al Coated with Silicon Oxide of Various Thicknesses

t of Silicon Oxide (μ)	0.43	0.61	0.7	0.88	1.1	1.4	2.8
% ϵ_n at 27°C	5.0	8.1	10.1	16.8	24.2	35.5	61.0
% ϵ_n	11.6	11.8	11.7	11.9	11.8	11.9	12.0

By the use of silicon oxide films of various thickness on top of aluminum, surfaces with any desired ϵ/ϵ ratio from about 6 to 0.2 can be prepared.

Of the more than 30 satellites, coated for temperature control in orbit with aluminum and silicon oxide, one reported temperature data for a time period of almost 3 years. The data showed that the coating did not change, and that the satellite temperature remained in the desired range all through this time.

Films of Al + Al₂O₃ are also very suitable as satellite coatings, since Al₂O₃ evaporated by electron bombardment is free of absorptance in the solar region and shows strong absorption bands in the infrared.

Evaporated SiO₂ and Al₂O₃ can also be used in combination with opaque and semitransparent metal films for producing surfaces with high solar absorptivity and low thermal emissivity for solar energy converters and for preparing coatings to study erosion effects in outer space.

6. Coatings for Solar-Powered Thermal Systems

Various types of solar energy converters, such as fluid cycle engines, thermoelectric devices, and thermionic emitters require surfaces with high solar absorptance and low thermal emittance to operate efficiently. Feasible collector coatings of this type have been developed and described by several investigators (13-15). The so called "dark mirror coatings" which absorb visible light by optical interference, and which were originally developed for reducing visible stray light in infrared equipment, represent one of the approaches for obtaining selective energy absorption for

solar-powered thermal systems. They consist of a highly reflecting metal surface which is overcoated with absorbing and nonabsorbing film materials to produce very low reflectance over a rather extended wavelength region. Two film combinations which result in extremely low visible but high infrared reflectance have been described by Hass et al (13). Both of these film designs which consist of Al-Ge-SiO₂ and Al-SiO₂-Al-SiO₂ use opaque aluminum as the highly reflecting base layer. This limits their use to temperatures less than 450°C. In many applications the coatings have to stand higher temperatures. This can be achieved by modifying the above basic film designs with respect to the film materials. Films of opaque and semitransparent Pt or Rh can be used in combination with evaporated SiO₂ or Al₂O₃ for producing very temperature resistant coatings with low visible and high infrared reflectance. Such coatings can be produced by high vacuum deposition in an electron gun evaporator. Fig. 10 shows the reflectance of a Pt-SiO₂-Pt-SiO₂ coating in the wavelength region from 0.4 to 10 μ . To prepare such a coating, SiO₂ is evaporated onto opaque Pt until the reflectance, controlled at a certain wavelength in the visible, decreases to a minimum. Then, Pt is deposited until the reflectance, after passing through zero, reaches a value of about 20%. Another film of SiO₂ on top brings the reflectance again down to almost zero. The coating of Fig. 10, whose preparation was controlled at $\lambda = 500 \text{ m}\mu$, has a normal solar absorptivity of about 80% and a thermal emissivity of less than 5%, and its reflectance remained unchanged during a one-hour heat treatment at 600°C.

7. Surface Coatings for Studying Erosion Effects in Outer Space

The optical properties of thin evaporated films and film combinations used in the space environment may change drastically if they are eroded away by dust, micrometeorites, or by sputtering. To study whether such a surface degradation takes place in outer space, a film combination which is extremely temperature sensitive to wear-off was developed for use as a test coating on the surface of satellites. It consists of an opaque Al film coated with 0.83 μ of SiO₂ and 100Å of Ge. Such a coating has a solar absorptivity about 40% and an α/ϵ ratio of about 3. If the 100Å of Ge are removed the solar absorptivity of this coating decreases to about 11% and its α/ϵ value drops to about 0.8, and this causes a large decrease in surface temperature of a test coating of this type mounted on the surface of a satellite. Fig. 11 shows the layer arrangement of the erosion sensor and the reflectance of Al + SiO₂ with and without 100Å of Ge for the wavelength region from 0.3 to 1.3 μ . Ge is used as the surface film because it absorbs only in the visible and not in the infrared. Therefore, the removal of such a thin Ge film changes only the solar absorptivity and not the thermal emissivity of the Al + SiO₂ coating.

COX, HASS, and RAMSEY

A test coating of the above described type was installed by NASA, Ames Research Center on the Orbiting Solar Observatory-I (S-16 Satellite) to provide a sensitive method of detecting erosion. The satellite was launched on March 7, 1963, and temperature data have been obtained for a period of over 16 months. Over this period the coating showed no evidence of deterioration by erosion or sputtering. Its solar absorptance is still at its initial value. It may, therefore, be concluded that for earth satellites erosion effects on thin films are insignificant. Similar experiments will be reported with satellites orbiting the moon.

REFERENCES

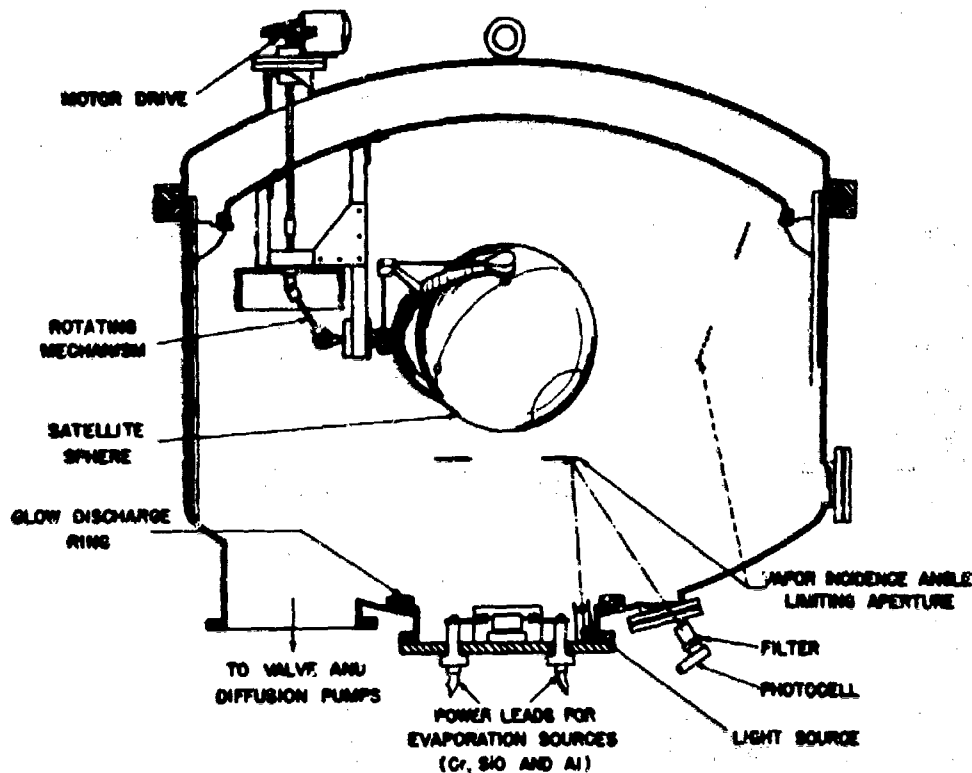
- (1) L. Holland, Vacuum Deposition of Thin Films, Wiley, New York, N. Y. 1956
- (2) G. Hass, J. Am. Ceramic Soc. 33, 353 (1950)
- (3) G. Hass and C. D. Salzberg, J. Opt. Soc. Am. 44, 181 (1954)
- (4) E. Cromer, Th. Kraus, and E. Ritter, Z. Elektrochem. 62, 939 (1958)
- (5) A. P. Bradford and G. Hass, J. Opt. Soc. Am. 53, 1096 (1963)
- (6) E. Cromer and H. Pulker, Monatshefte f. Chemie 93, 491 (1962)
- (7) L. Hiesinger and H. Koenig, Festschrift 100 Jahre Heraeus, p. 376, Hanau, Germany 1951
- (8) G. Hass, J. Opt. Soc. Am. 39, 532 (1949)
- (9) H. Schroeder, Optica Acta 9, 249 (1962)
- (10) R. S. Sokolova and T. N. Krylova, Opt. i Spektroskopiya 6, 788 (1959)
- (11) R. S. Sokolova and T. N. Krylova, Opt. i Spektroskopiya 14, 401 (1963)
- (12) G. Hass, L. F. Drummer and M. Schach, J. Opt. Soc. Am. 49, 918 (1959)
- (13) G. Hass, H. H. Schroeder, and A. F. Turner, J. Opt. Soc. Am. 46, 31 (1956)

COX, HASS, and RAMSEY

(14) H. Tabor, Bull. Res. Council Israel 5A, No. 2, 119 (1956)

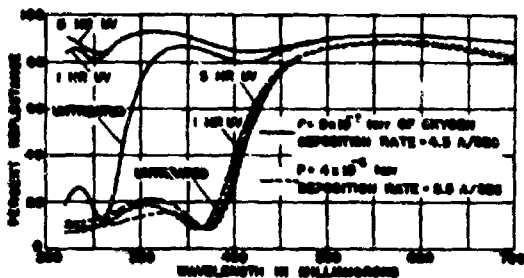
(15) R. B. Gillette, Solar Energy IV, 24 (Oct. 1960)

COX, HASS, and RAMSEY



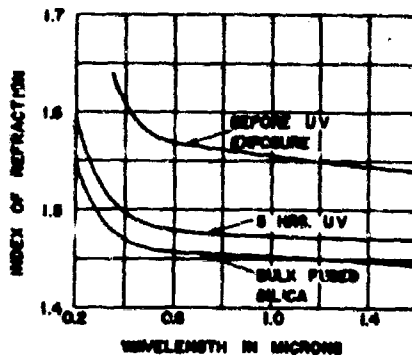
72° EVAPORATOR FOR PRODUCING UNIFORM FILMS ON SPHERES

FIGURE 1



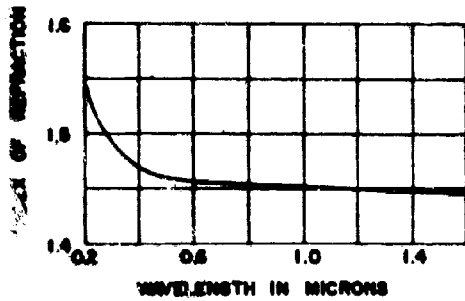
EFFECT OF ULTRAVIOLET RADIATION FROM A QUARTZ MERCURY QUENCH ON THE REFRACTIVE INDEX OF EVAPORATED ALUMINUM PROTECTED WITH UNDEPOSITED AND STRONGLY OXIDIZED SiO COATINGS, EFFECTIVELY $\frac{1}{2}$ THICK AT $\lambda = 600$ TO 650 m μ .

FIGURE 2

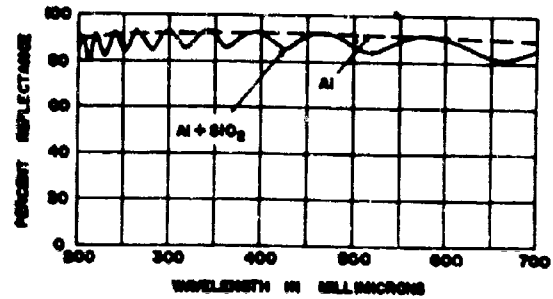


REFRACTIVE INDEX OF A SILICON OXIDE FILM PRODUCED BY EVAPORATING SiO SLOWLY IN THE PRESENCE OF O_2 , BEFORE AND AFTER ULTRAVIOLET IRRADIATION, FROM $\lambda = 0.2$ TO 1.4μ . DEPOSITION RATE $3A/SEC.$ AT 8×10^{-6} TORR.

FIGURE 3



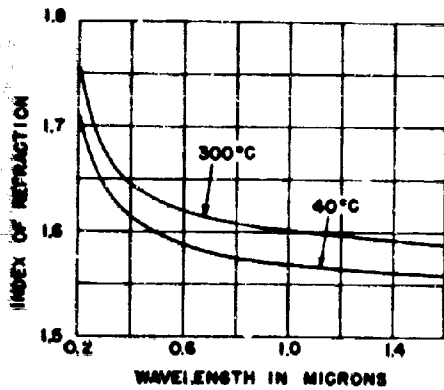
REFRACTIVE INDEX OF EVAPORATED SiO_2 PRODUCED BY ELECTRON BOMBARDMENT, IN THE WAVELENGTH REGION FROM 0.2 TO 1.6 μ .



REFLECTANCE OF ALUMINUM COATED WITH AN EVAPORATED SiO_2 FILM (1-7800 Å), AS A FUNCTION OF WAVELENGTH FROM 200 TO 700 $m\mu$.

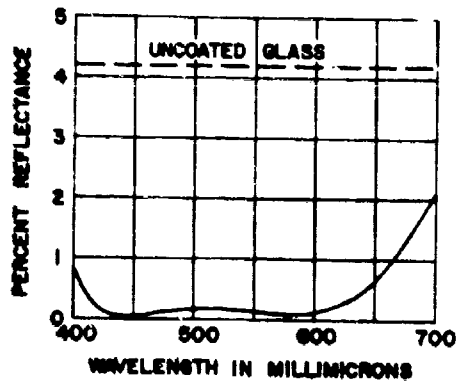
FIGURE 4

FIGURE 5



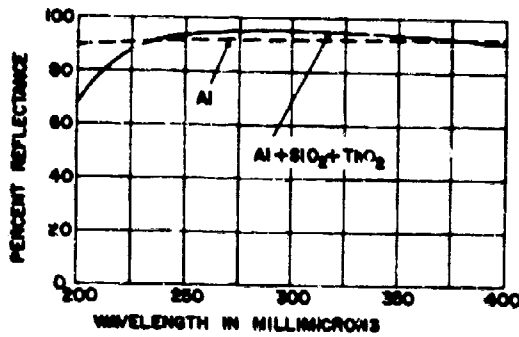
REFRACTIVE INDEX OF EVAPORATED Al_2O_3 PRODUCED BY ELECTRON BOMBARDMENT, IN THE WAVELENGTH REGION FROM 0.2 TO 1.6 μ . SUBSTRATE TEMPERATURE DURING DEPOSITION: 40°C AND 300°C.

FIGURE 6



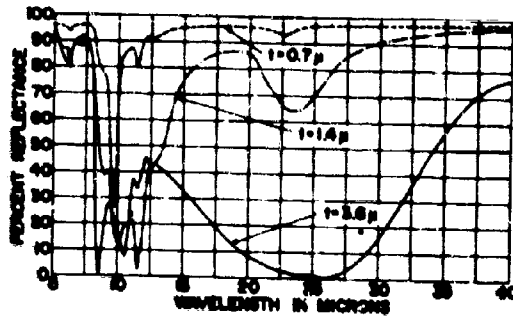
REFLECTANCE OF GLASS ($n_g=1.51$) COATED WITH A TRIPLE LAYER ANTIREFLECTION COATING ($\lambda/4 - \lambda/2 - \lambda/4$) OF $\text{Al}_2\text{O}_3 + \text{ZrO}_2 + \text{MgF}_2$, FROM λ : 400 TO 700 $m\mu$.

FIGURE 7



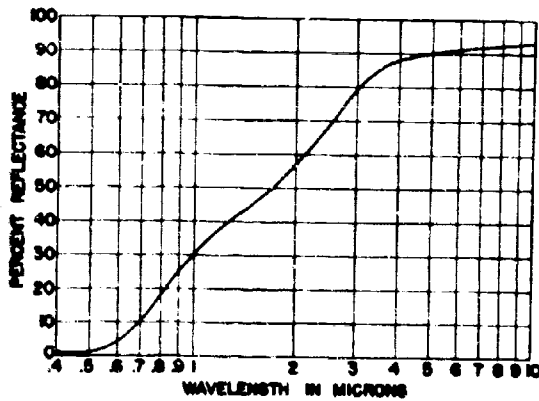
REFLECTANCE OF ALUMINUM COATED WITH A REFLECTANCE INCREASING FILM PAIR OF SiO₂+TiO₂, AS A FUNCTION OF WAVELENGTH FROM 200 TO 400 mμ.

FIGURE 8



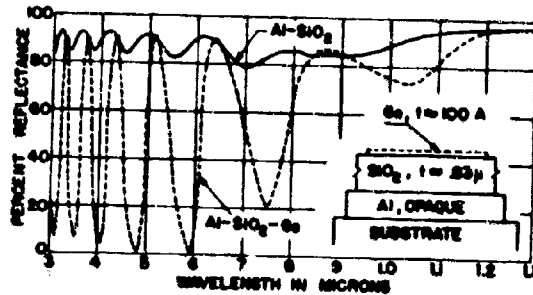
INFRARED REFLECTANCE OF EVAPORATED Al COATED WITH 0.7, 1.4, AND 3.6 μ THICK FILMS OF STRONGLY OXIDIZED SILICON OXIDE AS A FUNCTION OF WAVELENGTH FROM 5 TO 40 μ

FIGURE 9



REFLECTANCE OF A Pt-SiO₂-Pt-SiO₂ COATING AS A FUNCTION OF WAVELENGTH FROM 0.4 TO 10 μ. LAYER CONTROLLED AT λ = 800 mμ

FIGURE 10



LAYER ARRANGEMENT AND REFLECTION CHARACTERISTICS OF AN Al-SiO₂-Ge EROSION SENSOR

FIGURE 11

A MECHANISM OF ACTION
OF STAPHYLOCOCCAL ENTEROTOXIN POISONING

GERALD J. CRAWLEY, 1ST LT., VC; JOHN N. BLACK, CAPTAIN, VC;
IRVING GRAY, COLONEL, MSC; and WAYNE A. LEBLANG, 1ST LT., MSC
WITH TECHNICAL ASSISTANCE OF JACK W. BLANCHARD, PFC
U. S. ARMY MEDICAL UNIT, FORT DETRICK, MARYLAND

Although food poisoning has been a significant medical problem since biblical times, Albrecht von Haller (1) apparently conducted the first scientific study in the early 1800's. In 1894 Van de Velde (2,3) published the first reference to staphylococcal food poisoning and Denys (4) incriminated a yellow micrococcus in an outbreak. There followed many case reports and epidemiologic studies of outbreaks of food poisoning with several excellent clinical reviews (3,5,6,7,8,9).

Since the mid-thirties, basic investigations have employed classical pharmacologic and physiologic preparations such as isolated intestinal systems, muscle with and without attached nerve, uterine strips, perfused heart, frog skin, and many others. Although crude preparations of toxin possessed some physiologic action, as the material was purified, activity was lost. Bayliss (10), using young and adult cats and a variety of surgical techniques, concluded that the action of the enterotoxin was primarily on the peripheral sensory structures. In opposition to this concept, Sugiyama and his colleagues (11) stated that a critical factor in the action of the enterotoxin is its effect on the chemoreceptor trigger zone located in the area postrema. Not only have they ablated this section of the brain, preventing the emetic reaction, but they have also shown a synergism between the emetic effect of dihydroergotamine, which is said to act through the chemoreceptor trigger zone, and that of staphylococcal enterotoxin. In addition, Sugiyama, et al (12) have shown that perphenazine and reserpine have a definite anti-emetic effect. While the reserpine must be used as a pre-treatment, the perphenazine was effective even when given 45 minutes after an intravenous dose of the enterotoxin. Inasmuch as perphenazine is said to be anti-emetic for those substances acting upon the chemoreceptor trigger zone, these workers believe that this observation provides additional support for the conclusions stated above. Orlov (13) also concluded that there was little or no action on the gastrointestinal tract but on a

CRAWLEY, BLACK, GRAY, and LEBLANG

specific part of the CNS.

With the isolation of a reproducibly pure preparation of staphylococcal enterotoxin, type BII (14), a detailed study of its distribution, metabolism, and attendant physiologic changes was begun.

METHODS

Labeling of Toxin: In order to carry out the distribution studies, the toxin has been labeled with Iodine-131. The procedure is essentially as follows: 2 ml of a 1% solution of the enterotoxin in pH 7.4 phosphate buffer is incubated with a mixture containing 0.4 ml of 127 mg I₂ in 0.2M KI and 1 ml of carrier-free I¹³¹ obtained from a commercial source* averaging 1.0 mc per ml of radioactivity. The mixture is allowed to react for 30 minutes when it is separated from the unreacted iodide I¹³¹ by passing it successively through 2 separate columns of amberlite IRA-400 resin. This procedure yielded a product that had about 20 atoms of I¹³¹ per molecule of toxin and less than 5% unbound I¹³¹. When tested in dogs, the toxicity was identical with the starting material. This labeled product was used in the distribution studies to be reported below. Human serum albumin was labeled in the same way at the same concentration and used in control studies.

Disappearance Times: One ml blood samples were taken from monkeys at 10, 20, 30, 45 and 60 minutes and at 2, 3, 5, 8, 12 and 24 hours post-challenge in order to follow the rate of loss of the labeled toxin and albumin from the blood.

Blood Binding Studies: The in vitro distribution of staphylococcal enterotoxin was studied by collecting 10 ml of blood from normal monkeys. Amounts varying from 20 to 1,000 µg of I¹³¹ labeled toxin were added to various samples of blood. The mixture was incubated either at room temperature for 1 hour or at 37°C for 30 minutes. Following incubation, 1 ml of the mixture was assayed for radioactivity. The whole blood was then separated into platelet, leukocyte, plasma, and erythrocyte fractions. Formed elements were washed with normal monkey serum. Radioactivity of each fraction was determined and the per cent of activity in all fractions computed. One ml samples of plasma were treated with an equal volume of 20% TCA. The supernate was separated and radioactivity determined. The precipitate was washed 3 times with 5% TCA and the radioactivity of each wash, as well as the precipitate, was determined.

In vivo binding of the toxin within the blood was studied by challenging 5 monkeys with varying dose levels of labeled toxin (20 to 200 µg per kg). Ten ml samples of blood were taken at 15 minutes and 3 hours post-challenge. Following determination of

*Na Radio-Iodide (Oriodide) Abbott Laboratories, Oak Ridge, Tenn.

radioactivity in whole blood, the buffy coat was removed and radioactivity of the platelets, leukocytes, erythrocytes, and plasma was determined in the same manner as described above. It should be pointed out that total blood volume is a limiting factor in all studies in which blood must be removed from monkeys.

Distribution and Physiologic Studies: The distribution of and physiologic response to I¹³¹ labeled staphylococcal enterotoxin was followed in several species of animals. Additional animals of the same species were treated with a similar quantity of I¹³¹ labeled human albumin to serve as controls. The various species used and the number of animals in each group are summarized in Table I. Food and water were available ad libitum throughout the course of the experiment for all animals.

TABLE I. NUMBER OF ANIMALS USED IN SERIAL SACRIFICE STUDIES

<u>SPECIES</u>	<u>TYPE</u>	<u>TOXIN</u>	<u>ALBUMIN</u>
Rats	Sprague-Dawley	30	10
Rabbits	Mixed	30	9
Monkeys	<u>Macaca mulatta</u>	28	10
Previously challenged monkeys	<u>Macaca mulatta</u>	9	

The rats and rabbits were treated per os with Lugol's solution 24 hours prior to challenge to block the uptake of the I¹³¹ by the thyroid. Rats were challenged with 20 µg/kg or with 40 µg/kg of toxin or albumin via the penal vein. Three rats from each of the toxin groups were sacrificed at 1, 3, 6, 12 and 24 hours post-challenge. One each of the albumin controls were killed at the same time. Eighteen of the rats receiving toxin were placed in individual metabolic cages to collect urine. Rats were euthanized with chloroform and the remainder of the animals with "Lethol"^{*} given either iv or directly into the heart.

Rabbits were similarly divided and challenged at the 2 dosage levels via an ear vein. Three animals of each toxin dose level and 1 control (40 µg/kg) were killed at 1 hour. Subsequently, of each dose level, 4 of the principals and 1 of the controls were killed at 6, 12 and 24 hours. No urine was collected from rabbits.

Three groups of monkeys were used, those which had no previous experience with the enterotoxin, hereafter referred to as clean monkeys; those which had been previously exposed, hereafter termed re-challenge monkeys; and controls. In all studies reported here, monkeys were placed in specially designed metabolic chairs and a

^{*}26% pentobarbital Na and 13% amobarbital Na in a base of isopropanol and polyethylene glycol, Cora State Laboratories, Omaha, Nebraska.

CRAWLEY, BLACK, GRAY, and IJELANG

polyethylene cannula was placed in a peripheral vein into the posterior vena cava to facilitate blood sampling. Urine collections were made from 3 monkeys in each group. All monkeys were given either a 20 $\mu\text{g}/\text{kg}$ dose of toxin or albumin iv. Two of the principals were sacrificed at 1 and 3 hours post-challenge. Subsequently, 3 of the clean monkeys, as well as 3 of the re-challenge monkeys, were killed at 6, 12 and 24 hours. Two of the controls were killed at each time interval.

Tissues from all animals were removed immediately following sacrifice. Whole organs were weighed to the nearest 0.05 gm and placed in plastic tubes for assay of radioactivity. In cases where the organs were large, a representative sample approximating 1 gram was taken, weighed and counted. Calculations were made on a per gram, as well as a per organ, basis. The following tissues were taken: liver, lung, kidney, spleen, heart, brain cortex, pancreas, adrenals, pituitary, thyroid (from monkeys only), large intestine, small intestine, stomach, and skeletal muscle. In addition, urine, feces, stomach contents, and blood samples were obtained from all animals, as well as bile from the gall bladder of monkeys.

All samples were counted in a well-type crystal scintillation counter to an error of less than 5%.

To further examine the organ distribution and the nature of the material as to whether or not it was protein bound, a second group of 15 clean monkeys were handled as described previously. However, these monkeys were challenged with 100 $\mu\text{g}/\text{kg}$ of toxin, iv. Three of these animals were killed at 1, 3, 6, 12 and 24 hours. The same organs and tissue were removed, weighed and radioactivity ascertained as in other monkeys. In addition, homogenates were made of lung, liver and spleen and the concentration of radioactivity was determined. A volume of homogenate was treated with an equal volume of 20% TCA. The radioactivity of both the precipitate and supernate was assayed. From these results the per cent of the radioactivity still protein bound was determined.

Studies of Body Spaces: Body water compartments were determined by using tritiated water as a measure of total body water, and radiosulfate or inulin as measures of extracellular water in nephrectomized monkeys.

Inulin was assayed using Roe's resorcinal method as modified by Schreiner (15). Tritiated water and radiosulfate were assayed using a liquid scintillation β -particle spectrometer (16). Because previous work with staphylococcal enterotoxin challenged monkeys indicated renal impairment (17), the variable of renal excretion was eliminated by bilaterally nephrectomizing each monkey. Six or 7 hours following surgery, the inulin and tritiated water or radiosulfate and tritiated water were injected intravenously. Five

CRAWLEY, BLACK, GRAY, and LEBLANC

blood samples for pre-challenge values were taken during a 3-hour period, which started 8 to 9 hours after injection of the tracers. Immediately after the last pre-challenge blood samples were drawn, the monkeys were challenged iv with the toxin.

The monkeys in which inulin was used received 25 $\mu\text{g}/\text{kg}$ of toxin and the monkeys in which the radiosulfate was used received 200 $\mu\text{g}/\text{kg}$. At 0.5, 1, 2, 3 and 4 hours post-challenge, blood samples were collected using Sequester-Sol^{*} as an anticoagulant and a protein-free filtrate (Folin-Wu) was made from the plasma.

In other non-nephrectomized monkeys, RISA^{**} and hematocrits were used as indicators of vascular water shift. For each monkey 1.5 μc of RISA was injected into a superficial arm vein. Control blood samples were collected at 24, 25, 27, 28, 29, 45, 46, 47 and 48 hours following RISA injection to insure that the metabolic phase of the RISA distribution curve had been reached. Radioactivity per ml of serum was measured in a well-type crystal scintillation counter. Micro-hematocrits were done in duplicate at the 45, 46, 47 and 48 hour sampling times.

Immediately following the 48-hour control sample, each monkey was challenged intravenously with 25 $\mu\text{g}/\text{kg}$ of the staphylococcal enterotoxin. Blood samples were collected at 0.5, 1, 1.5, 2, 3, 4, 5, 6, 7, and 8 hours post-challenge. RISA content and hematocrits were measured as described above.

RESULTS

Symptomatology: Rats appeared to be resistant to the effects of the toxin at the dose administered. Rabbits exhibited a variable response, but generally symptoms were mild and short-lived, and, as with the rats, none died during the 24-hour period of study. Nearly 100% of the clean monkeys showed the usual clinical signs of illness after the first hour. These included vomiting, diarrhea, anorexia, pyrexia, pale mucous membranes, and mild to severe depression leading to death. None died in the first 24 hours of the study. However, many animals used in other concurrent studies have died within 24 hours following this dose. Only 2 of 9 monkeys, both having the lowest anti-enterotoxin titer of the group, showed any symptoms upon re-challenge during the 24-hour period of observation. Both were moderately depressed and vomiting was noted in one.

Disappearance Times: Table II summarizes the disappearance rates of the injected macromolecules. The initial half-life of I¹³¹ labeled toxin in the blood of 20 clean monkeys ranged from 10 to 45

*Dipotassium ethylene diamine tetra-acetate, Cambridge Chemical Products, Inc., Dearborn, Michigan.

**Radio-iodinated (I¹³¹) serum albumin, Abbott Laboratories, Oak Ridge, Tennessee.

CRAWLEY, BLACK, GRAY, and LEBLANG

minutes, while that of 6 re-challenged monkeys varied from 10 to 60 minutes and that of albumin under these conditions ranged from 200 to 400 minutes. However, the continued rate of loss of the albumin and toxin are similar.

TABLE II. RATE OF LOSS OF LABELED STAPHYLOCOCCAL ENTEROTOXIN AND ALBUMIN FROM THE BLOOD OF MONKEYS*

	CLEAN ENTEROTOXIN <u>T 1/2 IN MINUTES**</u>	RE-CHALLENGE ENTEROTOXIN <u>T 1/2 IN MINUTES</u>	ALBUMIN <u>T 1/2 IN MINUTES</u>
Average	18	28	275
Range	10-45	10-60	200-400

*Allowing 10 minutes for mixing.

**T 1/2 = time required for the concentration to decrease by a factor of 2.

Blood Binding Studies: The in vitro and in vivo distribution of the labeled enterotoxin within the blood has been studied. In vitro studies have shown that 15 to 20% of the label is bound to the buffy coat and the remainder to the albumin fraction of the plasma. All radioactivity is removed from the erythrocytes with 3 washes of serum. However, results from in vivo studies indicate that at 15 minutes following administration of toxin less than 1% of the dose is combined to the buffy coat, and by 3 hours there is no detectable buffy coat activity remaining. Following washes by serum, there is no activity on the erythrocytes at any time. At 15 minutes post-challenge, 90% of the label is protein bound within the plasma. Later, at 3 hours, as much as 70% of the amount within the plasma is protein free.

Tissue Distribution Studies: The results from studies of urine demonstrated considerable differences between rats (resistant) and monkeys and between clean and re-challenged monkeys in their ability to clear the I¹³¹. The results of urinary I¹³¹ clearance are indicated in Figure 3. It appears that those animals having the greatest resistance to the effects of the toxin excrete the greatest amount of the label in the urine. This observation has also been noted in clean monkeys, i.e., those showing minimal clinical symptoms voided larger amounts of I¹³¹ in the urine. It was further found that the accumulation of I¹³¹ in the thyroid of the re-challenged monkeys was 4 times greater than the uptake by the thyroid of clean monkeys (17.3% to 4.5%). This and the urine studies indicate that the I¹³¹ was being freed from the protein faster in previously challenged or more resistant animals and is probably an indication of a faster rate of metabolism of the toxin molecule. Limited data indicates that the liberated I¹³¹ is primarily in the form of inorganic iodide rather than as iodinated amino acids as is the case with the iodinated albumin.

CRAWLEY, BLACK, GRAY, and LEBLANG

The results of the tissue studies have shown minimal variation in the organ concentration from animal to animal within a species. Figures 4 and 5 illustrate the change in organ concentration with time after challenge. It is apparent that, except as noted below, the labeled material disappears from the tissue in a manner similar to that for albumin. There was no appreciable difference in the distribution of the enterotoxin in rats or rabbits at the higher dose level of either albumin or toxin.

There appears to be a selective concentration of the enterotoxin in the lungs of all species studied as compared with the albumin. This concentration remains elevated for a considerable length of time in clean monkeys at a time when the blood level is falling. Furthermore, this activity is largely protein bound. Radioactivity has been noted in the lungs of monkeys dying from 4 to 10 days following challenge, whereas activity in other organs was minimal or undetectable. A major difference seen among rats and rabbits (resistant animals) as compared to monkeys in the distribution of the toxin was this continuing high concentration in the lungs of the monkeys. The anomalous situation of the sustained high level of activity in the lungs of the re-challenged, resistant monkeys may be explained by the not unreasonable assumption that the labeled toxin combined with its antibody has been fixed by the phagocytes in the lungs.

It should be noted further that the toxin concentration rises in the heart and kidneys for the first 6 to 12 hours and then falls off in the same manner as albumin.

Body Space Studies: No significant differences between the pre-challenge and post-challenge total body water compartment (TBW) could be detected with T2O. Likewise, no changes in the extracellular fluid compartment (ECF) could be shown with either the inulin or the radiosulfate. These findings indicate that the intracellular water (ICW) (TBW-ECF) remains unchanged. An increase in RISA concentration occurred within 1 hour post-challenge as shown in Figure 1, and can best be explained by a loss of vascular water. This conclusion is supported by the rise in hematocrit, as shown in Figure 2.

TABLE III. INCREASE IN LUNG WEIGHT FOLLOWING ENTEROTOXIN CHALLENGE

<u>NO. OF MONKEYS</u>		<u>HOURS POST-CHALLENGE SACRIFICE</u>	<u>RATIO OF LUNG/HEART CONTROL/TOXIN</u>		<u>DIFFERENCE SIGNIFICANT AT 7 LEVEL</u>
<u>CONTROL</u>	<u>TOXIN</u>				
10	5	1	1.43	1.43	-
10	5	3	1.43	1.98	10
10	6	6	1.43	1.94	1
10	6	12	1.43	2.17	5
10	6	24	1.43	2.53	10

CRAWLEY, BLACK, GRAY, and LEE LANG

This rise occurs after a steady declining hematocrit due to the removal of red blood cells by sampling. Since there is no change in the extracellular or intracellular water, this loss of vascular water is best explained by a shift into the interstitial space. A very interesting observation was the increase in weight of the lungs of the clean monkeys at sacrifice as compared to the control monkeys as seen in Table III. This change might well reflect the shift of fluids seen in body water compartment studies.

DISCUSSION

From the above results a possible mechanism of action of staphylococcal enterotoxin may be postulated. In vitro studies have shown 15 to 20% of the labeled toxin combined with the buffy coat. However, less than 1% was found combined with the buffy coat from in vivo studies 15 minutes following iv administration of the toxic material. It is logical to assume that the material is being bound by the leukocytes which in turn are being cleared rapidly from the circulation by some organ. It has been shown that when leukocytes are altered in some manner by a variety of procedures, they are rapidly removed from the circulation by the lungs (18, 19). Furthermore, it has been shown that the disappearance of the toxin from the circulating blood is considerably faster than that of an equal amount of albumin. Therefore, it may be postulated that the lung is removing the leukocyte-bound toxin from the circulation. This theory is supported by (a) the high concentration of the label found within the lung at sacrifice and (b) by an increase in the number of circulating young and immature leukocytes, suggesting rapid replacement of mature cells. Other workers have also reported a high concentration of I¹³¹ within the lungs following challenge with a labeled exotoxin (20). Further, this high concentration within the lungs continues in non-resistant animals, whereas it falls rapidly in resistant animals (rats and rabbits). Animals which are resistant because of previous experience with the toxin also have continuing high levels of label in the lungs. As pointed out above, this may be explained by postulating that the antigen-antibody complex is being fixed by the phagocytes of the lung. A major part of the radioactivity found within the lung of clean monkeys following challenge is protein bound, suggesting the material is the same as that administered.

Excretion studies have demonstrated that the label is eliminated in the urine much more rapidly from resistant than from non-resistant animals. This label has been shown to be almost entirely inorganic iodide suggesting that it is a direct measure of the ability of the animal to metabolize the toxin. Further evidence of this is the large amount of label being concentrated in the thyroid of previously challenged, resistant, monkeys. Furthermore, it appears that even among non-resistant animals, those which are better able to eliminate the label in the urine are more apt to survive. It thus appears that toxin metabolism, as measured by the appearance of Iodide-131, is associated with decreased effectiveness. Since the

CRAWLEY, BLACK, GRAY, and LEBLANC

lungs of clean and re-challenged monkeys contain about the same amount of radioactivity, and since the re-challenged monkeys are eliminating much more of the label, it would seem that the lung may not be the major organ involved in breakdown of the toxin. This observation is supported by the large amount of radioactivity found in the livers of re-challenged monkeys at the 6-hour sacrifice and the rapid fall-off of radioactivity in subsequent sacrifices, suggesting the liver as a major organ of metabolism.

Results of water distribution within the various body compartments clearly show a loss of fluid from intravascular space, but not from the entire body. As pointed out above, the water is lost into the interstitial space of the extracellular fluid. Preliminary evidence, summarized in Table III, indicates that there is an accumulation of fluid in the lung. Furthermore, a significant histopathologic lesion seen in monkeys is an accumulation of fluid in the interstitial spaces of the lung.* Scattered alveoli may contain fluid but frank pulmonary edema is seldom seen. Likewise, several of the symptoms seen in animals suggest some type of interference with the exchange of gases, for example, dyspnea and cyanosis.

As yet the exact mechanisms involved in death from staphylococcal enterotoxin is merely speculation, but it appears to be closely related to a reaction within the lungs. Thus, illness and subsequent death may be due to interference with gas diffusion in the lungs resulting from an accumulation of fluids in the interstitial spaces. However, additional work must be and is being carried out to determine the details of this mechanism.

SUMMARY

These studies provide no information concerning the possible role of the CNS in the pathogenesis of acute staphylococcal toxin reactions, nor to the cause of vomiting and diarrhea. They do suggest that (a) the toxin is bound initially to WBC's and albumin, (b) that toxin-bound WBC's are trapped in the lung, (c) that the lungs develop increased weight due to interstitial fluid accumulation which occurs coincident with loss of vascular fluid, but with unchanged total body water, ICW and ECF, and (d) that re-challenged and resistant animals handle toxin in a different fashion, evidenced by an early increased accumulation of I^{131} label in the liver and an increased clearance of I^{131} free of protein by the kidney and thyroid.

REFERENCES

1. Murkerji, A., Organism of Summer Gastroenteritis in Lucknow, Dissertation, Doctor of Medicine, University of Zurich, 1955.

*We wish to express our appreciation to Lt. Col. Donald Yost, VC, USAMU, for histopathologic studies.

CRAWLEY, BLACK, GRAY, and LEBLANG

2. Van de Velde, La Cellule 10, 403, 1894.
3. Esselen, W.B., and Lavins, A.S., Bulletin No. 493, University of Massachusetts, College of Agriculture, Amherst, Mass., 1957.
4. Denys, Bull. Acad. Roy. de Med. de Belgique, 4th ser. 8, 496, 1894.
5. Tanner, F.W., and Tanner, L.P., Food Borne Infections and Intoxications, 2nd Ed. Garrard Press, Champaign, Ill., 1953.
6. Dach, G.M., Food Poisoning, 3rd Ed., University of Chicago Press, Chicago, Ill., 1956.
7. Dewberry, E.B., Food Poisoning, 4th Ed., Leonard Hill, Ltd., London, N.W. 1, England, 1959.
8. Blair, J.E., Bact. Rev. 3, 97, 1939.
9. Sugiyama, H., Summary of Work on the Site and Mode of Action of Staphylococcal Enterotoxin, Food Research Institute, University of Chicago, Chicago, Ill.
10. Bayliss, M., Jour. Exp. Med. 72, 669, 1940.
11. Sugiyama, H., Chow, K.L., and Dragstedt, L.R. II, Proc. Soc. Exp. Biol. & Med. 108, 92, 1961.
12. Sugiyama, H., Bergdoll, M.S., and Wilkerson, R.G., Proc. Soc. Exp. Biol. & Med. 103, 168, 1960.
13. Orlov, N.I., Food Poisoning and its Prophylaxis, Moscow, 1952.
14. Schantz, E.J., Roessler, W.G., Wagman, J., Spero, L., Stefanye, D., Dunnery, D.A., and Bergdoll, M.S., The Purification and Characterization of Staphylococcal Enterotoxin B, Army Science Conference, June 1964.
15. Methods in Medical Research, Vol. 5, A.C. Corcoran Ed., The Year Book Publisher Inc.
16. Publication No. 711580, Operating Instructions, Liquid Scintillation Systems 724 and 725, Nuclear Chicago Corp., Chicago, Ill.
17. Gray, I., Biochemical and Physiological Responses Associated with the Administration of Staphylococcal Enterotoxin, Executive Session, Commission on Epidemiological Survey of the Armed Forces Epidemiological Board, August 1963.
18. Weisberger, A.S., Heinle, R.W., Storaasli, J.P. and Hannah, R., J. Clin. Investigation 29, 336, 1950.
19. Weisberger, A.S., Guyton, R.A., Heinle, R.W., and Storaasli, J. P., Blood, J. of Hematology 6, 916, 1951.
20. Gennaro, J.F., and Ramsey, H.W., Nature 184, 1244, 1959.

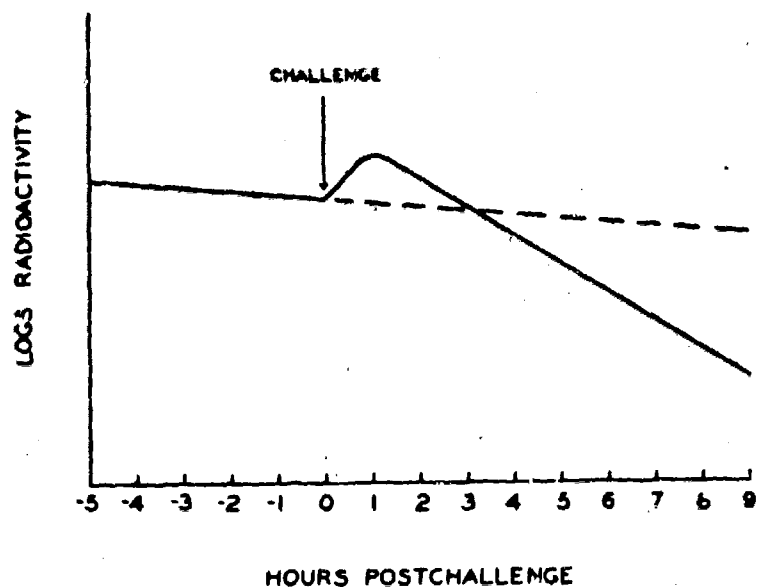


Figure 1: Changes in RISA Distribution

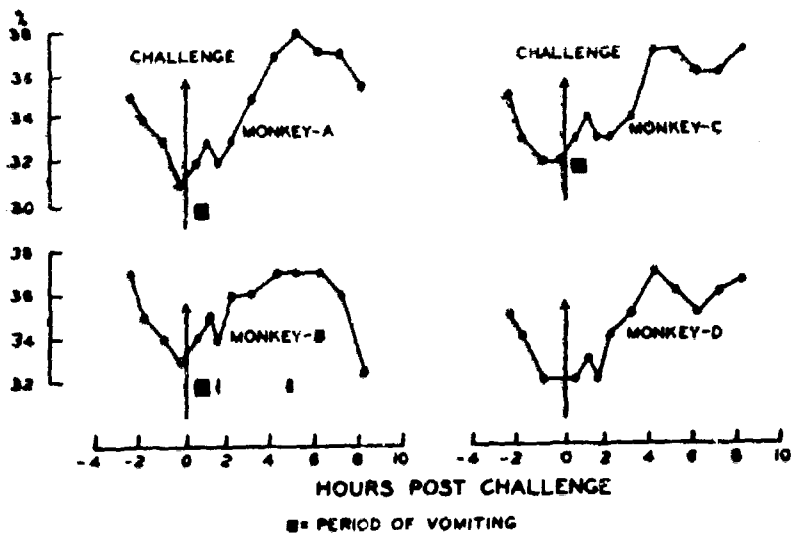


Figure 2: Changes in Hematocrits

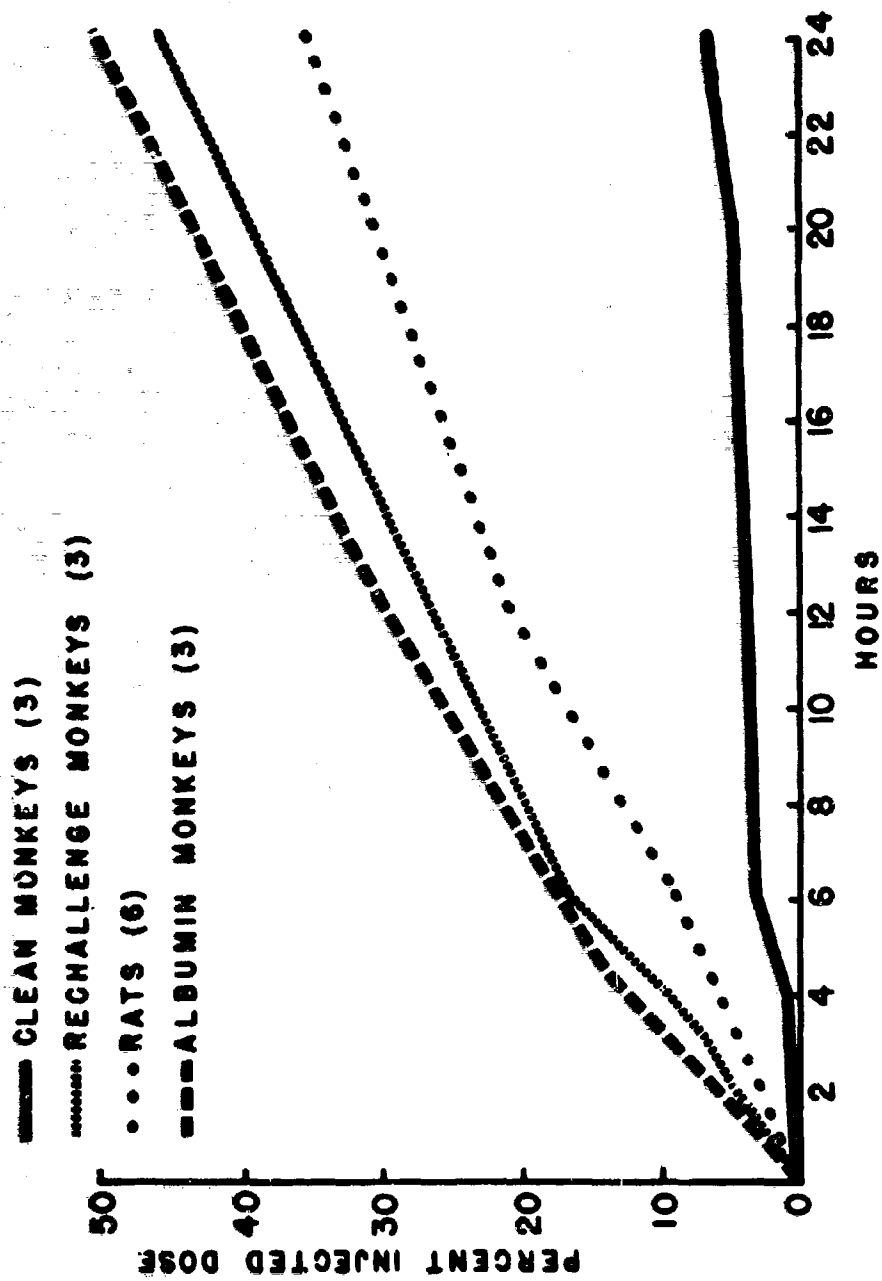


Figure 3: Urinary I¹³¹ Excretion

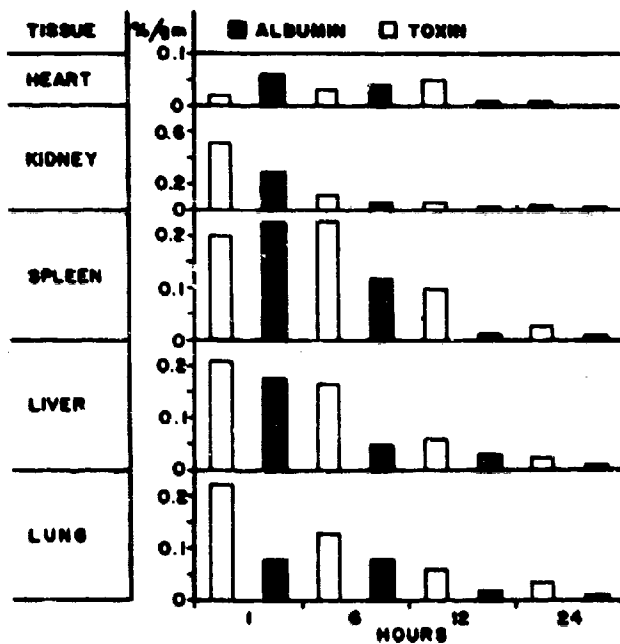
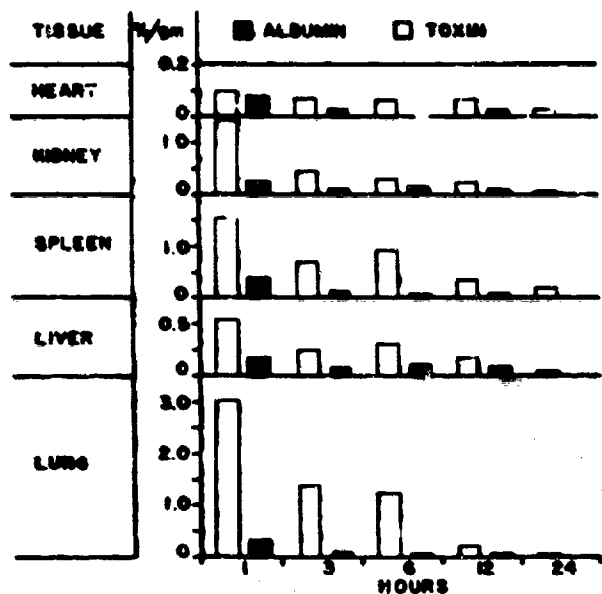


Figure 4: Distribution of ^{131}I in Tissues (wet) Relative to Time
 a. Rats
 b. Rabbits

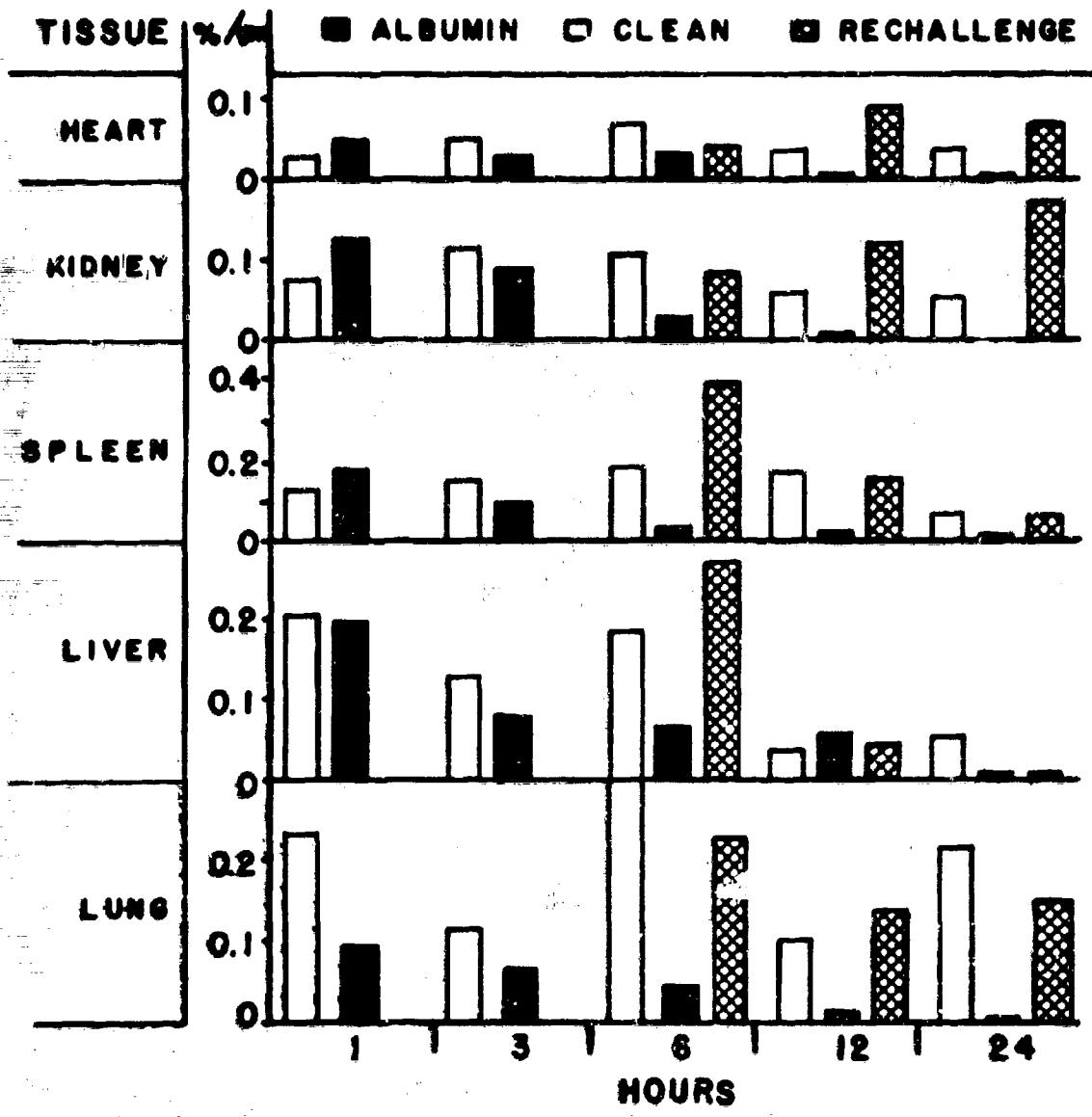


Figure 5: Distribution of I^{131} in Monkey Tissues (wet) Relative to Time

CRIM

CLOSED CYCLE GAS TURBINE

W. M. CRIM, JR.
NUCLEAR POWER FIELD OFFICE
FORT BELVOIR, VIRGINIA

The closed cycle gas turbine power plant has been under development by the firm of Escher-Wyss, Ltd. of Zurich, Switzerland since 1940. During this time some dozen or more central station type plants have been put into operation, in the 12000 KW class. This paper is presented based on the concept, theories explored, design, fabrication, construction, and initial testing in the first experimental compact closed cycle (AK process) thermodynamic system.

As far back as the early 1950's conceptual ideas had been forming around the use of the AK process for a compact, mobile nuclear power plant. Generally the process is referred to as the closed cycle gas turbine power plant. Dr. Keller and Professor Ackeret are credited with the initial patents on the process. The process shows how maximum work can be attained between given temperature limits by the internal transfer of heat. The Carnot and AK Cycles can be assumed to be identical in efficiency between given temperature limits. Both cycles require isothermal compression and expansion over a moderate pressure ratio, the major difference between them being that the Carnot Cycle requires the use of additional machinery and the AK cycle needs heat transfer apparatus to achieve the temperature range. Other advantages of the cycle are the low pressures and exotic gases that can be used, such as Helium, Nitrogen, Hydrogen, or mixtures of gases. Figure 1 shows the circulation pattern of the gas as it traverses the loop during the process.

The action on the fluid or gas in the cycle is one of compression and expansion. The gas is initially compressed and heat is then added by the gas traversing a recuperator or heat exchanger. Additional heat is added by a reactor or conventional heater, after which the gas expands through a turbine and gives up heat through the hot side of the recuperator to the gas just leaving the compressor on the cold side of the recuperator. The gas is then cooled further by traversing a precooler or heat exchanger and then the gas enters the

CRIM

compressor to repeat the cycle again. Power delivered by such a system is raised or lowered by increasing or decreasing the mass flow rate of the gas through the system.

The conceptual design of the original closed cycle nuclear power plant went through numerous optimizations and cycle studies until finally a system was conceptually put together, as shown in Figure 2. The system would use extremely high purity Nitrogen as the coolant, the basic rotating equipment would be of a moderately high speed, 18000 RPM, and the package containing the reactor and power conversion equipment could be held to a weight of 25 tons and a size of 8'X8'X18' which could be handled by the Army's modified 25 ton low-bed trailer. This would then only require a 2 ton truck to carry, in addition, the control shelter. The design went into such depth as to size the equipment and to establish the basis for the detailed design of the components. Table 1 lists those parameters as defined by the study.

It was then decided that an experimental test program should be set up in order to test the ideas of such a system up to that time. There was not adequate time for individual component development even though there was virtually nothing on the shelf that could be used in the system. At this time the Army Nuclear Power Program, under the joint auspices of the Department of Defense, represented by the Office of the Chief of Engineers, and the Atomic Energy Commission approved the design and construction of the Closed Cycle Gas Turbine Test Facility (GTTF) to be located at Fort Belvoir, Virginia and the Gas Cooled Reactor Experiment (GCRC) to be located at National Reactor Test Site in Idaho. The tests and experiments of these two facilities were to be coordinated so as to attain compatibility of system.

DESIGN OF THE GTTF

The GTTF was then initially designed so that as many components as possible could be spread apart in order to better evaluate the component during the forthcoming experimental and evaluation program. Figure 3 shows this spreading out far better than it can be said. This initial design was completed in early 1957.

The machinery set of the GTTF consisted of the closed cycle turbo-compressor set, recuperator, generator system, reduction gear and the appropriate couplings and pallet to assemble such items. The machinery set can be identified as the assemblage in the center of the machinery room in Figure 3. Probably the most critical item of the machinery set and the GTTF is the thermodynamic heart of the system, the turbo-compressor unit. The turbo-compressor is comprised of a two-stage axial turbine and two-stage centrifugal compressor mounted on a common shaft which is supported by two tilting pad radial bearings, and the shaft axial loading is taken up by a tilting pad type thrust bearing, as shown in Figure 4. It was felt that the tilting pad type

CRIM

bearing gave more assurance against oil whip which will occur at half rotational frequencies in lightly loaded high speed journal bearings. The turbo-compressor set is flange mounted to the gearbox at the compressor end and to the recuperator at the turbine end, the turbine being an overhung reaction type. The turbine nozzle rings and interstage diaphragm were supported by a conical member which was mounted on and located by the thrust bearing housing structure. The turbine rotors were bolted to the flanged end of the hollow rotor shaft. The seals at the turbine end are cooled by the Nitrogen gas bled from behind the second stage compressor wheel to the center of the main shaft. This gas flows through drilled holes in the first stage turbine wheel and through the turbine interstage labyrinth. The gas is directed alongside both wheels. Thus it serves to cool both the labyrinth and the wheels. A portion of the gas is fed between the first stage turbine wheel and the turbine shaft seal. Due to the short axial length available for the turbine exit diffuser, and as a result of tests, a set of vortex generator vanes were developed and located at the diffuser inlet, mounted on the turbine exit cone, between the three cone support strut. It should be noted that in order to meet the compactness requirement, the turbine exit cone protruded a few inches inside of the recuperator inlet flange. In order to prevent entry of oil into the main process gas, pressurized labyrinth seals were used at the ends of the bearings. All labyrinths were pressurized at their centers to a level equal to the Nitrogen pressure behind the second stage wheel. The GTTF recuperator was designed around an extended surface pin fin type core as shown in Figure 5. It was designed to use heat transfer passages of extremely small dimensions, since it was expected there would be no danger of fouling or plugging if the working fluid is kept clean. The design was made up of 38 low-pressure passages between which are sandwiched 37 high-pressure passages, block brazed into a single unit. This core was then welded into the exchanger as shown by Figure 5. The generator system including generator, exciter, and start motor was reasonably standard where extreme care had been taken in order to make them as light weight as possible. An actual generator system was used rather than a dynamometer system in order that transient tests could be performed in the future. The reduction gear was of a bull-pinion type due to the compactness of design which required that the generator be doubled back on the turbo-compressor as shown in Figure 6, rather than an inline design employing planetary gearing which would have been too long. Figure 6 shows the basic components of the original compact closed cycle machinery set.

The plate fin precooler originally designed for the roof of the nuclear power plant was moved to the top of the roof of the GTTF as it was felt that the increased pressure drop could be accounted for and handled. The precooler was of a brazed aluminum single pass cross flow exchanger made up of 12 blocks, each having 31 air flow passages between which are sandwiched 30 Nitrogen flow passages of the type shown in Figure 7, block brazed into a single unit. The complete exchanger comprises the twelve blocks, welded together with two joints

ORIN

paralleling the Nitrogen flow pass and two in the no-flow direction. The header "tanks" were welded on the ends of the nitrogen flow pass, making one integral unit of the exchanger. Air was circulated upward through the precooler by four propeller-type fans, each direct driven by its own constant speed electric motor.

The GTTF system piping created quite a challenge to the design engineer. As noted in Figure 3, it is easy to see that all of the piping ends up at the machinery set, and it is virtually impossible for all of the expansion and thermal movement to be taken up by the expansion joints. Design must reduce the majority of stress concentrations in the area of the machinery set and more specifically on the turbo-compressor set. The majority of piping connections were made light weight Marmon type flanges. The load changes in the closed cycle gas turbine power plant employing the AK process are met by changing the weight flow of the working medium in circulation in the system. The design parameters of the GTTF were that the volume and temperatures within the circuit remain the same, then the change in weight flow would be affected by changing the pressure level of the system, by the addition or withdrawal of working fluid. It had been shown by previous closed cycle power plants of a much larger power output that there was a "momentary effect" or almost instantaneous change in power, in the direction called for, if the working fluid was added or withdrawn on the high pressure side. This would then necessitate a supply of working fluid at a pressure greater than the then existing system high pressure. The fact that the system fluid was to be high purity Nitrogen, meant that it must be stored when not in the power plant circuit. In order to attain the compact parameters required of the system, it was necessary that the storage pressure be substantially greater than the maximum circuit pressure. This necessitated the use of a high pressure transfer compressor between the circuit and the storage vessel. As a reduction in power is achieved by withdrawal of working fluid from the circuit, the rate at which power could be reduced would be governed by the capacity of this transfer compressor and the volume of the receiver if it were not possible to temporarily unbalance the power relation between the main compressor and turbine. This was accomplished by means of a bypass valve around the heater, turbine and recuperator which, when open, effectively maintains the load on the compressor while simultaneously unloading the turbine. Therefore, by opening the bypass valve, the net turbine output can be immediately reduced to any desired level while the working fluid is being withdrawn from the circuit over some reasonable length of time.

EXPERIMENTAL TESTING IN THE GTTF

Prior to the final assembly of the turbo-compressor a series of open cycle component tests were run at the factory. These tests included compressor first and second wheel and diffuser calibrations, then calibration of both stages, turbine exit diffuser tests in eight configurations, and then turbine calibration. The compressor stages

GRIN

were assembled separately on a special rig in order to determine the precise configurations of the diffusers and a special test ducting was connected to the turbine tail cone for the vortex generator tests. The remainder of the testing at the factory was accomplished with the turbo-compressor set in normal assembly. The compressor was calibrated by varying the back pressure while holding constant speed. A butterfly valve at the compressor discharge provided the control. The valve was varied from wide open to a position near what was thought to be compressor surge. Several points were taken at each of the following speeds; 12000, 14000, 16000 and 18000 RPM. The only point taken at the surge line was at 14000 RPM. Multiple runs were made with the compressor incorporating indicated required changes in design as the testing proceeded. The limited turbine performance data allowed some estimated curves to be generated and at that time indicated that the turbine could be expected to operate in the range of 85% efficiency at design conditions. The turbo-compressor set was then delivered to the GTTF and installed on the machinery bedplate.

To gain operating experience and accumulate performance data on the turbo-compressor set, free of complications introduced by the recuperator and pre-cooler operation and of unknowns associated with the compact closed cycle system, the open loop was utilized for the initial tests. The loop included compressor inlet line and filter, compressor, line to heater, heater, line to turbine, turbine, and turbine exhaust duct. For these open cycle tests, the TC set, reduction gear, generator, exciter and startmotor were assembled on the machinery set bedplate. Installed in the loop were a control system bypass, overspeed trip valves, system for introducing boost air from an external compressor in the compressor discharge line, and the associated piping for such systems. When self-sustaining operation could not be attained at the speed of the four-pole startmotor, other methods were sought in an attempt to raise the turbine speed. It was decided to admit high pressure air between the compressor discharge and heater inlet. The air was introduced through a "boost" nozzle. This would in effect deliver greater mass flow to the turbine without requiring additional work from the compressor. The quantity of boost air required was not attained with the equipment on hand and self run in the open cycle configuration was never attained at the GTTF.

There was very little component testing of the heat exchangers to be used in this first compact closed cycle loop as the loop itself was the only available facility that could adequately generate the conditions. Standard static pressure and leak checks were made on the recuperator and pre-cooler and both seemed to stand the pressure checks with no problems. The recuperator however appeared to have a very small internal leak from the high pressure pass to the low pressure pass which constituted the same effect to the thermodynamic system as bypassing the turbine.

Initial self-sustained operation was attained at a compressor inlet pressure of 61 psia and a turbine inlet temperature of

CRIM

1100°F. After reaching self run, the turbo-compressor set accelerated to 15200 RPM, then decreased back to self sustaining speed. The system was then shut down due to bearing lubricating oil leakage into the loop. After much investigation and lubrication system experimentation it was determined that the lubrication problem was an operational one. The operational procedures were then adjusted so that the array of conditions which would allow the oil leakage into the gas stream could not occur during startup, experimentation, and shutdown of the system; basically this was accomplished by adjusting the oil flows to the bearings more commensurate with the required flows, thereby providing conditions which would allow the oil drains in the bearings to take care of the required oil flows.

At this point it is important to review some of the specific objectives of the initial testing to be performed in the GTTF, which were to: determine the compressor and turbine performance characteristics in a closed cycle configuration; determine the optimum starting pressure, speed, and temperature combinations; determine the flow factors for various cycle conditions; establish optimum startup and shutdown procedures; determine the effectiveness of both the precooler and recuperator; provide some endurance testing of those rotating components to insure the integrity of the rotor dynamics of the system; and to evaluate the use of gas injection and bleed-off as a means of starting the system.

The facility in the original GTTF configuration was experimentally operated for over 1600 hrs. During this period, the majority of planned testing was accomplished. A brief resume of the testing that was performed is listed in the following Table 2 entitled GTTF summary of testing.

TABLE 2 GTTF SUMMARY OF TESTS PERFORMED

1. Component test and investigations
2. Lubrication system experimentation
3. Turbine calibration including performance curves
4. Compressor calibration and mapping
5. Heat exchanger performance curves development
6. Impingement starting tests
7. Thermodynamic flow experiments
8. Endurance testing
9. Startup and shutdown of complete system

During this period of experimental testing the system produced up to 106 KW net power from the generator. Under the performance evaluation testing of the system at least 90 separate runs were made even though for the most part the net power produced was less than 10 KW. The testing was interrupted many times in order to correct, repair, or clean the heat exchanger equipment. Following this phase of experimentation, the turbo-compressor set was modified to reduce the effective turbine flow area and, thus improve the match

CRIM

between the turbine and compressor. At the same time it was decided to provide separate and independent lubrication systems for the turbine and compressor bearings. After these and other minor modifications the system was reassembled and the silver seal run-in was commenced. Eleven runs were conducted in order to run in the seal. Numerous runs were then conducted in order to further evaluate the turbo-compressor set. Self-sustained operation was easily attained and a gross output of approximately 130 KW was indicated. The turbine was further calibrated and the investigation of compressor surge was completed for this configuration. The next step was to evaluate the system transfer functions. Significant information was developed in this area but the investigation was not fully completed due to further opening up of the recuperator allowing gross internal leakage. Through all of this testing the bearing and seal system, after debugging appeared to be functioning in an acceptable manner, and after the transfer function tests were terminated, additional endurance testing was continued until it was felt that the bearing and seal system of the turbo-compressor was optimized for this particular system.

During the experimental test program the following major modifications were made to the GTTF and its components: 1. The gas duct from the heater to the turbine was replaced; 2. The recuperator and precooler were thoroughly cleaned and repaired; 3. Additional flow measuring devices were installed in the loop; 4. A fixed over-speed device was added to the turbo-compressor set; 5. The turbo-compressor set was modified in order to improve the match between the turbine and compressor.

ANALYSIS OF RESULTS

In analyzing the results to date on the compact closed cycle system it is appropriate to comment on four general areas. They are sizing of equipment, matching of the equipment, performance of the components and system, and the analytical methods used in studying the system and its components.

The sizing of the equipment in a compact closed cycle system is extremely critical. These turbo-compressor sets, recuperators, and the like are, by design, light weight and must stand extreme temperature differences. The initial TCS-560 was limited to a weight of 750 pounds which only allowed relatively fragile construction. Most of the problems of this turbo-compressor set were because it was designed too light weight and could not stand the thermal distortion. The heat exchangers both were of the type construction where thin-strip metal was in contact with much heavier sections and the thermal distortion caused excessive leakage.

The matching of the equipment can be made much easier if detailed design analysis is performed prior to design, during design, and even continued during fabrication, installation, and up to final tests.

CRIM

The performance of the turbine and compressor is best indicated by the curves of Figures 8 and 9. It is important at this time to generally discuss the net power generated by the system and the many factors which experimentation has taught effect the performance of the compact closed cycle. As noted in the equation of Table 3, the effect of pressure loss over and above the design ΔP 's of the system is devastating on the net KW output. It is noted that the greatest pressure loss effect is in the low pressure side of the system which effectively increases the compressor pressure ratio and power required to drive the compressor. The losses in the high pressure side of the loop further reduce net work by lowering the expansion ratio and gross turbine power generated. It is obvious that any decrease in expected efficiency will adversely effect the net power generated. One significant contributor to a net power deficiency is reduced flow through the turbine. Any internal leakage in the recuperator will have the same effect as a turbine bypass which again reduces the gross output of the turbine. Another type bypass of the turbine was seen when thermal distortion caused excessive leakage around the turbine stages. It has been demonstrated that the actual windage and parasitic losses were reasonably higher in the closed cycle system than was expected, again reducing the net power of the system. The closed cycle gas turbine power plant may be considered to have an infinite number of stable operating points, and a change in any system parameter stimulates the system to readjust any necessary pressures, temperatures, flows, speeds, and so on, until some new stable operating point is found where all equations of balance are again satisfied. It has therefore been realized that each component in a compact closed cycle loop affects, and is affected by, the performance of every other component in the loop.

The analytical methods used in studying the system and its components have proved very informative when the appropriate parameters could be measured to any degree of accuracy. Although a minor amount of analysis was conducted during the performing of an experiment in order to establish that the objectives of the experiment were being satisfied, in general the reduction of data and analysis of data is started at the completion of the experimental run. Data reduction was performed with an RCA 301 computer using a special data reduction code developed for this purpose. The conclusions and performance analysis were then made from a review of the data reduction process.

CONCLUSION

In conclusion it should be brought out that considerable experimental testing should be performed on the Compact AK process, but the results of the initial experimental work to date have shown that the compact closed gas cycle employing the basic aerothermodynamic principles, as discussed in the paper, can provide the military with an efficient compact reactor system.

CRIM

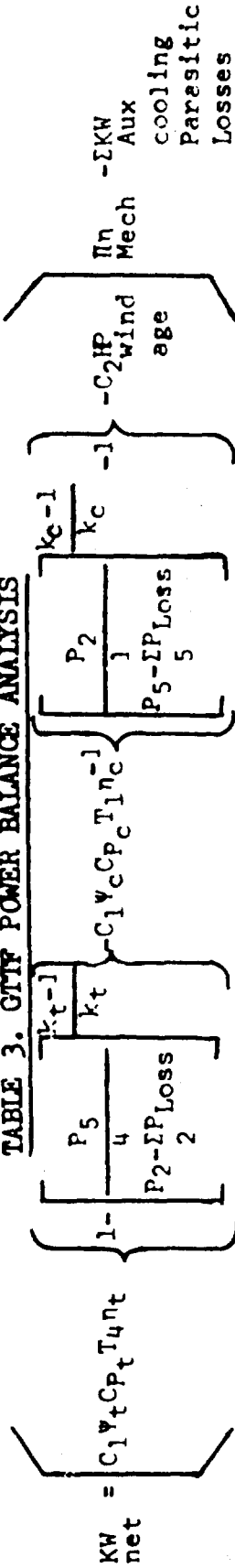
BIBLIOGRAPHY

1. Swain, R. K., Mader, G. F. and Crim, W. M., Design and Testing Aspects of the ML-1 Mobile Nuclear Power Plant, ASME Paper 61-WA-316.
2. Sanderson and Porter, Closed Cycle Gas Turbine Test Facility, Design Report, 1957.
3. Fairchild-Stratos, Closed Cycle Gas Turbine Compressor, Design Report, 1961.
4. Aerojet-General Corporation, Experimental Testing and Maintenance of the Gas Turbine Test Facility, Report, 1963.

TABLE 1. GTTF SYSTEM CHARACTERISTICS

General	80	Heater p (in) psia	186	Nr Blades stator/stator	79
Ambient temp (°F)	469	Turb t (in) of	1150	Reduction gear	Spur Gear-Pinion
Gross elec (KW)	415	Turb p (in) of	183	Type	18000:1200
Net elec (KW)	24.5	Recup T (in) of	847.5	Ratio (rpm)	2637
Cycle eff (thermal)	16.6	Recup p (in) psia	73	WR2 (#ft ²)	3100
Flow rate (#/sec)	2.88	Precooler t (in) of	420	Weight (#)	(Ext surf) Pin Fin
Mach set dim. (FT)	3.5x4.5x2.5	Precooler p (in) psia	70.7	Recuperator	
Temperature ratio	N2	Turbo-compressor set		Type	
Thermo Fluid	High Pur	Radial Bear (2)	Tilting Pad	Envelopes (Nr)	69
Generator	(AC) Synchronous	Thrust Bear (1)	Tilting Pad	Pins (Nr)	8x106
Type		Seals	Labyrinth	Linear Ft strip	5956
Voltage-Output (KW)	480 V-400KW	Housing	321SS	Heating Surface (Ft ²)	4360
Power factor	.8	Weight (#)	1100	Surf Compactness	(Ft ² /Ft ³)
RPM	1200	Speed (RPM)	18,000	Extended Surf (%)	72
WR2 (#ft ²)	1050	Nr Turb Blades (Stage)	79	Fasting Surf (Ft ³)	19.66
Cycle	60	Turb stages (Nr)	2	Int Vol w/manifolds	(Ft ³)24.5
Heater characteristics		Type (overhung)(Axial)(Reaction)		Heating Surf wt (#)	1530
Flow rate (#/hr)	70,509	Turb rotor mat'l	SAE 4340	Wt heating surf/ft ³ (#)	77
Heat Load (Btu/hr)	7.35x10 ⁶	Turb blade mat'l	Inco 713	Wt Hx without insul	2355
Pressure drop (ΔP/P)%	2	Turb Stat mat'l	Inconel	Wt Hx with insul	2500
Size (FT) ³	5x5x13	Exp ratio	2.56:1	Pin mat'l	Armaco Iron
Burners (Nr)	3	Turb Isen Eff (%)	85	Sheet mat'l	430SS
Tube mat'l	SS	Comp stages Nr	2	Marmon Flange mat'l	347SS
Power Cycle		Type	shrouded-centrifugal	Brazing mat'l	Copper
Type	AK process	Comp mat'l	cast AL alloy 355-T71	Effectiveness (H.P.)%	83.2
Total Vol Fluid (FT ³)	85	Nr Blades (Stage)	12	Heat Load (Btu/hr)	7.75x10 ⁶
Total wt Fluid (#)	30	Nr Half Blades (Stage)	12	Total Pr. drop (%)	3.3
Fluid transit time (sec)	1.5	Nr. Diffuser Blades (Stage)	14	Flow rate (rated) #/hr	70,500
Net power, KW	415	Comp ratio	2.75:1	Precooler	
Comp t (in) of	100	Comp isen eff (%)	80	Type (Serpentine)	Plate Fin
Comp p (in) psia	68.6	Norm Breakaway torque (in#)	95	Flow (#/hr)	66,500
Recup t (in) of	325	Norm Coast down (t) sec	52	Heat Load (Btu/hr)	5.3x10 ⁶
Recup p (in) psia	189	Turb Volute mat'l	Inconel	Relative Pr drop (ΔP/P)%	1.33
Heater t (in) of	760.5	Turb disc mat'l	Inc Forg 901	Effectiveness (Hot stream)%	94

TABLE 3. GTTF POWER BALANCE ANALYSIS



- C_1 - Constant = $.745 \times 778 \div 550$
- C_2 - Constant = $.746$
- v_t - Turbine flow in #/sec
- C_{pt} - Av spec heat (turb)
- T_4 - Total temp turb in OR
- n_c - Adiab Comp eff
- k_c - Ratio sp heats (comp)
- n_t - Adiab turb eff
- P_5 - Total press turb out
- P_2 - Total press comp out psia
- k_T - Ratio sp heats (turb)
- $\frac{k_T - 1}{k_T}$ - .259
- v_c - Compressor flow in #/sec
- C_{pc} - Av spec heat (comp)
- T_1 - Total temp comp in OR
- IP - Windage = Rotational losses HP
- P_n - Product Mech eff
- $\frac{k_c - 1}{k_c}$ - .286

	v_t	C_{pt}	T_4	n_t	$1 - \left(\frac{P_5}{P_4}\right)^{\frac{k-1}{k}}$	v_c	C_{pc}	T_1	n_c^{-1}	$\left(\frac{P_2}{P_1}\right)^{\frac{k-1}{k}}$	HP_w	Πn	KW Aux	KW Cool	KW parasitic Loss	KW T	KW C	KW Net
Typ	16.2	.273	1610	.84	.23358	16.6	.25	560	1.22	.33147	47.25	.89	14	22.4	3.75	1389	930	307.35
Est Des (orig)	16.6	.273	1610	.85	.23358	16.6	.25	560	1.25	.33147	52.3	.905	2	12	49	1520	1020	400
Exp Run 301 Prob	15.7	.273	1660	.81	.22	16.6	.25	560	1.285	.345	70	.905	Base Power		73	1230	981	106
Pre-diction	16.2	.273	1610	.83	.22358	16.6	.25	560	1.27	.33147	70	.905	5	22.4	73	1310	930	209.6

CRDM

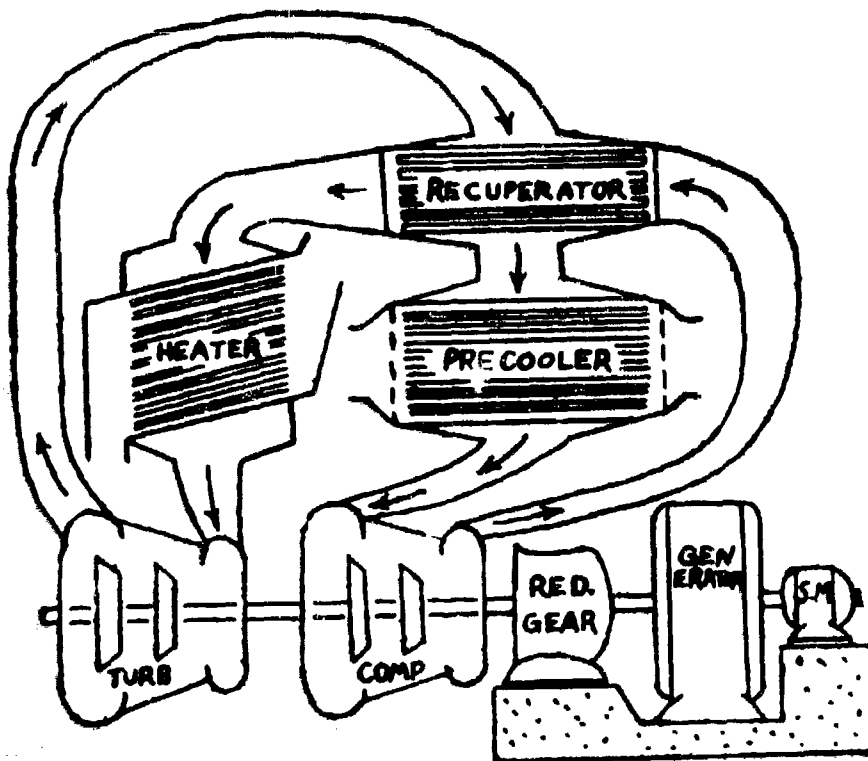


FIG. 1 GTF SCHEMATIC

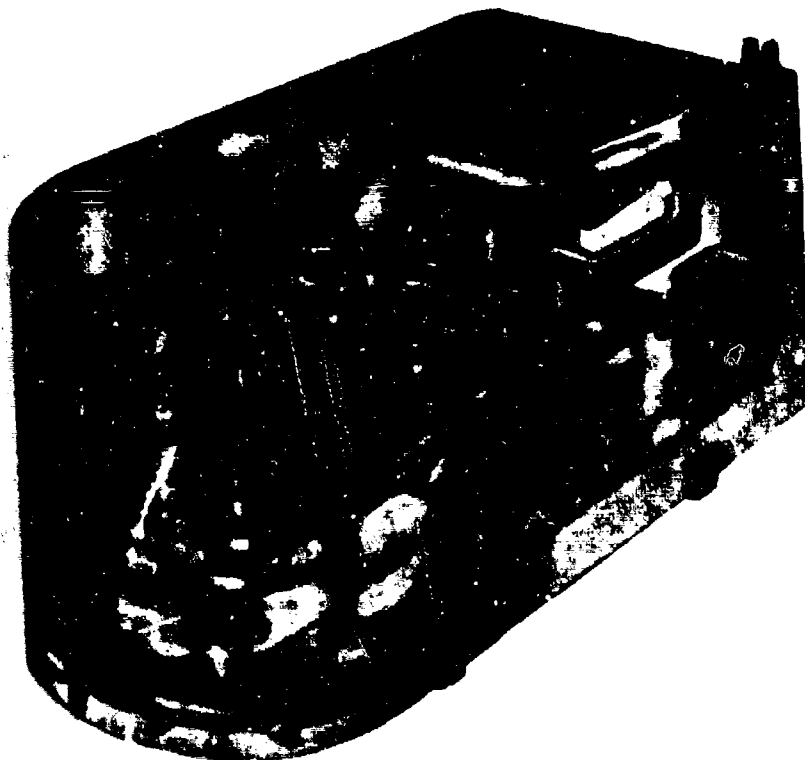


FIG. 2 INITIAL COMPACT GAS COOLED REACTOR SYSTEM CONCEPT

ORDM

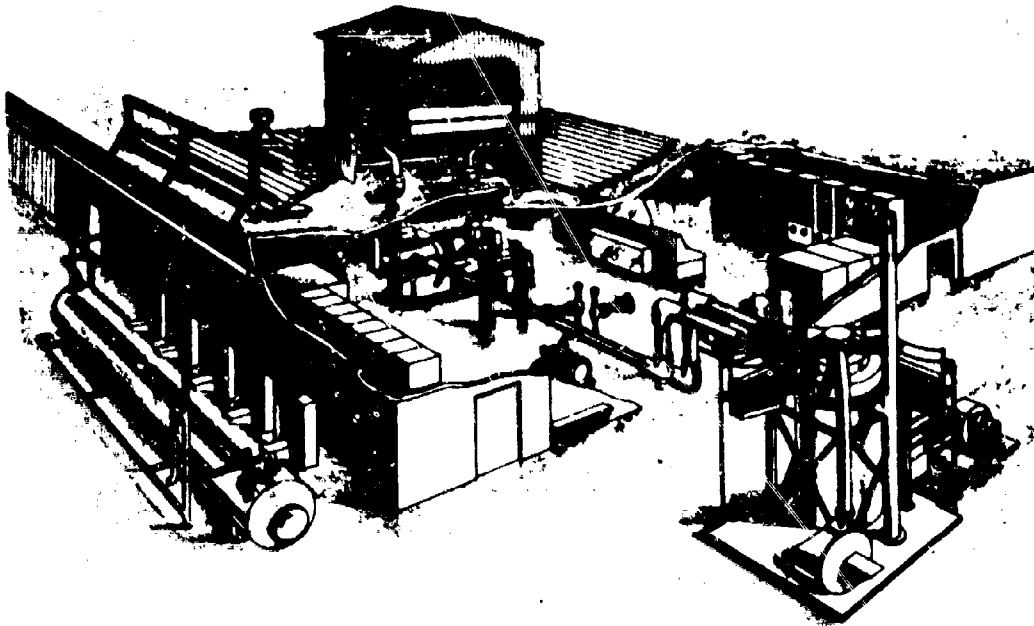


FIG. 3 CLOSED CYCLE GAS TURBINE TEST FACILITY - FT. HELVOIR, VA.

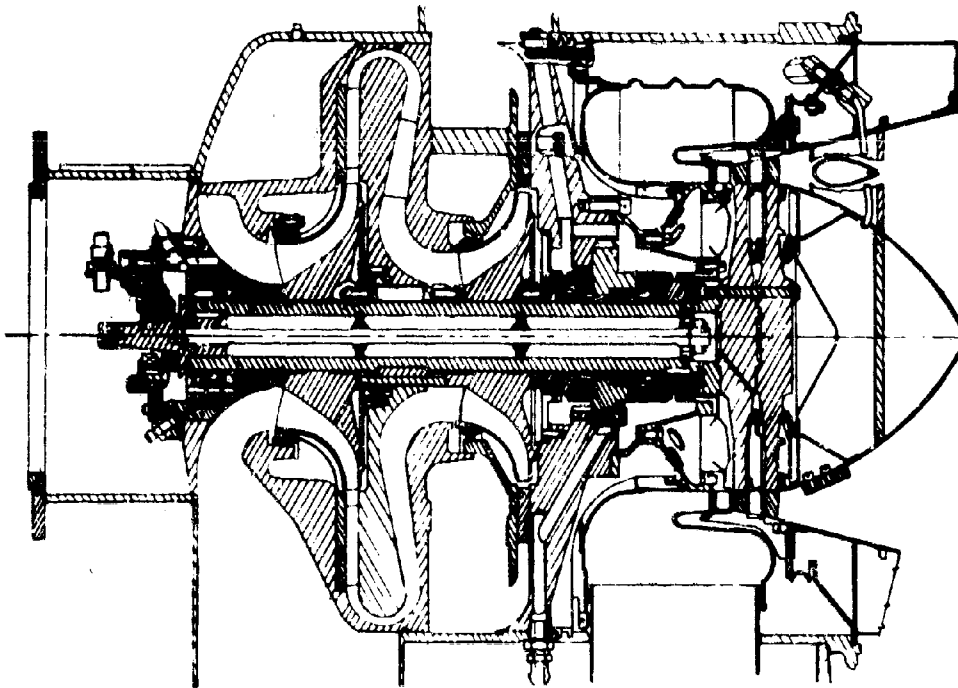


FIG. 4 INITIAL CLOSED CYCLE GAS TURBINE - TCS 560

ORIN

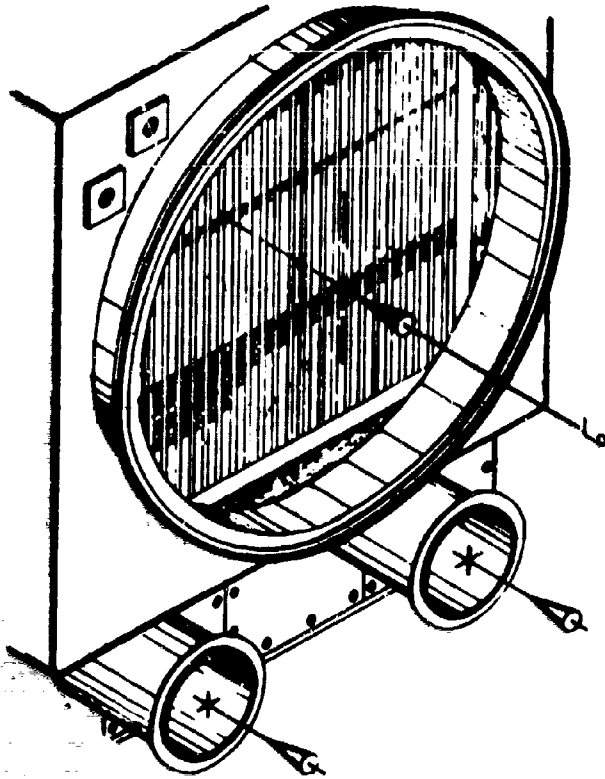


FIG. 5 EXTENDED SURFACE
RECUPERATOR (HP & LP) INLET

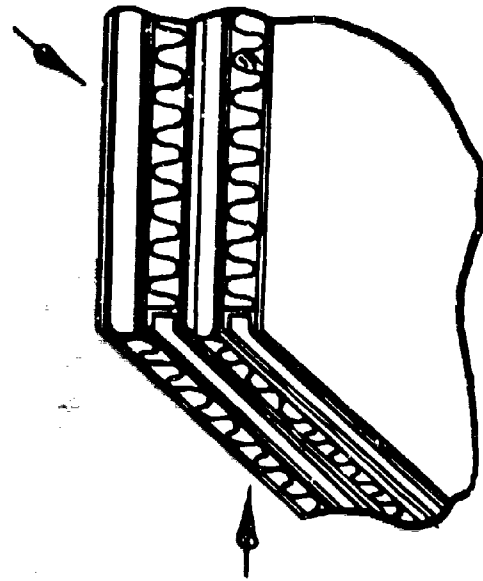


FIG. 7 SERPENTINE
PLATE FIN PRECOOLER

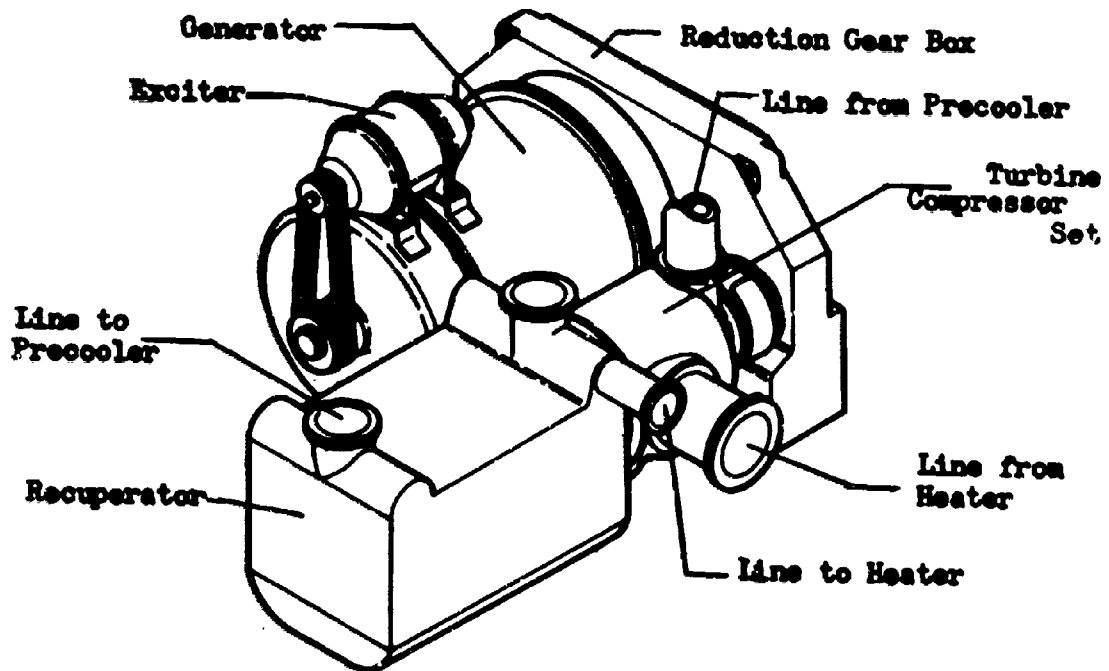


FIG. 6 GTTF MACHINERY SET

GRIN

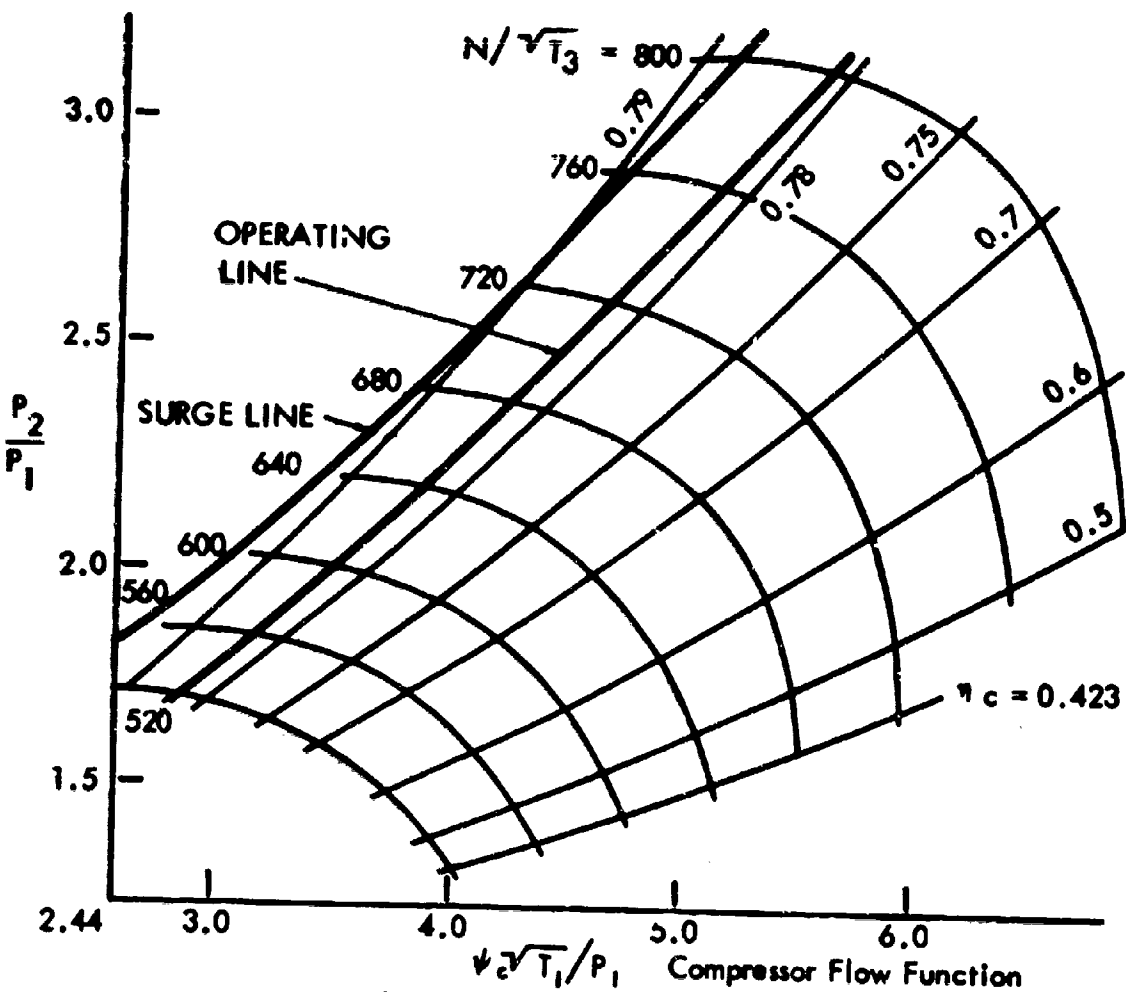


FIG. 8 COMPRESSOR MAP

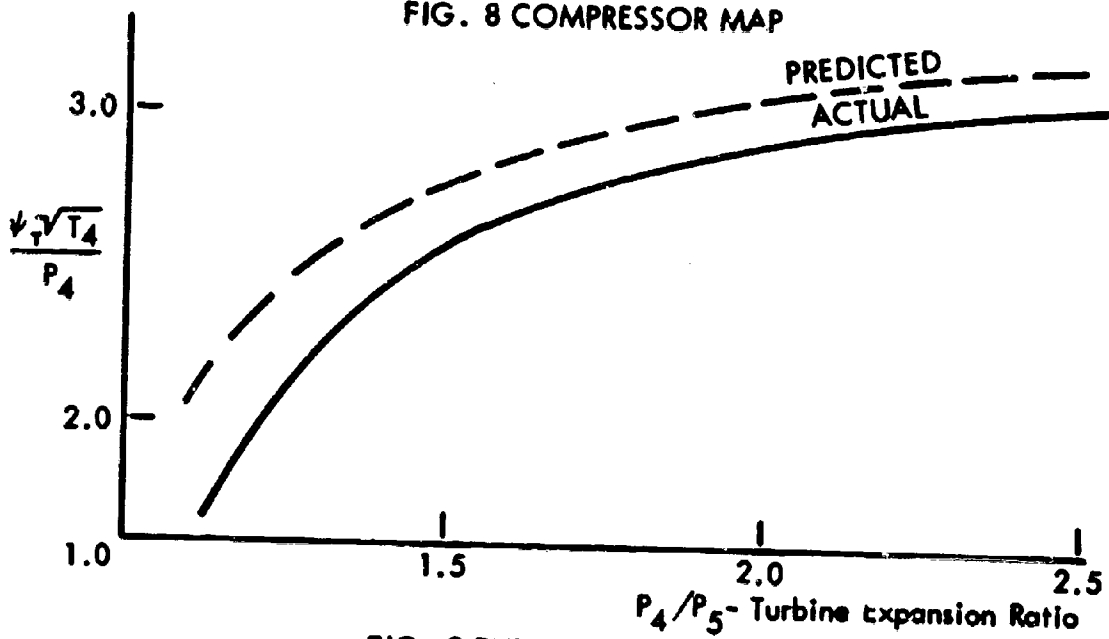


FIG. 9 TURBINE PERFORMANCE

THE EFFECT OF PRESSURE ON THE STRUCTURE AND PROPERTIES OF METALS

T. E. DAVIDSON, J. C. UY, and A. P. LEE
WATERVLIET ARSENAL
WATERVLIET, NEW YORK

INTRODUCTION

The classical theory of elasticity and plasticity predict that hydrostatic pressure, regardless of magnitude, will not cause plastic deformation or affect yielding and flow characteristics. However, the predictions of elasticity and plasticity are predicated upon ideal materials or continuum which often differ markedly from real materials. In contrast to theory then, it may be possible that sufficiently high pressures have some permanent effect on the mechanical and structural characteristics of metals. It is the purpose of this investigation to examine two aspects of the effects of large hydrostatic pressures on metals — the effects on structure and ductility.

In considering how hydrostatic pressures might affect structure, where the structure referred to is microstructural as would be visible through an optical microscope, one is reminded that the physical and mechanical properties of metals are often anisotropic. The elastic property linear compressibility, which is defined by the crystallographic relationships:

$$S_{11} + 2S_{12}$$

$$S_{11} + S_{12} + S_{13} - \gamma^2 (S_{11} + S_{12} - S_{13} - S_{33})$$

for the cubic and non-cubic cases respectively, can vary widely from one crystallographic direction to another for non-cubic metals depending upon how greatly the coefficient of the direction cosine γ deviates from zero. Since the direction cosine does not appear in the cubic relationship, the cubic metals are isotropic with respect to linear compressibility.

There are no macroscopic shear stresses associated with the exposure of metals to hydrostatic pressure, thus no macroscopic plastic flow. However, in the case of the polycrystalline form of those non-cubic metals exhibiting anisotropy in the linear compressibility, an imposed hydrostatic pressure will induce microscopic shear stresses at the grain boundaries. If the degree of anisotropy is great enough and the grain-boundary angle large, sufficiently high hydrostatic pressures may result in shear stresses in the boundaries exceeding the critical resolved shear stress, in which case, microscopic plastic deformation will result.

Johannin and Vu(1) observed the occurrence of such deformation in cadmium at low pressures. Davidson and Homan(2) observed microscopic deformation in polycrystalline bismuth to pressures of 20 kilobars. In this present investigation, several metals of varied degrees of anisotropy in the linear compressibility are examined to very high pressures in order to establish under what conditions hydrostatic pressure-induced deformation can occur and its form as a function of structure and mode of deformation.

Based on the work of Bridgman(3) and Pugh(4), the ductility of several metals is known to increase when extension is performed under a hydrostatic pressure. Usually, the ductility increases more or less linearly with superimposed pressure. Pugh(4) observed in the case of zinc and bismuth that the ductility increased abruptly over a very narrow pressure range. As part of a broad program associated with the study of the effect of pressure on deformation characteristics, in this current work the effect of pressure on the ductility of three highly brittle metals is examined.

PROCEDURE

The hydrostatic pressure system utilized in this investigation is shown schematically in Figure 1. It is a modified Bridgman-Birch type device capable of operating to pressures of approximately 40 kilobars. Pressure is produced by driving the piston into the 3/4 inch cavity by means of a hydraulic jack. The inner pressure cylinder is tapered on the exterior surface. Therefore, as the internal pressure is increased, the lower jack forces the inner cylinder into the tapered cavity of the restraining jacket. In this manner, the large pressure stresses in the inner cylinder are counteracted and the cylinder remains elastic up to the pressure limit. The bottom closure permits the introduction of seven leads into the pressure cavity for various pressure, temperature and electrical measurements. Pressure is measured by means of a manganin coil in conjunction with a Foxboro recorder.

The study of pressure-induced microscopic deformation was accomplished by metallographically preparing the surface of small samples prior to pressurization. Since the pressure medium used was

liquid, the prepared surface was not damaged during pressurization and could be examined directly after removal from the pressure cell without further preparation. For establishing the role of the degree of anisotropy, polycrystalline Zn, Cd, Sn, Al, Zr, Mg, and Cu were examined progressively to pressures of 20 kilobars except where otherwise noted.

To determine any residual effects on mechanical properties resulting from the pressure cycling of an anisotropic polycrystalline metal, a group of zinc tensile specimens were pressurized to 20 kilobars with subsequent testing in an Instron tensile machine. The zinc utilized contained small additions of Fe and Cu to raise the recrystallization temperature above ambient.

The conduct of the tensile tests under a superimposed pressure in order to measure the effects of pressure on ductility was by means of the segmented cylindrical yoke assembly shown in Figures 2A and 2B in the disassembled and assembled condition respectively. This tensile fixture fits into the pressure cell and bears against the bottom closure as shown in 2B. Two of the fixture segments are supported by the closure and grip the top shoulder of the tensile specimen. The remaining two segments bear against the bottom shoulder of the specimen and are forced downward by the forward motion of the pressure piston. In this manner, a tensile stress is exerted on the sample. The test is conducted at a constant displacement rate of 0.050 in/min.

To minimize pressure rise during the conduct of a tensile test, a hydraulic piston-cylinder intensifier device is inserted between the pressure piston and the moving fixture segments. This intensifier, which is shown resting on top of the tensile fixture in Figure 2B, amplifies the piston displacement by a factor of 10. This amplifier was not used in this current work at pressures above 5 kilobars.

Although not of particular importance to the experiments associated with this work, specimen strain and load can be measured internally. Strain is measured by a slide wire mounted on one segment of the fixture which forms two sides of a Wheatstone bridge with an external helipot completing the bridge. Load can be measured by a strain gage bridge mounted on the reduced section of the movable fixture segments.

The materials utilized for the conduct of the experiment were as follows:

- a. Zinc, .2% Cu, .005% Fe - hot rolled and annealed plate.
- b. Tungsten - 3/8 inch swaged and stress relieved rod.
- c. Magnesium - 3/4 inch hot extruded rod.
- d. 1045 Steel - water quenched from 1550°F and untempered.

RESULTS AND DISCUSSION

The magnitude of the localized deformation induced by the hydrostatic pressure varied widely for the metals investigated, ranging from none at the peak pressure to very severe at only a few kilobars. Figure 3 is typical of polycrystalline zinc as a result of exposure to pressures to 20 kilobars. As can be noted, the deformation, originally localized along the grain boundaries, initiates below 5 kilobars and increases in intensity with increasing pressure. At the higher pressures, substantial twinning is also observable. It is interesting to note that, based on prior work(2), no twinning was induced in bismuth as a result of pressure exposure although it is the primary mode of deformation at atmospheric pressure. Thus, it is possible that a superimposed pressure may have an effect on twinning propensity in some metals.

Typical appearances of the pressure-induced deformation in cadmium and tin at the peak pressure are shown in Figures 4A and 4B respectively. In the case of cadmium, the deformation is quite severe with extensive twinning. Tin, however, required a pressure exposure of 26 kilobars in order to initiate the deformation shown, which is boundary migration. No deformation was observed in the remaining metals investigated.

The magnitude of the deformation observed as a function of the degree of anisotropy in the linear compressibility is summarized in the following table.

<u>MATERIAL</u>	<u>CRYSTAL STRUCTURE</u>	<u>LINEAR COMPRESSIBILITY RATIO k_c/k_a</u>	<u>DEGREE OF DEFORMATION</u>
Cadmium	H.C.P.	11.25	Severe
Zinc	H.C.P.	7.50	Severe
Magnesium	H.C.P.	1.04	None
Zirconium	H.C.P.	1.05	None
Bismuth(2)	Rhom.	2.30	Severe
Tin	Tetra.	1.12	Slight
Copper	F.C.C.	1.00	None
Aluminum	F.C.C.	1.00	None

The measure of the degree of anisotropy is the ratio of the linear compressibility in the "c" direction to that in the "a" direction. As can be noted, the propensity to deform under pressure relates directly to the degree of anisotropy. No deformation was observed in the case of the cubic metals which are totally isotropic and in magnesium which is nearly isotropic. Tin, which is only slightly anisotropic, required a pressure of 26 kilobars to produce only slight deformation. The highly anisotropic bismuth(1), zinc and

cadmium exhibited severe deformation initiating at low pressure levels. Based on the results in tin, most probably magnesium would also deform if it were exposed to substantially higher pressures than used in this experiment.

As far as the effect of pressure cycling on residual mechanical properties is concerned, one would not expect to see any property change in a single phase isotropic or nearly isotropic material since there would be no microscopic shear stresses to cause structural modifications. This has been borne out by Ferron's⁽⁵⁾ results on magnesium and aluminum to pressures of 13.8 kilobars. However, when one considers a multi-phase material in which there is a substantial difference in the elastic properties of the two phases, or a polycrystalline anisotropic metal, pressure cycling might have an effect. Radcliffe⁽⁶⁾ found a loss in the sharp yield point in annealed mild steel pressure cycled to pressures of 25 kilobars at ambient temperature. This lowering of the yield strength has been attributed to shear strains in the region of the carbide and ferrite phase boundaries due to differences in compressibility.

Summarized below are the average and data-spread values for flow stress at 1 percent strain, ultimate tensile strength, and elongation for each six pressure and un-pressure cycled polycrystalline zinc tensile specimens.

	<u>FLOW STRESS</u> (PSI)	<u>ULTIMATE TENSILE STRENGTH</u> (PSI)	<u>% ELONGATION</u>
Pressure Cycled (20 kilobars)	14,329 (14,000-15,025)	18,183 (17,800-18,600)	65.8 (57.5-76.0)
Un-pressure Cycled	13,160 (13,000-13,500)	17,667 (17,500-18,010)	63.0 (61.5-65.0)

As can be noted from the above, there appears to be a slight tendency for pressure cycling to increase the 1% flow stress and ultimate tensile strength by 8.9 and 2.8 percent respectively based upon the average value. However, the change is small and could be readily accounted for by experimental error, particularly in the case of the ultimate tensile strength. The only real difference in elongation appears to be increased data scatter in the case of the pressure cycled material.

One might expect that the pressure induced strains in the region of the grain boundaries might result in the occurrence of the Bauschinger Effect, i.e. the lowering of flow stress in the loading direction opposite to that associated with initial straining. The fact that pressure cycling did not decrease flow stress indicates

DAVIDSON, VI, and LEE

that the strains are not such as to yield an average strain that would manifest itself by a reduction in the tensile flow stress.

In contrast to pressure cycling, a superimposed hydrostatic pressure can have a pronounced effect on some of the mechanical properties, the principal one being ductility. In Figure 5 is shown the percent reduction in area as a function of pressure for the three metals investigated in this current work. Also shown for comparison purposes is a curve from Pugh⁽⁴⁾ for zinc, along with some data for the zinc used in other phases of this current work.

It should be noted that the pressure varies somewhat during the conduct of a test due to the specimen displacement being induced by movement of the main pressure piston. Below 5 kilobars, the displacement amplifier retains the pressure effectively constant. Above 5 kilobars, the pressure may vary by 5-15 percent depending upon the magnitude of the strain to fracture for the material. In the case of the more brittle materials, i.e. steel and tungsten, the pressure did not vary by more than 1-2 kilobars, even at the highest pressure, due to the low total elongations involved. The data points shown correspond to the pressure at the onset of specimen strain.

As can be seen, the reduction in area for magnesium increases more or less continuously with pressure from 6% at atmospheric to 88% at approximately 20 kilobars. There is some tendency for the pressure effect to level off at the higher pressure, but this may be due to the difficulty of precisely measuring very large reductions in area.

In contrast to magnesium, zinc, tungsten, and the 1045 steel with an untempered martensitic structure undergo discontinuous changes in ductility with pressure. Increasing pressure has no effect on ductility up to a given level after which there is an abrupt increase. For the case of zinc⁽⁴⁾, the reduction in area changes from an atmospheric pressure value of 8 percent to 100 percent over a very narrow pressure range at approximately 1 kilobar. Tungsten and 1045 steel show no pressure effect on ductility up to in excess of 5 and 10 kilobars respectively. Above these pressures, there is a sudden increase in ductility to 48 percent reduction in area for the tungsten at 15 kilobars and 62 percent for the 1045 steel at 20 kilobars. Just how abruptly the transition actually occurs is the subject of current study.

The systematic increase in ductility with pressure has been attributed to the effect of a superimposed hydrostatic pressure on the extension of voids found in the region of the neck of a ductile type fracture. Why certain metals, such as those cited above, undergo abrupt transitions in ductility is uncertain. Pugh⁽⁴⁾ explains the phenomenon on the basis of the shape of the true stress-strain curve by postulating that if the plastic region is flat, i.e. low

strain hardening coefficient, then an abrupt change in ductility over some narrow pressure region will occur. However, all of the materials exhibiting abrupt pressure induced ductile-brittle transitions, i.e. bismuth⁽⁴⁾, zinc⁽⁴⁾, tungsten, and 1045 steel, also undergo temperature induced brittle-ductile transitions at some temperature above ambient. It is proposed, therefore, that the observed pressure induced transitions are some manifestation of the well-known transitions that occur at atmospheric pressure as a function of temperature; the superimposed pressure serving to lower the transition temperature to ambient. Galli and Gibbs⁽⁷⁾ observed that the brittle-ductile transition temperature of molybdenum was actually lowered 50°C by a superimposed hydrostatic pressure of 1.4 kilobars, which lends some credence to the latter argument. However, whether pressure has induced a new phenomenon or whether it is directly related to the well-known brittle-ductile transitions observed in many body-centered cubic and hexagonal close-packed metals at atmospheric pressure must await the completion of current pressure-temperature studies on ductility.

The drastic enhancement of the ductility of otherwise brittle materials by a superimposed hydrostatic pressure is also of great practical significance. It has formed the basis for a hydrostatic fluid extrusion process⁽⁴⁾ by which it has been possible to extrude many materials successfully at ambient temperature that normally require high temperatures using conventional ram type techniques. In addition, some investigators are also examining the possible enhancement of metal flow characteristics in closed-die forging and drawing by using a superimposed hydrostatic pressure.

CONCLUSIONS

1. Under certain conditions, sufficiently high hydrostatic pressures can induce localized microscopic deformation in polycrystalline metals. Based on the examination of the effect of hydrostatic pressures to 26 kilobars on several cubic metals, the occurrence of such deformation, which in the early stages is localized along grain boundaries and increases in severity with increasing pressure, results from microscopic shear stresses arising from anisotropy in the linear compressibility. A direct relationship exists between the severity of the deformation and the degree of anisotropy. Based on this relationship, metals exhibiting even slight deviation from total isotropy in linear compressibility, which is the case for most non-cubic metals, will undergo such deformation if the superimposed pressure is high enough.

2. Pressure cycling to 20 kilobars slightly increases the flow stress of polycrystalline zinc. The tensile strength and ductility are not reproducibly affected. The fact that the flow stress is not decreased indicates that the average microscopic strain associated with the anisotropy induced shear stresses is not such as to result in the occurrence of the Bauschinger Effect.

DAVIDSON, UT, and LEE

3. As has been observed in other materials, the ductility of magnesium, tungsten and 1045 steel with an untempered martensitic structure is drastically enhanced when fracture under a superimposed hydrostatic pressure. In the case of magnesium, the reduction in area increases continuously from the atmospheric pressure value of 8 percent to 28 percent at 19 kilobars. In contrast, the reduction in area of tungsten and 1045 steel does not increase continuously, but is unaffected up to a critical pressure level after which it abruptly increases to 50 and 60 percent respectively. These materials are also known to undergo abrupt brittle-ductile transitions as a function of temperature, thus indicating that the pressure induced transition may be directly related to the atmospheric pressure phenomenon.

REFERENCES

- (1) Johamin, P., and Vu, H., C. R. Acad. Science (Paris) 242, 2579 (1956)
- (2) Davidson, T. E., and Homan, C. G., "Some Observations on the Effects of Hydrostatic Pressures to 20,000 Atmospheres on the Structure of Polycrystalline Bismuth", Trans. ASM, 227, 1, Feb. 1964, pp. 167-176.
- (3) Bridgman, P. W., "Studies in Large Plastic Flow and Fracture", McGraw-Hill Book Co., N. Y., 1952
- (4) Pugh, H. L. D., "The Mechanical Properties and Deformation Characteristics of Metals and Alloys Under Pressure", ASTM International Conference on Materials, Philadelphia, Feb. 3-6, 1964
- (5) Ferron, J.R., "The Effects of High Hydrostatic Pressure on Subsequent Physical Properties of Metals", U. S. Govt. Research Report, Vol. 37, p. 93 (A), May 5, 1962. AD-272 011
- (6) Radcliffe, S. V., "Effects of High Pressure and Temperature on the Mechanical Properties of Metals and Alloys", ASTM International Conference on Materials, Philadelphia, Feb. 3-6, 1964
- (7) Galli, J. R., and Gibbs, P., "The Effect of Hydrostatic Pressure on the Ductile-Brittle Transition in Molybdenum", Based on Ph.D. Thesis of J. R. Galli, U. of Utah, Mar. 1963.

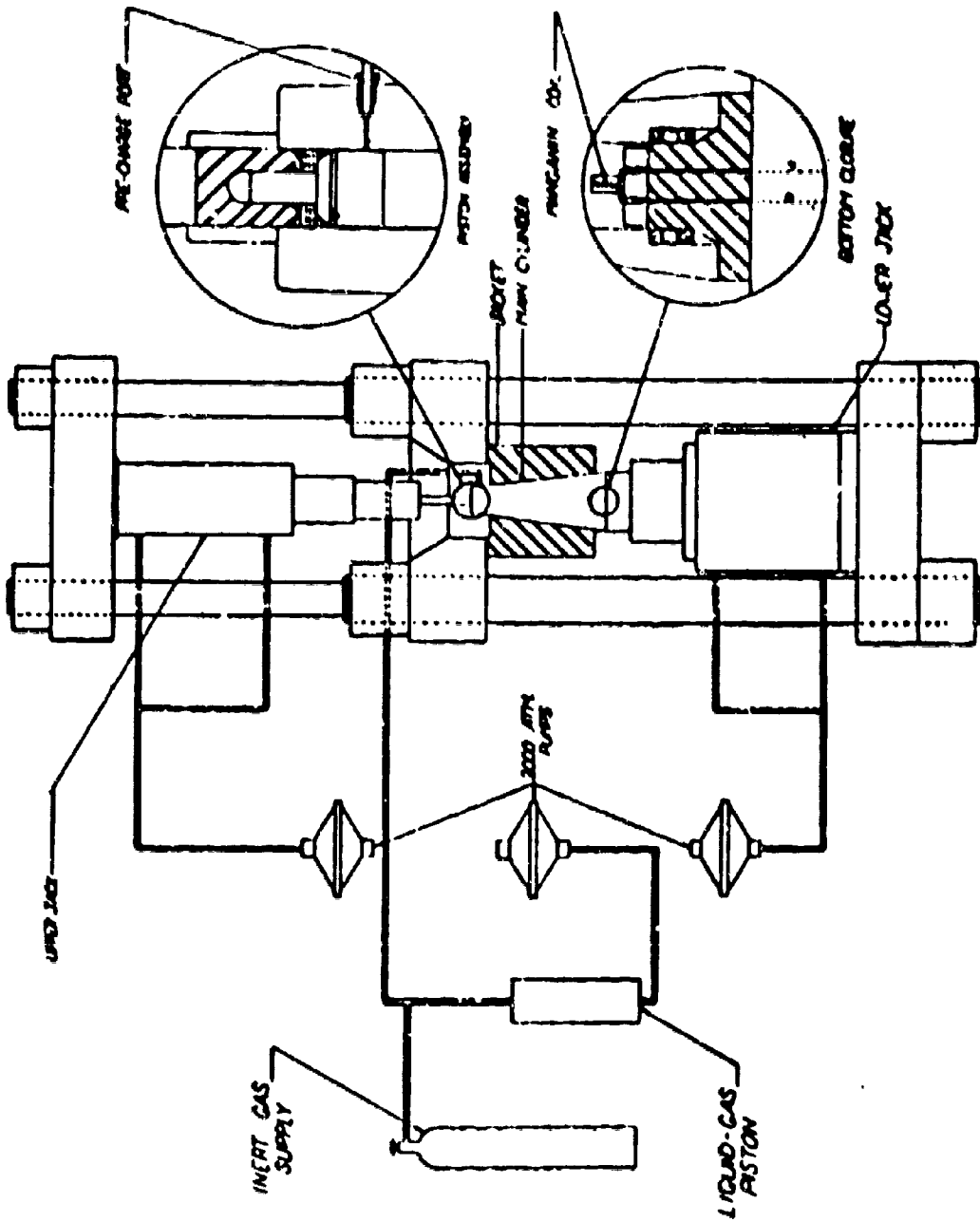


Figure 1 SCHEMATIC OF HYDROSTATIC PRESSURE SYSTEM



Figure 2A Disassembled Tensile Testing Fixture



Figure 2B Assembled Tensile Testing Fixture



A - Original Structure 1100



B - After 5000 BARS 1100



C - After 10,000 BARS 1100



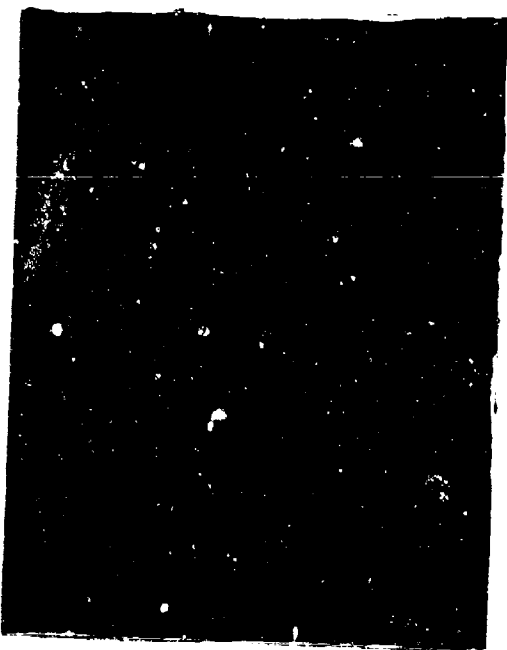
D - After 20,000 BARS 1100

FIGURE 3 STRUCTURAL CHANGES IN POLYCRYSTALLINE ZINC INDUCED BY HYDROSTATIC PRESSURE



A - Original Structure 2100

TIM



B - After 26,000 BARS 2100



A - Original Structure 2100

CADTON



B - After 20,000 BARS 2100

FIGURE 4. STRUCTURAL CHANGES IN TIM AND CADTON INDUCED BY HYDROSTATIC PRESSURE

$$\frac{A_0 - A}{A_0} \times 100$$

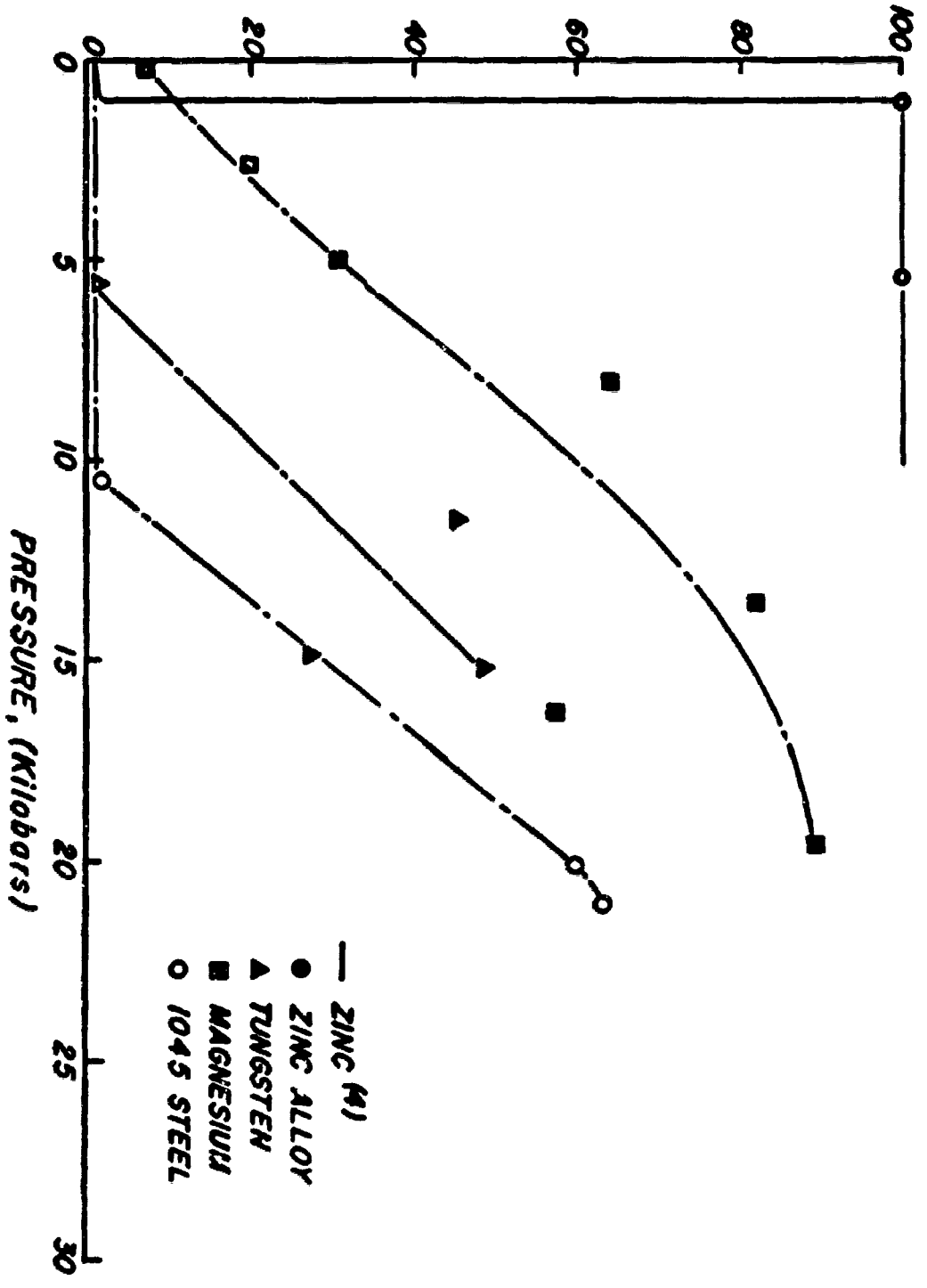


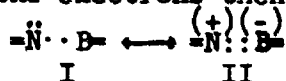
Figure 5 PRESSURE vs PERCENT REDUCTION IN AREA

HINDERED ROTATION IN AMINOBORANES

JOHN W. DAWSON and KURT NIEDENZU
 U. S. ARMY RESEARCH OFFICE-DURHAM at DUKE UNIVERSITY
 DURHAM, NORTH CAROLINA

Aminoboranes are conventionally illustrated by the basic structure $\text{N}=\text{B}$; hydrogen and/or organic groups are attached to the nitrogen and to the boron. To understand the chemical and physical behaviour of these compounds, the fundamental character of the B-N bond must be considered.

Primarily, there exists only one normal covalent bond (I), providing an electron deficiency on the boron. Naturally, this presents a very reactive state, which can explain most of the chemical reactions of aminoboranes. Participation of the free electron-pair of the nitrogen in the B-N linkage, however, promotes double bond character (II) as the electrons then reside in π -orbitals.

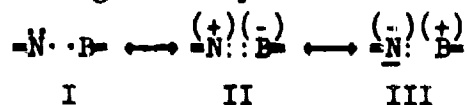


A special kind of double bond, where one partner is responsible for the contribution of three electrons to the bond, is highly important in such situations. A convincing experimental proof of the double bond in most cases depends upon recognition of the parts played by bond distances, bond forces and the mechanism of chemical reactions, in consonance with certain empirical rules.

Planar arrangement of its three bonds, according a sp^2 -hybridization, enables the free electron-pair of the nitrogen to participate as π -electrons in the double bond of aminoboranes. Such a planar arrangement of the nitrogen bonds in boron compounds has been established for aminoboranes¹ as well as for borazine² and boron nitride³. Furthermore, spectroscopic investigations of aminoboranes⁴ provide evidence for the double bond character of the B-N linkage.

In reference to the bond moment,⁵ however, three electronic structures are necessary to describe the B-N bond of aminoboranes

according to Pauling's theory.



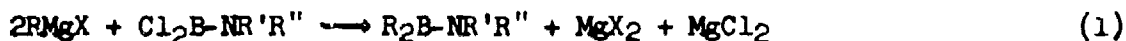
Which of these structures predominates should depend on the nature of the substituents bonded to the basic structure.

Since a boron-nitrogen grouping is isosteric and isoelectronic with a carbon-carbon grouping,⁶ aminoboranes have been compared with analogous ethylene compounds. More than a quarter of a century ago Langmuir⁷ suggested the correspondence of melting and boiling points as a sensitive criterion for the similarity of molecules. Indeed, physicochemical properties of alkylated aminoboranes are in close agreement with those of the corresponding ethylene derivatives. This relationship suggests a consistent tendency of the π -electrons of the nitrogen to participate in boron-nitrogen bonding.

Assuming it is reasonable to compare the structures of aminoboranes and ethylenes and assuming a B-N double bond is existent, cis-trans isomerism should be possible in this particular system, since, under these circumstances, rotation around the B-N axis would be hindered. The minimum condition for such an occurrence resides in the presence of two different substituents attached to the basic structure in the proper spatial arrangement, i.e., two groups bonded to the boron which are different between themselves but which are identical to the two substituents on the nitrogen, e.g., $\text{R}^1\text{R}^2=\text{N}-\text{B}=\text{R}^1\text{R}^2$.

The preparation of organic substituted aminoboranes through the reaction of Grignard compounds with (amino)-dichloroboranes has been reported.⁸ The disadvantage of this method resides in the consistently simultaneous replacement of both chlorine atoms by identical organic groups. A stepwise replacement, affording aminoboranes with either an organic group and a chlorine atom or two different organic groups attached to the boron, was not realized.

The general equation may be expressed as



The reaction, however, seems to be more complex as is indicated by the formation of small amounts of unidentified volatile boron compounds which may ignite spontaneously or explode. Since amino-dichloroboranes are easy to prepare, the reaction permits great variation of synthesis.

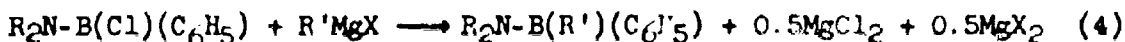
Aminoboranes with two different substituents attached to the boron can be prepared by an alternative procedure: Dichloro-organoboranes react with secondary amines to yield (amino)-monochloroboranes⁹



and it was also found possible to disproportionate equimolar amounts of a bisaminoborane with a dichloroborane¹⁰.



Treating the (amino)-monochloroborane with Grignard or organometallic reagents then produces the desired unsymmetrically substituted aminoborane. If an unsymmetrical secondary amine is utilized in this reaction sequence a completely unsymmetrical tetraorgano substituted aminoborane can be obtained.



The availability of a great variety of substituted aminoboranes¹¹ permits an evaluation of the influence of substituents on the B-N linkage:

In the infrared spectra of aminoboranes it is relatively simple to assign the B-N stretch.¹² The position of this absorption can generally be utilized for a rough evaluation of the bond character.

To demonstrate the sensitivity of the infrared spectra as a criterion of the bond character, some calculations of the boron-nitrogen stretching frequency and wave length were made using Gordy's rule.¹³ The rule is stated in mathematical form by equation 5

$$k = 1.67n[X_B X_N / d^2]^{0.75} + 0.30 \quad (5)$$

These values of the variables were used to calculate the force constants: (a) bond orders, $n = 1, 2$ or 3 ; (b) electronegativity of boron, $X_B = 2.0$; electronegativity of nitrogen, $X_N = 3.0$; (c) bond distance of B-N, $d = 1.44 \text{ \AA}$. and $d = 1.36 \text{ \AA}$.

The values for the electronegativity were obtained by consulting Pauling's table of electronegativities¹⁴. The bond distance "d" has not been accurately measured for a boron-nitrogen double bond; therefore approximate experimental values or computed estimates were used for this parameter. A value of 1.44 \AA . was assumed since this value was obtained experimentally for the boron-nitrogen distance in borazine.² If boron-nitrogen systems are analogous to carbon-carbon systems, this bond length is probably slightly exaggerated, just as the carbon-carbon distance in benzene is slightly longer than the ethylene carbon-carbon double bond. If Pauling's double bond atomic radii¹⁵ are used to estimate the boron-nitrogen double bond, a value of 1.36 \AA . is obtained. Both values were used in the computation.

The computed values of the force constants were substituted into the familiar equation 6

$$\nu = \frac{1}{\lambda} = 1307 \sqrt{\frac{k}{\mu}} \quad (6)$$

in order to obtain the frequencies and wavelengths. In this equation ν is the frequency, λ the wave length and μ the reduced mass of boron and nitrogen atoms. Calculations were made for boron isotope 11 and nitrogen 14, since these isotopes are the most common. The wave length due to the stretching of the boron 10 and nitrogen 14 bond was calculated to be approximately 0.974 of the computed boron 11-nitrogen 14 wave length. The presence of nitrogen isotopes other than nitrogen 14 can be neglected. The results of these approximate calculations are found in Table I.

TABLE I

Computed Wave Lengths of B-N Stretch Frequencies Using Gordy's Rule

Bond order	Interatomic distance, Å	λ, μ
1	1.44	9.6
2	1.44	6.8
2	1.36	6.56
2	1.20	6.0
3	1.27	
3	1.44	11.6

The wave lengths found for both bond distances $d = 1.44 \text{ \AA}$. and $d = 1.36 \text{ \AA}$. when $n = 2$ are close to a bond wave length which has been observed experimentally in the 6.8μ region. This bond was assigned to the boron-nitrogen stretching frequency.

It is quite obvious that the weakness of this treatment lies in the lack of a fixed experimentally determined B-N bond distance for the compounds studied. A more subtle weakness could lie in the use of an empirical rule to draw theoretical conclusions about the nature of the B-N bond. Furthermore, Gordy¹³ has defined the bond order used in his rule as "the effective number of covalent (electron pair) or electrovalent (ionic) bonds acting between two atoms." Hence, the agreement between experiment and theoretical results tells us only that two electron pairs have interacted in some way to form a bond. It does not tell us whether the bond is ionic or covalent.

In the absence of an accurate boron-nitrogen bond distance, it is necessary to examine the sensitivity of bond order to bond distance if any valid conclusions are to be drawn. To do this one must assume that the experimentally observed bands near 6.8μ reflect the B-N stretching frequency. If one works backward from this band at 6.8μ , which has been shown to be in close agreement with a calculated value, to compute the force constant, a value of 7.7 dynes/cm . is found. Further substitution of this value into equation 5 provides an expression showing the variation of bond order "n" with bond distance "d". This expression is found to be

$$n = 1.16d^{1.5}$$

(7)

From this equation, it is evident that the character of a B-N bond absorbing at 6.8μ does not vary drastically with the bond distance. In order to obtain single or triple bond orders, i.e., $n < 1.5$ or $n > 2.5$, it is necessary to have single bond lengths less than 1.18 \AA . or triple bonds greater than 1.68 \AA ., respectively. Either case seems highly 'mprobable, so one may assume that, if the bond at 6.8μ indicates a stretching of boron-nitrogen, there must be a double bond involved.

The B-N stretching frequencies of some aminoboranes are listed in Table II. For the compounds studied in this work, the bands were assigned by comparison of the original compounds with their respective hydrolysis products.

TABLE II

B-N Stretching Bands of Aminoboranes

Substituents		Frequency (cm^{-1})	Substituents		Frequency (cm^{-1})
N-	B-		N-	B-	
2Me	2H	1194	Ph, Me	2Me	1388
Ph, H	2Me	1332	2Me	2p-Me-C ₆ H ₄	1410
2p-Me-C ₆ H ₄	2Ph	1361	Me, Bu	Ph, Et	1412
2Ph	2Ph	1372	2Me	Ph, Me	1417
Ph, Me	Ph, Me	1381	2Et	2 Et	1490
Ph, Et	Ph, Et	1383	Me, H	2 Me	1525
2Ph	2p-Me-C ₆ H ₄	1385	2Me	2 Me	1530

As outlined above, the B-N bond in aminoboranes may have considerable double-bond character due to resonance of the nitrogen's free electron pair and thus cause restricted rotation. However, the degree of double bond character will depend mainly upon the nature of the substituents bonded to the basic structure $\text{N}=\text{B}$. The observations illustrated in Table II indicate that the introduction of an N-phenyl group brings about a considerable lowering of the double-bond character of the B-N linkage by resonance interaction of the unshared electron pair of the nitrogen atom with the adjacent phenyl group. Since such interaction results in increased electron deficiency at the boron atom, the latter is more subject to nucleophilic attack (e.g., hydrolysis).

The effect of B-aryl groups seems less pronounced, although here again a tendency toward lower frequencies is apparent (cf. Table II). Resonance interaction is again possible (since the phenyl group is "amphoteric" in this respect) leading to increased double-bond character for the B-phenyl bond and a corresponding decrease in the double-bond character for the B-N link. It should be noted, however, that the electron density around the boron atom is

not diminished and, consequently, increased reactivity toward nucleophilic attack is not to be expected.

The wide range of B-N stretching frequencies shown in Table II is a good indication of the considerable variation in the bond orders of the B-N bonds in these compounds. While, in general, the bond order (or double-bond character) increases as the B-N stretching frequency shifts to higher frequencies, an exact determination of the bond orders will have to await the availability of further infrared data (mainly on aminoboranes) and the determination of the B-N bond distances for some of the aminoboranes. However, on the basis of a rough, semi-quantitative calculation, following Gordy's equations³ (neighboring-group mass effects and coupling between the B-N and the adjacent B-C and N-C stretching modes being neglected) and the best available estimates for the B-N bond distances being used, a range of 1.30-1.60 is obtained for the bond orders (corresponding to 30-60% double-bond character) for the aminoboranes listed in Table II, which appears to be consistent with their chemical behavior.

On preparation of unsymmetrically substituted aminoboranes only rarely could sharp melting or boiling points be recorded. Moreover, in selected cases it was possible to observe crystalline and liquid species of one and the same compound. On attempts to separate the two phases no uniform material could be obtained. Crystals immediately partially liquified and vice versa.

The following data are at hand: (a) Among unsymmetrical tetraorgano substituted aminoboranes there exist two, and only two, different "modifications" of one single compound, showing differences in their physical behavior. (b) Molecular weight determinations and analysis confirm that the two different "modifications" have identical empirical formulas. (c) Rupture of the B-N bond in any case always produces the secondary amine, used as the starting material. (d) "Modification"-transformation, i.e., equilibration, occurs very rapidly without addition of further energy at room temperature.

On the basis of the above cited data, certain conclusions can be drawn illustrating the existence of two isomers having identical empirical formulas. There are, however, three possible methods for obtaining two isomers of unsymmetrical tetraorgano substituted aminoboranes: (i) intermolecular disproportionation, (ii) intramolecular disproportionation, (iii) rotation about the B-N bond (i.e., cis-trans isomerization).

Since the molecular weights of each fraction are identical, there cannot be intermolecular disproportionation except in those rare cases where the disproportionating groups have the same weight. This does not apply to the examples at hand.

Intramolecular disproportionation cannot occur except in two cases: (1) Exchange of one or more groups of the boron substituents with those on the nitrogen. This is not possible since only the original secondary amine was recovered. (2) Exchange in position of the B-attached group. This, however, would have the same effect as rotation about the B-N bond. Therefore, the spatial arrangement of the groups within the isomers would not be different from that obtained by rotation; the mechanism of isomer formation would be the only variation.

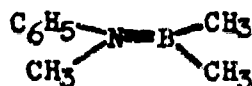
Since isomer equilibration occurs at room temperature without the application of additional energy and, since rotation about the B-N bond (excluding the existence of steric hindrance) requires little energy, the latter is the more probable mechanism. Furthermore, if extensive bond breaking did occur, it is almost certain that a degree of intermolecular disproportionation would occur with the formation of many isomers.

Therefore, since both types of disproportionation have been eliminated, only rotation about the B-N bond is left as an explanation of the existence of these isomers. Further, infrared investigations when analyzed in terms of the relative degree of electronegativity of the substituents tend to confirm the hypothesis of electron delocalization being the primary factor in explaining the appearance of double bond character in the B-N system and the resultant cis-trans isomerization.

In tetrasubstituted aminoboranes such as (ethylphenylamino)-ethylphenylborane, one isomer will have trans configuration, the other being the cis form. The very rapid rearrangement of either form into a mixture of both could be attributed to the presence of the electron structure III, as described previously. Since, in that structure, no double bond but rather a quasi ionic character is involved, its participation in the electronic equilibrium would be likely to be responsible for the nearly spontaneous isomerization. In the same way the diffuse boiling points of several of the aminoboranes might be explained as being due to the fact that the lower boiling cis and higher boiling trans form (or vice versa) are in equilibrium. The mixture shows a range of boiling points, but the first fractions may be overly enriched with one form. However, a very fast isomerization might occur, due to equilibration of the three electron structures, thus demonstrating the bond character of an aminoborane.

An independent proof for this concept of hindered rotation in aminoboranes was recently furnished by NMR spectroscopy of some unsymmetrically substituted derivatives. It could be demonstrated, for instance, that at room temperature the two B-attached methyl groups in (phenylmethylamino)-dimethylborane, $(C_6H_5)(CH_2)N-B(CH_3)_2$ are not equivalent, giving rise to two sharp separate NMR signals for the methyl protons¹⁶. Only at higher temperatures the two signals

broaden and finally merge and at high temperatures (near 125°) one single sharp B-methyl signal is observed. This can be interpreted only in terms of hindered rotation about the B-N linkage: At lower temperatures rotation is so slow that two unequivalent methyl groups are existent. At higher temperature rotation is enhanced and finally the two different environments of the methyl groups can no longer be recorded since rotation has become too fast.



B-N double-bonding causes different magnetic environments for the two B-attached methyl groups. Consequently at lower temperatures the following NMR spectrum for the B-methyl protons is recorded:



At higher temperatures rotation about the B-N axis becomes less restricted. The B-methyl protons become more equivalent with the increase in rotation, as evidenced by the following spectrum, in which the proton resonances have merged to a certain degree.

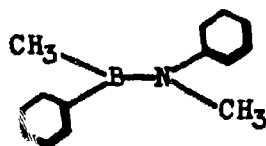


At very high temperatures rotation about the B-N linkage is essentially unrestricted. Hence the B-methyl protons are equivalent and the NMR spectrum shows only one sharp single peak.

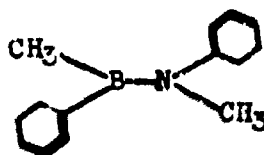


However, the ready acceptance of NMR data as sole proof for restricted rotation appears to be unreasonable. For instance, two NMR signals were recorded for both methyl groups of (methylphenyl-amino)-methylphenylborane at low temperatures. But the observation of only one species in the Raman and infrared spectra of (methylphenyl-amino)-methylphenylborane suggests the presence of only one rotational species with respect to the B-N linkage. Examination of molecular models illustrates that the trans configuration is favored, although even this form is probably not planar, and that both phenyl groups should be skewed out of plane. (This is consistent with the results reported for dimethylstilbene¹⁷.) If this is the situation it is then possible to explain the observation of two NMR signals for

both the boron and the nitrogen-attached methyl groups by the indicated structures (a) and (b). Preference of one species (namely, b) should be expected (and is in agreement with the experimental results) if steric effects produce not only nonplanar arrangement of the phenyl groups but also cause a minor tilting about the B-N linkage. This tilting then can lead to different magnetic environments for the corresponding methyl groups in the two forms.



(a)



(b)

In two independent molecular orbital calculations added proof for the described experimental data was furnished. J. J. Kaufman¹⁸ confirmed the influence of substituents on the B-N bond order by a Huckel calculation of charge distribution at the boron and the nitrogen sites. Agreement of the experimentally derived data as described above¹² with the calculations¹⁸ was essentially complete. The mathematical data illustrate that shifts of the B-N frequency, depending upon the nature of the organic substituents, indeed conform to a change in the bond character of the B-N linkage. Hence, IR spectral data can be used for the evaluation of the nature of the boron-nitrogen bond.

R. Hoffman¹⁹ calculated the total charge distribution in aminoboranes and showed that in these compounds the distribution of charges can quite possibly result in a neutral molecule which nevertheless shows B-N double bond character. The barrier of rotation in such a neutral molecule should be in the order of 10 Kcal/mole, which value is in good agreement with the values obtained experimentally by n.m.r. spectroscopy. Of course, the value is considerably less than that of corresponding olefines, but nevertheless it is substantial enough to justify acceptance of a B-N double-bond in aminoboranes.

Summarizing, the concept of restricted rotation in aminoboranes due to B-N double bonding seems well founded, as illustrated by chemical, physicochemical and mathematical evidence.

References:

1. J. Goubeau and H. J. Becher, Z. anorg. allg. Chem. 268, 133(1952)
2. S. H. Bauer, J. Amer. Chem. Soc. 60, 524 (1938)
3. A. Brager and H. Shdanov, C.R. (Doklady) Acad. Sci. USSR. 29, 629 (1940)
4. H. J. Becher, Z. anorg. allg. Chem. 289, 262 (1957)
5. H.J. Becher, *ibid.* 270, 273 (1952)
6. J. Goubeau, Naturwissenschaften 35, 246 (1948)
7. I. Langmuir, J. Amer. Chem. Soc. 42, 274 (1920)
8. K. Niedenzu and J.W. Dawson, *ibid.* 81, 5553 (1959)
9. K. Niedenzu and J. W. Dawson, *ibid.* 82, 4223 (1960)
10. H. Jenne and K. Niedenzu, Inorg. Chem. 3, 000 (1964)
11. K. Niedenzu, H. Beyer, J.W. Dawson, and H. Jenne, Chem. Ber 96, 2653 (1963)
12. G. M. Wyman, K. Niedenzu and J.W. Dawson, J. Chem. Soc. 1962, 4068
13. W. Gordy, J. Chem. Phys. 14, 305 (1946)
14. L. Pauling, "The Nature of the Chemical Bond", Cornell University Press, Ithaca, N.Y., 1946, Table 11-3, p. 64
15. *ibid.*, Table 21-2, p. 164
16. G. E. Ryschkewitsch, W.S. Brey and A. Saji, J.Amer.Chem.Soc.83, 1010 (1961)
17. O. Simamura and H. Suzuki, Bull. Chem. Soc. Japan 27, 231 (1954)
18. J. J. Kaufman, Advances in Chemistry 42, 000 (1963)
19. R. Hoffman, *ibid.*, 42, 000 (1963)

HIGH POWER MICROWAVE DEVICE APPLICATIONS FOR
NEW NARROW RESONANCE LINEWIDTH PLANAR HEXAGONAL FERRITES

SAMUEL DIXON, JR., ROBERT O. SAVAGE, JR., AND ARTHUR TAUBER
U. S. ARMY ELECTRONICS RESEARCH AND DEVELOPMENT LABORATORIES
FORT MONMOUTH, NEW JERSEY

INTRODUCTION: To meet future device requirements in radar, countermeasures, and satellite systems between 9 and 140 gigacycles (Mc), a wide variety of nonlinear ferromagnetic components such as power limiters, frequency multipliers, tuneable filters, and pulsed ferromagnetic generators, presently restricted because of an inherent frequency limitation or the prohibitively large applied d.c. field requirement, must be designed and constructed. These magnetic devices require a material with unique directional properties (high magnetic anisotropy), permitting device operation with small applied d.c. fields. A family of ferromagnetic hexagonal ferrites, with an easy plane of magnetization, shows considerable promise in meeting these requirements. The large magnetic anisotropy (9900 oersteds in $\text{Ba}_2\text{Zn}_2\text{Fe}_{12}\text{O}_{22}$, commonly called ZnY) binds the spin magnetization vector to a preferential plane and thus endows these materials with unique microwave properties. When the r.f. magnetic field is applied perpendicular to the d.c. field, nonlinear effects can occur in the usual fashion as with isotropic ferrites.⁽¹⁾ However, when using planar hexagonal ferrites the large planar anisotropy (built-in magnetic field) decreases the critical power level at which spin-wave instability sets in and extends the frequency range in which certain non-linear ferromagnetic devices are feasible. As a further advantage, the applied d.c. magnetic field required for resonance is considerably reduced.

The high-power nonlinear effects to be expected at ferromagnetic resonance are very critically dependent on the availability of spin waves with suitable resonant frequencies. A spin wave may be pictured as a ripple distribution or a time lag in precessional motion of adjacent spins. The location of the uniform mode (exciting frequency) with respect to the spin-wave spectrum will determine which spin waves are generated. The excitation of spin waves leads to a reduction in the susceptibility of the ferrite at resonance and, under certain conditions, to an increase in absorption at a d.c.

field below resonance (subsidiary resonance). At low power level, when the ferrite sample is magnetized perpendicular to the signal field direction, the magnetization vector fans out, dissipates its energy, and relaxes to its equilibrium position. A long relaxation time indicates that a large amount of energy is stored in the spin system of the ferrite sample, analogous to a high-Q resonant circuit. A narrow-linewidth (ΔH) ferrite will have a high peak absorption per unit volume, indicating that a large amount of r.f. energy is being stored by the ferrite. As the input power is increased, the magnetization vector fans out to what is called the critical angle. At this point the material is said to be saturated. Any further increase in signal power is coupled into the creation of spin waves. Two possible mechanisms should be distinguished. For one of them, the spin waves generated will have half the frequency of the applied signal; for the other, they have the same frequency as the applied signal. It is convenient to refer to these two mechanisms as nonlinearities of the first and second order, respectively. The incident microwave power required for spin-wave instability is called the critical field (h_c). The threshold field for the first order effect is inherently lower than that of the second order effect, provided both are allowed. This is because the first-order effect depends only on the amplitude of the uniform precession, while the second-order effect depends upon the square of the amplitude of the uniform precession.(2)

In planar hexagonal ferrites the critical field for the second-order effect is given by:(3)

$$(I) \quad h_c (4\pi M_s)^{\frac{1}{2}} / \Delta H (\Delta H_K)^{\frac{1}{2}} = (2/\delta)^{\frac{1}{2}}$$

where $4\pi M_s$ is the saturation magnetization, ΔH is the measured linewidth, ΔH_K is the intrinsic linewidth characteristic of the unstable spin wave mode, and δ equals $H_a / 4\pi M_s$ (H_a is the anisotropy field). As equation I indicates, the magnitude of the anisotropy field will markedly affect the critical field at which spin-wave instability sets in.

In the design of some nonlinear ferromagnetic devices, the insertion loss and efficiency of operation depends upon the narrowness of the resonance linewidth. The first single crystals of ZnY ($Ba_2Zn_2Fe_{12}O_{22}$) had impractical linewidths of 30 to 50 oersteds. Refinement of the crystal growth procedures yielded crystals of 13 to 16 oersteds linewidth.(4) High-power experimental measurements at that time indicated potential device feasibility,(5) although the resonance linewidth was still sufficiently broad to make device applications prohibitive. When a new solvent, BaO-B₂O₃, and an annealing procedure were devised, crystals were obtained with a linewidth of 8 oersteds.(6) Before evaluation of these crystals was complete, it was found that substitution of manganese for a portion of the zinc gave crystals with a resonance linewidth of 3.8 oersteds

at 9.0 Mc and room temperature. This represents the lowest value reported for the planar hexagonal ferrites and permits the design of nonlinear ferromagnetic devices utilizing these materials.

EXPERIMENTAL DETAILS: Single-crystal hexagonal ferrites with the chemical composition $Ba_{1-x}Zn_xMn_{1-x}Fe_{1-x}O_{10}$ were grown from melts using a BaO/B_2O_3 flux. The constituent oxides or carbonates were mixed together and packed into platinum crucibles. These were heated in electrical resistance furnaces to 1300°C , held at temperature until solution was complete, and cooled between 1 and 2°C/hr . to 1000°C . The furnace was then cooled more rapidly to 800°C , at which temperature the crucibles were removed and allowed to cool to room temperature in air. Crystals were freed from the crucible by leaching with dilute nitric acid.

Crystals obtained by this method were typically black, hexagonal in shape, with highly reflecting natural faces. They were smaller than those normally attainable from $MgFeO_3$ flux, never exceeding $3/8$ in. long or $1/8$ in. thick. The larger crystals were typically $1/8$ in. long and $1/8$ in. thick. Spectrochemical analysis of several crystals from each batch indicated that the amount of Mn present in the crystals was close to the starting composition. The crystals were examined by x-ray diffraction methods to establish that they were single and had the Y structure.

Irregularly shaped samples were ground into spheres using the method described by Carter et al.⁽⁷⁾ The samples were then polished with successively finer grit diamond abrasives. The final samples used in making high-power measurements had a 0.1 micron finish. Previous investigations⁽⁸⁾ have shown the critical effect of surface polishing on the resonance linewidth of ZnY spheres. High-power magnetic susceptibility measurements were made at 17.3 Mc, using the cavity perturbation technique. The final diameter of the samples ranged from 0.018" to 0.030". A tuneable magnetron was used as the r.f. power source. A transmission cavity⁽⁹⁾ was used to measure the linewidth and evaluate the high-power properties. The length, or mode number, of the cavity was adjusted to give an appropriate filling factor. Small tuning screws placed along the length of the cavity allowed retuning of the cavity at each one-half susceptibility point, so that the frequency of the power source was not changed during the course of a measurement. A block diagram of the equipment used in making the measurements by the cavity technique is shown in Fig. 1. A waveguide switch with high isolation between ports permitted either the reflex klystron or the magnetron output to be incident upon the cavity. A waveguide power divider allowed the incident power level to be varied over a sufficiently wide range. Ferrite isolators were used to isolate the cavity from the rest of the microwave circuitry and to insure matched conditions for stable magnetron operation. The transmission type of cavity permits high-power susceptibility measurements to be made at a constant r.f.

magnetic field. As the energy absorbed in the ferrite sample increased, the monitored power at the output termination of the cavity decreased. Increasing the input power until the output power reached an appropriate reference level thus enabled a constant r.f. magnetic field to be maintained in the cavity over the complete range of measurements. The samples were placed in the center of the cavity at a position of maximum r.f. magnetic field. The orientation of the sample was such that the r.f. magnetic field and the applied magnetic field were perpendicular to each other and in the easy plane of magnetization.

The d.c. magnetization was measured with a vibrating sample magnetometer⁽¹⁰⁾ as a function of field at room temperature. The crystals were mounted so that a rotation of 90° would bring the easy or the hard direction parallel to the applied field. The saturation magnetization ($4\pi M_s$) determined along the easy direction is given in Table I.

Table I
Saturation Magnetization ($4\pi M_s$) as a Function of
Substituted Mn in $\text{Ba}_2\text{Zn}_{2-x}\text{Mn}_x\text{Fe}_{12}\text{O}_{22}$

x	$4\pi M_s$	x	$4\pi M_s$
0	2700	0.6	2300
0.1	2600	0.7	1800
0.2	2570	0.8	1700
0.3	2400	0.9	1700
0.4	2360	1.0	1650
0.5	2330	1.1	1600

EXPERIMENTAL RESULTS: The saturation magnetization is seen to slowly decrease as substituted Mn increases. This is to be expected if Mn replaces Zn. Near the composition $x = 0.7$ the magnetization drops sharply to below the value for pure MnY. This may be related to the size of the Mn^{2+} ion, the largest divalent substituent to enter the structure. Distortion of magnetic interaction due to changing interatomic distance gives rise to a reduced magnetization. Variation in cation distribution as a function of increasing Mn substitution would also decrease the magnetization. These possibilities will be distinguished by careful crystallographic studies. The magnetization continues to decrease with increasing Mn substitution until $x = 1.1$. Beyond this point planar crystals are no longer stable and axial crystals are obtained thru $x = 2.0$.

Ferromagnetic resonance linewidth measurements were made on compositions from $x = 0.1$ to $x = 1.1$. Figure 2 shows the linewidth as a function of substituted manganese. The indicated linewidths refer to the measured properties of the unpolished samples. As Mn substitution increases, the rough linewidth decreases to a minimum of 17.5 oersteds between $x = 0.5$ and 0.7. Additional manganese increased the linewidth to 21 oersteds, after which it leveled off.

The specimens with a linewidth of 17.5 oersteds were selected for further investigation. These crystals were polished with successively finer grit abrasives, to a final diameter between 0.018" and 0.030". An exponential decrease in the linewidth was noted with decreasing grit size of the polishing compound. A cyclic effect in linewidth as a function of polishing time, as found in lithium ferrite single crystals,⁽¹¹⁾ was not observed. The minimum linewidth obtained with substituted ZnY crystals is 3.8 oersteds at 9.0 kMc. The observed linewidth increases to 4.9 oersteds when measured at 17.3 kMc.

A comparison of the normalized magnetic susceptibility (decline of the main resonance) of lithium ferrite, yttrium iron garnet (YIG), and Mn-ZnY as a function of incident power at 17.3 kMc is shown in Fig. 3. The low-power magnetic susceptibility is used as a reference for normalizing the susceptibility at high-power levels. As Fig. 3 indicates, there is a substantial decline in the main resonance as the incident power level is increased. These nonlinear effects are particularly significant in determining the feasibility and performance of nonlinear ferromagnetic devices.

At 17.3 kMc the second-order process is responsible for the observed nonlinear effects. The decline in the main resonance curve may be used to determine the critical field⁽¹²⁾ for the onset of spin-wave instabilities by extrapolating back to the low-power magnetic susceptibility. From the critical field value, the spin-wave linewidth (ΔH_K) may be calculated using equation I. The critical field (h_c) for manganese substituted ZnY is approximately 6×10^{-2} oersteds, which compares very favorably with that of YIG, whose linewidth is lower (2.35 oersteds). This critical field value leads to a calculated spin-wave linewidth of 0.74 oersteds. As indicated in Fig. 3, the measured linewidth of lithium ferrite and substituted ZnY are approximately equal, but the critical field for spin-wave instabilities in ZnY is an order of magnitude lower.

DEVICE APPLICATIONS: The results of this investigation show that single-crystal ZnY materials substituted with manganese are promising for use in several microwave devices. Such devices include:

Pulsed-Field Millimeter-Wave Generator. The ability of single-crystal ferrites to act as high-Q resonators and to store energy for periods of time of the order of their relaxation time has been used as the basis for a pulsed millimeter-wave generator. Operation of the pulsed generator has been described by Elliott, Schaug-Patterson, and Shaw,⁽¹³⁾ who obtained an output frequency of 32 kMc with an input frequency near 4 kMc using single-crystal YIG. More recently, Shaw⁽¹⁴⁾ has obtained an output frequency of 50 kMc with no input frequency, by pulsing the ferrite at an angle to a small saturating d.c. field. In this generator the energy is stored in the spin system which is forced to precess around the resultant of the pulsed magnetic field and the small applied d.c. field. [The

energy is subsequently generated as the spin system relaxes back to its equilibrium position.] The output frequency depends on the magnitude of the pulsed field. With the use of substituted ZnY, the large anisotropy (9900 oersteds) will substantially reduce the magnitude of the pulsed magnetic field required for the generation of millimeter-wave energy. Using YIG spheres the pulsed magnetic field required is determined by

$$(II) \quad \omega = \gamma H_0,$$

where ω is the frequency of the energy generated, γ is the gyromagnetic ratio (usually 2.8 Mc/oersted), and H_0 is the pulsed magnetic field. A pulsed magnetic field of approximately 18 kilogauss would be required to generate a frequency of 50 Mc. The size and weight of the line required to produce such a pulse is rather large and bulky. Using ZnY spheres the pulsed magnetic field required is given by

$$(III) \quad (\omega/\gamma)^2 = H_0(H_0 + H_a),$$

where H_a is the anisotropy field, H_0 is the pulsed magnetic field, γ is the gyromagnetic ratio, and ω is the frequency generated. The pulsed field required to generate a frequency of 50 Mc, using ZnY, with an anisotropy field of 9900 oersteds, would be approximately 14 kilogauss. This pulsed field may be reduced even greater by using a material that has a larger anisotropy field. In addition, since the required pulsed magnetic field decreases as the anisotropy field is increased, miniaturization of the pulse line becomes feasible. It is important in the operation of such a device that the magnetic field pulse have a rise time which is small compared to the relaxation time of the material. At present, ZnY is marginal for this application, and investigations are required to further reduce the linewidth.

Power Limiters. The basic physical principle of the ferrite limiter involves the nonlinear behavior of the r.f. magnetization at the critical power level (h_c), as shown in Fig. 3 (decline of the normalized magnetic susceptibility). At this point, a transfer of energy takes place from the uniformly precessing spin mode to other spin waves. This phenomenon, and others which are similar, provides for limiting the level of r.f. power transmitted in a microwave component containing a ferrite. The manifestations of these nonlinear responses give rise to gyromagnetic coupler limiters and the subsidiary resonance limiter. The gyromagnetic-coupling limiter utilizes the main resonance instability (decline of the main resonance) for its operation. The coincident limiter, one type of gyromagnetic coupler limiter, where the subsidiary resonance coincides with the main resonance, has the lowest critical power level. However, this type of limiter utilizes both the first- and second-order process for its operation. Using spherical isotropic ferrites, a rather low

DIXON, SAVAGE AND TACHER

upper frequency limitation is imposed on this limiter and is given by:

$$(IV) \quad \omega < \frac{2}{3} \gamma 4\pi M$$

The lower frequency range is determined by the field required to saturate the material and can be written as

$$(V) \quad \omega > \gamma 4\pi M/3 .$$

For pure YIG, the saturation magnetization is 1750 oersteds. This gives an operational frequency range of 1.6 to 3.2 Mc. Nonmagnetic ions, such as gallium or aluminum, can be substituted into YIG to lower the magnetization and lower the operational frequency range of the coincident limiter. No practical method of increasing the saturation magnetization of YIG has been reported. Single-crystal lithium ferrite, with a $4\pi M_s$ of 3900 oersteds and a resonance linewidth of 2 oersteds at 6.0 Mc, has been reported.⁽¹⁵⁾ This material has a coincident limiter frequency range of 3.64 to 7.82 Mc. The frequency range has been extended to 8.8 Mc by tightly coupling the uniform precession to orthogonal strip-line resonators.⁽¹⁶⁾ This approach to increasing the frequency range of coincident limiting is limited by the resulting excessive losses and the magnitude of the dipolar coupling.⁽¹⁷⁾ The frequency range for the coincidence of the first- and second-order nonlinear processes in planar hexagonal ferrites is given by:⁽³⁾

$$(VI) \quad \frac{\gamma 4\pi M}{3} < \omega_c < \gamma 4\pi M \sqrt{4N_{\perp}^2 \left[1 + \left(\frac{\delta}{9}\right) \left[\left(1 + \frac{2}{16} \delta^2/N_{\perp}^2\right)^{\frac{1}{2}} - 1\right]\right]}$$

where N_{\perp} is the transverse demagnetizing factor, and δ equals $H_0/4\pi M$. Using planar hexagonal ferrites that are presently available, coincident limiting is feasible up to a frequency of 15.6 Mc.

The second-order process limiter, another kind of gyro-magnetic coupler limiter, may be used in frequency ranges where coincident limiting is not possible. In this limiter, the slope transition above the threshold is not as abrupt, the limiting characteristic is not flat, and the thresholds are much higher than with coincident limiters. Accordingly, the frequency range of the coincident limiter can be significantly extended by utilizing planar hexagonal ferrites. At 17.3 Mc, as shown in Fig. 3, the second-order process is responsible for the decline in the main resonance. The critical field in manganese substituted ZnY is comparable to that of YIG, even though the linewidth is twice as large. The experimental data indicates the effect of the large anisotropy in reducing the threshold field, and also the feasibility of using Mn-ZnY in the second-order process limiter.

DIXON, SAVAGE AND TAUBER

In ferromagnetic power limiters, the subsidiary resonance limiter has the highest threshold power level. However, the limiting point is abrupt, giving a well-defined plateau in the output. This type limiter can be used as a pre-limiter in combination with a low-level limiter. Figure 4 shows the development of the subsidiary resonance with an increase in input power. At low power levels, no subsidiary resonance is present and the signal power propagates with very little loss. At threshold power, the development of the subsidiary resonance limits the level of the output power. The applied d.c. field is set at a fixed value approximately half that required for the main resonance. Figure 5 shows an experimental model of a subsidiary resonance limiter designed for operation at 16.6 Mc using substituted ZnY. This type limiter utilizes the first-order nonlinear process for its operation. Because the limiter incorporates a resonant cavity, it serves an additional filter function of rejecting undesired signals which might cause instability. The limiter is constructed in reduced-height brass material. The cavity is composed of two waveguide half-sections joined along the broad dimension as shown in Fig. 5. Each half-section is of unit construction; the iris half-plates and the waveguide flanges at each end are an integral part of the machined piece to eliminate lower Q's resulting from soldered iris plates. The cavity was excited in the TE₁₀₅ mode with an Mn-ZnY sphere placed near the side wall and biased along the narrow dimension. Using a 0.025" diameter sphere, experimental data have indicated that the filling factor (ratio of sample volume to cavity volume) is too small. The cavity shown in Fig. 5 is being redesigned to be excited in the TE₁₀₁ mode and large ZnY spheres are being processed. This should give the necessary larger filling factor so that the losses in the sample will represent a sufficiently large proportion of the total cavity losses above the threshold power level.

CONCLUSIONS: A ferromagnetic resonance linewidth of 3.8 oersteds at 9.0 Mc and room temperature represents a considerable improvement in material technology for hexagonal ferrites with planar anisotropy. The linewidth compares favorably with that found in single-crystal YIG and in lithium ferrite. This investigation has shown that the large planar anisotropy endows ZnY materials with several distinct advantages in enhancing the design of microwave devices. The anisotropy field increases the frequency range for the coincidence of the first- and second-order processes, which inherently gives a lower threshold power level. In addition, the anisotropy field tends to decrease the critical field (power level) at which spin-wave instabilities set in. From the measured threshold field, a spin-wave linewidth of 0.74 oersteds has been calculated for manganese-substituted ZnY.

DIXON, SAVAGE AND TAUBER

A variety of investigations on the feasibility of microwave devices using manganese-substituted ZnY are currently in the design stage. Among them are an orthogonal resonant strip-line gyromagnetic limiter at X-band frequencies, a parallel pumped shunt arm cavity limiter at K_u-band frequencies, and a TE₁₀₁ cavity subsidiary resonance limiter.

Further experimental work involving the growth of large crystals and the substitution of other elements, such as aluminum, should make a host of magnetic devices feasible in the millimeter-wave frequency range.

ACKNOWLEDGMENTS: The authors are grateful to Dr. J. A. Kohn and D. W. Eckart for their crystallographic support and to J. L. Carter and I. Reingold for their many helpful discussions.

REFERENCES:

- (1) Suhl, H., Proc. IRE 44, 1270 (1956).
- (2) Lax, B. and K. J. Button, Microwave Ferrites and Ferrimagnetics, McGraw-Hill, New York, 1962, p. 675.
- (3) Schlomann, E., R. I. Joseph and I. Bady, J. Appl. Phys. 34, 672 (1963).
- (4) Tauber, A., R. O. Savage, R. Gambino and C. G. Whinfrey, J. Appl. Phys. 33, 13818 (1962).
- (5) Dixon, S. Jr., J. Appl. Phys. 33, 13688 (1962).
- (6) Tauber, A., S. Dixon, Jr. and R. O. Savage, J. Appl. Phys. Suppl., in press, March 1964.
- (7) Carter, J. L., E. V. Edwards and I. Reingold, USAELRDL Tech. Report. 2049 (1959).
- (8) Dixon, S. Jr., J. Appl. Phys. 34, 3441 (1963).
- (9) Carter, J. L., S. Dixon, Jr., and I. Reingold, P.G.M.T.T. MTT-9, 195 (1961)
- (10) Foner, S., Rev. Sci. Instr. 30, 548 (1959).
- (11) Nielsen, J. W., D. A. Lepore, J. Zneimer and G. B. Townsend, J. Appl. Phys. 33, 13798 (1962).
- (12) Schlomann, E., J. Saunders and M. Servetz, Tech. Memo. T-167, Raytheon Company, Waltham, Mass. (1959).

DIXON, SAVAGE AND TAUBER

- (13) Elliott, B. J., T. Schaug-Pettersen and J. E. Shaw, P.G.M.T.T. MIT-9, 92 (1961).
- (14) Contract DA36-039 AMC-00041 (E), Rept. #1, 1 Oct.-31 Dec. 1962 USARLKD, Ft. Mon., N.J.
- (15) Comstock, R. L. and J. P. Remeika, J. Appl. Phys. Suppl., in press, March 1964.
- (16) Comstock, R. L. and M. Losvy, Proc. DEKE 51, 1769 (1963).
- (17) Comstock, R. L. and W. A. Dean, J. Appl. Phys. 34, 1275 (1963).

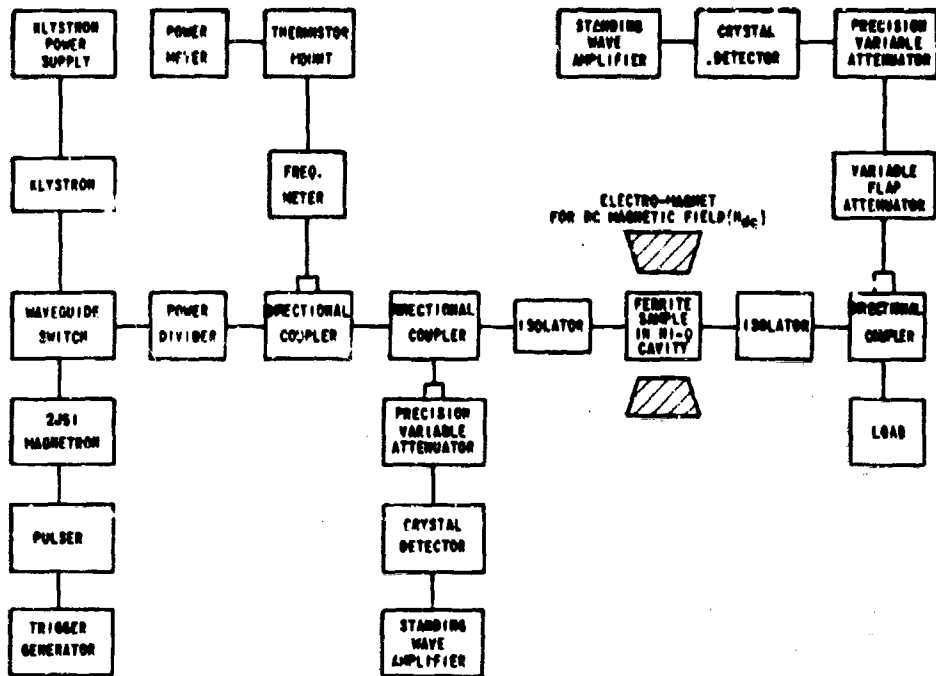


Figure I - Block diagram of test equipment used in making the high power susceptibility measurements.

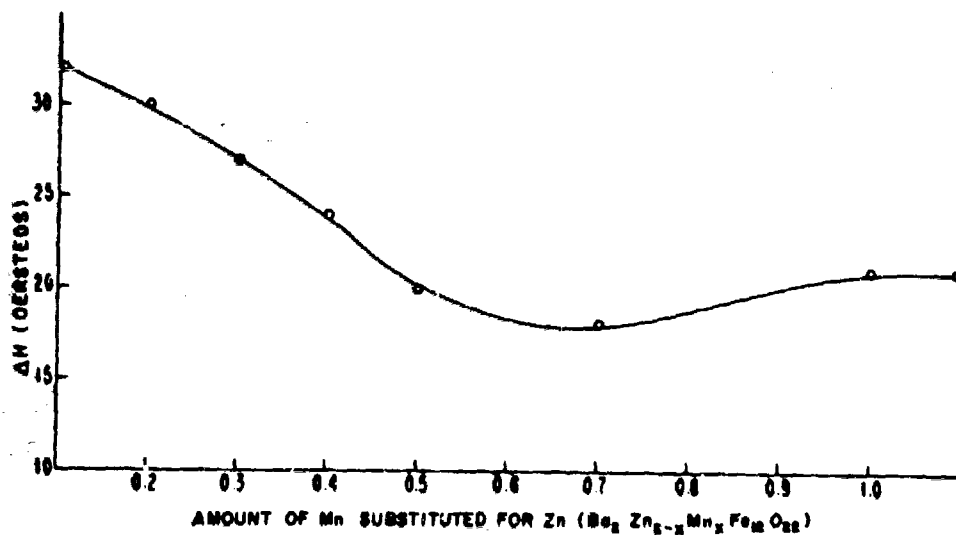


Figure II - Variation of the unpolished resonance linewidth (ΔH) as a function of the amount of manganese substituted for zinc. ($Ba_2Zn_{2-x}Mn_xFe_{12}O_{22}$)

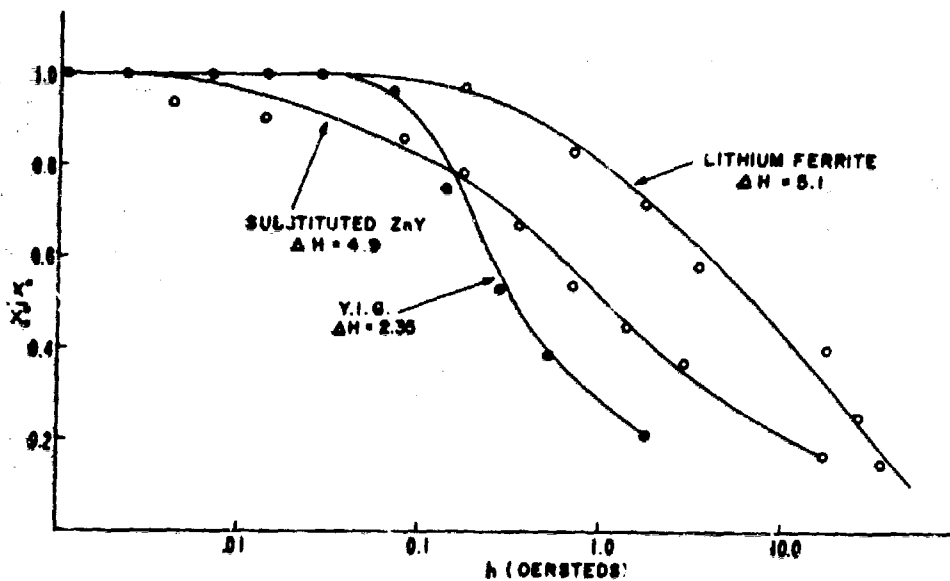


Figure III - Comparison of the normalized susceptibility of lithium, Y.I.G., and substituted ZnY single crystal ferrites as a function of incident power at 17.3 Mc.

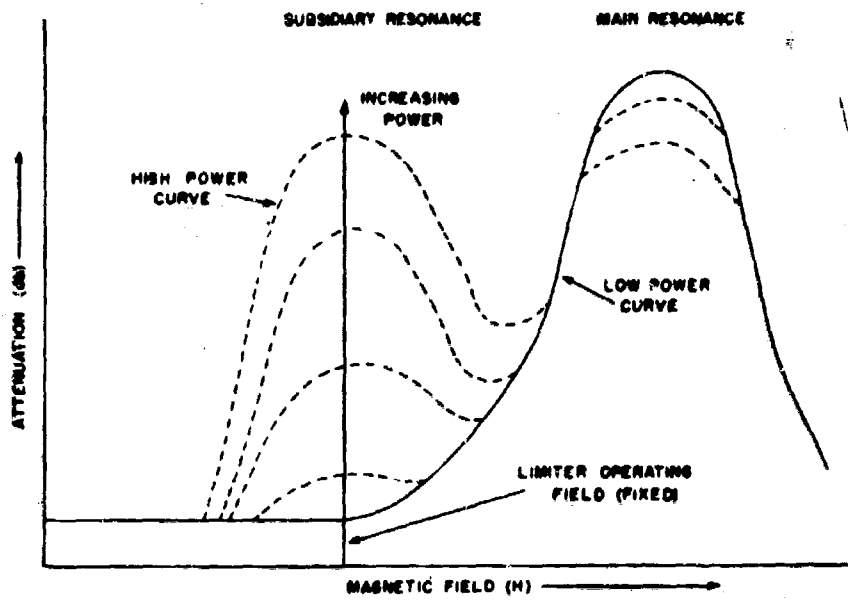


Figure IV - Development of subsidiary resonance with increase of input power.

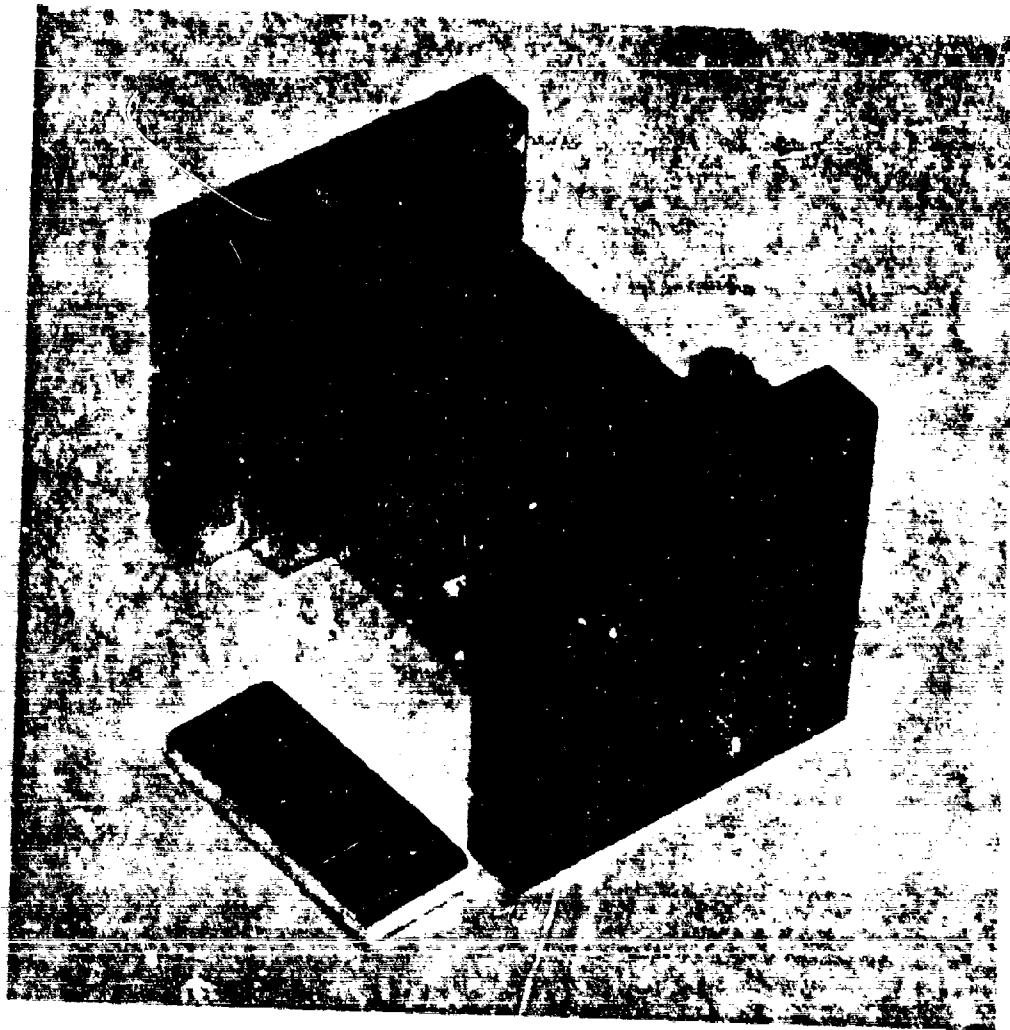


Figure V - Experimental subsidiary resonance limiter at 16.6 Mc.

RETROFUGAL ELECTRON FLUX
FROM MASSIVE TARGETS IRRADIATED
WITH A MONOENERGETIC PRIMARY BEAM

RALPH W. DRESSEL
NUCLEAR EFFECTS LABORATORY
WHITE SANDS MISSILE RANGE, NEW MEXICO

I. INTRODUCTION

A collimated beam of monoenergetic electrons directed at a thick target will sustain a variety of interactions with constituent atomic nuclei and their associated atomic electrons to produce a large dispersal of the original beam. Part of this dispersal is due to single coulomb scattering together with the simple diffusion expected from multiple nuclear and electron-electron scattering. But, in addition, secondary processes transmit energy to other resident electrons of the target medium. Thus, the total number of energetic electrons in the system may exceed the number introduced by the incident beam. For example, some secondary electrons may appear as the recoils from direct electron-electron scattering or perhaps as recoils from Compton-scattered bremsstrahlung previously generated within the target by the primary beam. Photoelectrons and positron-electron pairs may appear as well. The relative contribution from each of these numerous components strongly depends upon the primary electron energy and also upon the atomic number of the target material. As part of the total transport within the medium some electrons will migrate close enough to the target surface to escape. These constitute an electron flux directed backward with respect to the incident beam and often have been called "backscattered" electrons. However, from the previous discussion it is apparent that the term, "backscattering" is not sufficiently descriptive. Therefore, a new term, "retrofugal flux" is introduced to include within its denotation the secondary electrons as well as the scattered primary electrons. "Backscattering" is more appropriately used to describe the effect associated with incident electrons having energies considerably less than 1 MeV. In this case, the backward flux from the target surface consists of two distinguishable components. The one, identified by its higher energy, consists principally of singly- or multiply-scattered primaries while the other

is a low-energy component composed of secondary-emission electrons ejected in large numbers from the target surface, but with energies less than 50 eV. However, if the energy of the incident electrons is increased, a shower begins to develop within the target so that primary and secondary electrons no longer are distinguishable on the basis of energy. Accordingly, all electrons leaving the target surface are equivalent. These constitute the retrofugal flux.

Earlier investigations, in general, have been conducted at energies less than 1 MeV. These have been collected and summarized by Bothe (1933). A more recent review has been presented by Birkhoff (1958) covering work prior to 1953. Since then, detailed and careful studies of the energy and angular distribution of backscattered electrons from monoenergetic primaries have been performed by Kulenkampff and Spyra (1954) and by Kulenkampff and Rüttiger (1954) using thick targets of Al, Cu, Ag, and Pt and initial energies from 20 to 40 keV. Kanter (1957) has made similar measurements in an energy range from 30 to 100 keV. Using 1.75 MeV electrons extracted from a betatron, Frank (1959) has investigated both the angular distributions and the energy spectra of the retrofugal electron flux from Al, Cu, and Pb targets at various angles of incidence for the primary beam. Wright and Trump (1962) have reported experimental results for the total retrofugal flux from a variety of target materials at incident electron energies of 1, 2, and 3 MeV. The most recent work has been done by Bienlein and Schlosser (1963) who performed measurements on Al, Ag, Sn, and Au targets at energies of 60, 75, and 100 keV. Although a large amount of data has been collected for incident electron energies below 500 keV, there is sufficient scatter in the reported results to open serious questions about which values to accept. Measurements by independent investigators at the higher energies appear to agree, but there is reason to suspect that reported values all are too low due to various losses encountered in determining the electron flux. The measurements reported here have been designed to avoid such systematic errors and extend the observations to 10 MeV.

Numerous attempts have been made to calculate the angular distribution and the energy spectrum of the retrofugal flux. Bothe (1949) has proposed a diffusion theory applicable to energies near 370 and 680 keV. Everhart (1960), incorporating simplifying assumptions, has obtained results applicable to energies below 50 keV while others including Archard (1961), Engelmann (1961), Bödy (1962), and Tomlin (1963) have worked with diffusion theories appropriate for energies below 100 keV. Calculations employing the Monte Carlo technique have been performed by Sidei, Higashimura, and Kinoshita (1957) for 2-MeV electrons on Al. MacCallum (1960) has also applied the Monte Carlo method to calculate backscattering coefficients for energies from 10 to 500 keV and more recently Green (1963) has run a calculation to determine the coefficient for a Cu target at 29 keV. The only calculations applicable to higher energies have been performed by Perkins (1962) who obtained values for Al and Cu at energies from 0.4 to 4 MeV. In each of these calculations, approximations were employed to reduce the

DRESSEL

computational labor; therefore, a comparison of computed results against experimental measurements is desirable to evaluate the validity of approximations used.

II. EXPERIMENTAL SYSTEM

The system employed for measurement is illustrated in Fig. 1. A large cylindrical scattering chamber approximately 12 inches in diameter and 12 inches high has been constructed of aluminum to reduce the background currents due to electrons scattered from its interior walls. Targets are solid right cylinders approximately 2 inches in diameter and 2 inches long supported in an aluminum holder mounted so that the center of the plane end-face lies precisely at the center of the scattering chamber. This holder permits target rotation about a vertical axis to position it at any desired angle with respect to the incident beam. The target surface exposed to the beam is machined plane within ± 0.001 in. Ten measurement ports are spaced around the scattering chamber periphery. By rotating the chamber it is possible to select appropriate pairs of ports that provide regularly increasing deflection angles from 0° to 155° at 5° intervals. Thus, 32 angular intervals are available with only 10 ports. Target materials have been chosen to span the scale of atomic number from $Z = 4$ for beryllium to $Z = 92$ for uranium at reasonably uniform intervals. Only metallic conductors have been selected for targets to avoid accumulation of electric charge and also to conduct away heat generated by the incident beam.

Electrons leaving the LINAC pass through a system of aligned collimating apertures in a drift tube. This same drift tube contains a beam current monitor described by Dressel (1963). It is calibrated by removing the target from the beam which is then caught in a Faraday cup mounted at the straight-through port. Although the LINAC current is reasonably steady so that measurements of relative currents could be made, greater precision is obtained by comparing the integrated monitor current against the total charge collected by the Faraday cup. A sliding bias is used on the Faraday cup during charge collection so that it is always within a fraction of a volt from ground potential. This technique reduces to a negligible amount possible losses due to leakage currents. An evaluation of the errors and losses expected from this Faraday cup by Dressel (1961) indicates they are less than 0.2 percent. Immediate repetition of monitor calibration readings gives results consistent within 0.5 percent. Unfortunately, due to random fluctuations in the many parameters that determine the LINAC operating conditions, this high precision cannot be held over a period longer than a few minutes. However, these fluctuations have a standard deviation of about 5 percent so that averages over several replications yield results reliable within 2 or 3 percent.

Although the primary beam formed by the LINAC has a cross section not greater than 2 mm in diameter, the beam itself migrates during the irradiation period and executes an excursion over a region about 5 mm in diameter. The collimators have been designed with apertures 1 cm in diameter so that when properly aligned, the electron beam always remains within these apertures. Another collimator having a conical taper is used in the measurement port to define the solid angle through which the retrofugal flux is sampled. This collimator views the entire region on the target face from which retrofugal electrons may be emitted. When measurements of the energy spectra are desired, the analyzer magnet and deflection chamber are installed as shown. Alternatively, when only angular distribution measurements are made, the Faraday cup is coupled directly to the appropriate measurement port. The entire assembly, including scattering chamber and Faraday cup, is evacuated and connected directly to the LINAC system. A vacuum gate valve permits isolation of the LINAC whenever the scattering chamber is opened to change targets or perform other system modifications.

Test measurements have been made to check out the system for electron background and other spurious effects. The scattering chamber itself appears to be quite free of such troubles and contributes a background signal level substantially less than 1 percent of the normal signal. The Faraday cup is quite insensitive to any radiation except those electrons entering through its aperture but, interestingly, under some conditions a small positive signal may appear during operation while the target is removed. It has been determined that this positive signal is associated with secondary electron emission from the outer walls of the lead cup due to broadside irradiation with stray bremsstrahlung from the LINAC. The effect is small and has been further reduced by appropriate shielding of the LINAC from the rest of the system. Another effect that is slightly larger in magnitude is due to polarization of the coaxial cables in the radiation field. This contributes a small negative signal. All of these background signals are negligible for measurements on targets having atomic numbers greater than 13. However, they have been evaluated and are applied as corrections to the measurements at lower atomic numbers. A Faraday cup was chosen specifically for the detector in this application, partly because of its relative insensitivity to spurious background effects, and also because it is an absolute instrument. Thus, many problems such as calibration, dead time, and coincidence losses commonly associated with other detectors are successfully avoided.

III. RESULTS

An electron beam of finite cross section, upon entering a solid target, spreads due to scattering. This behavior can be made visible by exposing a block of PLEXIGLAS to the beam. After sufficient exposure, a green-orange color is registered in the block recording the electron density. The resultant image resembles the flair of the bristles

on a dry shaving brush. Apparently, most of the electrons that work back out of the target face to form the retrofugal flux issue from a region having a diameter less than 1/2 the electron range in the material. This conclusion is supported by PLEXIGLAS exposures made at the target surface. Photographs of gas fluorescence induced by the retrofugal electrons from targets bombarded with a 45-keV electron beam have been obtained by Schumacher (1962). These give a visual indication of the presence and also the angular distribution of the electrons in space.

If a target surface were uniformly bombarded with electrons over an extended region, the angular distribution, expressed as the number of retrofugal electrons per steradian per second, would be independent of the position on the surface, but when a narrow beam is used position with respect to the beam becomes important. Suppose, as in Fig. 2, that an electron beam of vanishing cross section is incident normally upon a plane target at a position \bar{r} . The retrofugal electrons escaping from the surface will be distributed symmetrically about the incident beam axis. Let this distribution be described by a function $k(|\bar{p}-\bar{r}|, \theta, E, Z)$. Then the total number of retrofugal electrons received in a detector subtending a finite solid angle $\Delta\Omega$ at an angle θ measured from the target normal will be given by integrating the contributions from all parts of the finite incident beam. For a given energy E for the incident primary electrons, this number may be expressed

$$N(\theta, E, Z) = N_0 \int_{\Delta\Omega} \int_{S_T} \int_{S_T} q(\bar{r}) k(|\bar{p}-\bar{r}|, \theta, E, Z) da_{\bar{r}} da_{\bar{r}} d\Omega$$

Here the label S_T designates the total area covered by the incident beam while S_p is the area surrounding the point of incidence from which retrofugal electrons escape. It is expected that $k(|\bar{p}-\bar{r}|, \theta, E, Z)$ will be approximately a gaussian function of $|\bar{p}-\bar{r}|$ having a full width at half-maximum less than the electron range in the target medium. The total incident beam current expressed in electrons/sec is denoted by N_0 while the function $q(\bar{r})$ describes the beam distribution over its cross section. A differential retrofugal flux coefficient $\rho(\theta, E, Z)$ may be defined as

$$\rho(\theta, E, Z) = \frac{N(\theta, E, Z)}{N_0 \Delta\Omega}$$

so that

$$\rho(\theta, E, Z) = \frac{1}{\Delta\Omega} \int_{\Delta\Omega} \int_{S_T} \int_{S_T} q(\bar{r}) k(|\bar{p}-\bar{r}|, \theta, E, Z) da_{\bar{r}} da_{\bar{r}} d\Omega$$

DRESSEL

Evidently $\rho(\theta, E, Z)$ is the average value of $A(\beta-F1, \theta, E, Z)$ obtained under the experimental conditions and θ is the average value of Θ . Since all distributions are symmetrical about the beam axis, θ is simply the angle between the target normal and the detector axis. It is clear that the finite beam cross section only decreases the resolution slightly, but introduces no other error providing the detector can receive retrofugal electrons from the entire region over which they issue from the target. The angular resolution expected from the experimental configuration is $\pm 3^\circ$. This is quite adequate since all angular distributions are rather broad.

Measurements of $\rho(\theta, E, Z)$ have been performed using targets of Be, C, Al, Cu, Sr, Mo, Ag, Ba, W, Pb, and U. The data cover an angular range from $\theta \approx 90^\circ$ to 180° and an energy range from $E \approx 0.5$ to 10 MeV. Limited space does not permit a display of all of these data. Accordingly, only the extremes are presented. The family of curves in Fig. 3 show the trends obtained with a Be target for a primary beam at normal incidence. The magnitude of $\rho(\theta, E, Z)$ drops rapidly with increasing energy reaching minimum near 7 MeV and then rises slightly again with further increase in energy. This trend is typical of the light-target materials; but as the atomic number Z advances, the drop with increasing energy is less pronounced and the minimum tends to disappear. For $Z \geq 13$ there is no minimum at all. All curves have been extrapolated to 100 keV to afford comparison with the work of others at the lower energies. Values indicated at $\theta = 180^\circ$ were obtained by a technique that circumvents the obvious problem created by the primary beam. When a target is rotated so that the incident beam no longer strikes at normal incidence, the angular distribution pattern for the retrofugal flux shifts. The maximum for this distribution moves in angle more rapidly than the angular displacement of the target. In addition, the maximum also increases a few percent. The amount of angular shift and the magnitude increase has been determined for the targets investigated. Thus, for Be, the normal to the target must be set at 5° from the incident beam to direct the maximum of the retrofugal flux pattern into the measurement port at 155° . At this angular offset the maximum is 8 percent greater than what it would have been at $\theta = 180^\circ$ and the primary beam at normal incidence. Similar settings for a U target require an offset angle of 17.5° and a magnitude correction of 1.5 percent.

Data obtained with a depleted uranium target are displayed by the family of curves in Fig. 4. The slower decrease in $\rho(\theta, E, Z)$ with energy is immediately apparent, but an interesting inflection occurs in all of the curves corresponding to a primary electron energy near 5 MeV. This coincides with the threshold for photofission of ^{238}U at 5.1 MeV. Since such an inflection does not occur for other target materials, it appears that energetic process of uranium fission adds a significant contribution to the total retrofugal electron flux from this target.

The development of $\rho(\theta, E, Z)$ as a function of the target atomic number Z is most conveniently displayed at $\theta = 180^\circ$. These

curves are shown in Fig. 5. The points have been taken from curves similar to those of Figs. 3 and 4 for each of the targets. For small values of Z these curves rise rapidly and then, as Z continues to increase, they tend to level off. Such behavior is expected since most of the interaction cross sections for electrons with the target atoms increase with Z . The tendency to level off, on the other hand, is due to the fact that only a finite number of electrons can work back out of the target for a finite number incident. This ratio would tend toward a constant even if the interaction cross sections were infinite.

In Fig. 6, the family of curves shows $\rho(185^\circ, E, Z)$ as a function of energy for various values of Z . Here the general development from small to large Z is evident with interesting changes in the structure of the curves. The distinctive appearance of the curve for Z differing from all the others at the high energy end is very clear. Effects of bremsstrahlung generation within the target are visible in the progressive upward bow of the curves in the vicinity of 3 MeV.

Changes in the angular distribution of the retrofugal electrons from low Z targets are indicated in Fig. 7. Once more the progressive development is apparent. At low energies, the angular distributions tend toward cosine distributions but become sharper with increasing Z . However, at the higher energies very material changes occur. For target materials having $Z > 13$ the angular distributions are all the same and are effectively independent of energy. They are sharper than cosine distributions and tend to follow the values expressed for Al at 8 MeV.

A comparison of the results obtained from these measurements with those obtained by other investigators is important because of significant discrepancies that exist. For the present measurements, the earth's magnetic field deflects the very low-energy electrons sufficiently to establish a low-energy cutoff at about 800 eV. This is consistent with the low-energy discrimination employed by others although the cutoff value is a little higher. A summary of results is presented in Fig. 8 where data and calculated values reported by previous investigators are superimposed upon curves derived from the present measurements. Since much of the literature reports only the total retrofugal flux coefficient $R(E, Z)$ obtained by integrating $\rho(\theta, E, Z)$ over the entire backward hemisphere, the data in Fig. 8 are displayed in this form. The logarithmic scale for energy was chosen partly for the convenience it affords in plotting the extended range in energy and also to emphasize the low-energy values where most of the previous work has been done. It is apparent immediately that nearly all of the previously reported results lie below the present values although the general trend with energy appears to be similar. The range of energies over which measurements have been performed has been indicated by a horizontal line and the uncertainties, where they have been given, are depicted by a vertical line. At the lower energies, the present results are consistent with previous reports that R is roughly independent of energy.

DRESSEL

For Al the present value falls close to the general average, but for the targets of higher Z the present values correspond more closely to the upper-limit of previous data. Measurements reported at higher energies agree moderately with one another but fall substantially lower than the present data. The discrepancy increases with increasing Z. A Monte Carlo calculation of R by Green shows very good agreement with the present data. However, the values obtained by Perkins are quite below the present curves. To illustrate that all calculations do not agree with each other, a value determined by Sidei for Al at 2 MeV falls considerably above the current data.

The calculations performed by Perkins may be expected to yield values that are too low, since the computation involved only the effects of scattering with energy loss but did not include the regeneration of electrons due to secondary processes. One may take the results of Perkins to illustrate how much of the total retrofugal flux is composed of scattered incident primaries. The difference between present data and Perkins results then may be regarded as due to the additional high-energy secondary electrons generated within the target. Under this interpretation the discrepancy, which increases with larger Z, becomes understandable. Sidei's calculations were performed manually and treated the effects of Coulomb scattering only. For the calculations carried out by Green, however, contributions due to secondary effects are not important because of the low electron energy treated; hence, the observed good agreement with present data is significant.

Low values reported by previous workers are not surprising for as Kulenkampff and Ruttiger (1954) have pointed out, "Every operating influence serves to impose a reduction in the measured electron current." Recognizing this, they confidently reported results which were larger than other values accepted at the time but which agree quite well with the present data. Measurements reported by Frank (1959) are probably low for two reasons. First, Frank argues, quite plausibly, that a target thickness equal to one-half the electron range is an infinite target for all practical purposes. Any electron scattered backwards could not escape from the initial surface if it penetrated deeper into the target than this. But such an argument overlooks the bremsstrahlung generation which contributes significantly to the retrofugal electron flux. Since the bremsstrahlung is quite penetrating, much thicker targets than Frank used are needed to approximate an infinite medium. In fact, if Frank's own data, obtained with targets of increasing thickness, are plotted against the target thickness it becomes obvious that the saturation thickness had not been reached. As a result, the reported values are too low. A second reason for suspecting a low result is associated with the fact that Frank used geiger tubes for detectors and a betatron operating at a frequency of 500 Hz. Although he was careful to hold down the counting rate there is a distinct possibility that there may have been coincidence losses, especially if the electron flux were greater than that expected. Measurements performed by Wright and Trump (1962) employed a collector for retrofugal electrons in the form of an aluminum cylindrical cup having a 0.25-inch wall thickness. It turns out that 0.25 inch of aluminum is

DRESSEL

approximately the range for 3-MeV electrons. This was the highest energy employed in their investigation. However, a significant charge loss will occur from such a thin-walled collector due to ejection of medium-energy electrons from its exterior walls as a result of photo-emission and Compton collisions by photons of scattered bremsstrahlung. Such charge loss will lead to a low measured value. For the reasons given, it appears that the present measurements which were carefully designed to avoid losses and other systematic errors are more likely to be correct.

Energy spectra of the retrofugal electron flux at selected incident electron energies from 0.5 to 10 MeV and for a number of targets have also been obtained. These interesting results will not be reported here since they would introduce a whole new topic.

ACKNOWLEDGEMENT

Appreciation is expressed for the assistance of Mr. Anthony Grijalva for keeping the LINAC in good operating condition and to the Nuclear Effects Laboratory for the extended use of its instruments.

DRESSEL

BIBLIOGRAPHY AND REFERENCES

- G. D. Archard, Journ. Appl. Phys. 32 1505 (1961).
- J. V. Bienlein and G. Schlosser, Zeits. für Physik, 174 91 (1963).
- R. D. Birkhoff, Encyclopedia of Physics, Vol. 34 pp. 131-132 S. Flügge (Springer-Verlag, Berlin 1958).
- W. Bothe, Handbuck der Physik, Bd. 24 Teil 2 S.1 (Berlin 1933).
- _____, Zeits. für Naturforschung, 4a 542 (1949).
- _____, Ann. der Physik, 6 44 (1949).
- T. E. Everhart, Jour. Appl. Phys. 31 1483 (1960).
- H. Frank, Zeits. für Naturforschung, 14a 247 (1959).
- M. Green, Proc. Roy. Soc., 82 204 (1963).
- H. Kulenkampff and W. Spyra, Zeits. für Physik, 137 416 (1954).
- _____ and K. Rüttiger, Zeits für Physik, 137 426 (1954).
- H. Kanter, Ann. der Physik, 20 144 (1957).
- _____, Phys. Rev., 121 681 (1961).
- C. Mac Callum, Bull. Am. Phys. Soc., 5 379 (1960).
- J. F. Perkins, Phys. Rev., 126 1781 (1962).
- Sidei, Higashimura, and Kinosita, Mem. Fac. Eng. Kyoto Univ. 19 220 (1957).
- B. W. Schumacker, Canad J. Phys., 40 376 (1962).
- S. G. Tomlin, Proc. Phys. Soc., 82 465 (1963).
- K. A. Wright and J. G. Trump, Journ. Appl. Phys., 33 687 (1962).
- R. W. Dressel, Evaluation of the Faraday Cup, NEL Report (October 1961).
- _____, Electron Energy Spectrum of the Linac, NEL Report (August 1962).
- _____, Nuclear Inst. and Methods, 24 61 (1963).

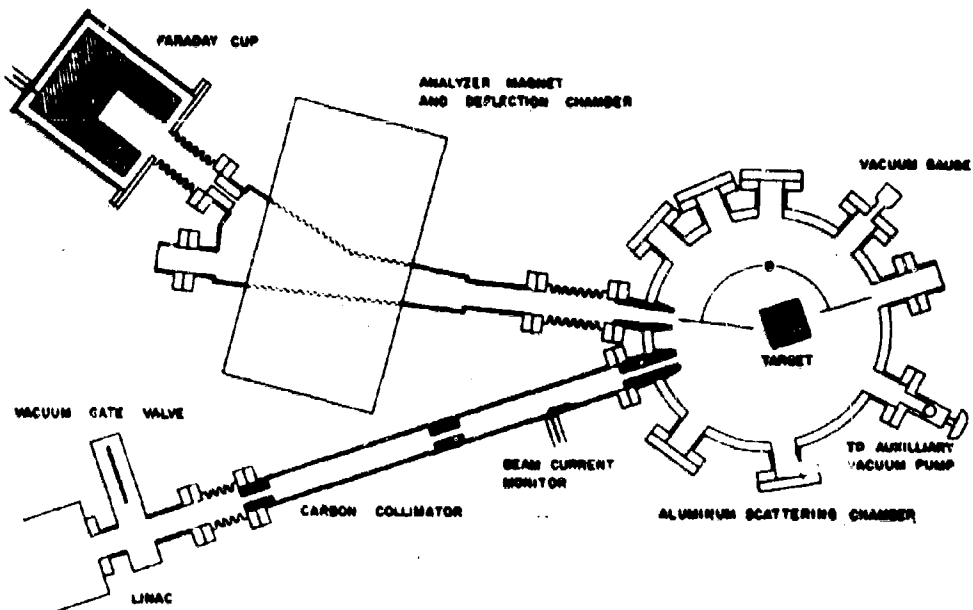


FIG 1 ARRANGEMENT OF APPARATUS FOR MEASURING THE RETROFUGAL FLUX OF ELECTRONS FROM THICK TARGETS

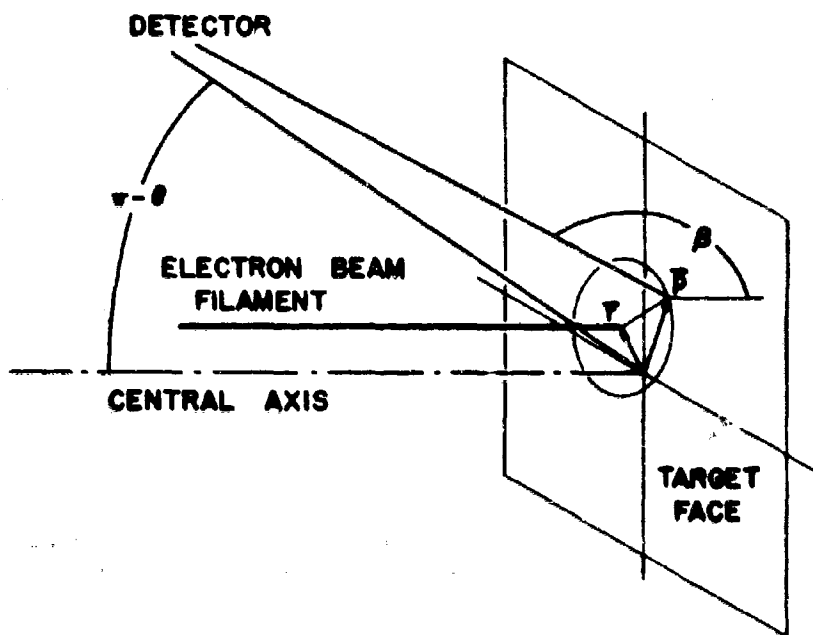


FIG. 2 DIAGRAM DEFINING GEOMETRICAL RELATIONSHIPS AT THE TARGET.

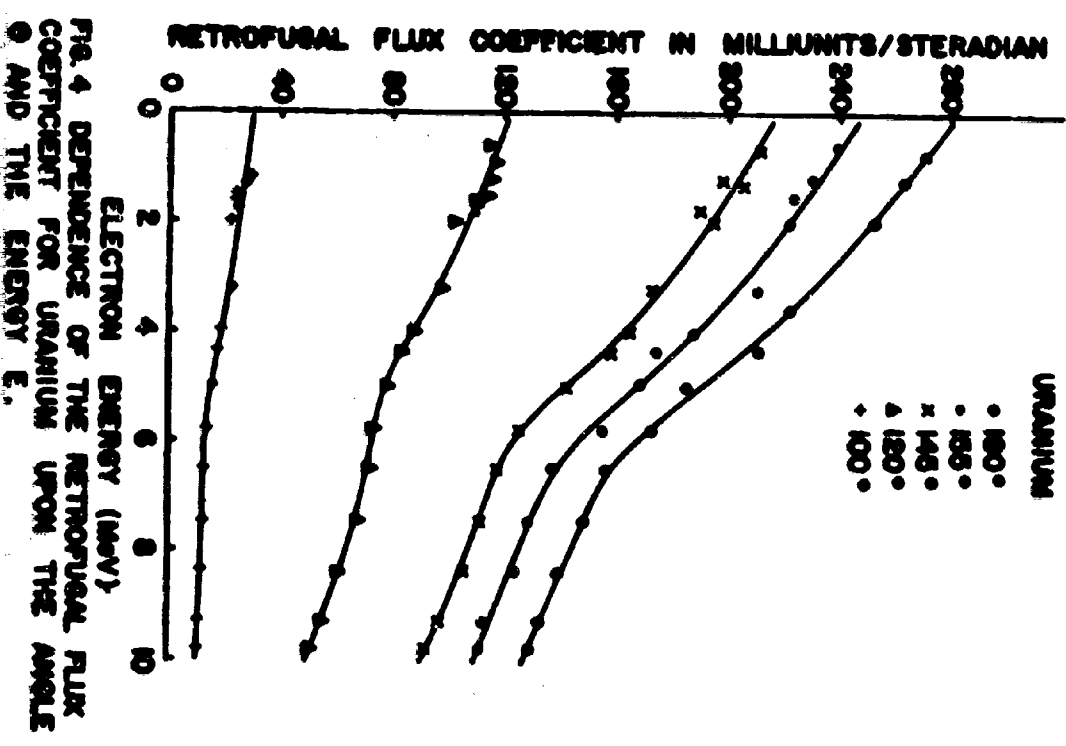
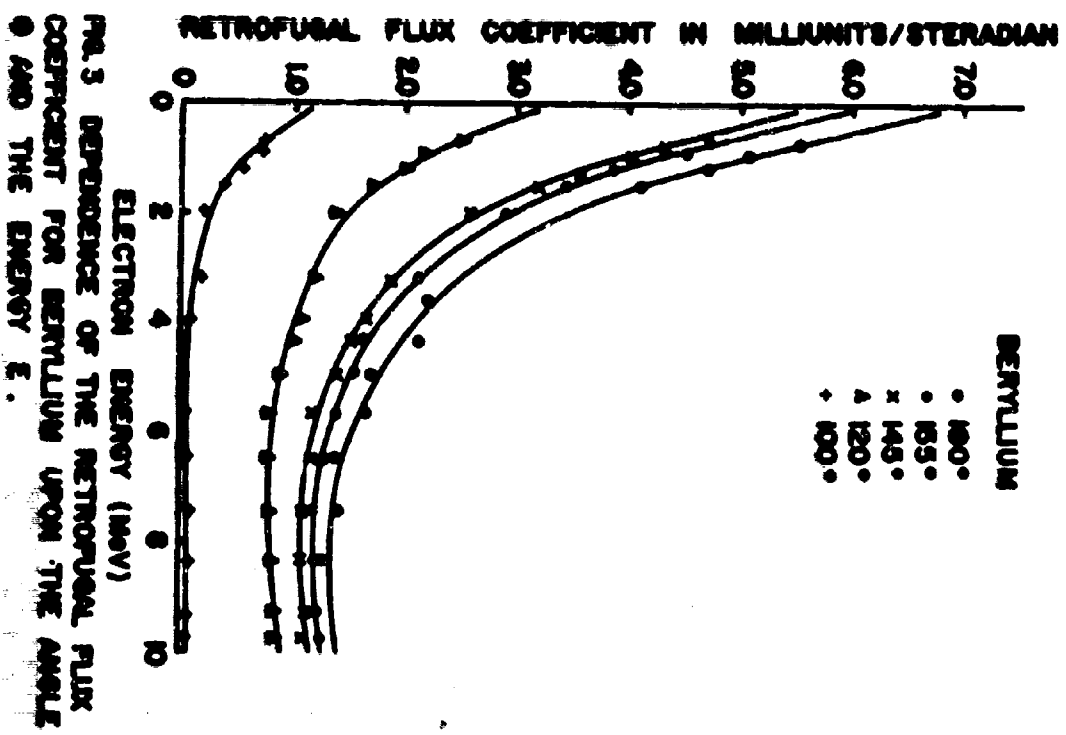
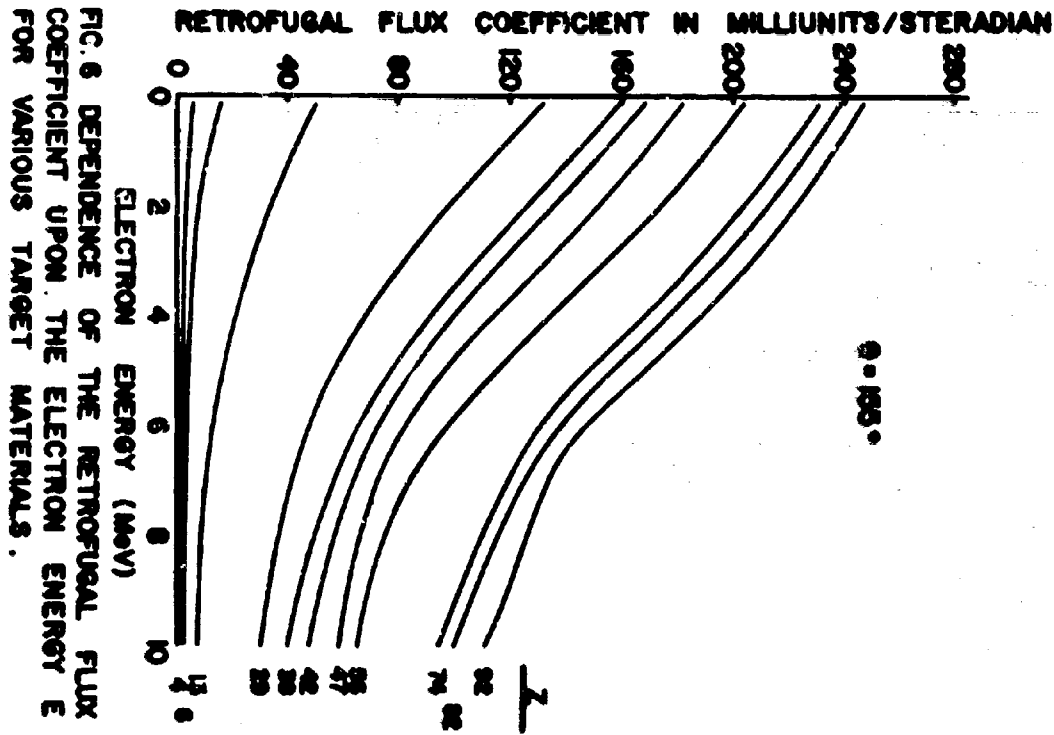
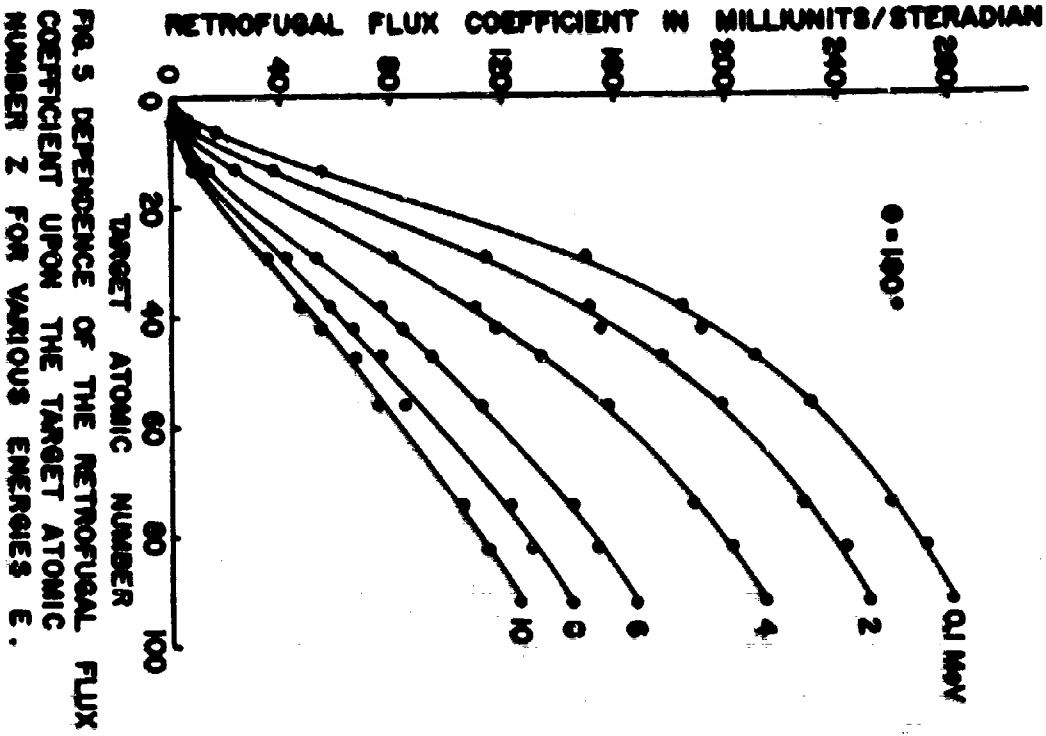


FIG. 3 DEPENDENCE OF THE RETROFUSAL FLUX COEFFICIENT FOR BERYLLIUM UPON THE ANGLE AND THE ENERGY E.

FIG. 4 DEPENDENCE OF THE RETROFUSAL FLUX COEFFICIENT FOR URANIUM UPON THE ANGLE AND THE ENERGY E.



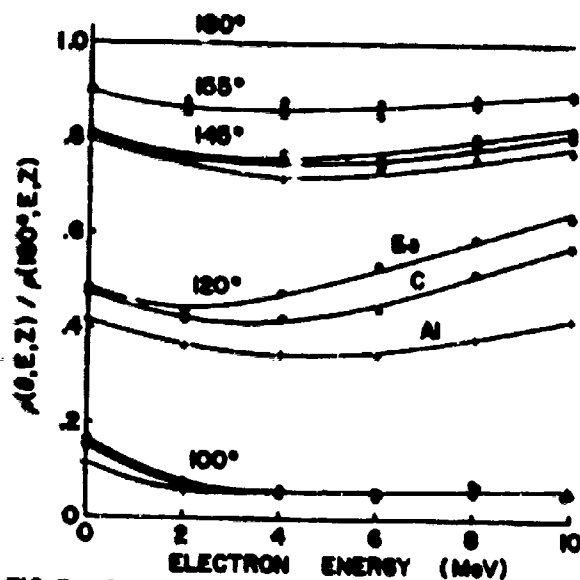


FIG. 7 ENERGY DEPENDENCE OF THE ANGULAR DISTRIBUTION OF RETROFUSAL FLUX FROM TARGETS OF Be, C, AND Al.

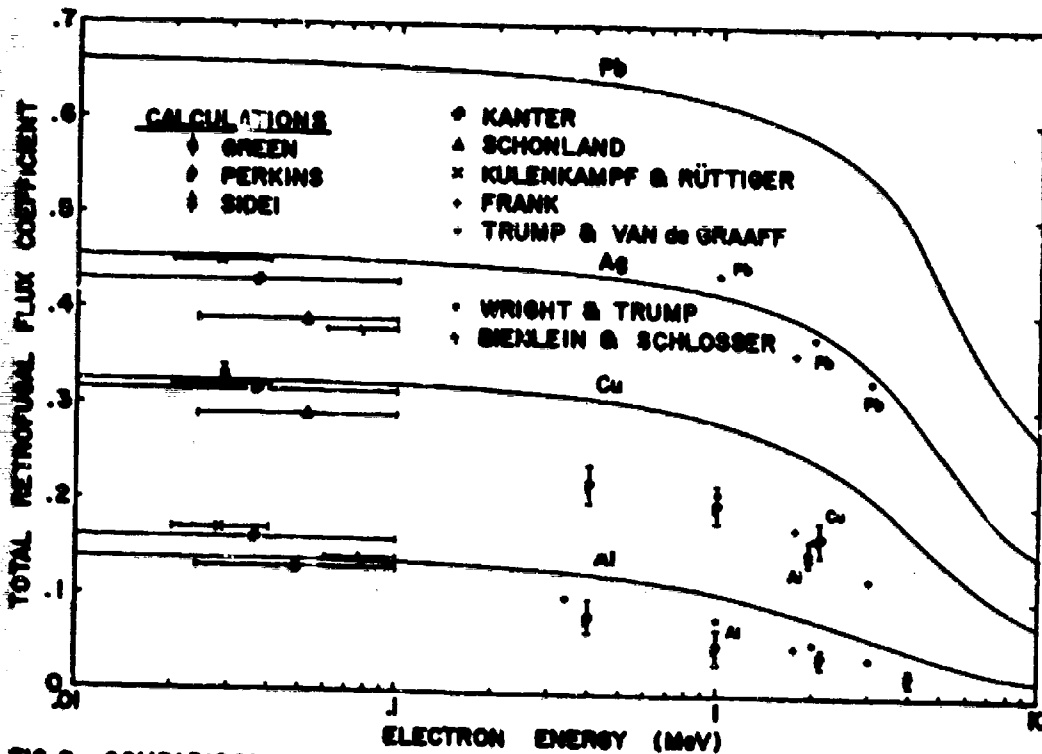
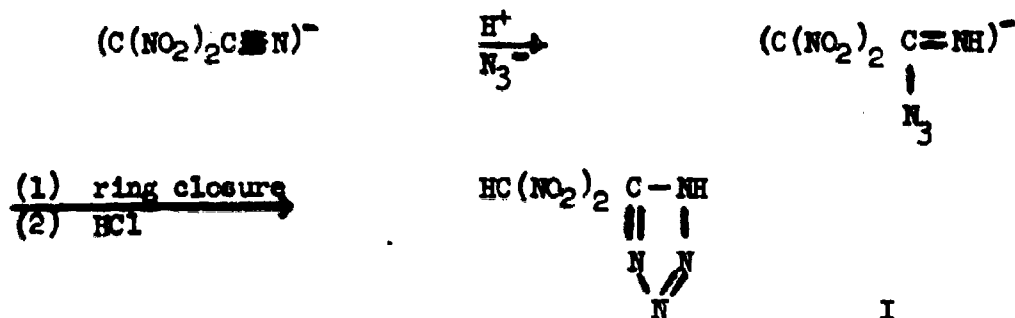


FIG. 8 COMPARISON OF RESULTS. CURVES ARE FROM PRESENT DATA.

PREPARATION OF 5-DINITROMETHYLTETRAZOLE FROM SALTS
OF DINITROACETONITRILE

F. EINBERG
FRANKFORD ARSENAL
PHILADELPHIA 37, PA.

A general method for preparing tetrazoles by the reaction of nitriles with sodium azide in the presence of ammonium or tri-alkylammonium ion in dimethylformamide has recently been described.¹ It was more recently reported² that 5-trifluoromethyltetrazole could be prepared by an extension of this method and that the electronegative trifluoromethyl group promoted the reaction to such an extent that the use of an ammonium catalyst was unnecessary. We attempted to prepare the previously unknown dinitromethyltetrazole by the reaction of dinitroacetoneitrile and sodium azide in water but with negative results probably due to generation of the sodium salt of dinitroacetoneitrile and the highly associated hydrazoic acid. This led us to examine the reactions of salts of dinitroacetoneitrile with sodium azide for the formation of dinitromethyltetrazole. It has now been found that the product can be obtained by the reaction of sodium or ammonium dinitroacetoneitrile (but not the potassium salt) with sodium azide and ammonium chloride (mole ratio 1.00: 1.02: 1.02, respectively) in refluxing water but not in dimethylformamide. This appears to be the first example of the use of an organonitrile salt³ to form a tetrazole and is the first example of a 5-nitroalkyltetrazole. The reaction may be represented as follows.



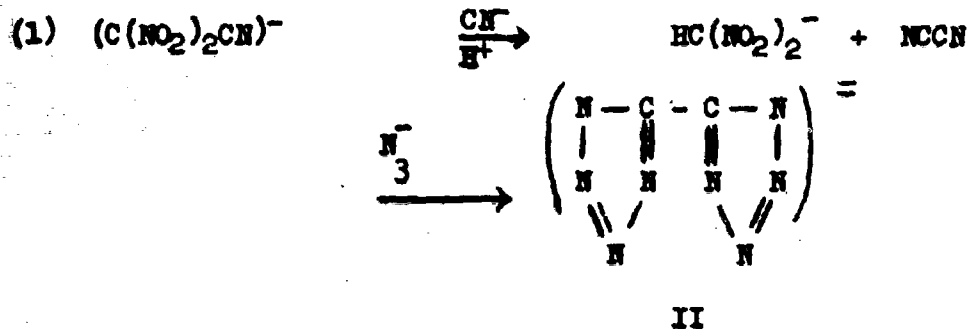
EINBERG

It would appear that the strong electron withdrawing effect of the gem-dinitro group would greatly enhance nucleophilic attack by the azide ion on the carbon of the nitrile group.⁴ The product, however, was obtained in only 8-12% yields after purification. This may have been partly due to hindrance of the reaction by repulsion of the negative azide ion by the dinitroacetonitrile anion, or sterically by the bulky dinitromethyl group.

Over seventy percent of the dinitroacetonitrile salt was recovered after 24 hours of refluxing, starting with either the sodium or ammonium salt. In the reaction of sodium dinitroacetonitrile, however, most of this substance was recovered as the ammonium salt, which means that the ammonium chloride was depleted. This may account for the somewhat lower yield starting with the sodium salt. That this was the case, was indicated when a much lower yield of 5-dinitromethyltetrazole was obtained using one-tenth the original amount of ammonium chloride catalyst in the reaction of azide ion with ammonium dinitroacetonitrile. Available ammonium ion may also be depleted by formation of ammonium salts of the products.

The yield of dinitromethyltetrazole obtained after 48 hours of refluxing was nearly the same as that obtained after 24 hours, while that of diammonium bitetrazole (II), a by-product of the reaction, increased threefold. Decomposition products also increased considerably. Cyanide ion formed as a breakdown product was shown to be present in the reaction mixture after 24 hours of refluxing (see experimental). Cyanide ion, in turn, must react to produce bitetrazole.

Bitetrazole likely is formed by nucleophilic attack of cyanide ion on the nitrile carbon of dinitroacetonitrile anion. The dinitromethyl group would be displaced by cyanide ion and the product formed by azide ion attack on the nitrile carbons, followed by ring closure.



F. TNERG

It is not clear how the cyanide ion arises. Methyl 4,4-dinitro-4-cyanobutyrate has been reported⁵ to react with methanolic potassium hydroxide to give potassium methyl 4,4-dinitrobutyrate and potassium cyanate. This result was attributed to a nucleophilic attack by hydroxyl ion on the nitrile displacing the dinitrobutyrate fragment.

The reaction of sodium dinitroacetonitrile with sodium azide and ammonium chloride in dimethylformamide appeared to produce only decomposition products. No product was obtained with potassium dinitroacetonitrile in refluxing water or in dimethylformamide at 110-120° during 48 hours. It has been observed that potassium dinitroacetonitrile is unchanged in alkylation, addition, and other reactions contrary to the more normal behavior of the sodium and other salts of dinitroacetonitrile.⁶

Mono- and disodium, monoammonium, and monoguanidine salts of 5-dinitromethyltetrazole also were prepared from the free dibasic acid. The disodium salt was found to be much more thermally stable than the monosodium salt. This may be attributed to a pronounced resonance stabilization of the tetrazole ring in the dianion by release of an electron pair to the ring.

Infrared bands of dinitromethyltetrazole and its salts attributed to NH stretching, the tetrazole ring, and the nitro groups are summarized in Table I.

Table I

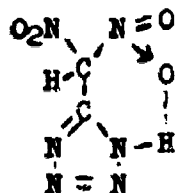
Infrared Bands of 5-Dinitromethyltetrazole
And Its Salts^a

Compound	$\bar{\nu}^b$ cm ⁻¹	Tetrazole Ring ^c		Gem-Dinitro Group ^d		Others ^e	
		cm ⁻¹	cm ⁻¹	cm ⁻¹	cm ⁻¹		
5-Dinitromethyl- tetrazole	3130	1070, 1048, 1016, 997	1590, 1330	1518, 1407, 1350			
Monosodium-5-Dinitro- methyltetrazole	3170	1670, -	1018, 1005	1570, 1350	1485, 1412, -		
Disodium-5-Dinitro- methyltetrazole	-	-	1032, 1015	1530, 1365	-	-	-
Monosodium-5-Dinitro methyltetrazole	3230	-	1045, 1022	1530, 1360	-	1400, -	

- a. All spectra were obtained on NaJol mills (Perkin-Elmer Spectrophotometer Model 321).
- b. Bands corresponding to hydrogen bonded NH (3230-3030 cm⁻¹ 7).
- c. In general, the tetrazole ring absorbs between 1100-1000 cm⁻¹ 8
- d. In general, gem-dinitro absorbs near 1580 cm⁻¹ and 1330 cm⁻¹ 9
- e. Possibly due to nitro.

EINBERG

The shift in frequency for the nitro groups from 1590 cm^{-1} for dinitromethyltetrazole, to 1570 cm^{-1} for the monosodium salt, to 1530 cm^{-1} for the disodium and monoammonium salts, may be due to a change in character from nitro to nitronate¹⁰. The absence of the NH band in the disodium salt, and its presence in the monosodium salt of dinitromethyltetrazole shows that the stronger of the two acidic hydrogens of dinitromethyltetrazole is in the dinitromethyl group. The occurrence of N-H stretching below 3300 cm^{-1} is indicative of hydrogen bonding and is probably due to intramolecular hydrogen bonding. For the free acid this may be illustrated as follows.



Intramolecular hydrogen bonding also has been observed for 2,2-dinitroethyl amine.¹²

Experimental¹³

Potassium dinitroacetonitrile¹⁴ was treated with decolorizing charcoal and recrystallized from water. Sodium and ammonium dinitroacetonitrile were prepared from the potassium salt according to a previously described method⁶. The salts were dried at 100° for 2 hrs., and kept in a desiccator over Drierite¹⁵; sodium dinitroacetonitrile, m.p. $224-226^\circ$ (lit.⁶, m.p. $224-226^\circ$); ammonium dinitroacetonitrile, m.p. $175-176$, dec. (lit.⁶ m.p. 182° , dec.). Sodium azide was treated with decolorizing carbon and crystallized from water. Ammonium chloride and guanidine carbonate were reagent grade and were used without further treatment.

Reaction of Sodium Dinitroacetonitrile, Sodium Azide and Ammonium Chloride. - A solution of sodium dinitroacetonitrile (15.30 g., 0.100 mole), sodium azide (6.65 g., 0.102 mole), and ammonium chloride (5.50 g., 0.102 mole) in 50 ml of water in a round bottom flask fitted with a condenser and magnetic stirrer was stirred and refluxed for 24 hr.¹⁶.

Isolation of Diammonium Bitetrazole And Its Conversion To Bitetrazole. - The solution was cooled and kept at 5° for several hours and the separated yellowish solid was collected on a filter, washed with cold water, and dried at 110°. The yield of the salt was 0.59 g. (7.0%). In another run using half the quantities, the yield after 48 hours of refluxing was 0.90 g. (21.2%). Crystallization from water gave white needles which did not melt up to 300°.

Anal. Calcd. for $C_2H_8N_{10}$: C, 13.95; H, 4.65; N, 81.40.
Found: C, 13.78, 13.98; H, 4.83, 5.02; N, 81.50, 82.03.

The diammonium bitetrazole was dissolved in an excess of hot 10% hydrochloric acid (1 g. of salt in 10 ml. of acid). The solution was filtered hot and the filtrate cooled to 5° causing bitetrazole to precipitate as a white, crystalline solid. The product was highly acidic in aqueous solution. For analysis, a sample was recrystallized twice from water and dried at 110°; m.p. 255.0-255.5, dec. (lit. 17, m.p. 254-255, dec.).

Anal. Calcd. for $C_2H_2N_8$: C, 17.38; H, 1.45; N, 81.16;
neut. eq. 69.0. Found: C, 17.33, 17.36; H, 1.46, 1.51; N, 82.40,
82.43; neut. eq. 69.1.

Isolation of 5-Dinitromethyltetrazole. - The pH of the ammonium bitetrazole filtrate was adjusted to less than one (measured with a pH meter) with conc. hydrochloric acid. Hydrazoic acid and water were removed under water aspiration on the steam bath. The remaining solids were extracted with boiling acetone and the insoluble salts were removed by filtration and discarded. The acetone was evaporated under reduced pressure, the remaining dry solid was extracted several times with ethyl ether totalling several hundred milliliters, and the insoluble solids reserved for recovery of dinitroacetonitrile salt. Removal of the ether gave 2.26 g. of crude product. Crystallization from methyl alcohol gave 1.39 g. (8.0%) of white, crystalline product which produced intensely yellow, highly acidic aqueous solutions. The yield after 48 hours of refluxing using half the quantities was 0.77 g. (8.8%). For analysis a sample was recrystallized from methyl alcohol.

Anal. Calcd. for $C_2H_2N_6O_4$: C, 13.80; H, 1.16; N, 48.28;
 neut. eq. 87.0. Found: C, 14.01, 14.04; H, 1.50, 1.24; N, 48.10,
 48.20; neut. eq. 87.7; μK_1 , 1.65, μK_2 = 3.60 (water/25°).¹⁹

The free acid, on heating fairly rapidly in a capillary tube produced nitrogen dioxide fumes at about 100°, gradually darkened above 150°, and went abruptly from dark brown to black, without melting, at about 180°.

Recovery of Ammonium Dinitroacetoneitrile¹⁸. - The ether insoluble solids were extracted several times with small quantities of hot ethyl acetate leaving 5.3 g. of ammonium dinitroacetoneitrile which was collected on a filter. The filtrate was evaporated to a small volume (approx. 25 ml.), cooled to 5°, and an additional 3.9 g. of ammonium dinitroacetoneitrile was collected. Evaporation of the final filtrate left 2.1 g. of ammonium dinitroacetoneitrile containing a small amount of the sodium salt. The total recovered salt was 11.3 g. (76.4%). Total recovered salt from the reaction mixture refluxed 48 hours using half quantities was 5.2 g. (70.7%). The ammonium salt was recrystallized twice from alcohol giving white crystals whose infrared spectrum was identical to that of a known sample prepared as described above.

Preparation Of Monoammonium 5-Dinitromethyltetrazole. - 5-Dinitromethyltetrazole was dissolved in sufficient 10% aqueous ammonia to give a solution having a pH greater than 9. The water and excess ammonia were evaporated and the product was crystallized from ethyl alcohol. The monoammonium salt darkens above 195°, and abruptly decomposes without melting at 214°. Several attempts to prepare the diammonium salt gave monoammonium salt.

Anal. Calcd. for $C_2H_5N_7O_4$: C, 12.56; H, 2.62; N, 51.36
 Found: C, 12.69; H, 2.70; N, 51.55.

Preparation of Monosodium and Disodium 5-Dinitromethyltetrazole. - 5-Dinitromethyltetrazole was titrated electrometrically to either the first or second end points with 0.1N sodium hydroxide. The water was removed and the dry salts were washed with alcohol, then with acetone, and then dried. The disodium salt appeared to be unchanged up to 310°, and the monosodium salt exploded violently at approximately 160°.

EINBERG

Anal. Calcd. for $C_2H_6O_4Na$: C, 12.21; H, 0.76; N, 42.73; Na, 11.76. Found: C, 12.24; H, 0.51; N, 42.86; Na, 11.73.

Anal. Calcd. for $C_2N_6O_4Na_2$: C, 11.13; N, 38.39; Na, 21.30. Found: C, 11.01; H, 0.00; N, 38.53; Na, 21.10.

Preparation of The Monoguanidine Salt of 5-Dinitromethyltetrazole. - 5-Dinitromethyltetrazole (0.075 g.) and guanidine carbonate (0.098 g.) were dissolved in two ml. of water. The pH of the solution was adjusted to about 3 with concentrated hydrochloric acid. The light yellow, crystalline precipitate was filtered, after cooling to 5°, and recrystallized twice from water. It decomposed without melting at 164.0-164.5°.

Anal. Calcd. for $C_3H_6N_9O_4$: C, 15.52; H, 2.61; N, 54.30. Found: C, 15.56, 15.26; H, 3.57, 3.37; N, 54.41, 54.20.

Reaction of Ammonium Dinitroacetonitrile With Sodium Azide And Ammonium Chloride. - Ammonium dinitroacetonitrile (7.40 g., 0.050 mole), sodium azide (3.33 g., 0.051 mole), and ammonium chloride (2.75 g., 0.051 mole) in 25 ml. of water were refluxed 24 and 48 hr. Both reaction mixtures then were treated as described for the reactions with sodium dinitroacetonitrile except that the ether-insoluble solids were not extracted with ethyl acetate. Yields are summarized below:

Reflux Time hr.	Bitetrazole g. (%)	5-Dinitromethyltetrazole g. (%)	Recovered Ammonium Dinitroacetonitrile g. (%)
24	0.68 (8.0)	1.07 (12.3) from 2.1 g. crude	5.37 (71.8)
48	1.80 (21.9)	0.80 (9.2) from 2.7 g. crude	2.64 (35.2)

A reaction using 0.05 mole of ammonium dinitroacetonitrile, 0.051 mole of sodium azide and 0.0051 mole of ammonium chloride (10 mole % of the sodium azide) in 25 ml. of water refluxed 24 hours gave 0.34 g. (4.0%) of bitetrazole and 0.2 g. (2.3%) of crude 5-dinitromethyltetrazole.

After 24 hr. of refluxing, the reaction mixture gave a positive test for the presence in large concentration, of cyanide ion

EIDBERG

with benzidine-copper acetate reagent.²⁰

Elementary analyses were by Schwarzkopf Microanalytical Laboratory, Woodside, N. Y.

Acknowledgement

The author gratefully acknowledges the many helpful suggestions made by Dr. G. P. Sollott of this laboratory in reviewing this paper, and helpful discussions with Drs. R. A. Henry and W. G. Finnegan of the Naval Ordnance Test Station.

Footnotes

1. W. G. Finnegan, R. A. Henry, and R. Lofquist, J. Am. Chem. Soc., 80, 3908 (1958).
2. W. P. Norris, J. Org. Chem. 27, 3248 (1962).
3. Trimethyl ammonium cyanide has been used to prepare tetrazole.¹
4. Electronegative groups attached to the nitrile facilitates tetrazole formation.¹
5. C. O. Parker, W. D. Emmons, A. S. Pagano, H. A. Rolewicz and K. S. McCallum, Tetrahedron 17 89, (1962).
6. C. O. Parker, W. D. Emmons, H. A. Rolewicz, and K. S. McCallum, Ibid., 79 (1962).
7. R. A. Braun and W. A. Mosher, J. Am. Chem. Soc. 80, 4919 (1958).
8. E. Lieber, D. R. Levering, and L. J. Patterson, Anal. Chem. 23, 1594 (1951).
9. J. F. Brown, J. Am. Chem. Soc. 77, 6341 (1955).
10. A similar shift is shown by α,α -dinitronitriles compared with salts of 1,1-dinitroparaffins.¹¹
11. L. W. Kissinger and H. E. Ungrade, J. Org. Chem. 25, 1471 (1960).
12. M. J. Kamlet and J. C. Dacons, J. Org. Chem. 26, 1471 (1960).
13. All melting points are uncorrected.
14. Obtained from the Rohm and Haas Company, Redstone Arsenal Research Division.
15. Sodium dinitroacetonitrile absorbs moisture from the atmosphere and forms a stable monohydrate.⁵

EINERBO

16. Highly alkaline fumes escaped from the condenser after approximately one half hour of refluxing. The absorbed fumes in dilute hydrochloric acid gave a negative test for ammonium ion with Nessler's reagent.
17. F. R. Benson, Chem. Revs. 41, 1 (1947).
18. A. Weissberger, "Physical Methods of Organic Chemistry", Volume I, Part II, Interscience Publishers, Inc., New York, 2nd ed., 1946, pp. 1747-1748.
19. Salts of dinitroacetonitrile are recovered unchanged from concentrated hydrochloric acid and, therefore, were recovered from the reaction mixture after adjusting the pH to less than 1.
20. N. D. Cheronis and J. B. Entrikin, "Semimicro Qualitative Organic Analysis", Interscience Publishers, Inc., New York, 2nd ed. 1957, pp. 174-175.

MODAL ANALYSIS OF TRANSIENT VIBRATION PROBLEMS
IN LINEARLY VISCOELASTIC SOLIDSALEXANDER S. ELDER
USA BALLISTIC RESEARCH LABORATORIES
ABERDEEN PROVING GROUND, MARYLAND

INTRODUCTION

Spurred by the urgent need for analyzing the stresses in solid propellant rocket motors, in order to provide a more rational basis for design, various investigators have developed general methods of analysis based on the similarities exhibited by the basic equations in linear elasticity and viscoelasticity. E. H. Lee and J. R. M. Radok (1) developed a correspondence principle for quasi-static problems. A more general approach based on the classical stress functions for an elastic solid has been suggested by the author (2), and also by Gurtin and Sternberg (3). A new correspondence principle established by D. S. Berry (4) provides a general method for solving steady state vibration problems.

Formal solutions of transient vibration problems may be obtained by means of the Laplace transform, but inversion of the resulting transform by analytical methods has not been possible in the general case (5). D. R. Bland has given solutions in terms of one-sided Fourier transforms. Evaluation of these transforms by numerical methods is difficult because the integrand is highly oscillatory. E. Volterra used separation of variables in an early and important paper (6), but his treatment of initial and boundary conditions is not entirely satisfactory. In this paper, it is shown that his method is valid only if the boundary conditions can be reduced to a form that does not involve the viscoelastic properties of the material. If this condition is satisfied, the boundary conditions will generally determine a sequence of real characteristic numbers, and Sturm-Liouville methods may be used to obtain the complete solution. Cauchy contour integrals provide a more general type of solution when the Sturm-Liouville theory does not apply.

THEORY

The Basic Equations. The basic equations are expressed in spherical coordinates r, θ, α . This coordinate system is used as the vehicle for developing the theory because a variety of displacement fields may be expressed in terms of Bessel functions of half-integral order. Computation of these functions for both real and complex arguments may be carried out in terms of elementary functions. The equations of motion are (?)

$$\begin{aligned} \frac{\partial \sigma_r}{\partial r} + \frac{1}{r \sin \theta} \frac{\partial \tau_{r\alpha}}{\partial r} + \frac{1}{r} \frac{\partial \tau_{r\theta}}{\partial r} + \frac{2\sigma_r - \sigma_\alpha - \sigma_\theta + \tau_{r\theta} \cot \theta}{r} &= \rho \frac{\partial^2 u}{\partial t^2} \\ \frac{\partial \tau_{r\alpha}}{\partial r} + \frac{1}{r \sin \theta} \frac{\partial \sigma_\alpha}{\partial \alpha} + \frac{1}{r} \frac{\partial \tau_{\alpha\theta}}{\partial \theta} + \frac{3\tau_{r\alpha} + 2\tau_{\alpha\theta} \cot \theta}{r} &= \rho \frac{\partial^2 v}{\partial t^2} \quad (1) \\ \frac{\partial \tau_{r\theta}}{\partial r} + \frac{1}{r \sin \theta} \frac{\partial \tau_{\alpha\theta}}{\partial \alpha} + \frac{1}{r} \frac{\partial \sigma_\theta}{\partial \theta} + \frac{3\tau_{r\theta} + (\sigma_\theta - \sigma_\alpha) \cot \theta}{r} &= \rho \frac{\partial^2 v}{\partial t^2} \end{aligned}$$

The displacements and strains are related by the equations

$$\begin{aligned} \epsilon_r &= \frac{\partial u}{\partial r} \\ \epsilon_\theta &= \frac{1}{r} \frac{\partial v}{\partial \theta} + \frac{u}{r} \\ \epsilon_\alpha &= \frac{1}{r \sin \theta} \frac{\partial w}{\partial \alpha} + \frac{u}{r} + \frac{v \cot \theta}{r} \\ \gamma_{r\alpha} &= \frac{1}{r \sin \theta} \frac{\partial u}{\partial \alpha} - \frac{w}{r} + \frac{\partial v}{\partial r} \\ \gamma_{r\theta} &= \frac{1}{r} \frac{\partial u}{\partial \theta} - \frac{v}{r} + \frac{\partial w}{\partial r} \\ \gamma_{\alpha\theta} &= \frac{1}{r} \frac{\partial w}{\partial \theta} - \frac{w \cot \theta}{r} + \frac{1}{r \sin \theta} \frac{\partial v}{\partial \alpha} \end{aligned} \quad (2)$$

Two stress-strain laws are required to express the mechanical behavior of an isotropic, linearly viscoelastic solid. Let $G(t)$ be the response to a unit step function of strain applied at $t = 0$; then the response to an arbitrary strain is given by

$$\tau(t) = \gamma(t) G(0+) + \int_0^t \gamma(t_1) G'(t - t_1) dt_1 \quad (3)$$

ELDER

provided the stresses and strains are zero prior to $t = 0$ (8). The term $\gamma(t) G(0+)$ represents the purely elastic response; the hereditary integral represents the delayed response. A functional notation of V. Volterra will be used to abbreviate the right-hand side of Eq. (3).

$$G[\gamma(t)] = \gamma(t) G(0+) + \int_0^t \gamma(t_1) G'(t - t_1) dt_1 \quad (4)$$

The stress-strain laws involving the shear modulus are

$$\begin{aligned} \gamma_{r\theta} &= G [\gamma_{r\theta}(t)] & \sigma_r - \sigma_\theta &= 2G [\epsilon_r(t) - \epsilon_\theta(t)] \\ \gamma_{r\alpha} &= G [\gamma_{r\alpha}(t)] & \sigma_\theta - \sigma_\alpha &= 2G [\epsilon_\theta(t) - \epsilon_\alpha(t)] \\ \tau_{\alpha r} &= G [\gamma_{\alpha r}(t)] & \sigma_\alpha - \sigma_r &= 2G [\epsilon_\alpha(t) - \epsilon_r(t)] \end{aligned} \quad (5)$$

The bulk modulus $K(t)$ is a measure of the hydrostatic pressure produced by an infinitesimal change in volume.

$$\Theta = 3K [e(t)]$$

$$\text{where } \Theta = \sigma_r + \sigma_\theta + \sigma_\alpha, \quad e = \epsilon_r + \epsilon_\theta + \epsilon_\alpha.$$

Applications of the Sturm-Liouville Theory. Consider a rod bounded by the conical surface $\Theta = \Theta_0$ and the spherical surfaces $r = a$ and $r = b$, where $b > a$. We assume the rod is fixed at $r = a$. At $r = b$ it is attached to a rotating disc, as shown in Fig. 1. The motion of the system produced by a step function of torque applied at the disc is required. The solution is based on H. Lamb's analysis of free torsional vibrations of a solid elastic sphere (9). In the solution considered here, concentric spherical surfaces rotate about the axis $\Theta = 0$ without distortion. We assume that

$$u = 0, \quad v = 0, \quad w = (r \sin \Theta) \psi(r, t) \quad (6)$$

where $\psi(r, t)$ is the angle of rotation. Then

$$\gamma_{r\alpha} = (r \sin \Theta) \frac{\partial \psi}{\partial r}, \quad \tau_{r\alpha} = (r \sin \Theta) G \left[\frac{\partial \psi}{\partial r} \right] \quad (7)$$

The remaining stresses and strains are zero. The first and third equations of motion are satisfied identically; the second equation reduces to

$$G \left[\frac{\partial^2 \psi}{\partial r^2} + \frac{4}{r} \frac{\partial \psi}{\partial r} \right] = \rho \frac{\partial^2 \psi}{\partial t^2} \quad (8)$$

ELDER

Separable solutions suitable for analyzing standing waves are obtained by assuming that

$$\psi(r, t) = R(r) T(t) \quad (9)$$

Then

$$T''(t) + \Omega^2 G [T(t)] = 0, \quad rR''(r) + 4R' + \lambda^2 rR = 0 \quad (10a), (10b)$$

where

$$\lambda^2 = \rho \Omega^2$$

The twisting moment $M(r, t)$ is required in the analysis of the boundary condition at $r = b$. We have

$$\Delta M(r, t) = (r \sin \theta) (\tau_{r\alpha} \Delta S)$$

where $r \sin \theta$ is the moment arm and ΔS is an element of surface area. We also have

$$\Delta S = (r^2 \sin \theta) \Delta \theta \Delta \alpha, \quad \tau_{r\alpha} = (r \sin \theta) G \left[\frac{\partial \psi}{\partial r} \right]$$

We obtain the formula

$$M(r, t) = f(\theta_0) r^4 G \left[\frac{\partial \psi}{\partial r} \right] \quad (11)$$

where

$$f(\theta_0) = \frac{2\pi}{3} (2 - 3 \cos \theta_0 + \cos^3 \theta_0)$$

We consider free vibrations first. The boundary conditions at the stress-free conical surface $\theta = \theta_0$ are satisfied since $\sigma_\theta = \tau_{\theta r} = \tau_{\theta \alpha} = 0$. The bar is fixed at $r = a$, so that

$$\psi = 0 \text{ at } r = a \quad (12)$$

The boundary condition at $r = b$ may be regarded as the equation of motion of the disc.

$$J_d \frac{\partial^2 \psi}{\partial t^2} + M(r, t) = 0 \text{ at } r = b \quad (13)$$

In Eq. (9), we set

$$R(\lambda, r) = C_1(\lambda r) S_1(\lambda a) - S_1(\lambda r) C_1(\lambda b),$$

where the functions $S_n(x)$ and $C_n(x)$ are given by (14).

$$S_n(x) = \left(\frac{1}{x} \frac{d}{dx} \right)^n \left(\frac{\sin x}{x} \right), \quad C_n(x) = \left(\frac{1}{x} \frac{d}{dx} \right)^n \left(\frac{\cos x}{x} \right) \quad (14)$$

On combining these results, we obtain the characteristic equation

$$\Delta(\lambda) = 0 \quad (15)$$

where

$$\Delta(\lambda) = \ell \lambda^2 R(\lambda, b) - \frac{\partial R(\lambda, b)}{\partial r}, \quad \ell = J_d / \rho b^4 r(\theta_0)$$

This characteristic equation involves only the inertial and geometric properties of the system, and not the viscoelastic properties of the rod.

Eq. (15) is satisfied by a sequence of real characteristic roots $\lambda_1, \lambda_2 \dots \lambda_n$. The functions $r^2 R(\lambda_n, r)$ form an orthogonal set. We can prove that

$$\int_a^b r^4 R(\lambda_n, r) R(\lambda_m, r) dr + \ell b^4 R(\lambda_n, b) R(\lambda_m, b) = 0, \quad n \neq m \quad (16)$$

According to the Sturm-Liouville theory (11,12), a suitably restricted function may be represented by a generalized Fourier series of the

form $\sum_{n=1}^{\infty} c_n R(\lambda_n, r)$. The coefficients are calculated by a generalization of Fourier's rule. Consider, for instance, the function

$R_0(r) = \frac{1}{3}(1 - \frac{a^3}{r^3})$, which is a particular solution of Eq. (8). We assume that

$$R_0(r) = \sum_1^{\infty} c_n R(\lambda_n, r) \quad (17)$$

where

$$c_n = \frac{\int_a^b r^4 R_0(r) R(\lambda_n, r) dr + \ell b^4 R_0(b) R(\lambda_n, b)}{\int_a^b r^4 R^2(\lambda_n, r) dr + \ell b^4 R^2(\lambda_n, b)} \quad (18)$$

The numerator may be evaluated by multiplying Eq. (10b) by $R_0(r)$ and integrating; a limiting process described by Ford (13) is used for the denominator. The results are further simplified by using the boundary conditions satisfied by $R_0(r)$ and $R(\lambda_n, r)$. We obtain finally

$$c_n = \frac{2a^3}{b^3} \cdot \frac{1}{\{\lambda \Delta'(\lambda)\}_{\lambda=\lambda_n}} \quad (19)$$

ELDER

The normal coordinates $T_n(t)$ are obtained from Eq. (10) by means of the Laplace transform. We assume that $T(0+) = 1$, $T'(0+) = 0$; then

$$\bar{T}(s) = \frac{1}{s + \Omega^2 \bar{U}(s)} \quad (20)$$

The forced vibrations produced by a torque which is proportional to $G(t)$ may be expressed in terms of the normal modes and normal coordinates derived above. The response to a step function of torque is then obtained by solving a Volterra integral equation. Finally, the response to an arbitrary torque may be calculated by using Duhamel's integral, since the mechanical system under consideration is linear. Consider the series

$$\psi(r, t) = R_0(r) - \sum_{n=1}^{\infty} c_n R(\lambda_n, r) T_n(t) \quad (21)$$

which formally satisfies Eq. (8) and the initial conditions

$$\psi(r, t) = 0, \quad \frac{\partial \psi(r, t)}{\partial t} = 0, \quad \text{when } t = 0+, \quad a \leq r \leq b$$

The torque is obtained from Eq. (11). We find after some reduction that

$$\frac{\partial^2 \psi}{\partial t^2} + M(b, t) = M_0(t) = a^3 f(\theta_0) G(t) \quad (22)$$

Let $\phi(r, t)$ be the response to a step function of torque. Then the response to the torque $M_0(t)$ is given by Duhamel's integral:

$$\psi(r, t) = M_0(t) \phi(r, 0+) + \int_0^t M_0(t_1) \frac{\partial}{\partial t_1} \phi(t - t_1) dt_1 \quad (23)$$

As the function ψ is known and ϕ is unknown, this equation will be regarded as a Volterra integral equation for ϕ . To solve, write $\psi(r, t)$ in the alternate form

$$\psi(r, t) = \sum_{n=1}^{\infty} c_n R(\lambda_n, r) \{1 - T_n(t)\}, \quad (24)$$

calculate the Laplace transform of the last four equations, and then solve these transformed equations for $\bar{\psi}(r, s)$. We obtain

$$\bar{\psi}(r, s) = \sum_{n=1}^{\infty} \frac{\Omega_n^2 c_n \bar{T}(s) R(\lambda_n, r)}{a^3 f(\theta_0) s^2} \quad (25)$$

If a spring-dashpot model is used to represent the viscoelastic properties of the rod, the stress-strain laws may be

represented in differential operator form, and each term of the above series will be a rational function in the transform variable s . We have

$$\sum_{j=0}^m p_j \frac{d^j x}{dt^j} = \sum_{j=0}^m q_j \frac{d^j y}{dt^j} \quad (26)$$

The Laplace transform of $G(t)$ is $q(s)/sp(s)$ where

$$p(s) = \sum_{j=0}^m p_j s^j, \quad q(s) = \sum_{j=0}^m q_j s^j$$

Consequently

$$\bar{T}(s) = \frac{sp(s)}{s^2 p(s) + \Omega^2 q(s)} \quad (27)$$

If $q_0 > 0$, the roots of $p(s) = 0$ and $q(s) = 0$ are negative, real, and simple, and are interlaced along the negative real axis. If $q_0 = 0$, then $q(s)$ has a simple root at the origin. The poles and zeros of the transfer function $h(s) = p(s)/q(s)$ alternate along the negative real axis, as shown in Fig. 2. The equation

$$s^2 p(s) + \Omega^2 q(s) = 0 \quad (28)$$

has $m + 2$ roots since the polynomials $p(s)$ and $q(s)$ are of degree m . This equation may be written in the form

$$h(s) = -\Omega^2/s^2$$

A total of m real, simple roots are obtained at the intersections of the curves $y = h(s)$ and $y = -\Omega^2/s^2$; the remaining two roots are complex. We can prove that the real part of these complex roots is negative, as would be expected on physical grounds. The response $\phi(r, t)$ is obtained by means of the Heaviside inversion formula. We find that

$$\phi(r, t) = \sum_1^{\infty} R(\lambda_n, r) F_n(t)$$

where

$$F_n(t) = D_0 + D_1 t + D_2 e^{\mu_1 t} \cos \omega_1 t + D_3 e^{\mu_1 t} \sin \omega_1 t + \sum_{i=2}^{m+2} D_{i+2} e^{\mu_i t} \quad (29)$$

If $q_0 = 0$, $F(s)$ has a double pole at the origin, and $D_1 > 0$; if $q_0 > 0$, this pole is simple, and $D_1 = 0$.

The response to an arbitrary torque may now be found by a direct application of Eq. (23), since ϕ is now known.

The Sturm-Liouville theory may be applied to any dynamic problem in which all the characteristic numbers are real and the characteristic functions form a complete orthogonal set. Consider for instance the vibrations in a hollow viscoelastic sphere in which the outer surface is bonded to a rigid case and the inner surface is subjected to an arbitrary radial displacement. We assume that v and w vanish, and u is independent of θ and α . Then

$$\epsilon_r = \frac{\partial u}{\partial r}, \quad \epsilon_\theta = \epsilon_\alpha = \frac{u}{r}, \quad \gamma_{r\theta} = \gamma_{\theta\theta} = \gamma_{r\alpha} = 0 \quad (30)$$

The equation of motion in terms of the radial displacement is

$$\left(K + \frac{4}{3}G\right) \left[\frac{\partial^2 u}{\partial r^2} + \frac{2}{r} \frac{\partial u}{\partial r} - \frac{2u}{r^2}\right] = \rho \frac{\partial^2 u}{\partial t^2} \quad (31)$$

We assume a separable solution of the form

$$u(r, t) = U(r) T(t)$$

Then

$$r^2 U'' + 2rU' + (\lambda^2 r^2 - 2)U = 0, \quad T'' + \sigma^2 \left(K + \frac{4}{3}G\right) [T(t)] = 0 \quad \begin{matrix} (32a) \\ (32b) \end{matrix}$$

The first equation has the general solution $U(\lambda, r) = rR(\lambda, r)$. We assume that $U(r, t) = 0$ when $r = a$ and $r = b$. Then

$$U(\lambda, r) = C_1(\lambda b) rS_1(\lambda r) - S_1(\lambda b) rC_1(\lambda r) \quad (33)$$

Let

$$\Delta(\lambda) = U(\lambda, a); \text{ then } \Delta(\lambda) = 0$$

is the characteristic equation for these boundary conditions. The functions $rU(\lambda_n, r)$ form a complete orthogonal set if the λ_n are roots of this characteristic equation.

The transient vibrations produced by the radial displacement of a boundary can be expressed in terms of these normal modes and the function $U_0(r)$ defined below. Assume that

$$u(r, t) = 0, \quad \frac{\partial u(r, t)}{\partial t} = 0, \quad t < 0, \quad a \leq r \leq b$$

$$u(r, t) = 1, \quad t > 0, \quad r = a$$

The function $U_0(r) = a^2(b^3 - r^3)/r^2(b^3 - a^3)$ is a particular solution of Eq. (31) and satisfies the boundary conditions $U_0(b) = 0$, $U_0(a) = 1$. It may be expanded in series

$$U_0(r) = \sum_{n=1}^{\infty} c_n U(\lambda_n, r)$$

ELDER

where

$$c_n = \frac{\int_a^b r^2 U_0(r) U(\lambda_n, r) dr}{\int_a^b r^2 U^2(\lambda_n, r) dr} = \frac{2U_0(a)}{\lambda_n \Delta'(\lambda_n)}$$

after evaluation by the limiting process described previously. The displacement $u(r, t)$ is given by either of the series

$$u(r, t) = U_0(r) - \sum_{n=1}^{\infty} c_n U(\lambda_n, r) T_n(t) = \sum_{n=1}^{\infty} c_n U(\lambda_n, r) F_n(t) \quad (34)$$

where

$$F_n(t) = 1 - T_n(t), \quad \bar{F}_n(s) = \frac{1}{s} - T_n(s) \quad (35)$$

The function $F_n(t)$ is obtained by means of the Laplace transform. We find

$$\bar{T}(s) = \frac{1}{s + \Omega^2 \left\{ \bar{K}(s) + \frac{4}{3} \bar{G}(s) \right\}} = \frac{s r''(s)}{s^2 p''(s) + \Omega^2 q''(s)}$$

where $q''(s)/sp''(s)$ is the Laplace transform of $K(t) + \frac{4}{3} G(t)$. On combining the preceding results, we find

$$\bar{F}_n(s) = \frac{\Omega_n^2 q''(s)}{s^2 p''(s) + \Omega_n^2 q''(s)} \quad (36)$$

The function $F(t)$ may be evaluated by Heaviside's expansion formula. The pole at the origin is simple, so that the linear term is absent.

Application of Contour Integration. The Sturm-Liouville method used in the preceding analysis is effective only if separation of variables in the boundary conditions leads to a characteristic equation with real roots only. The characteristic equation must not involve the normal coordinate $T(t)$ or the transform variable s . If the boundary conditions are homogeneous and linear, but do not satisfy the above conditions, we assume separable solutions with $\exp(st)$ as the time-dependent factor. A frequency equation for real and complex values of s is obtained. In general, the associated modes of displacement do not form an orthogonal set, so that coefficients in the expansion of an arbitrary function cannot be calculated by Fourier's rule. A formal expansion based on contour integration is suggested. Formal solution of two dynamic problems are obtained by this method.

If we add a torsional spring to the torsional pendulum, as indicated in Fig. 1, the equation of motion of the disc becomes

ELDER

$$J_d \frac{\partial^2 \psi}{\partial t^2} + k\psi + M(r, t) = 0 \text{ at } r = b \quad (37)$$

We assume that $\psi(r, t) = R(\lambda, r) T(t)$. We find after some reduction that

$$\Delta(\lambda) T''(t) + \beta \lambda^2 R(\lambda, b) T(t) = 0 \quad (38)$$

where $\beta = k/J_d$. Clearly we cannot separate variables in this equation if $T(t)$ is arbitrary and $\beta \neq 0$. We obtain a complex frequency equation by assuming that $T(t) = \exp(st)$; then

$$s^2 = -\lambda^2 \beta R(\lambda, b) / \Delta(\lambda) \quad (39)$$

The right-hand side of this equation is an even function of λ . The variables λ and s are also connected by the relation

$$\lambda^2 = -\frac{\rho s^2 p(s)}{q(s)} \quad (40)$$

The real roots satisfying these equation may be located by finding the intersection of curves defined by Eqs. (39) and (40), plotted in the λ, s plane.

As the functions $r^2 R(\lambda, r)$ no longer form an orthogonal set, a more general procedure is required in order to find the response to a step function of torque. Assume that

$$\begin{aligned} J_d \frac{\partial^2 \phi}{\partial t^2} + k\phi + M(r, t) &= 1, & t > 0, & r = b \\ &= 0, & t < 0, & r = b \end{aligned} \quad (41)$$

Then

$$J_d s^2 \bar{\phi} + k\bar{\phi} + \bar{M}(b, s) = \frac{1}{s} \quad (42)$$

Assume that $\bar{\phi} = z(s) R[\lambda(s), r]$, and solve for $z(s)$. Then evaluate $\phi(r, t)$ formally by the inversion integral. We obtain

$$\phi(r, t) = \frac{kJ_d}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{R[\lambda(s), r] e^{st} ds}{s^3 \Delta(\lambda) - \beta s \lambda^2(s) R[\lambda(s), b]} \quad (43)$$

This expression may be reduced to a Fourier integral by a change of variable, and evaluated numerically for various values of r and t .

ELDER

However, this form of the solution does not exhibit the damped harmonic nature of the solution. A formal solution of the form

$$\phi(r, t) = F(r) + \sum_1^{\infty} c_1 R \{ \lambda(s_1), r \} e^{s_1 t} \quad (44)$$

may be obtained from the Heaviside expansion formula, provided the roots satisfying both Eqs. (39) and (40) can be located. The practical difficulty of locating these roots may limit the usefulness of this method.

The occurrence of two distinct time-dependent moduli in the boundary conditions also prevents separation of variables for arbitrary $T(t)$. Consider the radial vibrations of a sphere subject to the boundary conditions

$$u(r, t) = 0, \quad r = b; \quad \sigma_r(r, t) = 0, \quad r = a \quad (45)$$

We assume $u(r, t) = U(\lambda, r) T(t)$, where $U(\lambda, r)$ is given by Eq. (33). The radial stress is calculated from Eqs. (2), (6), and (7). We obtain

$$\sigma_r = \frac{4}{3} \left\{ \frac{\partial U}{\partial r} - \frac{U}{r} \right\} G[T(t)] + \left\{ \frac{\partial U}{\partial r} + \frac{2U}{r} \right\} K[T(t)] \quad (46)$$

We cannot satisfy the boundary condition $\sigma_r = 0$ at $r = a$ for all values of t unless $K(t)/G(t)$ is constant. E. Volterra apparently overlooks this point in his analysis of a related problem. In formulating the characteristic equation,* he neglects the hereditary portions of $G(t)$ and $K(t)$ in calculating σ_r , so that the results are exact only at $t = 0+$. In general, if a viscoelastic material is in the rubbery state, $G(t)$ decreases more rapidly than $K(t)$ as t increases, and consequently $G(t)/K(t) \neq G(0+)/K(0+)$ when t is large.

The response to a step function of pressure may be expressed as an inversion integral. Assume that

$$u(r, t) = 0, \quad r = b \quad (47)$$

$$\begin{aligned} \sigma(r, t) &= 0, \quad t > 0, \quad r = a \\ &= 1, \quad t > 0, \quad r = a \end{aligned} \quad (48)$$

Then

$$\sigma_r(s, r) = \frac{1}{s} \quad \text{when } r = a.$$

*See Eq. (25) of his paper and the unnumbered equation which precedes it.

ELDER

We find after some manipulation that

$$2\pi i \sigma_r(r, t) = \int_{c-i\infty}^{c+i\infty} \frac{z(\lambda, r)}{z(\lambda, b)} \cdot \frac{e^{st}}{s} ds \quad (49)$$

where

$$z(\lambda, r) = 4 \left\{ \frac{\partial U(\lambda, r)}{\partial r} - \frac{U(\lambda, r)}{r} \right\} \bar{G}(s) + 3 \left\{ \frac{\partial U(\lambda, r)}{\partial r} + \frac{2U(\lambda, r)}{r} \right\} \bar{K}(s) \quad (50)$$

Eq. (40) shows that λ^2 is a single-valued function of s . The radial stress may be expressed as a bilinear series of the type given in Eq. (44); however, the calculation of the coefficients and characteristic frequencies would be difficult.

Status of the Calculations. Calculations are in progress for the two problems solved formally by the Sturm-Liouville theory. The calculations are based on the dimensions given in Fig. 1. N.B.S. polyisobutylene is used as a representative viscoelastic material. Viscoelastic model constants were calculated by Capt. W. Goldberg by the method indicated in Ref. (14). These model constants are given in Ref. (15). The real and complex roots of Eq. (28) have been calculated by the method given in Ref. (16). The functions $F^n(t)$ are being calculated by means of the Heaviside expansion formula.ⁿ

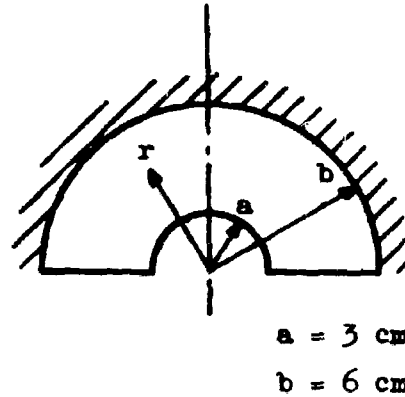
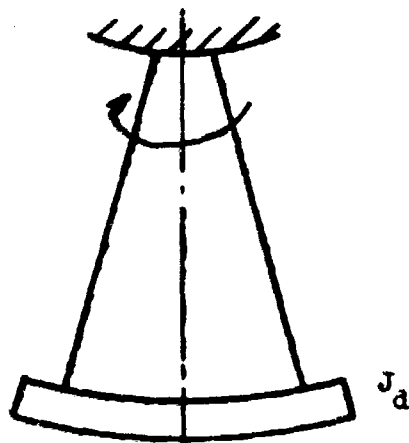
Conclusions. Transient vibration problems in linear viscoelasticity may be solved by means of the Sturm-Liouville theory provided the homogeneous boundary conditions for the free vibrations can be combined to give a characteristic equation with real roots. This characteristic equation will depend only on the geometric and inertial characteristics of the system, and not on the viscoelastic properties of the solid.

If all the conditions of the preceding paragraph are not satisfied, it is still generally possible to reduce the homogeneous boundary conditions to a complex frequency equation by assuming separable solutions with $\exp(st)$ as the time dependent factor. The associated modes of deflection will generally be non-orthogonal. Transient vibration problems of this type may be solved by contour integrals, which may be evaluated by residue theory or reduced to a Fourier integral by a change of variable.

Acknowledgment. The calculations have been programmed for the BRLECC Computer by Mrs. E. Wineholt and Mrs. A. DePue.

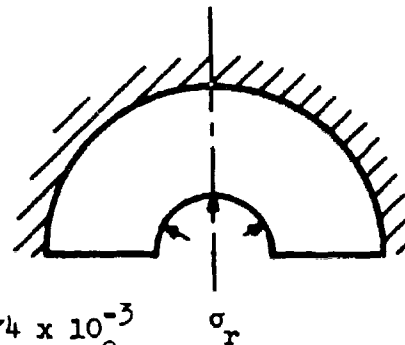
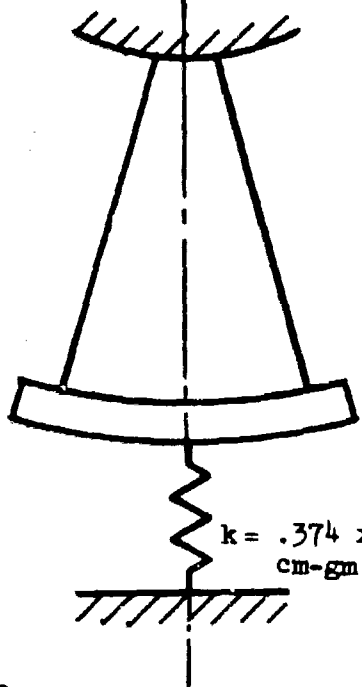
References.

1. Lee, E. H., Radok, J. R. M., and Woodward, W. B. Stress Analysis for Linear Viscoelastic Materials. Brown University, Sep 59
2. Elder, A. S. General Solution of Certain Integro-Differential Equations of Linear Viscoelasticity. 20th Meeting Bulletin, Physical Properties Panel, SPIA, Oct 61
3. Sternberg, E. and Gurtin, M. E. On the Linear Theory of Viscoelasticity. Brown University, June 62.
4. Barry, D. S. Stress Propagation in Visco-Elastic Solids. Journal of the Mechanics and Physics of Solids, Vol. 6, No. 3, May 58, Pergamon Press.
5. Hunter, S. C. Viscoelastic Waves, Chapter I in Progress in Solid Mechanics, Sneddon, I. N. and Hill, R., editors. Interscience Publishers, New York, 1960.
6. Volterra, E. On Elastic Continua with Hereditary Characteristics. Jour. App. Mech., Vol. 18, No. 3, Sep 51, pp 273-279.
7. Sokolnifoff, I. S. Mathematical Theory of Elasticity. McGraw-Hill Book Co., 1946.
8. Mandel, J. Sur les Corps Viscoelastiques a Comportement Linéaire. IX Congres International de Mecanique Appliquee, 1957, See Equation 2, p. 401, Tome V.
9. Lamb, H. London Math. Soc. Proc., Vol. 13, 1882.
10. Love, A. E. H. The Mathematical Theory of Elasticity. Fourth Edition, Dover Publications, New York, 1944.
11. Ince, E. L. Ordinary Differential Equations. Dover Publications, New York.
12. Churchill, R. V. Modern Operational Methods in Engineering. McGraw-Hill Book Company, 1944.
13. Ford, W. B. Studies on Divergent Series and Summability. Michigan Science Series, Vol. II, The Macmillan Company, 1916, p. 130.
14. Goldberg, W. and Dean, N. W. Determination of Viscoelastic Model Constants from Dynamic Mechanical Properties of Linear Viscoelastic Materials. BRL Memo Report No. 1180, Nov 62.
15. Elder, A. S. Derivation of Formulas for Calculating the Transient Response of a Viscoelastic Torsional Pendulum. Bulletin of the 2nd Meeting, Working Group on Mechanical Behavior, CPIA, Nov 63.
16. Leser, T. Routine for Finding Roots of Polynomials with Real Coefficients. BRL Memo Report No. 1467, Apr 63.



$$J_d \frac{\partial^2 \psi(b, t)}{\partial t^2} + M(b, t) = 1, \quad t > 0 \quad u(a, t) = 1, \quad t > 0$$

Sturm-Liouville Systems



$$J_d = .374 \times 10^{-3} \text{ cm-gm-sec}^2$$

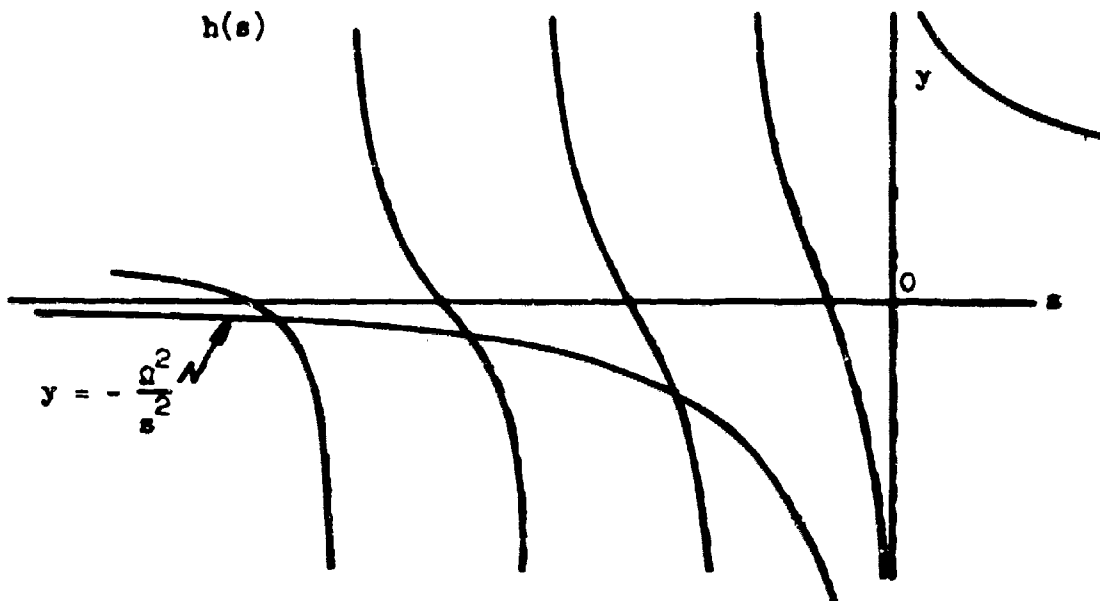
$$k = .374 \times 10^3 \text{ cm-gm}$$

$$\sigma_r(a, t) = 1, \quad t > 0$$

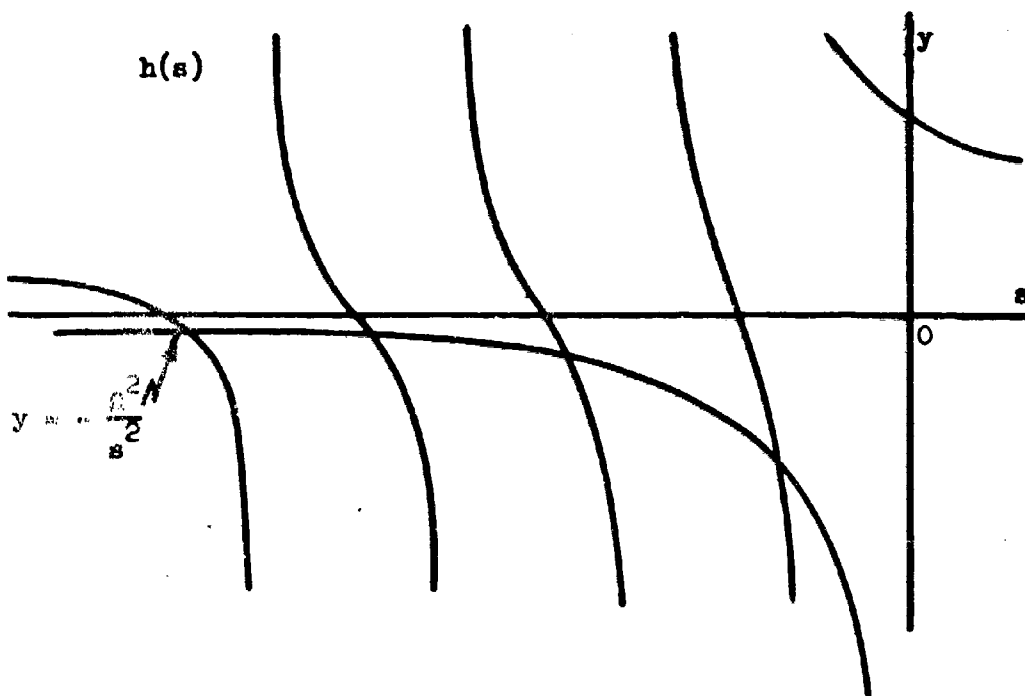
$$J_d \frac{\partial^2 \psi(b, t)}{\partial t^2} + k\psi(b, t) + M(b, t) = 1, \quad t > 0$$

Systems Solved by Contour Integrals

Figure 1. Mechanical Systems Considered in Text



Type 2 $q_0 = 0$



Type 1 $q_0 > 0$

Figure 2. Location of the Real Roots of $s^2 p(s) + \Omega^2 q(s) = 0$

THE RECOVERY OF THE ULTRASONIC ATTENUATION
IN COPPER SINGLE CRYSTALS
FOLLOWING SMALL PLASTIC DEFORMATION

JULIUS FRANKEL
WATERVLIET ARSENAL
WATERVLIET, NEW YORK

When a metal undergoes small plastic deformation, there is a change in its apparent Young's Modulus and internal friction. Both of these quantities recover following the deformation. (1,2)

It is the aim of this experiment to study the decrease of ultrasonic attenuation in copper single crystals following small plastic deformation.

THEORY:

Ultrasonic attenuation in metals is strongly affected by dislocation density and dislocation loop length. (3) Granato and Lucke (4) developed a theory based on a model by Koehler, which predicts that the strain independent decrement of an ultrasonic wave in a single crystal due to dislocations is:

1. (a)

$$\Delta_I = \frac{120 \Omega \Lambda B \omega L^4}{\pi^3 C}$$

and the change in modulus;

(b)

$$\left(\frac{\Delta E}{E}\right)_I = \frac{6 \Omega \Lambda L^2}{\pi^2}$$

E is Young's Modulus,

Δ_I is $\propto (.712/\omega)$, where α is the attenuation, Ω is the orientation factor, which takes into account that the shear stress on the slip plane is less than the applied longitudinal stress, Λ is the dislocation density, L is the average loop length, B is a damping constant, ω is the frequency, and C is the dislocation line tension.

L , the loop length, is the average length of the edge component of a dislocation network which lies between two pinning points.

The behavior of this length L towards an elastic wave of small amplitude traveling in the metal is that of the string vibrating in a viscous medium. The conditions imposed on such a string are:

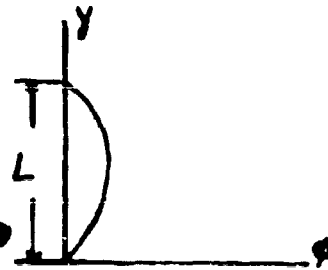
- (a) its end points are fixed;
- (b) it can only vibrate in slip planes.

FRANKEL

The equation for such a dislocation is:

2.
$$A \frac{\partial^2 \phi}{\partial t^2} + B \frac{\partial \phi}{\partial t} - C \frac{\partial^2 \phi}{\partial y^2} = a \sigma$$

where
$$\phi = \phi(y, t)$$



DISLOCATION LINE UNDER TENSION →

Here the plane of the paper is the slip plane, A is the effective mass per unit length of the string, a is the spacing between lattice points along dislocation, and σ is the force per unit length exerted in the slip plane on the dislocation by the external stresses.

Small plastic deformation of a metal is associated with dislocation formation and dislocation motion(5-7). This motion of dislocations and their interaction during plastic deformation, cause point defects in the single crystal. The point defect most predominately formed in f.c.c. metals is the vacancy.

It is assumed that these vacancies are the only mobile point defects in the lattice at the temperature considered.

A dislocation in the metal acts as a sink for vacancies and other impurities in the lattice. These impurities or point defects are said to pin the dislocation.

Let n = number of pins on a dislocation of length

$$n = \sum \text{vacancy pins} + \sum \text{impurity pins}$$

and L_N = network length

and \mathcal{C}_i = concentration of the i^{th} type pin on dislocation = $\frac{\text{number of pins}}{\text{number of lattice sites at dislocation}}$

then for n pins
$$L = \frac{L_N}{n+1}$$

$$\frac{1}{L} = \frac{1}{L_N} + \frac{n}{L_N} = \frac{1}{L_N} + \sum_{i=1}^n \frac{n_i}{L_N}$$

for k types of pins.

but $\mathcal{C}_i = n_i / \frac{L_N}{a}$

so that
$$\frac{1}{L} = \frac{1}{L_N} + \sum_{i=1}^k \frac{\mathcal{C}_i}{a}$$

3.
$$\frac{1}{L} = \frac{1}{L_N} + \frac{\mathcal{C}_v}{a} + \frac{\mathcal{C}_i}{a}$$

for only two types of pins in addition to network pins.

The recovery of the ultrasonic attenuation is based on the above equation since \mathcal{C}_i is a time dependent quantity. \mathcal{C}_v is the concentration of vacancies which have diffused to dislocations on the dislocations after plastic deformation, and \mathcal{C}_i is the impurity concentration at dislocation.

FRANKEL

ρ_d is given by the Cottrell-Bilby strain aging law written in terms of concentrations for f.c.c. metals.

4.
$$\rho_d(t) = \rho_{d0} \frac{4\alpha}{\alpha} \left(\frac{ADt}{kT} \right)^{2/3}$$

where ρ_{d0} is the concentration of deformation produced vacancies in lattice, α is a constant with value of about 3, A is a parameter measuring the strength of the Cottrell attraction, and D is $D_0 \exp\{-U/kT\}$ the diffusion coefficient where U is the activation energy for migration of the defect.

The equilibrium concentration of these point defects on the dislocations is given by the Cottrell relation.

5.
$$C_d = C_0 \exp\{Q/kT\}$$

for vacancies, C_0 is the vacancy concentration in lattice (total), C_d is the vacancy concentration on dislocation and Q is the interaction energy between the vacancy and dislocation.

Substituting equations 4 and 3 into equation 1 will obtain the following results:

6. (a)
$$\Delta L = \frac{A_1 \rho_d^2}{(\rho_{d0} + \rho_{i0})^2} \cdot \frac{1}{[1 + \beta t^{-1/3}]^2}$$

(b)
$$\left(\frac{\Delta L}{L} \right)^2 = \frac{A_2 \rho_d^4}{(\rho_{d0} + \rho_{i0})^4} \cdot \frac{1}{[1 + \beta t^{-1/3}]^2}$$

(c)
$$\beta = \frac{\rho_{d0}}{\rho_{d0} + \rho_{i0}} \cdot \frac{4\alpha}{\alpha} \left(\frac{AD_0}{kT} \right)^{2/3}$$

$$A_1 = \frac{6\Delta\Omega}{\pi^2} \quad A_2 = \frac{120\Delta\Omega\beta\omega}{\pi^2 C}$$

and ρ_{i0} is the impurity concentration in lattice.

Equation 6 (c) predicts that a graph of $\log \beta$ vs. $1/kT$ gives a straight line whose slope equals $-\frac{2}{3}U$, where U equals the activation energy of migration of the deformation caused point defect.

7.

$$\log \beta = \log \frac{\rho_{d0}}{\rho_{d0} + \rho_{i0}} \cdot \frac{4\alpha}{\alpha} + \frac{2}{3} \log \frac{AD_0}{kT} - \frac{2}{3} \frac{U}{kT}$$

TABLE I

T °K	β calc (min) ^{-2/3}	β Smith (min) ^{-2/3}	t _{1/2} min
373	10 ⁻¹	1.6x10 ⁻¹	2.6
353	27.7x10 ⁻³	41x10 ⁻³	18.5
340	14.3x10 ⁻³	20x10 ⁻³	48.5
325	5.91x10 ⁻³	8.2x10 ⁻³	180
320	3.08x10 ⁻³	4.4x10 ⁻³	500

t_{1/2} is the time required for the attenuation to decrease by a factor of two.

β calc from Eq. 6c for samples 99.9% pure, deformed about 1%. $\epsilon_{10} = 10^{-2}$, $\epsilon_{20} = 10^{-3}$,
 $A = 10^{-21}$ dyn/cm²,
 $D_0 = 10^{-2}$ cm²/sec, and
 $U = 1\text{ev} = 1.6 \times 10^{-12}$ erg.
 values were assumed.

β Smith - Values listed in Ref. 2, obtained by curve fitting from Smith's data, taken at KC frequencies (resonant bar method) for specimens of same purity as calculated above.

DISCUSSION AND BACKGROUND:

In copper the only recovery measurements available in which the temperature was varied systematically are from a paper by A.D.N. Smith⁽⁸⁾. Smith measured the modulus after deformation as a function of time and temperature, for specimens of two purities and two different deformations, using the resonant bar method.

Table I shows the values of β obtained in the Grnato, Hikata and Lücke paper from Smith's experiment, for the 99.98 purity specimens deformed 1%. These values were plotted on a $\log \beta$ vs. $1/RT$ plot to give an activation energy of 1 e.v. for the migration of the deformation caused defect. The calculated values are in all cases greater than the listed ones by a factor of approximately 1.5. β was calculated so that the time required for a substantial recovery in the attenuation may be estimated at various temperatures.

Two mentions were found in the literature where the recovery of the attenuation or modulus was seen in an experiment using ultrasonics; one is by Ying and Truell and is given as private communications in Reference 2. It is mentioned there that the change in the ultrasonic attenuation following extremely small plastic deformation has a finite slope at $t = 0$, but the authors did not expect that many point defects were produced during this small deformation.

The other mention was in Reference 1, where Alers and Thompson note but do not measure the recovery of the Young's Modulus and attenuation following a 1% permanent deformation.

EXPERIMENTAL DISCUSSION:

1. Electronics

The ultrasonic pulse method is used in this experiment (Fig. 6). A pulse modulated R.F. signal is applied to a piezoelectric crystal mounted on a test specimen to set up sound pulses which reflect back and forth inside the specimen. The amplitude of these sound pulses is converted back into electrical energy by the crystal and applied to the input of an R.F. amplifier for observation of time

FRANKEL

and amplitude. The insertion loss can be taken visually as the difference in the readings giving equal heights of the delayed and direct signals on the C.R.O. The pulse oscillator used here has a 0-600 volt peak to peak output and covers .5 to 110 M.C.

There are many difficulties involved with measurements of this type. The major ones are: nonlinear amplification in receiver; attenuation which is R.F. pulse amplitude sensitive; and overloading the receiver by R.F. pulse so that its circuits will not recover during the early part of the echo train.

The nonlinear amplification of the receiver is side-stepped by monitoring one echo of the echo train continuously in time, on a strip chart recorder. This pulse is calibrated in terms of db. losses every minute so that not only a continuous reading is available on a chart but also a complete calibration. This pulse is monitored by means of a phantastron type circuit, consisting of two one-shot multivibrators (o.s.m.v.). The pulsed oscillator which triggers the oscilloscope also triggers the first o.s.m.v. which provides a pulse which then can trigger the second o.s.m.v. after a time which is adjustable. The second o.s.m.v. provides a 7 volt gate which can be superposed on any one of the echoes of the echo train. This superposed echo is now passed through the last stage, which becomes conducting and charges a capacitor resistor network. The average charge is then displayed on a strip chart recorder.

Attenuation which is R.F. pulse amplitude sensitive is not really understood; but in order to avoid these effects, the experiment is conducted with a R.F. input to transducer which gives an unchanging attenuation of the echo train while the energy fed into the transducer is varied by ± 10 db.

The receiver overloading has been eliminated by use of a balancing network in the form of a complex impedance Wheatstone Bridge. The transducer is part of one of the arms of the bridge. The object is to balance the bridge with respect to the R.F. pulses coming from the pulsed oscillator. This pulse however energizes a transducer and the transducer output is fed from the bridge to the preamplifier and amplifier⁽⁹⁾.

2. Sample Preparation

Copper single crystals of approximately $2\frac{1}{2} \times 1\frac{1}{2} \times 1\frac{1}{2}$ " with the normal to a $2\frac{1}{2}$ " side coinciding within 2° with the (111) or (100) axes and having the faces in these planes parallel and flat to .0001 inch/inch are prepared from cylindrically shaped copper single crystals $4\frac{1}{2}$ " long and $3/4$ " in diameter. The separate operations involved in the sample preparation are: a. mounting specimen in a frame which would hold it on the goniometer; b. adjusting goniometer so as to maintain its alignment with respect to x-ray machine and milling machine; c. x-raying; d. flycutting; e. electropolishing; f. etching; and g. lapping.

The typical sequence for specimen preparation will be given here:

1. Mount specimen in frame using a litharge and glycerine mixture which hardens after some time.
2. Screw frame on to goniometer, having specimen vertical.

FRANKEL

3. Take an x-ray picture of sample for orientation determination.
4. Place goniometer with sample in place on milling machine.
5. Adjust goniometer so that desired major axis points along axis of flycutter.
6. Cut a small flat on sample.
7. Take the single crystal out of frame and rest it with small flat on a flat surface and put frame on the same surface, then pour in the licharge and glycerine mixture.
8. Remount frame on goniometer, again keeping specimen vertical.
9. Cut the crystal with flycutter, until desired dimension of face is reached.
10. Take another x-ray picture to check orientation.
11. Cut the other three sides, by remounting crystal.
12. Etch crystal with a 50% nitric acid solution to remove cold work.
13. Lap each specimen on ground, cast iron plate to achieve parallelism of faces of .0001 inch/inch.

An appreciable amount of material was removed by electropolishing in a 10% phosphoric acid solution. The frames were made of copper $\frac{1}{4}$ " thick and were ground parallel.

A. Machining of Crystal

The crystal was machined by means of a flycutter. The amount of material removed varied from .002 inches at the beginning to .0005 inches/cut near the final surface. In order to decrease the amount of damage done by the cutter, the feed used was very slow and the cutter speed high, and a lead-glycerine lubricant was used. The cutter was also sharpened several times for one face. The amount of cold worked surface that had to be removed by etching was between .005 to .01 inches.

B. Lapping

Wet lapping takes place on a surface ground cast iron plate. Clever lapping and grinding compound 4A was used. It has 25μ particle size. Final lapping was done with 3μ particle size diamond paste on a piece of newspaper. The lubricant was carbon tetrachloride. The newspaper gives the sample a nice scratch-free finish. Lapping and etching was done intermittently.

C. Parallelism

The sample is placed on a surface plate which is flat to .00004 inches over $1\frac{1}{2}$ x $1\frac{1}{2}$ feet and checked by means of a dial indicator with .0001 inch per division. Parallelism was induced by pressing lightly on high points during lapping.

D. X-Raying

Standard x-ray methods involving the Wulff net and the Greminger chart were used to determine crystallographic orientation. The samples are within 2° of desired orientation. Back reflection Laue photographs were also taken of specimen after preparation and annealing to determine that single crystal structure was maintained.

E. Tension System

The tension system has to meet the following

FRANKEL

requirements: a. Tension must be performed in temperature baths; and b. the only cause of deformation should be a uniform axial stress.

(Fig. 7) To avoid any undue deformation due to clamp type grips generally used the grips were glued to the copper by means of high temperature cured epoxy. The grips are then linked to the tension system by means of a dowel and universal on one side and a dowel and ball and socket type joint on the other. This is meant to avoid any bending and any twist on specimen. This system was tested at room temperature and less than 4% bending was seen by means of strain gages attached to the sides of a dummy specimen.

F. Recording System

The attenuation recording was done by means of an oscillo-record camera. Pictures can be taken at a maximum rate of 3 a second. The film is then read on an oscillograph analyser and reader. A pulse which may be 5cm. high on C.R.O. screen becomes over 6 inches high on the oscillograph analyser and reader. Relative heights can be read on it as accurately as the attenuation calibration on the scope. The major recording system is now that described under the electronic section.

DESCRIPTION OF EXPERIMENTAL PROCEDURE:

The transducer is mounted on (100) face of the specimen with a Nonaq vacuum bond. The specimen and transducer are then mounted inside a water-tight metal container which is placed in a temperature controlled oil bath. This whole device is built so that the specimen (Fig. 7) can be stressed in tension in an Instron tensile machine. The specimen is then stressed to about .4% strain and the stress released quickly. The change in ultrasonic attenuation of a given echo is then monitored continuously at the same temperature at which the specimen was stressed.

RESULTS AND DISCUSSION:

The normalized attenuation $\alpha_N = \frac{\alpha_f - \alpha_{f=0}}{\alpha_{f=0} - \alpha_{f=0}}$

here $f=0$ is the time at which stress is removed from the specimen.

If one assumes that the attenuation due to dislocation after complete pinning takes place is negligible then one can show that

$$\left[\alpha_N^{2/3} - 1 \right] = \beta f^{2/3}$$

A plot then of $(\alpha_N^{2/3} - 1)$ vs. $f^{2/3}$ will have a slope of β . From Equation 7 it can be seen that a plot of $\log \beta$ vs. $1/RT$ will yield a slope equal to $-2/3 U$ where U is the activation energy of migration of the defect which causes the pinning of dislocations.

Two runs are reported here. The recovery measurement at 80 C., Fig. 1, was recorded by means of film. During this run, a continuous calibration was not made, and the scatter was considerable. It is, therefore, difficult to assign β a specific value. All we can say is that β could be in the range of .02 min^{-2/3} to .07 min^{-2/3} (Fig. 2). From Table 1, one can see that at 80°C. (353°K.) the calculated value of β under the assumptions set forth there, is .027 min^{-2/3}. Run 2 at 95°C was made with different recording equipment. There a strip chart recorder was used, and it was continuously calibrated. From Fig. 3, one sees that the normalized attenuation α_N

FRANKEL

does not decay smoothly. A plot to find β , Fig. 4, reveals that the data can best be fitted by two straight lines. This would indicate that there are two recovery processes taking place. But in as much as the low β (implying a higher activation energy) occurs first and the high β (implying a lower activation energy) occurs second, it would seem that an interpretation of the two β must await further experiments. If one wants to consider the β_{max} only for runs 2 and 3 (Fig 2 and Fig 4) as the recovery parameter most clearly measured, Fig. 5 can give an indication of the slope which is 1.5ev.

A further worthwhile observation can be made from Fig. 3. One sees that the attenuation increases at 75 minutes after decreasing previously. This indicates that another process takes place which increases the average loop length L of the dislocations in the crystal. The accuracy to which the attenuation can be read is better than one tenth of the increase involved.

CONCLUSION:

The change in the attenuation of an ultrasonic wave in a stressed single crystal does not take place smoothly. This work indicates that more than one recovery process is involved, and that the most clearly measured process has an activation energy of 1.5ev.

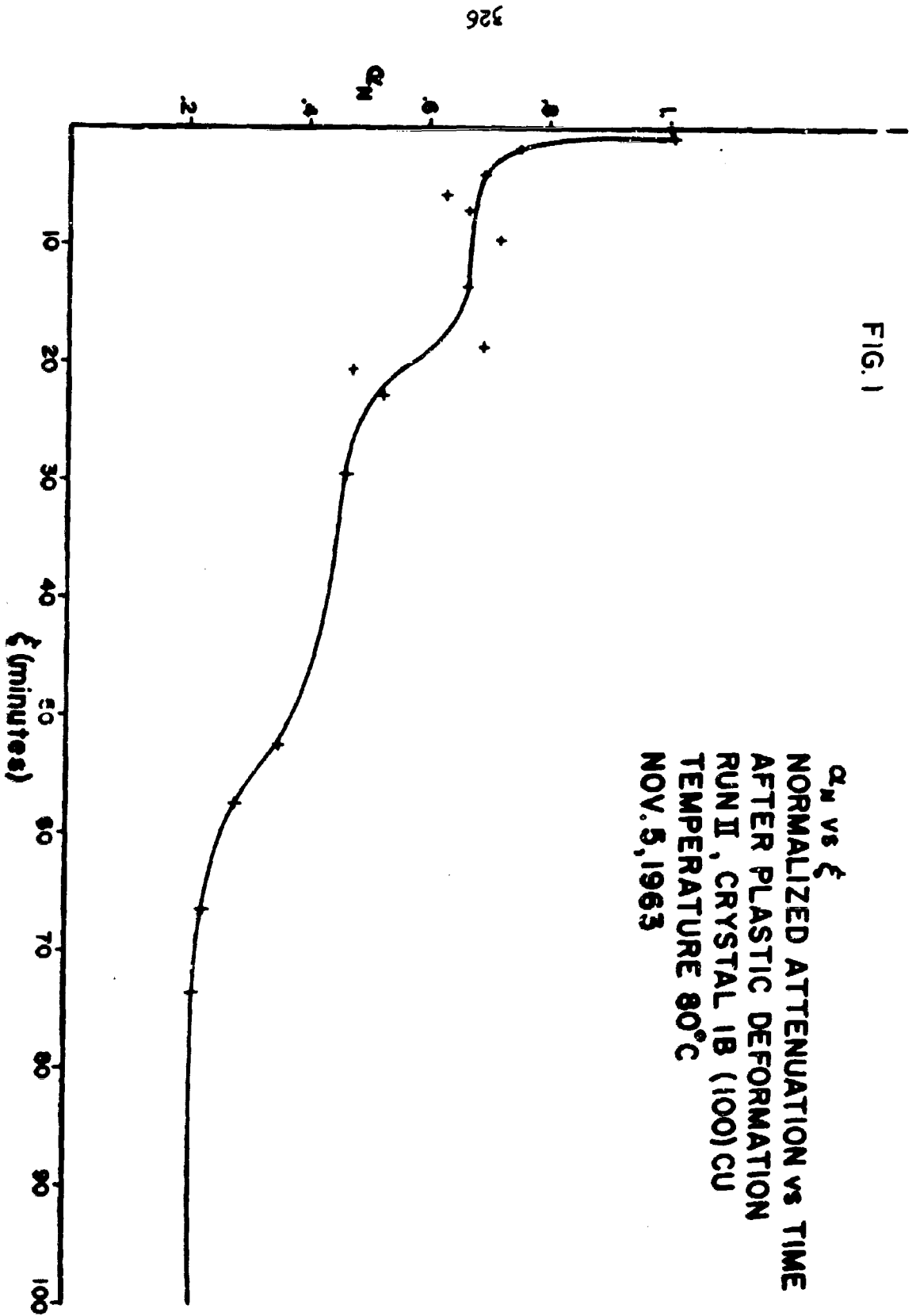
FRANKEL

REFERENCES:

1. G. A. Alers and D. O. Thompson, J.A.P., 32, 2, 283, 1961.
2. A. Granato, A. Hikata, K. Lücke, Acta. Met., 6, 470, 1958.
3. R.M. Stern, A.V. Granato, Acta. Met., 10, 353, 1962.
4. A. Granato, K. Lücke, J.A.P., 27.6, 1956, and 27.7, 1956.
5. Dislocations and Plastic Flow in Crystals, A.N. Cottrell, Oxford Press.
6. F.W. Young, Jr., J.A.P., 33, 3, 963, 1962.
7. H.G. VanBueren, Acta. Met., 3, 519, 1955.
8. A.D.N. Smith, Phil. Mag., 44, 453, 1953.
9. The electronics used in this experiment were manufactured by Arenberg Ultrasonic Laboratory. The operation of the balancing network is contained in their Bulletin T.D. (58) Rev.

FIG. 1

α_n vs ξ
NORMALIZED ATTENUATION vs TIME
AFTER PLASTIC DEFORMATION
RUN II, CRYSTAL 1B (100) CU
TEMPERATURE 80°C
NOV. 5, 1963



FRANKL

$a_N^{-1/4} - 1$ vs $\xi^{2/3}$
 RUN II, CRYSTAL IB
 TEMPERATURE 80°C
 NOV. 5, 1963

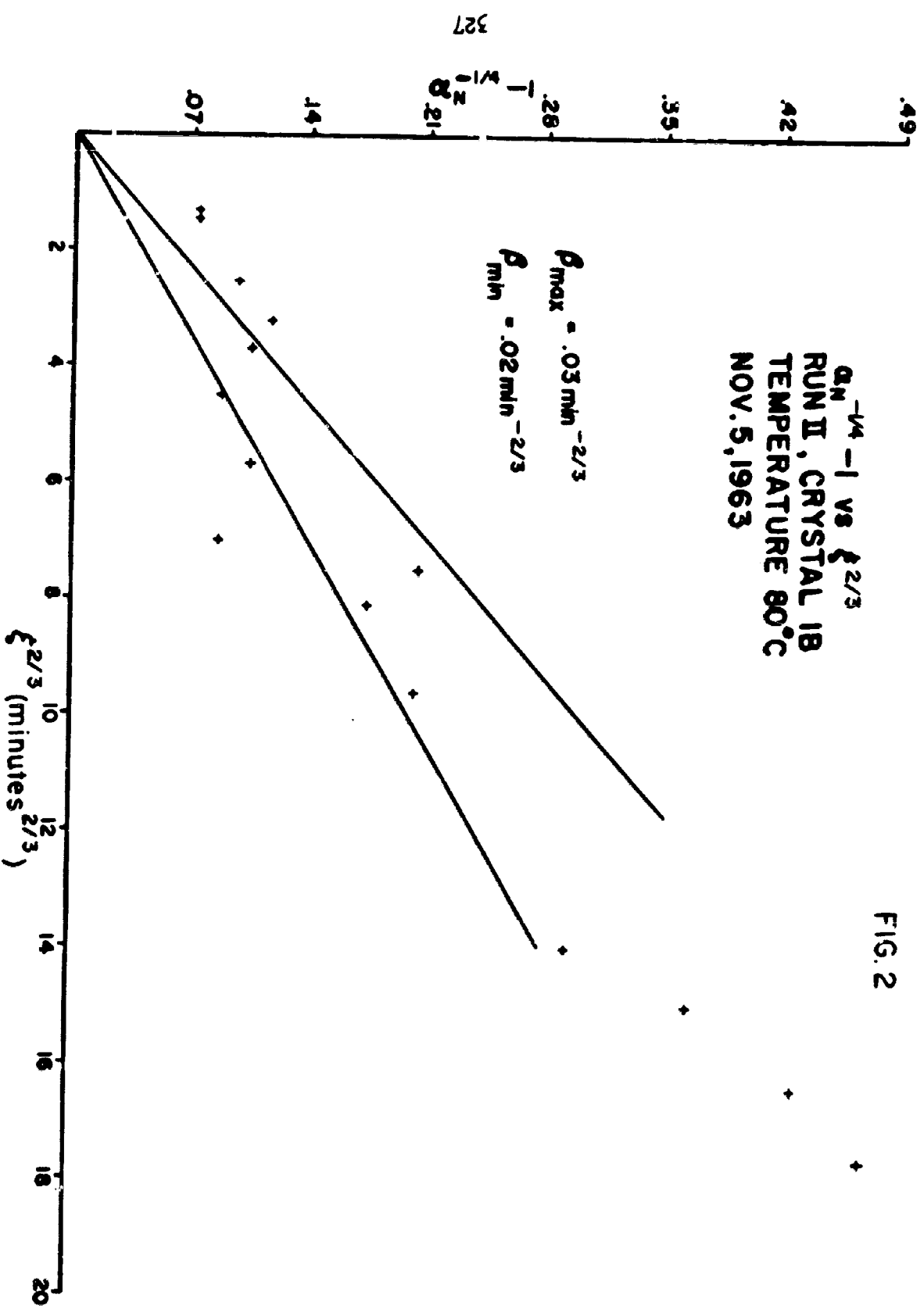
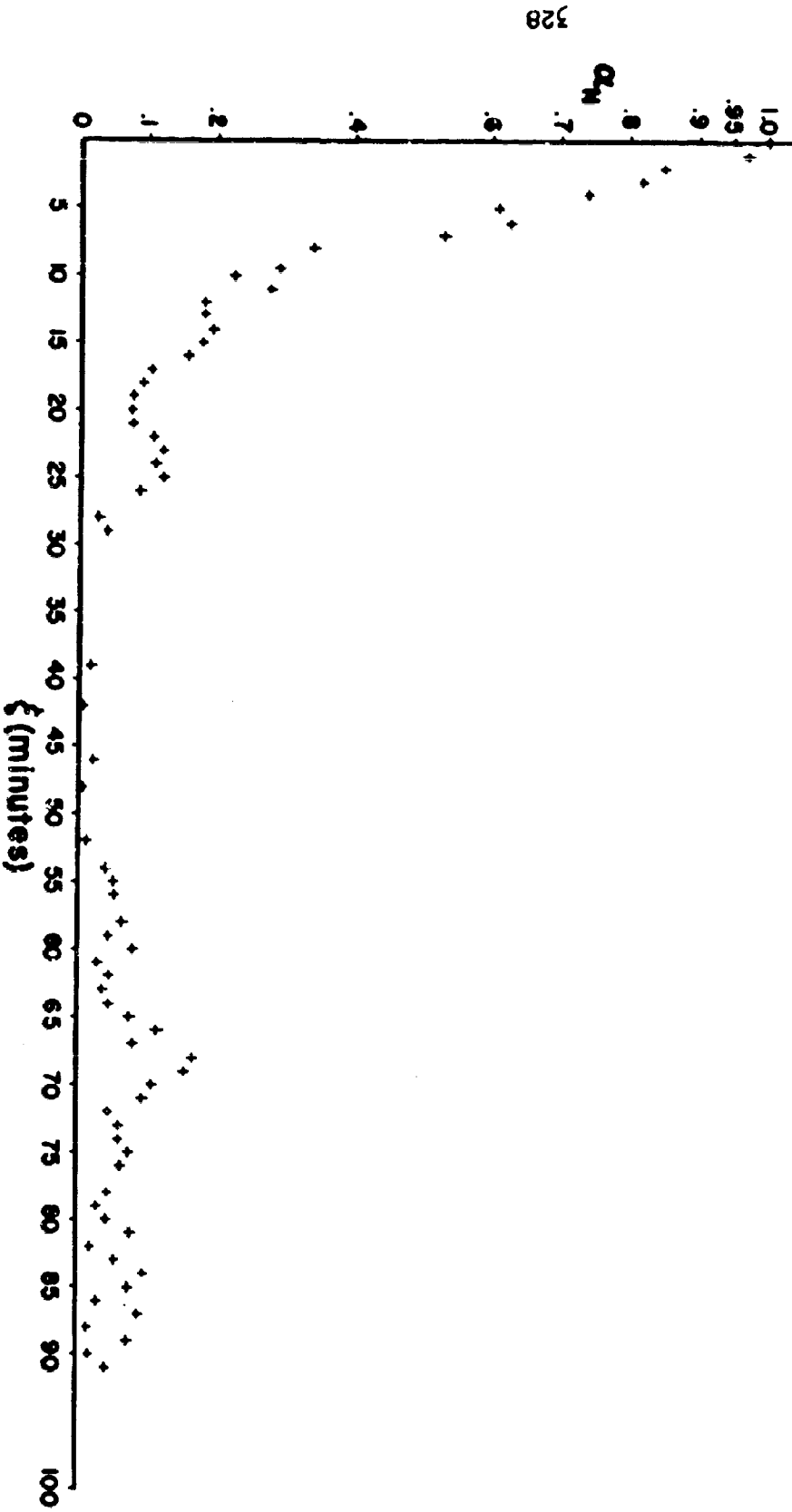


FIG. 2

FRANKEI

FIG. 3

α_M vs ξ
RECOVERY OF THE ATTENUATION.
RUN III, TEMPERATURE 95°C.
FEB 8, 1964

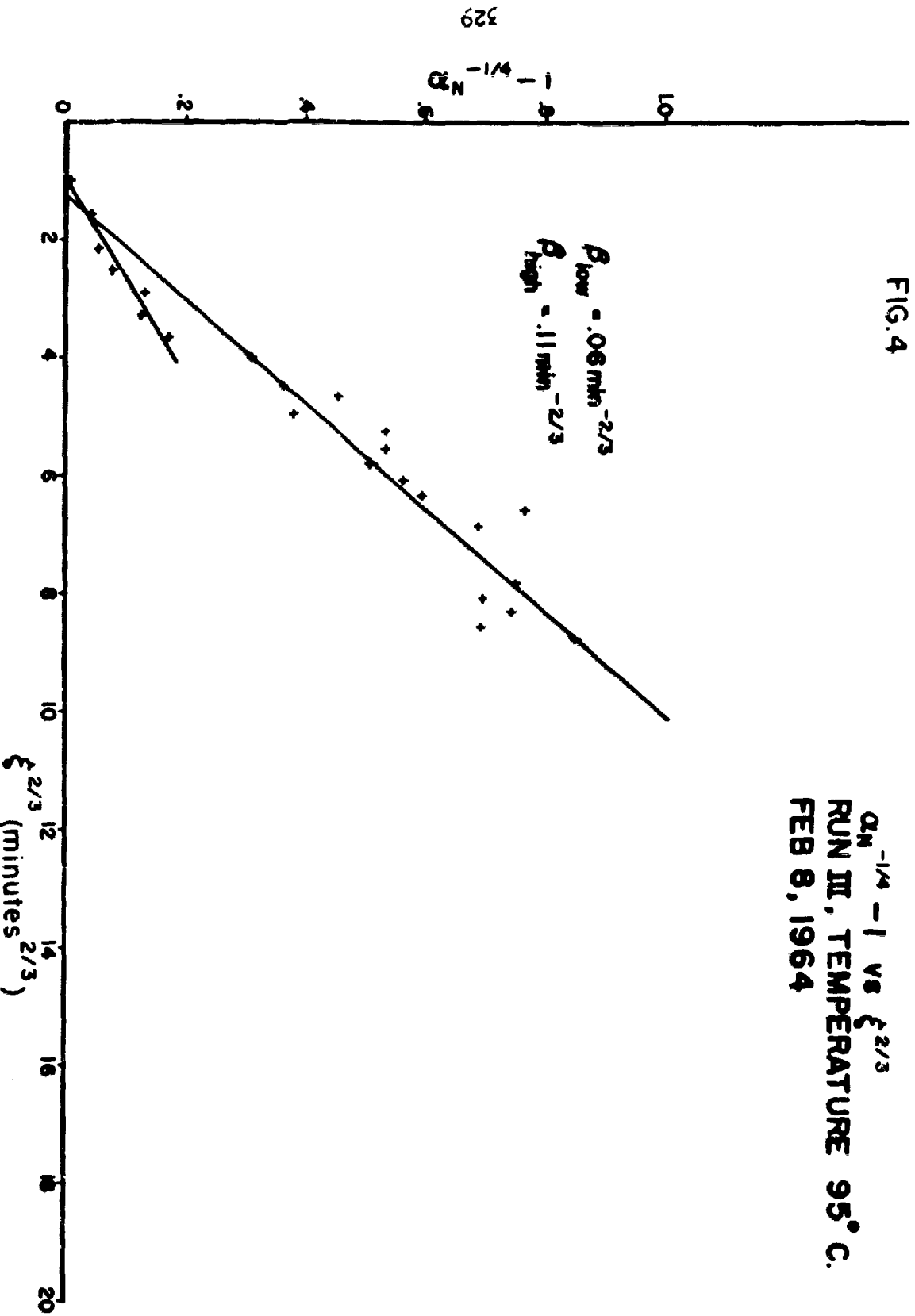


FRANKEL

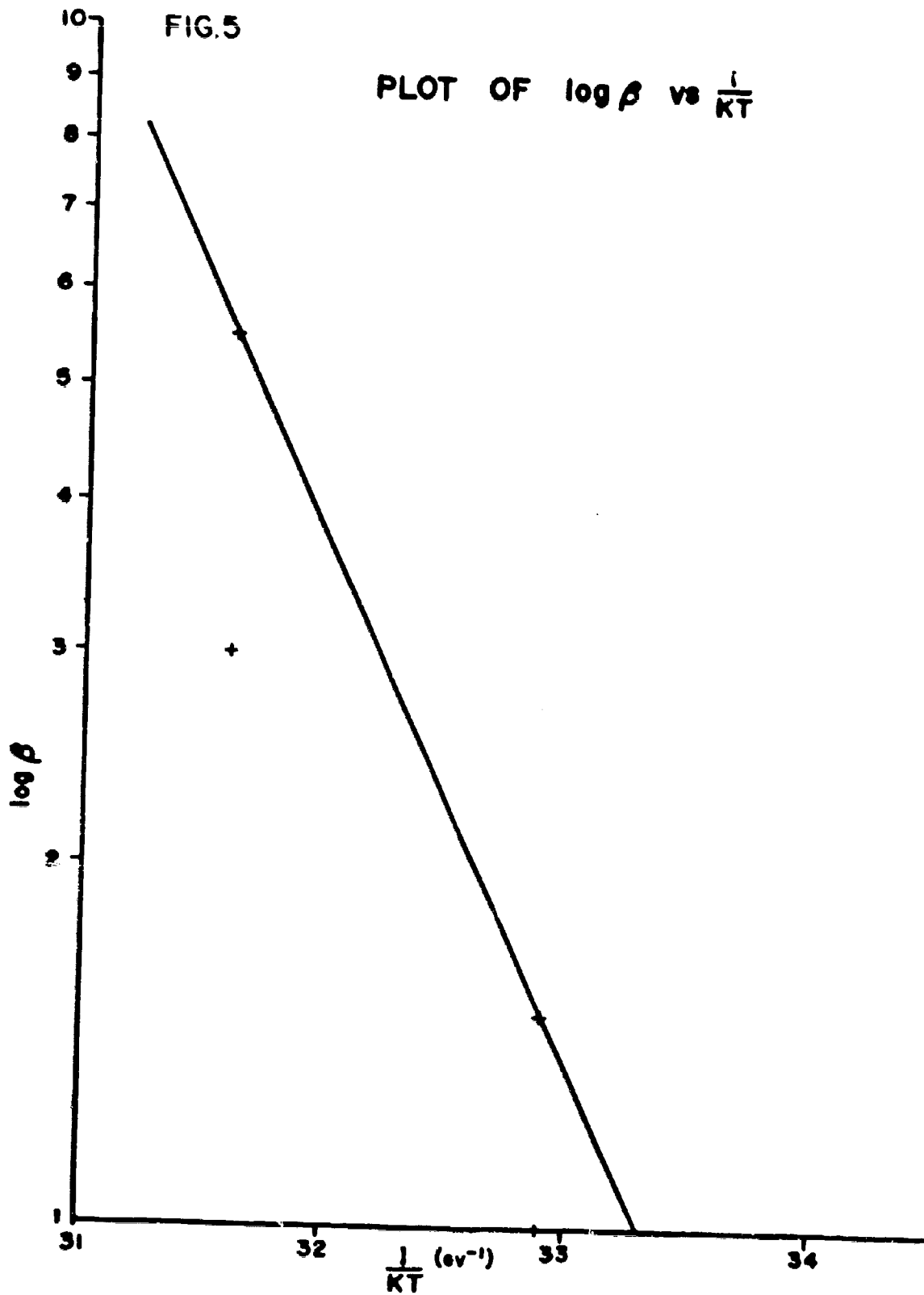
$D_N^{-1/4} - 1$ vs $\xi^{2/3}$
 RUN III, TEMPERATURE 95° C.
 FEB 8, 1964

FRANKEL

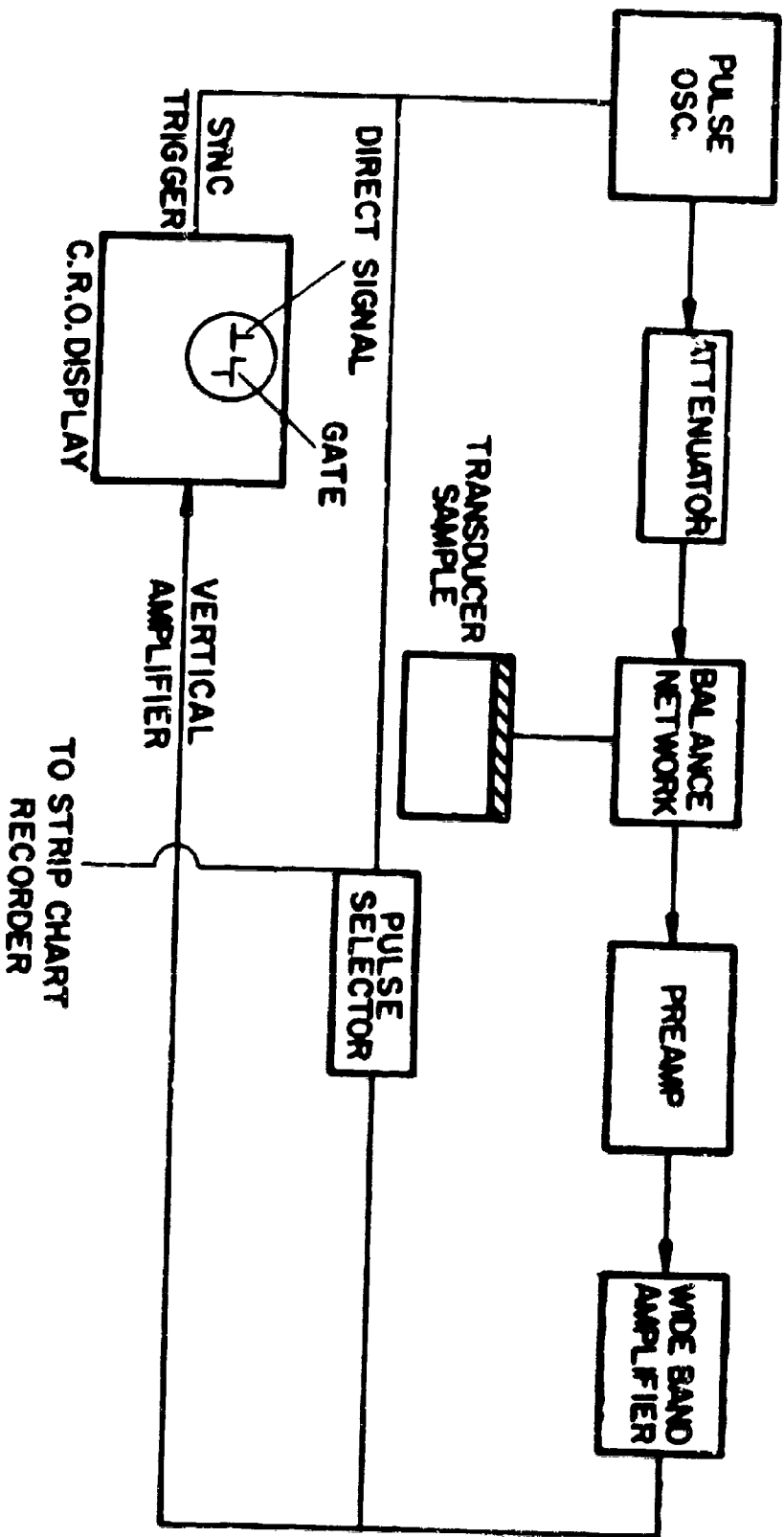
FIG. 4



FRANKEL

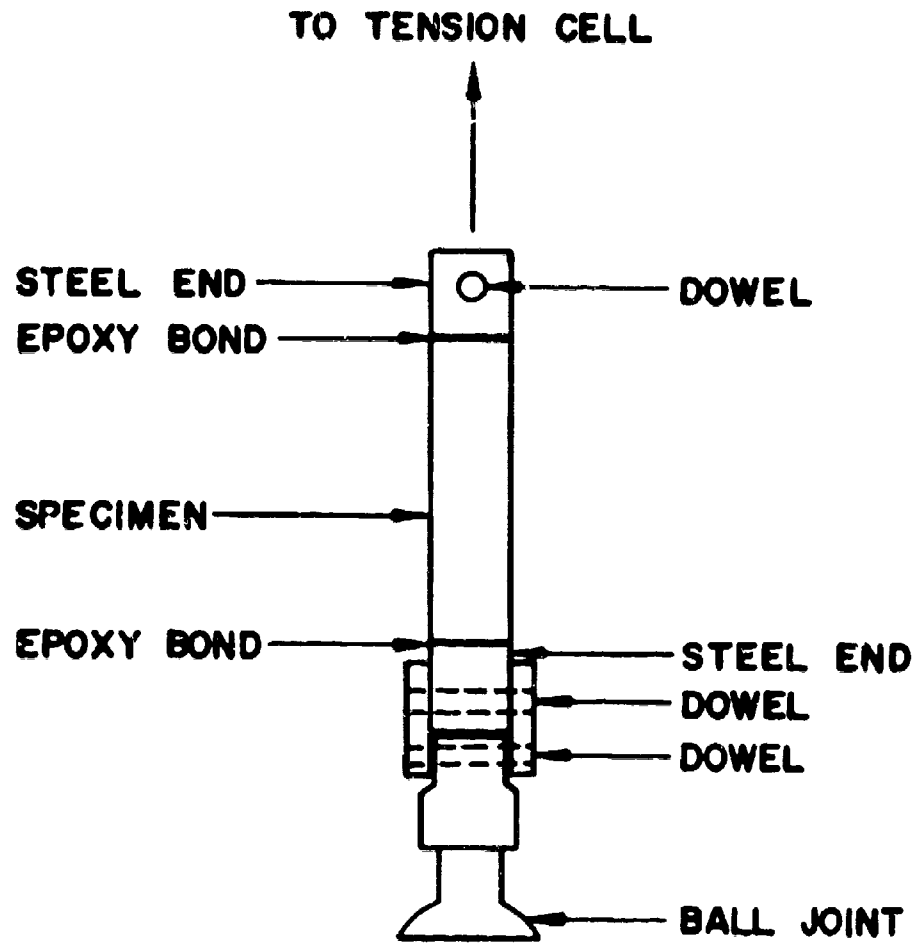


BLOCK DIAGRAM FOR INSERTION LOSS MEASUREMENTS



331

Fig. 6



**METHOD OF APPLYING UNIAXIAL
TENSION TO SPECIMEN
(NOT TO SCALE)**

Fig. 7

FREITAG, and POWELL

LABORATORY INVESTIGATION OF THE
MOBILITY OF PNEUMATIC TIRES IN
COHESIONLESS SOIL

D. R. FREITAG, and C. J. POWELL
U. S. ARMY ENGINEER
WATERWAYS EXPERIMENT STATION
VICKSBURG, MISSISSIPPI

1. In an age when supersonic flights are commonplace, and man's first explorations into space are receiving much attention in the headlines, it is difficult to accept the fact that ground vehicles still cannot strike out cross-country toward an objective with any assurance of being able to reach it. The vehicle may bog down because of insufficient bearing capacity of the soil; it may spin its wheels because of loss of traction; or it may be unable to pass over or around a wide variety of surface obstacles because of limitations imposed by its own geometry. Solutions for each component of the total ground mobility problem are being sought by various Department of Defense agencies. The study discussed in this paper is designed to shed light on just one area of the problem. Specifically, it is part of an investigational program intended to discover laws that govern the interaction between pneumatic tires and certain representative homogeneous soils.

BACKGROUND

2. In 1959, the Chief of Research and Development, Department of the Army, (now Army Materiel Command) directed the Army Mobility Research Center,* Soils Division, U. S. Army Engineer Waterways Experiment Station, through the Chief of Engineers, to undertake a study of selected pneumatic tires operating on homogeneous soils. The study, entitled "Performance of Soils Under Tire Loads," was begun in June 1960. The work has been guided by a plan of tests agreed upon by representatives of the then-existing technical services--Ordnance, Transportation, and Engineers. Briefly, it provided for testing a range of tire sizes, under various conditions of load and inflation pressure, on three representative types of soil. All testing was to be accomplished with treadless tires so that geometry effects could be evaluated independently of tread effects. The study of tire

* Now the Army Mobility Research Branch, Mobility and Environmental Division.

FREITAG, and POWELL

treads and of track configurations was to be deferred until a later date.

PURPOSE OF STUDY

3. The purpose of the portion of the study described in this paper was to define the relations that exist between the physical properties of a typical dry desert sand, the geometric characteristics of pneumatic tires of various sizes and shapes, and the performance of these tires under various loads.

DEFINITIONS

4. The following terms and symbols used in this paper are defined here as briefly as is possible without risk of misunderstanding.

<u>Carcass diameter (d)</u>	Outside diameter of tire, exclusive of tread, with a specified inflation pressure but with no load applied.
<u>Section width (b)</u>	Outside width of the cross section of the tire with a specified inflation pressure but no load applied.
<u>Deflection (δ)</u>	Change in height of the cross section of the tire, measured from the rim to the supporting surface, when the load is applied.
<u>Measured hard-surface rolling circumference</u>	Forward advance per revolution of the loaded tire when towed on a flat, level, unyielding surface.
<u>Cone index (CI)</u>	An index of soil strength. It is the unit load required to maintain movement of a cone-shaped probe into a soil. It has the dimensions pounds per square inch, and is usually given as an average value for a specified layer of soil several inches thick.
<u>Torque (M)</u>	Torque input at the axle.

<u>Pull (P)</u>	The component, acting parallel to the direction of travel, of the resultant of all soil forces acting on the wheel. It is considered positive when the wheel is performing useful work and negative when the component represents the force that must be applied to maintain motion.
<u>Maximum Pull (P_m)</u>	Maximum net horizontal force developed by the powered wheel. It is usually in a positive force.
<u>Towed force (P_t)</u>	Horizontal force required to tow the wheel when no torque is applied. It is always a negative force.
<u>Load (W)</u>	Vertical force applied to the wheel.
<u>Sinkage (z)</u>	Maximum depth to which the wheel penetrates into the soil relative to the original soil surface.
<u>Slip</u>	Lost motion computed as follows: theoretical wheel advance (based on measured hard-surface rolling circumference) minus actual wheel advance all divided by theoretical wheel advance, usually expressed as a percentage.

TEST CARRIAGE

5. The primary research instrument used in determining the performance of soils under tire loads is the test carriage built especially for testing with an assortment of pneumatic tires. The test carriage is supported by solid rubber-tired rollers on a pair of accurately aligned and graded overhead rails that are, in turn, suspended from cantilever columns and cross-arms. The carriage is towed by an endless steel cable that is fastened fore and aft to the carriage, passes up over pulleys at the ends of the track system, and is driven by sheaves mounted on a platform above the overhead rails. Fig. 1 shows a section of the test lane with soil cars in place and the carriage positioned directly under the towing-cable drive mechanism. The speed of the towing cable can be varied continuously from zero to about 30 ft per sec, and the cable and pulleys can be shifted transversely across the width of the soil cars, along with the test carriage, so that the traffic lane can be established at any transverse

position in the cars.

6. The test wheel is mounted in the lower frame assembly, which consists of an inner and an outer frame. The entire assembly is fastened to the upper frame by columns and roller assemblies so that free vertical movement of the wheel is possible. Both the wheel and the transmission are suspended from the inner frame. The inner frame is fastened to the outer frame at all four corners by specially built load cells that are mounted vertically and also serve as hinges, allowing relative movement of the two frames longitudinally. These four load cells measure the vertical load on the wheel and are connected in a bridge circuit so that torque effects will be canceled. The relative longitudinal movement is opposed by a load cell mounted horizontally between the two frames so that the reading from this cell is a measure of the pull on the test wheel.

7. Load can be applied to the wheel by the addition of dead weights or by "Bellofram" pneumatic cylinders mounted between the upper and lower frames. The Bellofram assembly consists of an outer cylinder and loose-fitting piston joined together and sealed against air leakage. As the piston moves through the length of the cylinder, the Bellofram or membrane turns inside out, thus eliminating wall friction as experienced in an ordinary piston-and-cylinder arrangement. The travel of the piston is fairly long, but the system is single acting so that two cylinders are required, one for upward force and one for downward. The test carriage utilizes a pair of cylinders at the front and a pair at the rear. The air-storage tanks, which are visible on both sides of the upper frame in fig. 1, provide a reserve air supply to compensate for movement of the loading cylinders caused by increasing sinkage of the wheel as it progresses down the test lane.

INSTRUMENTATION

8. The carriage is instrumented to measure torque input to the wheel, vertical movement of the axle, dynamic tire deflection, wheel revolutions, and carriage position, in addition to the horizontal and vertical forces. Provision also is made for recording the profile of the soil surface before traffic is applied. Torque input for most of the tests is provided through a mechanical transmission mounted on the axle of the wheel and a hydraulic motor. A transmission driven directly by an electric motor has also been used, and for special applications a chain-driven auxiliary reduction gear can be mounted. The transmission is restrained from rotating about the axle by a lever arm of known length attached to a load cell that is anchored to the inner frame; consequently, the load cell output signal can be calibrated to indicate the reactive torque directly. Axle movement is measured by circular potentiometers mounted on the outer frame and geared to rack gears mounted on supports from the upper carriage. There is an assembly for this purpose at both the front and rear of the lower carriage. Tire deflection is measured by two

linear potentiometers that are mounted through the rim and are spaced 180 degrees apart radially around the wheel. The signals are transmitted through a slip ring mounted on the axle. Wheel revolutions are recorded by a stationary photoelectric cell and a perforated circular disc that rotates with the axle, and carriage position is indicated by a photoelectric cell on the upper frame, which is actuated by tabs spaced 1 ft apart on the overhead rails.

9. The operation of the test carriage during a test run is controlled by a specially constructed programmer. This unit causes the forward speed of the carriage to conform to a preselected value or to vary through a predetermined pattern. Another unit has been constructed to sense the actual speeds of the carriage and the wheel and to convert these values into the slip ratio.

TIRES

10. A variety of tire sizes has been used in this study, as shown in the following list. The sizes were chosen to cover a range of diameters, widths, ply ratings, and types of construction. Some of the tires were supplied without tread, and the rest were buffed smooth after delivery. All were operated tubeless with the exception of the 1.75-26 bicycle tire and the 6.00-16 solid rubber tire. The deflection values used throughout this paper were measured on a level, unyielding surface unless otherwise specified.

TIRES USED IN PROGRAM

- 1.75-26, bicycle tire, buffed smooth
- 4.00-18, 2-PR* buffed smooth (originally motorcycle tire tread)
- 4.50-18, 4-PR, buffed smooth (originally motorcycle tire tread)
- 6.00-16, 2-PR, buffed smooth (originally highway tread)
- 6.00-16, radial ply, buffed smooth (originally directional bar tread)
- 6.00-16, solid rubber tire, buffed smooth (originally nondirectional bar tread)
- 9.00-14, 2 PR, supplied without tread
- 9.00-14, 4-PR, buffed smooth (originally highway tread)
- 9.00-14, 8-PR, supplied without tread
- 5.00-12, 2-PR, buffed smooth (originally directional bar tread)

* PR indicates the ply rating specified by the manufacturer.

FREITAG, and POWELL

4.50-7, 2-PR, buffed smooth

4.50-18, 4-PR, buffed smooth, mounted in dual configuration

16-15-6R, 2-PR Terra tire, supplied without tread

TEST SOIL

Soil properties

11. The desert sand used in these tests was dug from the top 12 in. of sand dunes near Gray's Wells, California (17 miles west of Yuma, Arizona), by personnel of the Engineer Detachment at the Corps of Engineers Yuma Test Branch and was sent to the Waterways Experiment Station. The sand classified as SP-SM under the Unified Soil Classification System. Values of angle of internal friction, as determined by direct shear tests, averaged about 36 degrees. Specific gravity was determined to be 2.66. Average maximum dry density was determined in the laboratory to be about 104 lb per cu ft, and minimum density was about 87 lb per cu ft.

Preparation of test section

12. All the tests were performed with air-dried sand in which the actual moisture content was usually about 0.5 percent of dry weight. Different degrees of soil strength were achieved by varying the density of the sand. Depending upon strength conditions desired, the sand was placed in the cars either in uniform layers, the last of which was screeded and surface-vibrated, or in individual layers which were screeded and vibrated as the filling progressed. Uniform deposition of each layer was achieved by placing a large 1/4-in.-mesh screen on the tops of the cars and depositing the sand uniformly over the area of the screen with a tractor-mounted shovel. Compaction was accomplished by an electric vibrator mounted on a steel baseplate that was long enough to span the width of the soil car. It was propelled back and forth over the length of the car at a constant speed by an electrical towing unit built especially for this purpose.

Soil measurements

13. There is a wide difference of opinion on what constitutes a valid measurement of soil strength for use in a study such as this one. To insure that the performance data would be useful to the maximum number of researchers, soil strength data were obtained with a variety of instruments. Penetration tests were run with circular flat plates of at least three different sizes and with the standard trafficability cone penetrometer. In-situ shear tests were made with a ring shear device and a four-bladed vane-shear apparatus. Triaxial and direct shear tests were run in the laboratory on samples obtained from the test sections.

TESTING TECHNIQUE

14. The soil-tire performance tests were conducted by uniformly increasing the tire's slip throughout a pass down the test lanes. Power was applied to the axle of the wheel and the desired rotational speed of the wheel was attained almost instantaneously and held essentially constant throughout the pass. At the same time, power was applied to the towing cable so that the carriage was accelerated through the length of two approach cars and entered the actual test cars traveling faster than the wheel was propelling itself. This resulted in a skidding or negative-slip condition. At this point a photoelectric cell engaged the carriage-speed program unit, and the speed of the towing cable was reduced linearly with time until the carriage was brought to a standstill near the end of the test cars. This produced slip that varied linearly with time and ranged from some negative value at the beginning of the test cars through a range that included the towed point and the maximum pull point and then on to some large positive slip near the end of the cars.

ANALYSIS

15. Three major classes of independent variables are involved in the passage of a wheel through soil. They are: (a) wheel geometry, (b) load, and (c) soil strength. Time is not considered in this study because all tests were run at, or near, the same forward speed. The dependent variables that are measures of the performance of the tire include towed force, pull, torque input, sinkage, and slip. The independent variables constitute the "cause" and the dependent variables the "effect."

16. The simplest and most direct method of studying the relations between dependent and independent variables is through the medium of basic plots in which a single performance variable is plotted against a single test condition, with all other independent variables held constant. The introduction of one additional pertinent independent variable will normally produce a family of curves separated by that additional variable. In the analysis, the basic plots were examined first to establish clearly the trend of data in response to the several variables. Fig. 2 shows the interrelations between the dependent variables slip, sinkage, and pull at the towed point as these factors change in response to changes in the independent variable soil strength (here expressed in terms of cone index). It should be noted that the data represent only one tire at one percentage of deflection and one load. The sand strength range shown, from a low of about 25 to a high of about 70 CI, represents closely the total range commonly encountered in nature for dry sand.

17. In the upper left plot of fig. 2, it can be seen that

FREITAG, and POWELL

the slip experienced by the towed wheel is always negative. On the firmest sands the slip is small, but in soft sand the wheel slips or skids the equivalent of about a third of a revolution per revolution. Slip is related both to the sinkage, s , and to soil strength. Large sinkage and great slip occur in very soft soils. In the lower right plot the data show the strong relation between the force required to pull the wheel and the amount of sinkage it experiences. Finally, in the upper right plot the relation between pull at the towed point and slip is shown. This figure when viewed as a whole shows the complex interrelation of the pertinent variables as dependent upon soil strength for a particular case.

18. Fig. 3 is a continuation of the examination of basic data relations that are included in the test results. The left plots show the effect of load and strength on the performance of one tire at one deflection. The upper plot refers to the force required to tow the wheel while the lower part refers to the maximum pull the wheel was able to exert. It is of interest to note that the towed force is a large negative value and the maximum pull is a small positive (sometimes actually small negative) value in the softest soil. In the high strength soil the towed force is a small negative value while the maximum pull is a relatively large positive value. The curve families shown on the right side of fig. 3 reveal the effect of tire deflection on performance. Increased tire deflection, which for a constant load requires a decrease in inflation pressure, causes the towed force to be of a lesser absolute value and the maximum pull to be of a greater absolute value, other factors being equal. Figs. 2 and 3 together describe for one tire how the variables load, strength, and tire deflection (a tire shape factor) influence the dependent variables pull, slip, and sinkage. To complete the examination of basic relations it would now be desirable to view the effects of differences in tire diameter, width, construction, etc. However, the addition of these relations to the basic plots would make it unduly complex; also space does not permit this elaboration.

19. The obvious alternative to plotting individual variables against each other is to combine all of the pertinent independent variables into a single significant dimensionless term (or numeric) and plot each of the dependent variables, expressed as dimensionless numerics, against the independent numeric. The ideal dimensionless independent numeric must contain all of the significant variables in the proper proportions and cause the performance curves for all the tires, under all test conditions, to collapse into a single curve within the band of experimental scatter of the data. Dimensional analysis has proved to be a useful tool in developing the form of the dimensionless independent numeric. It must be recognized, however, that dimensional analysis has its limitations and usually serves only as an intermediate step in the process of producing a dimensionless numeric in its final form. Numerical coefficients, for instance, cannot be determined by dimensional analysis, but must be

obtained through experimentation and experience. Also, it is often necessary to perform mathematical manipulations on the several numerics resulting from a dimensional analysis in order to produce an optimum independent dimensionless numeric. Dimensional analysis, then, is a powerful aid in the study of physical relations, but is limited in itself to the production of qualitative rather than quantitative results. Several numerics will emerge from the dimensional analysis, and it is necessary to choose the dominant one. The numeric which appears to be dominant in the mobility field is of the general form:

$$\frac{W}{\tau \ell^2}$$

where

W = load, lb

τ = soil strength, psi

ℓ = a characteristic linear dimension (such as wheel diameter), in.

20. The form of the general dimensionless numeric, W/ℓ^2 , obtained through dimensional analysis, may be modified extensively so long as the dimensionless quality is not destroyed. Several variations of the numeric have been investigated, each utilizing cone index for the soil strength term. The linear dimension squared (ℓ^2) has been modified to bd (tire width times diameter), and δd (deflection times diameter). In the former case, there is no term to describe the deflected shape of the tire, and in the latter, there is no tire-width term. Since both deflection and width are known to influence results, it is not likely that numerics that fail to include them will provide a basis for correlation of the performance of all tire sizes and test conditions.

21. Several combinations of the three linear factors, diameter, width, and deflection, have been tried without achieving a satisfactory degree of data collapse for all tire sizes and shapes. Some of the difficulty obviously was due to the data scatter that is evident in figs. 2 and 3, but visual examination of the locations and relative positions of the data from some of the tires suggested that the numeric was influenced by tire size. At this point in the analysis adherence to the true dimensionless relation was relaxed and the best combination of the linear factors was sought by an empirical analysis. One of the best relations developed by this procedure is shown in fig. 4. It will be noted that the expression used is not dimensionless. Each line represents a visual average relation for all the data for one tire. The final relation shows a fair grouping of the data and similar trends for most of the tires, but it obviously still leaves much to be desired. This study is continuing,

particularly along lines that will explain the apparent need to include "extra" linear dimensions. One theory has been advanced that in a frictional soil such as the dry sand used, the strength term used must itself be scaled to dimensions that are consonant with the size of the tire. At this point it can be stated only that the idea has promise.

VEHICLE TESTS

22. There is another aspect of the study that must be resolved if the data are to be useful in a practical way. It must be shown that the results of tests conducted with a single wheel suspended in the special test rig with controlled input are in some way comparable to the behavior of an actual vehicle. To date, a few tests of this nature have been run. A test vehicle was made from a salvaged commercial 4x4 station wagon. The vehicle was stripped of all nonessentials to result in the minimum dead weight feasible. Mechanically, the major modification required was the elimination of differential action in both axles so that all wheels would rotate at the same speed. However, the spring system also was blocked out to minimize any dynamic oscillations or weight shifting. The vehicle was tested with all wheels driving at a constant slip rate. The amount of force the vehicle exerted was measured for several different levels of soil strength and tire deflection, but so far with only one load.

23. For comparison data were extracted from the results of the single-wheel tests with the test carriage in the following manner. It was assumed that the first pass of the single wheel in the test lane represented the passage of the front wheel of the vehicle and the second pass in the same path represented the passage of a rear wheel. Then the pull values at the appropriate slip for the first and second passes of one single wheel were added together and multiplied by two (for the two sets of vehicle wheels). This total pull was then compared with the results obtained with the actual 4x4 vehicle under similar conditions.

24. It was found that the combined single wheel results agreed quite closely with the vehicle test results. There was a tendency for the combined single-wheel pull to be a little higher than the vehicle pull, and it may be that this difference arose from rolling resistance losses of the vehicle. However, the agreement between the two sets of data was considered satisfactory. Many more correlation tests of this type will be made using other combinations of tires and loadings.

SUMMARY

25. From the experience to date it can be stated that the test equipment, particularly the single-wheel test carriage, has made it possible to conduct rapidly a large number of soil-tire tests

under a broad variety of controlled conditions. Some special tests have shown that the single-wheel data can be related readily to the performance of an actual 4x4 vehicle. The results of these tests show that a tire's performance varies in a regular and systematic way in response to the three primary independent variables of soil strength, tire load, and tire deflection. In a particular test condition, the amount that a tire sinks and slips is related to the other test variables. However, except in a very general sense of the word, it has not been possible to correlate the performance data for all tire sizes and shapes by means of a single general expression.



Figure 1 - Single-wheel test carriage in position over a line of test cars. Towing cable drive mechanism is visible at upper right and instrumentation cables can be seen to the left beside the test cars.

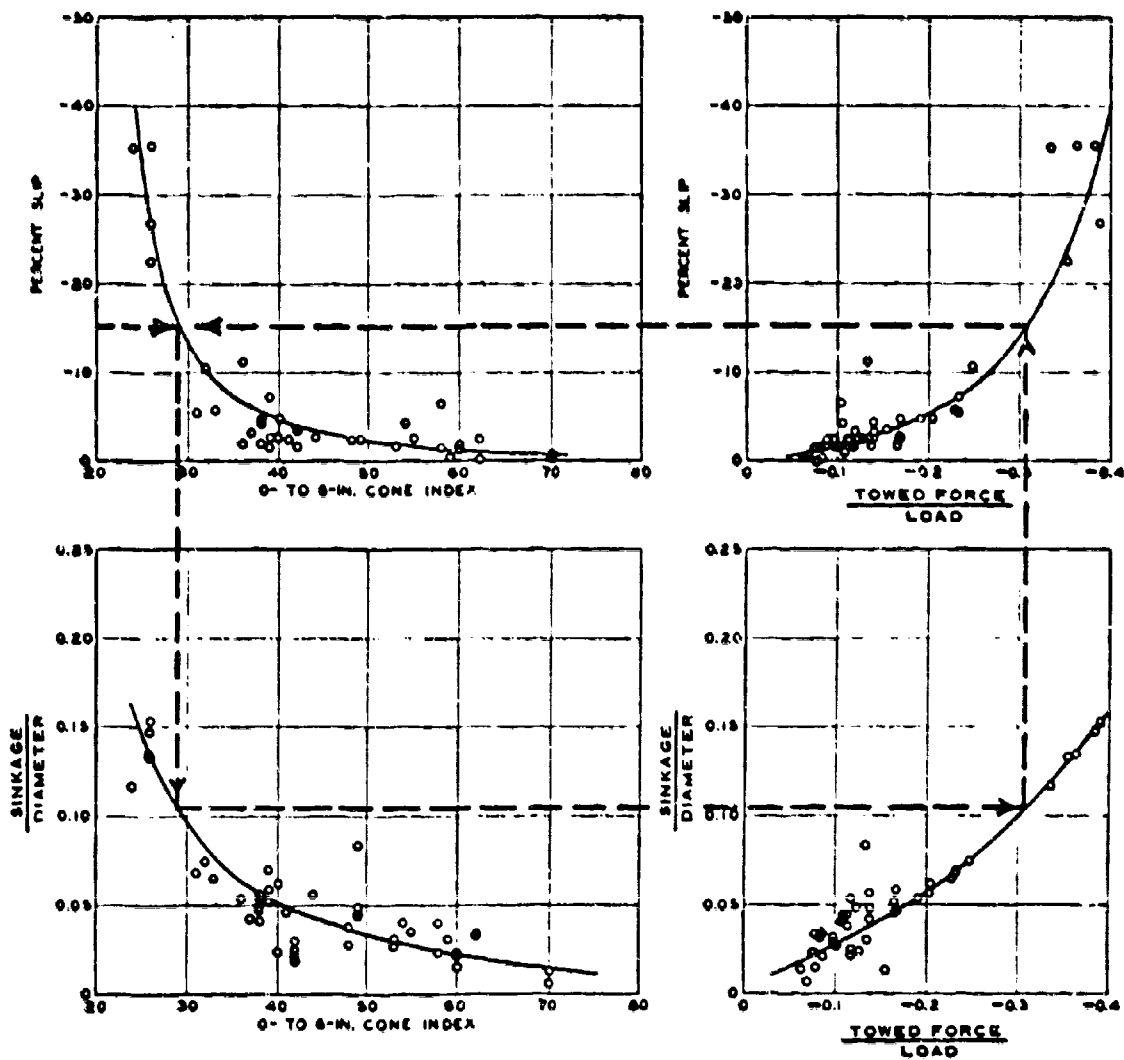


Figure 2 - Interrelations among the dependent variables slip, sinkage, and pull at the towed point for the 4.50-18, 4-PR tire, 35% deflection, 910-lb. load, Yuma sand.

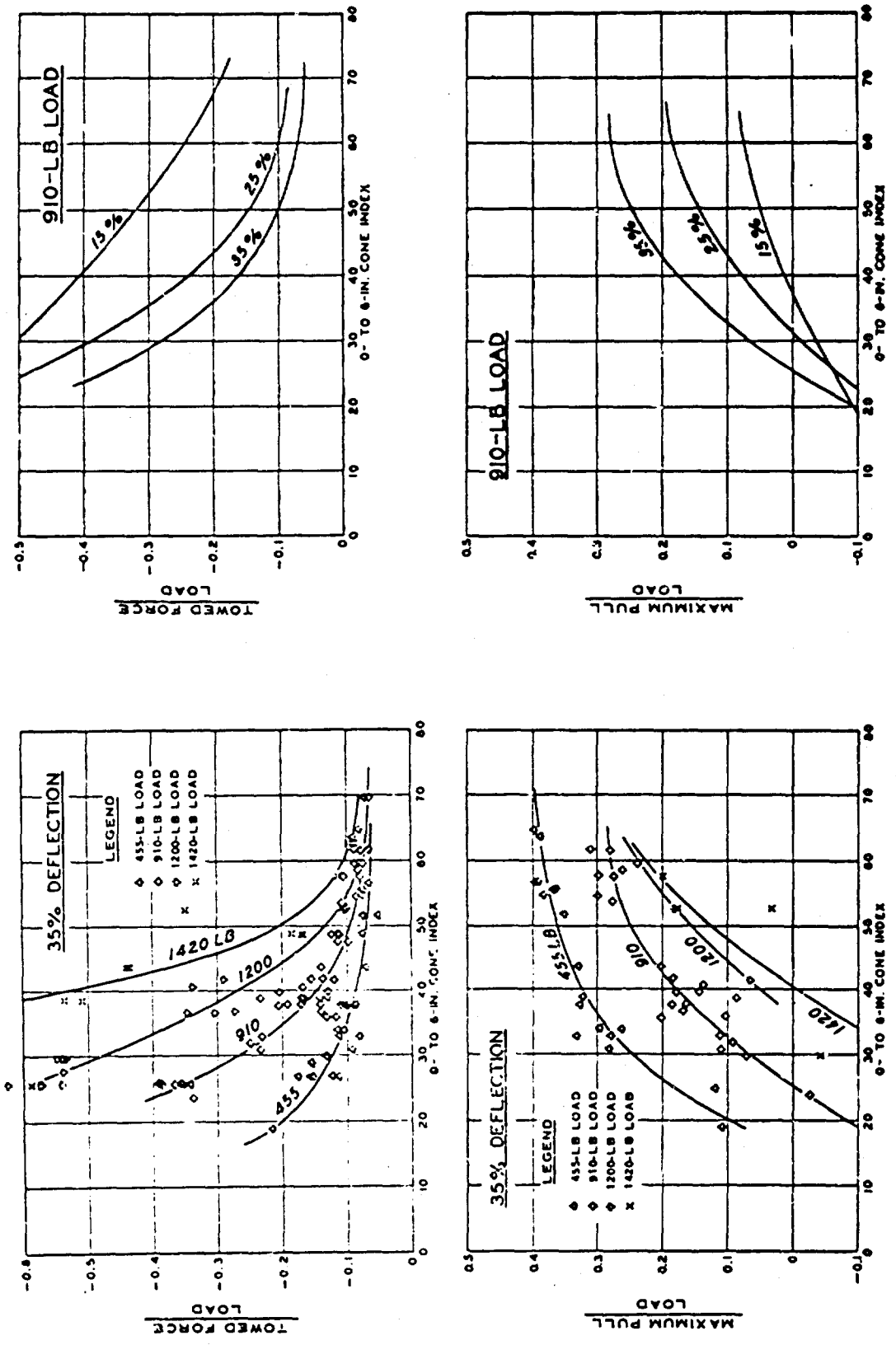


Figure 3 - Maximum pull and towed coefficients versus cone index for the 4.50-18, 4-PR tire.

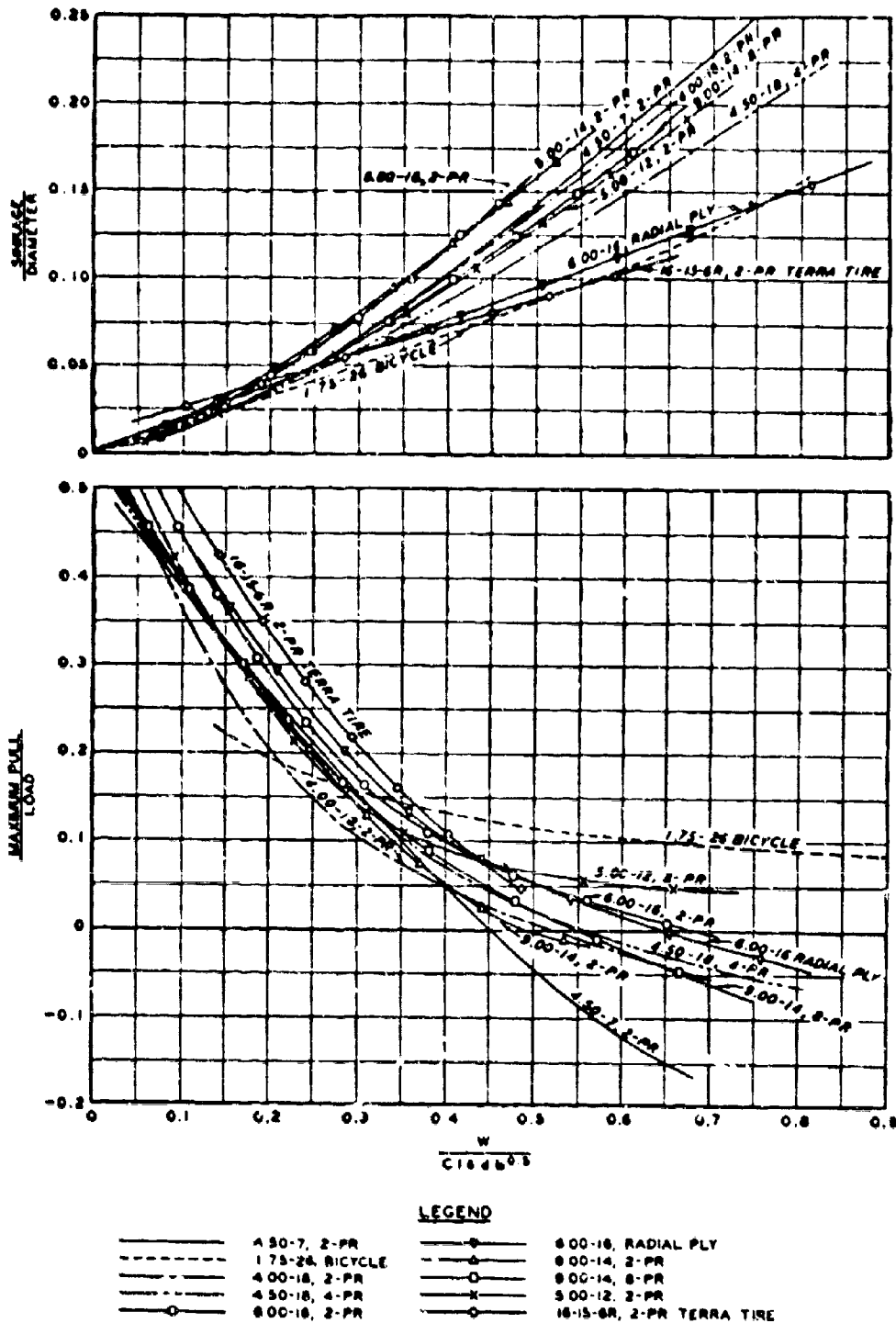


Figure 4 - Sinkage and maximum pull coefficients versus a numeric consisting of a suitable combination of pertinent independent variables.

ACCURACY AND INFORMATION
RATE STUDIES ON TARGET
CROSS SECTIONS UTILIZING THE
NIKE HERCULES RADARS

GABRIEL E. GALOS
WHITE SANDS MISSILE RANGE,
NEW MEXICO

I. INTRODUCTION

The characterization of the radar cross section of target missiles is essential in the development of anti-ballistic missile defenses and also to the enhancement of the penetration capabilities of the offensive ballistic missiles.

The gathering of the radar cross section information is accomplished by measurements on static model ranges and by actual in-flight measurements utilizing instrumentation radars such as the FPS-16. Since the tactical Nike Hercules target tracking radar (TTR) is a precision tracking radar it can be used to gather radar cross section data on short range ballistic missiles in flight. This paper discusses the accuracy and data rates available from the tactical Nike Hercules TTR.

The Nike Hercules TTR is a monopulse X-Band tracking radar. The basic characteristics of the Nike Hercules TTR are classified and are not given in this report. They are readily available to persons having a need to know.

The Nike system's event recorder could be used to gather the data required for a radar cross section measurement program. Generally it is more convenient to add instrumentation recorders in order to get higher resolution and higher data rates. The addition of these recorders does not interfere with the normal tactical operation of the radar.

A block diagram of the elements of the Nike TTR and the recording equipment is given in Figure 1. The recording can be performed on analog strip charts, cathode ray tube (CRT) oscillographs and digital magnetic tape. All three methods of recording were used in the study program discussed in this report. Automatic gain control

GALOS

(AGC) voltage of the intermediate frequency (IF) amplifier was recorded on the strip chart, the CRT oscillograph and the digital magnetic tape. The pulse output of the logarithmic IF amplifier was recorded on the CRT oscillograph and the digital magnetic tape. The video pulse output of the automatic gain controlled IF amplifier was recorded on the CRT oscillograph. This video pulse is a measure of the error signal of the AGC servo. Target position was recorded on digital tape.

The data rates of the digital magnetic tape and the CRT recordings were 500 per second. This is the pulse repetition rate of the Nike Hercules TTR. The analog recorders had a nominal frequency response in excess of 100 cycles.

The study discussed in this paper was performed at WSMR on tactical systems. Targets discussed in this paper are the 6-inch sphere, a sheet of tin foil and an Arcas meteorological rocket. Actual data recorded on targets of the short range tactical ballistic missile category are not included in this report due to the higher security classification of this data. The results of the accuracy and information rates study presented in this report apply directly to all cross section measurement programs utilizing the tactical Nike Hercules TTR.

II. THEORETICAL ACCURACY OF CROSS SECTION MEASUREMENTS

The basic equation used to solve for radar cross section is the radar range equation given below:

$$\sigma = \frac{(4\pi)^3 R^4 (S/N) kT (N.F.) BFL}{P_T G^2 \lambda^2} \quad \text{Eq 1}$$

Where:

- σ = Radar cross section in square meters
- R = Range in meters
- S/N = Signal to noise ratio
- k = Boltzman's constant
- T = Absolute temperature of receiver
- N.F. = Noise figure of receiver
- B = Bandwidth of receiver
- F = Propagation losses
- L = System losses
- P_T = Peak transmitted power in watts
- G = Antenna gain
- λ = Transmitted wavelength

The accuracy with which σ can be determined is dependent on the accuracy of measurement of each of the variables in the range equation and the degrees to which they are invariant during the time interval over which the measurements are made. Most of the variables involved are deterministic in the sense that they can be measured

GALOS

with sufficient accuracy and that they are basically invariant over the time involved in any given test. There are, however, some of the variables which are stochastic in nature such as S/N and F. It is therefore desirable to calibrate the radar in terms of a known cross section (e.g. a 6-inch sphere) rather than resolve the individual system parameters. The method used in this study was to calibrate all the variables in the radar range equation which are not range dependent. If these quantities are grouped together and called K we get:

$$\sigma = KR^4 (S/N) F \quad \text{Eq 2}$$

The value of K can be obtained by a sphere track. K is given in Equation 3.

$$K = \frac{\sigma_S}{R_S^4 \frac{S}{N_S} F_S} \quad \text{Eq 3}$$

Where the subscripts S are referenced to the sphere. Inserting the value of K in Equation 2 we get:

$$\sigma_T = \left(\frac{R_T}{R_S}\right)^4 \left(\frac{S/N_T}{S/N_S}\right) \left(\frac{F_T}{F_S}\right) \sigma_S \quad \text{Eq 4}$$

Where the subscripts T and S refer to the target and sphere respectively. The validity of Equation 4 is dependent on the quantities F_T , G_L , λ , K_T (N.F.) B and L not changing between the time that the sphere track was run and the time the target was tracked. Equation 4 can be written in logarithmic form to simplify error analysis. Equation 4 becomes:

$$\begin{aligned} \log \sigma_T &= 4 \log R_T - 4 \log R_S + \log S/N_T - \log S/N_S \\ &+ \log F_T - \log F_S + \log \sigma_S \end{aligned} \quad \text{Eq 5}$$

In terms of decibels (db) Equation 5 becomes:

$$\begin{aligned} \sigma_T &= 40 \log R_T - 40 \log R_S + S/N_T \text{ (db)} - S/N_S \text{ (db)} \\ &+ F_T \text{ (db)} - F_S \text{ (db)} + \sigma_S \text{ (dbsm)} \end{aligned} \quad \text{Eq 6}$$

The symbol dbsm references the cross section to 1 square meter. It follows from differential calculus that the maximum db error in σ_T is the sum of the maximum db errors in all the parameters listed. Assuming the errors to be randomly distributed and independent, then the mean and variance of σ_T is the sum of the means and variances of the individual variables. Since some of the errors are range dependent the most significant result will be that obtained by

GALOS

summing the contributions at each range increment. Plots of the standard deviation (S) of errors (in db) for all variables and the resulting total error are given as functions of range in Figure 2.

Values for the standard deviation of the variables listed in Figure 2 were obtained from consideration of the following error sources:

1) Errors in R_T and R_S are essentially determined by the N.H. TTR range accuracy. This is essentially independent of range as long as the $S/N > 15$ db. Then the variance (error) increases until the radar breaks track (generally in angle first). In this sense range errors are dependent on both range and target size.

2) Errors in target S/N are essentially reading errors. By calibrating the S/N_T recording in 5 db increments the standard deviation in reading strip charts (analog records and CRT recordings) has been shown to be of the order of 1.0 db with $S/N > 10$ db. This reading error increases with range due to the decreased S/N . For digital recordings this error is somewhat smaller especially in the region where $S/N > 10$ db. This reading error must not be confused with the actual variations in S/N due to changes in target cross section. This factor will be discussed later.

3) Errors in sphere track S/N are also reading errors. Several factors result in a lower overall error contribution. First, the sphere track is smoother and this results in a lower variance in readings. Secondly, the target cross section is assumed constant and hence an average of S/N readings over a one second interval (or longer) can be used.

4) Of all the error contributions those due to propagation (F) are probably the most difficult to establish. Propagation effects are dependent on atmospheric conditions. In the basic calculation of cross section a correction is made for attenuation by the dry atmosphere. The errors introduced by moisture can be quite large, e.g., 2 db per 100 Kyd for light rain. In general, the moisture content of the air is low at WSMR and target cross section measurements are made on clear dry days. For this reason an estimate of error is given as 0.5 db + 0.5 db/100 Kyds.

5) Similar propagation errors could exist in determining the sphere characteristics. But here the measurements generally are made at ranges under 30 Km. The estimate of total F_S errors is 0.5 db.

6) The estimate of error in determining σ_s is less than 0.1 db since a 6-inch sphere is used for calibration. The 6-inch sphere has a circumference to wavelength ratio of about 15. Under these conditions the radar cross section is within 0.1 db of the geometric cross section. General radar theory and the results of the studies

GALOS

described in the next section are the bases for the estimates of error given above.

III. INFORMATION RATES

The Nike Hercules has a basic information rate of 500 per second. This is set by the pulse repetition frequency (PRF) of the system. It was one of the objectives of this study to determine the information rate required to determine all significant parameters associated with target cross section.

Information rates in sampling and reduction will determine not only the granularity of the discrete data but also the presence of significant target signature characteristics. Methods of recording are either analog or digital. Generally all data will be reduced to digital form in the process of analysis. The Nike Hercules in its tactical configuration has only analog outputs which can easily be recorded on strip chart records. The information readily available that is of interest in target signature determination is: AGC, S/N, target position in X, Y, H and slant range. In addition the video pulse height from the AGG/IF and the log IF (in the improved Nike Hercules) is available for recording on video tape or by CRT photography. The analog strip charts can be recorded with galvanometer recorders which have frequency response characteristics in excess of the 500 pulse per second repetition rate of the radar. With the expanded scales of the order of 2 inches per second, the records can easily be read at 20 per second. Higher chart rates (up to 160 inches per second) are available but rarely used for long runs. Practical reading rates for the normal strip chart data are 10 or 20 points per second. The pulse height information was recorded and read at rates up to 500 pulses per second. The reading of the film is time consuming and hence this type of data often is not recorded.

When the higher information rates are required it is more expedient to digitize and record on a magnetic tape in a format that can be fed directly into a standard digital computer such as the IBM 7094. In the case of the Nike Hercules this is done at the pulse repetition frequency of the radar. The analog voltages for AGC and position (X, Y, H and slant range) are converted in a standard analog-to-digital converter and recorded in IBM format. The pulse heights are converted to pulse widths and used to gate a counter. The resultant data is converted to IBM format and recorded on the same digital tape as the digital voltages. Reference timing is also recorded on the tape. The spheric calibration is recorded on the tape. This one tape contains all of the data required to establish the cross section of the target as a function of time, range and aspect angle of the velocity vector to radar line of sight. The information rate of this data is 500 per second. The data is processed by the computer.

IV. RESULT OF THE STUDIES

1) Accuracy of calibration

The purpose of calibration is to establish the calibration curve for the AGC and pulse height as functions of power input or changes in power input. The Nike Hercules has a built-in calibration device in the form of an IF test set. The calibration is performed in 5 db steps. The attenuator used to step the signal level has a nominal accuracy of ± 1 db. Listed in Table I are the standard deviation on a typical set of point by point readings of AGC and pulse height (log amp) as presented on the CRT recording.

Table I

Reading Accuracy of IF Test Set Calibration (20 Readings per Sample)

S/N	Standard Deviation AGC	Standard Deviation Pulse Height (Log Amp)
59	0.62 db	0.82 db
49	0.89	0.64
39	0.64	0.55
29	0.72	0.67
19	0.52	0.47
9	0.69	1.33
0	> 1.5	> 1.7

It can be seen that reading error can be kept small and that the overall calibration accuracy would be set by the inherent accuracy of the IF attenuator.

In the Nike Hercules there is also an RF test set which can be used to calibrate the system. Calibration with the IF test set was compared with the RF test set. It was found that the incremental changes were essentially the same but that there was a shift in relative value of the AGC and the logarithmic pulse height by 3 or 4 db when the RF test set or a real target was used. This can be explained by the fact that the IF test set pulse has some ringing resulting in a distorted pulse in the logarithmic amplifier. This does not affect the calibration accuracy since the calibration is used only to scale the power levels. The actual reference level for all records was set by the 6-inch sphere which is a real target.

2) Accuracy of the Sphere Track

The purpose of the sphere track is to establish the system constants in the radar range equation. Ideally, the sphere would be tracked at ranges close to the range at which the actual target missile measurements are to be made. Often this is difficult to achieve and it is necessary to extrapolate from ranges of 10-30 Kyds to ranges up to 100 Kyds. Extrapolated data is corrected for

GALOS

known atmospheric attenuation. In Figure 3 are presented actual plots of the measured signal returned from 6-inch spheres and the theoretical curve which is used to extrapolate the data. Generally, the standard deviation of points around this curve is of the order of 1 db.

In the determination of the signal power returned from the sphere, various records were analyzed and compared for data rate and method of reading. The means (M) and standard deviations (S) of relative received power were determined for each sample. These are presented in Tables II and III.

Table II
AGC Recordings of Sphere Track (Balloon Lofted)

Range	Strip Chart (20/sec - 40 Points)	Cathode Ray Tube (70/sec-140 Pts)	Digital Tape (125/sec-20 Pts)
2000 yds	M = 44.7 db s = 1.4 db	M = 40.8 db s = 0.97 db	M = 42.8 db s = 1.6 db
4000 yds	M = 27.9 db s = 1.4 db	M = 30.2 db s = 1.05 db	M = 31.5 db s = 1.3 db
8000 yds	M = 18.2 db s = 0.82 db	M = 19.75 db s = 1.1 db	M = 19.2 db s = 1.03 db

Table III
Comparison of AGC (With and Without Video Correction) and Pulse Height from the Logarithmic Amplifier for Sphere Track

Range	Log Amp	AGC Plus		
	Pulse Height	AGC Only	Video Correction	Data Rate
2000 yds	M = 37.3 db s = 1.76 db	M = 40.8 db s = 0.97	M = 41.1 db s = 1.55 db	140/sec
4000 yds	M = 24.6 db s = 1.2 db	M = 30.2 db s = 1.05 db	M = 29.9 db s = 1.94 db	140/sec
8000 yds	M = 15 db s = 1.25 db	M = 19.75 db s = 1.1 db	M = 18.6 db s = 1.16 db	140/sec
9200 yds	M = 11.74 db s = 1.84 db	M = 9.32 db s = 1.06 db	M = 8.82 db s = 2.1 db	500/sec

The results indicate that there is no advantage in a high data rate, in fact if only 25 pulses at 500 per second are used, the resulting standard deviation of the data can be as high as 2 db while for the longer data samples the standard deviation is close to 1 db. It is also apparent that the standard deviation of the AGC data is lower than the standard deviation of the actual pulse height data from the logarithmic amplifier. This is attributed to the

GALOS

smoothing due to the response time of the AGC servo loop. When the AGC voltage was corrected by the video pulse height (error signal of the AGC servo) the standard deviation was comparable to that of the logarithmic amplifier.

The mean value of each piece of data is given in Table III along with the standard deviations. The means of the log data are generally several db below the AGC. This is attributed to calibration techniques and is of no consequence in the use of the sphere data since the mean value of the sphere return is used as a reference for each method of recording.

Ordinarily, sphere track calibration is performed with a 6-inch sphere lofted by a weather balloon. The balloon affects the sphere return. The magnitude of this error was evaluated by comparing the standard deviation of the balloon and sphere with a sphere dropped from a B-47. It was found that the sphere drop data had a slightly lower standard deviation (0.78 db vs 1 db). The effect is considered second order. It was impossible to determine any bias since the data was not run on the same day. Data run on different days compare within 1 db. This is shown on the plot in Figure 3.

The overall accuracy with which the return power from a sphere track as a function of range is estimated as 1.0 db when 10 or more range samples are used to determine the 40 db/decade sphere curve shown in Figure 3. Standard deviations of 1 to 2 db are representative of the spread of the range samples from the 40 db per decade sphere curve.

As a final note it may be stated that on a specific radar the signal returned from the 6-inch sphere does not vary by more than 2 or 3 db from the theoretical value calculated from the basic radar range equation when the measured values of transmitter power, antenna gain, wavelength and receiver sensitivity are used.

3) Accuracy of Rapidly Scintillating Targets

In order to evaluate a rapidly scintillating target, a piece of tin foil was balloon lofted to serve as a target for the radar. The information was recorded as mentioned above. The data was analyzed using different sampling rates and different information periods.

The analysis showed that in order to adequately characterize the cross section, the information rates and periods must be set to record the changes of signal. Analysis of the foil tracks indicates that the signal level can change by ± 10 db in $1/500$ sec., that is, at the pulse rate of the radar. Figure 4 gives plots of the signal returned from the foil at the prf of the radar. The data plotted indicates clearly that the AGC without video pulse correction (servo error correction) has greater excursions than the pulse height

GALOS

as recorded by the log amplifier. When video correction is added, the two curves are in reasonable agreement. The dynamic range of the video error signal is not sufficient to make corrections in excess of ± 10 db. When the 25 readings used in the sample shown in Figure 4 were analyzed the following statistics resulted:

	Mean	Standard Deviation
AGC Only	29 db	5.8 db
Log Pulse Height	32.6 db	4 db
AGC + Video Correction	31 db	4.08 db

It is evident that the mean and standard deviation of the AGC when corrected by the video error are in reasonable agreement with the log pulse height. The implications are that the AGC alone should not be used to develop short term (fine resolution) information on rapidly scintillating targets.

The Arcas missile, which is a meteorological sounding rocket, was used as an example of missile targets. The signal return fluctuations were similar to those of the foil. The resulting statistics on the comparison of the AGC alone, AGC with video correction and log pulse height were similar to those on the foil.

In order to determine the significance of the data rate and data interval a section of the Arcas track was analyzed by reading at different rates and for different time intervals. The following statistics were developed:

a) With 20 readings at 10 per second the mean was 23.4 db and the standard deviation was 6.2 db.

b) With 20 readings at 100 per second the means on consecutive samples in the same 2 second time interval covered in (a) above varied from 18.7 db to 27.7 db. The standard deviations varied from 2.4 to 5.6 db. Obviously, the long term statistics do not adequately describe the target cross section.

V. CONCLUSIONS

The above studies indicate that log pulse height information at the repetition rate of the radar must be used in order to determine the detailed statistical characteristics of missile targets. On the other hand long term (order of 1 second) average characteristics can be determined by use of AGC data alone. For slowly fluctuating targets, the AGC data is quite adequate. The overall accuracy of the determination of radar cross section is dependent on range and S/N. Figure 2 gives the best estimates of errors for a radar cross section measurement program utilizing the Nike Hercules target track radar.

GALCS

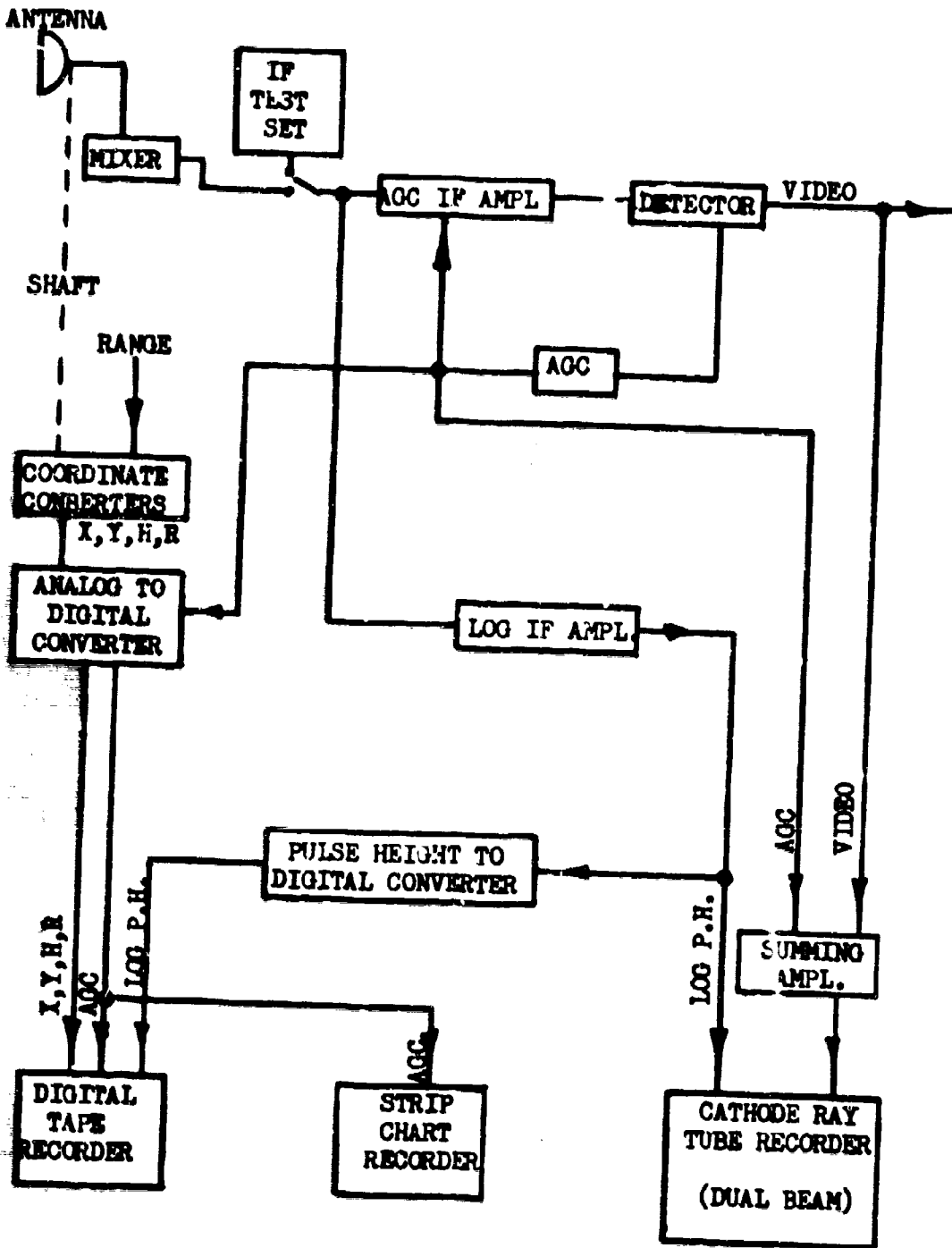
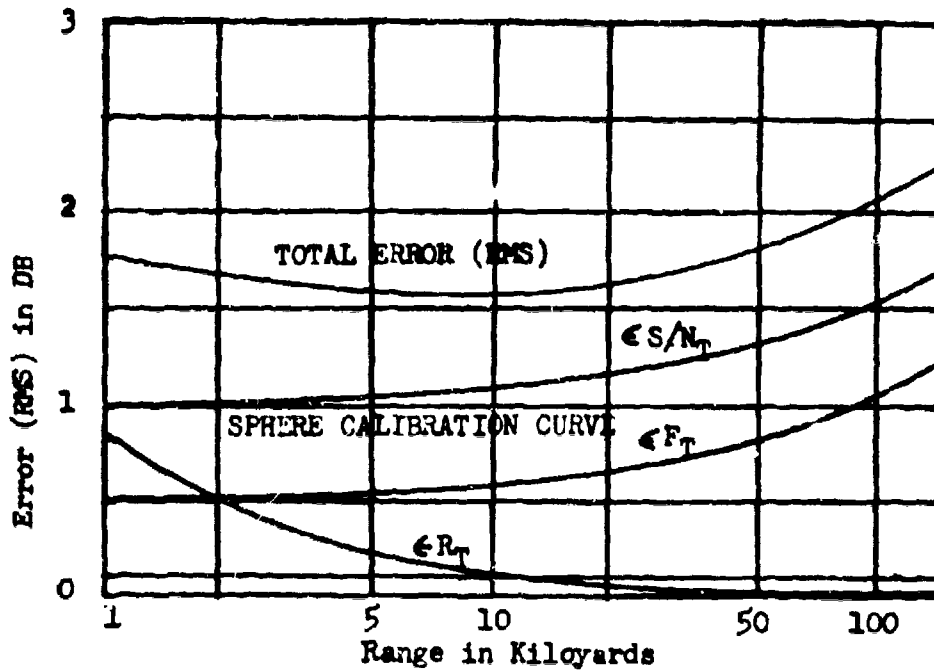


FIGURE 1

NIKE TTR ELEMENTS AND RECORDERS

GALOS



Where:

- ϵ_{S/N_T} = RMS TARGET S/N READING ERROR
- ϵ_{F_T} = RMS PROPAGATION ERROR FOR TARGET
- ϵ_{F_S} = RMS PROPAGATION ERROR FOR SPHERE
- SPHERE CALIBRATION CURVE-- RMS ERROR IN ESTABLISHING THE SPHERE CALIBRATION CURVE (95% CONFIDENCE)

FIG. 2

ERROR CONTRIBUTIONS FOR N.H. TTR
RADAR CROSS SECTION MEASUREMENTS
(FOR EACH SAMPLE OF TARGET DATA)

GALCS

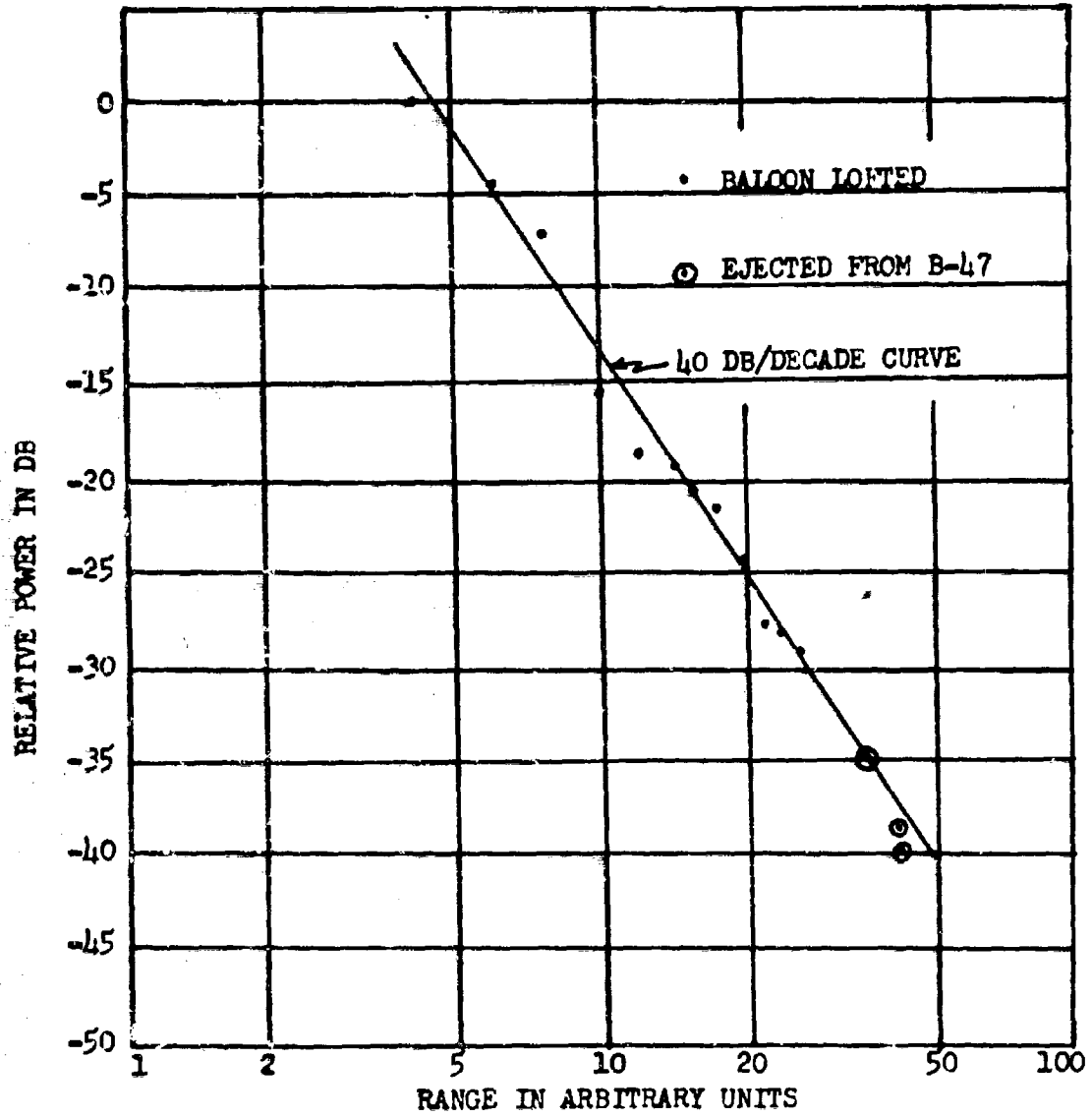


FIG. 3
SPHERE TRACK OF 6 INCH SPHERE

GAIOS

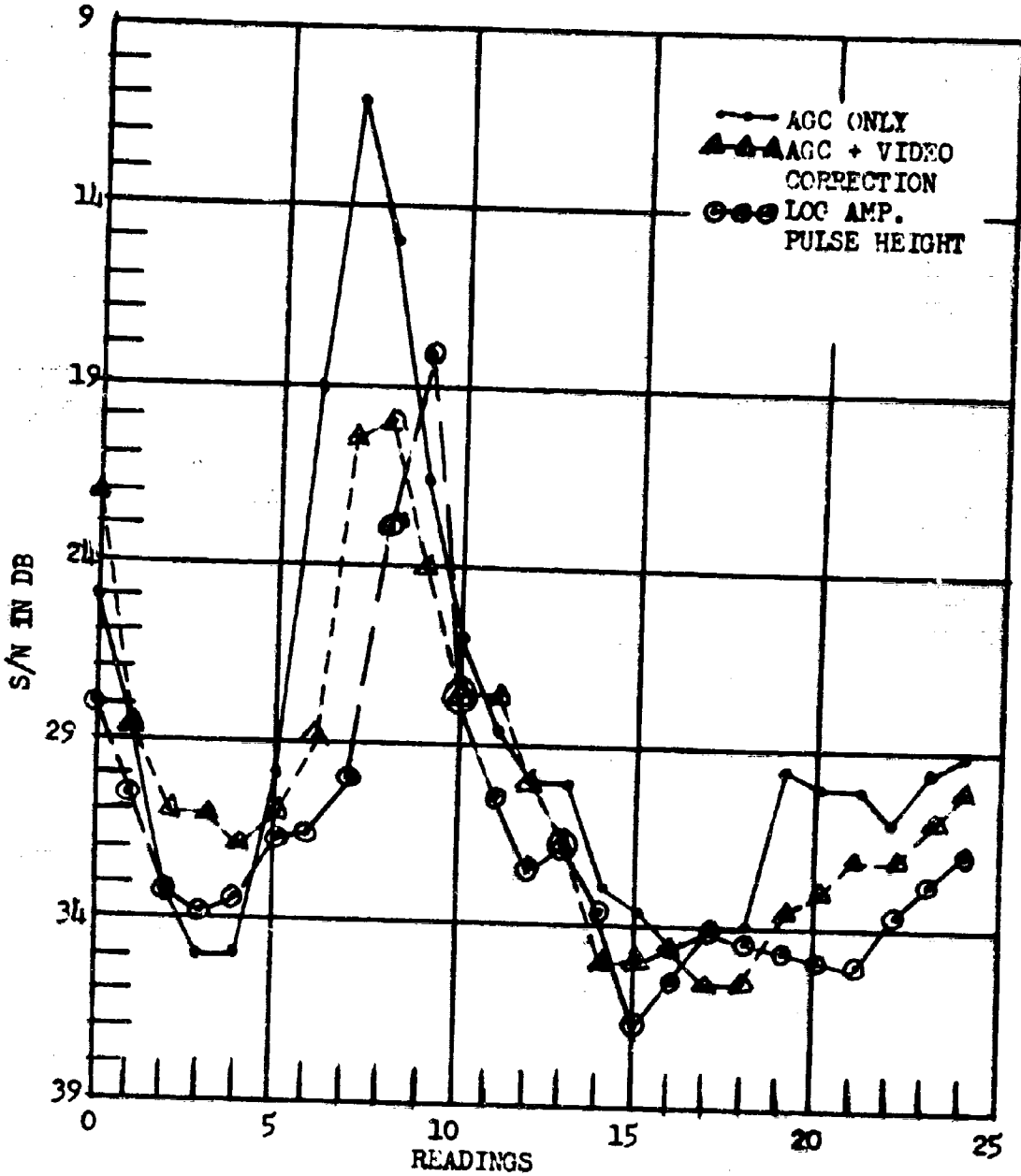


FIG. 4

S/N OF FCIL TRACK AT 500/SEC.
(25 CONSECUTIVE READINGS)

SUPERCONDUCTORS IN ADVANCED ELECTRONICS

GERHART K. GAULÉ, JAMES T. BRISLIN
RAYMOND L. BOSS and ROY S. LOGAN, 2d Lt, USAR
U. S. ARMY ELECTRONICS RESEARCH AND DEVELOPMENT LABORATORIES
FORT MONMOUTH, NEW JERSEY

1. INTRODUCTION

A superconductor is completely free of ohmic losses for a direct current and nearly free of such losses for alternating currents with frequencies up to Mc/sec and higher, depending on the circumstances. Circuits with superconducting inductors and capacitors, and superconducting interference shields for such circuits thus appear as a natural choice where a minimum of loss, noise, and interference from the environment is desired, as for example in the first stage of a communications receiver. Superconducting resonant circuits with very high Q values (10^6 and higher) have indeed been successfully operated in the Mc/sec range (1), (2), but the tuning of such circuits proved to be difficult. It was recognized at USAELRDL (3) that remotely controlled switches would solve the tuning problem provided the switches would introduce no loss into the circuit. The "cryotron," a superconducting computer element (4), has often been called a superconducting "switch," but it is really only a current gate with zero resistance in the "closed" position and a finite resistance--a few hundred ohms at most--in the "open" position. A superconducting switch with a true "open" position (virtually infinite resistance) was, however, deemed necessary for tuning and, in fact, for most purposes. It was predicted (5) that a novel combination of superconducting solenoids with superconducting contacts would constitute a remotely controlled superconducting switch with the desired properties. Experimental work was begun in fall 1963 which soon resulted in the construction of an actuating mechanism capable of giving contact forces up to 1000 g (Section 2). The search for a suitable superconducting contact material and a suitable contact configuration proved much more difficult, but contacts which would reproducibly control supercurrents of 16 A had been developed in spring 1964 (Section 3). This surprisingly high current value permits the use of the new relay switch not only for the low power applications mentioned in the beginning but also for high power applications, especially in

**GAULÉ, BRESLIN, ROSS and
LOGAN, 2d Lt, USAR**

connections with the now widely used superconducting solenoids for the generation of high magnetic fields. Using the analogy between a hypothetical lossless capacitor and a superconducting inductor or solenoid, basic circuits for the "charging" and "discharging" of a superconducting solenoid were developed (6) and some of them tested (Section 4).

The low operating temperatures of the presently used superconductors often prevent the application of superconducting devices and circuits in spite of their obvious advantages. Recent progress in this area as a result of simplified cooling methods and of improved materials is outlined in Section 5.

2. RELAY MECHANISM WITH SUPERCONDUCTING COILS

The mutual attraction and repulsion of two similar superconducting coils were used to actuate the first experimental relay switch built by the authors, shown in Fig. 1. Other possible magnet configurations have been discussed elsewhere (5). No soft iron return path was used for the magnetic flux, because this would increase the cooling requirements for the relay, and also because the high magnetic fields generated by superconducting coils tend to magnetically saturate the iron, thus making it useless.

Each magnet coil was made by winding approximately 1300 turns of molybdenum-rhenium wire on a nylon bobbin with 1.6 cm outer diameter. Molybdenum-rhenium is a "hard" superconductor, that is, a material remaining superconducting in a strong magnetic field and when carrying a large current.

The relay consists of two simple bakelite parts, one stationary and one movable, each carrying one coil and one contact, (Fig. 1). The magnetic force, magnified by lever action, acts on the two contacts. To measure the contact force, as a function of coil current, a nylon string was attached to the upper contact and connected to a balance on top of the cryostat in which the relay was kept at the temperature of 4.2°K. Some of the results are given in Table II, Section 3. The force increases strongly with the coil current and reaches ≈ 1000 g for the maximum current carrying capacity of the coils (≈ 6.4 A). A conventional copper wire of equal thickness (0.25 mm) in place of the Mo-Re wire would sustain no more than ≈ 0.1 A. It is this difference in operating current for a given winding which makes superconducting coils so superior to conventional ones in producing mechanical forces for the operation of contacts and other purposes.

3. CONTACTS FOR LARGE SUPERCURRENTS

In 1914 Kamerlingh-Onnes, the discover of superconductivity, made the first superconducting "switch" when he pressed two

GAILLÉ, BRISLIN, ROSS and
LOGAN, 2d Lt, USAR

pieces of superconducting lead together and found superconductivity across the contact (7). Later workers found that an oxide layer and even an insulating layer of moderate thickness (up to 10^{-7} cm) between the contact surfaces does not prevent superconductivity across the contact (7),(8). This surprising effect was explained theoretically only recently by Josephson (9) ("Josephson-Tunneling").

When the current through a superconducting switch is raised, a critical current value, I_{cr} , is finally reached at which the superconductivity across the contacts vanishes. This becomes evident as a voltage across the contacts. I_{cr} can be determined by observing this voltage with separate voltage leads, as shown in Fig. 1. All the switches reported in the literature had small I_{cr} values, usually much less than 1 A. Various contact pairs used in the initial phase of this work were found superconducting, but with disappointingly small I_{cr} . Results for three typical contact pairs are listed in Table I.

TABLE I

contact 1	contact 2	contact force (g)	I_{cr} (A)
multifinger arrangement of Nb wire (0.4 mm)	Nb - sheet	650	$2 \cdot 10^{-4}$
		750	$2 \cdot 10^{-4}$
Pb-Sn alloy (40/60), ball shaped	Nb - sheet	350	$2 \cdot 10^{-3}$
		400	$11 \cdot 10^{-3}$
		550	$73 \cdot 10^{-3}$
		800	$74 \cdot 10^{-3}$
Pb-Sn alloy (40/60), ball shaped	Pb-Sn alloy (40/60) ball shaped	550	$18 \cdot 10^{-3}$
		850	$18.5 \cdot 10^{-3}$

It was deduced from these data that (a) increasing the contact force beyond ≈ 500 g gives little improvement, and that (b) at least one of the contacts should be of a soft material, such as the lead-tin alloy used in two of the experiments. It is known, however, that even this alloy becomes rather hard when cooled to liquid helium temperature. Pure lead was chosen therefore for the next experiments. Lead retains a plasticity at 4°K comparable to that of copper at room temperature. The plasticity of a contact material appears to secure the formation of an "effective" contact area larger than that obtainable with a hard material (7). The contacts finally used in this work were of lead galvanically coated with a very thin layer of tin. The contacts were kept under helium gas during the cooling down period. These measures were aimed at preventing the formation

GATLIN, BRISLIN, BOSS and
LOGAN, 2d Lt, DEAR

of layers of oxide, ice, or frozen air on the contact surfaces. The results with tin coated load were excellent and vastly superior to any previous results, as shown in Table II. This table also gives the contact force as a function of coil current.

TABLE II

coil current (A)	contact force (g)	I_{cr} (A)
1.5	350	< 1
2.3	475	2.3
3	600	13
4	800	13.5

These results were quite reproducible. In one run an I_{cr} value of 16 A was attained. Technical refinements of the crude experimental set up, shown in Fig. 1, will undoubtedly improve the performance of the superconducting relay switch further.

4. APPLICATIONS

The successful operation of radio frequency and of microwave circuits made of superconducting components has been reported by other groups (1),(2),(10). The new superconducting relay switch described in the previous section now makes possible the tuning of such circuits by remote control. Several switches could serve, for example, to add capacitors or inductors to a superconducting, high Q resonant circuit. Since the switches have zero resistance in their "closed," and virtually infinite resistance in their "open" position, the switches will not introduce any losses into the circuit (3). An older switching device, the "cryotron," could not be used for this purpose, because it has no true "open" position, as discussed in Section 1.

The first step toward high power applications of the new superconducting switch was taken by the authors when they connected the new switch to the terminals of a high field (42 kOe) superconducting solenoid, and used it to control the solenoid current. The solenoid is of medium size (1.25 cm bore) and has an inductance of ≈ 5 Henry. The highest supercurrent which could be controlled so far was 12 A. One purpose of these experiments is to explore the storage of electrical energy with the solenoid. The storage of electrical energy in low loss capacitors has been used in electronics for many years for the operation of pulsed transmitters, of pulsed light sources, and also in plasma research. A superconducting solenoid can be considered as an inductor which is lossless in low frequency applications. The properties of such an inductor are set in analogy to those of a hypothetical leakage free capacitor in Fig. 2. The quantum of charge for the capacitor is of course e , the electronic

GAULÉ, BRUELLIN, BOSS and
LOZAN, 2d Lt, USAR

charge; its equivalent for the inductor is the newly discovered (11) quantum of magnetic flux, $hc/2e$ (h : Planck's constant, c : velocity of light). Since both "charge quanta" are very small, they are not observed except under very special conditions and the charging processes can be treated as continuous in the following. Charging the capacitor is accomplished through the transfer of electric charge between the capacitor plates, resulting in a voltage between the plates. "Charging" the inductor is accomplished by building up a current through the winding, resulting in a magnetic flux threading the hole of the inductor. The charged state of the capacitor is characterized by an open circuit, the analogous state of the inductor by a short circuit (produced by the closing of a superconducting switch). The maximum charge of the capacitor is limited by dielectric breakdown, the maximum magnetic flux of the inductor is limited by the field induced breakdown of the superconductivity of the winding. Technically attainable values for either kind of breakdown are used in Fig. 2 to demonstrate that the energy, w , which may be stored per unit volume of the device is much greater in the case of the solenoid. Formulas for the total amounts of energy, W , storable in either case are also given in Fig. 2. The largest superconducting solenoids which are now commercially available have bores greater than 15 cm and inductances up to 1000 Henry. Carrying a typical current $I \approx 20$ A, such a solenoid stores $W \approx 2 \cdot 10^5$ Joules. The corresponding value for a capacitor of comparable size and weight would be much smaller.

The bottom part of Fig. 2 illustrates the transfer of the stored energy into an outside load R' . For the capacitor, this is accomplished by closing a switch, for the inductor by opening a superconducting switch. Until now, only thermally activated cryotrons were used in connection with superconducting solenoids. The resistance of a "thermal" cryotron can be varied only slowly (within ≈ 20 sec) from zero to the maximum value of a few ohms. Most of the stored energy is thus consumed by the cryotron itself and transfer to an outside load is ineffective. The discharge process presented in the right bottom part of Fig. 2 is possible only with a superconducting switch which goes from "zero" to "infinite" resistance immediately, such as the relay switch built and used by the authors.

Figure 3 shows the two analogous basic charging circuits for alternating current. The timing for the periods of charging is indicated. The superconducting switch across the inductor is open only during the charging period. When it is closed, no voltage is applied to the inductor and its current not influenced by the power source. The charging is completed when the supercurrent through the inductor equals the peak current of the power source.

Figure 4 describes two circuits employing current transformers. One winding of a current transformer is outside of the cryostat. The energy transfer is through the wall of the cryostat by induction. This is superior to the direct transfer of energy through

GAULÉ, BRISLIN, ROSS and
LOGAN, 2d Lt, USAR

heavy current leads which are difficult to cool. Except for the current transformer, the first circuit in Fig. 4 operates like that in Fig. 3. The second circuit in Fig. 4 is analogous to a vibrator circuit with capacitors. It converts the stored direct current into alternating current by using a reversal switch S2, and by immediately storing energy in the superconducting transformer winding W2 combined with the closed superconducting switch S3. The timing of the switches for discharge is indicated. With different timing, the same circuit is used for charging.

Figure 5 indicates the storage of energy in a superconducting combination S and L as before, but a switching transistor T1 outside of the cryostat is connected parallel to S. In the "standby" position, S is opened and the solenoid current is forced to flow through T1. The storage is then of course no longer loss free, but the loss during a short standby period is tolerable. For operation, the switching transistors T1 and T2 are alternately opened and closed with the resonance frequency of the transmitting circuit, thus converting the stored energy into a very powerful radio frequency signal.

The equations at the bottom of Fig. 2 indicate that a "discharging" inductor behaves essentially like a constant current generator. This feature, together with others listed above, makes a superconducting solenoid combined with a superconducting switch particularly attractive for the operation of pulsed light sources. At the onset of the pulse the resistance of such a device is usually high; the solenoid then applies a high "igniting" voltage. Later on, when the resistance of the device drops, the voltage drops accordingly, and the initial current is never exceeded. This is in contrast to the action of a capacitor--essentially a constant voltage generator--which tends to apply too much voltage after "ignition," so that much of the pulse energy must be destroyed in a stabilizing resistor or the like. Superconducting solenoids and relay switches capable of delivering pulses of several hundred volts and many thousands of Joules should be obtainable. Fig. 6 shows a simple circuit for pulsed light sources, based on principles similar to those of Fig. 5.

5. COOLING PROBLEMS

In the usual laboratory operation of a superconducting circuit the circuit is housed in a cryostat which is cooled by transferring liquid nitrogen and liquid helium into appropriate compartments of the cryostat. The transfer of a cryogenic liquid is a difficult and time consuming manipulation which is necessary because the consumer of the cooling energy, the circuit, is separated from the source of the cooling energy, namely the liquefier for the helium or nitrogen. In this fashion several cryostats can be served from one liquefier, permitting simultaneous work on several experiments. The laboratory method of cooling, however, is neither necessary nor even suitable in the technical application of

GAUL, HREBLIN, BOSS and
LOGAN, 2d Lt, USAF

superconducting circuits. Here, the best approach is obviously through a combination of the cryostat with the liquefier (or another source of cooling energy), such that fully automatic operation without any external transfer of cryogenic liquids is accomplished. Refrigerators of this advanced kind, called "closed cycle refrigerators" are now commercially available in various sizes, even rather small ones. A closed cycle refrigerator has several "cooling stations," essentially cooled metal plates. The superconducting circuit is mounted of course at the station with the lowest temperature. Electrical leads, coaxial cables, and waveguides connected to the superconducting circuit are also thermally linked to the stations with intermediate temperatures. In this fashion, the heat flowing into the cryostat is "intercepted" at a relatively high temperature level. The thermodynamical efficiency of a closed cycle refrigerator is therefore excellent even when heavy cables or waveguides or the like are used.

Progress would be still greater, of course, if it were possible to operate the superconducting circuit itself at a higher temperature. Although many superconductors with transition temperatures, T_c , between 15°K and 18°K are known (12), In ($T_c = 7.2^\circ\text{K}$) and similar metals are predominantly used now in electronic circuits. As discussed in Section 5, lead is also the best contact material known. If a search for materials with satisfactory high frequency and contact properties among the superconductors with $T_c \geq 15^\circ\text{K}$ proves successful, the operating temperature could be raised accordingly. This would in turn permit at least a tenfold reduction in refrigerator input power.

6. CONCLUSION

The low power as well as the high power applications of superconductivity are increasing. Both are expected to benefit from the high current relay switch developed by the authors. In contrast to previous "switching" devices, the new switch has a true "open" position, and is thus loss free in either position. A novel method for the storage of electrical energy is given by connecting the new switch with a high current, high inductance superconducting solenoid. The basic circuits are derived from analogous capacitor circuits. The inductive energy storage is superior to the capacitive one in the operation of pulsed light sources, for example. A sharp distinction should be made between the cooling methods suitable for the laboratory and those for technical applications. For the latter, greater use should be made of the now available closed cycle refrigerators. Higher operating temperatures for superconducting circuits are sought through improved materials. The development of special low loss dielectric materials for these circuits is also suggested.

7. ACKNOWLEDGMENT

The authors wish to thank Dr. Hans Meissner, Stevens Institute of Technology, Mr. Charles Nixon, Army Ballistic Missile Agency, Huntsville, Alabama and Mr. J. Mellichamp, Mr. M. Woodruff and Mr. L. Dathe of USAELRDL for stimulating discussions. The authors wish to express their appreciation to Mr. Lester Wilcox for fabrication of the first experimental relay switch.

8. BIBLIOGRAPHY

- (1) W. H. Hartwig, "Superconductive Frequency Control Devices." Final report, DA-36-039-AMC-00036(E) August, 1963. W. H. Hartwig, paper presented to the Frequency Control Symposium, Atlantic City, N. J., May, 1963.
- (2) F. L. Arams, Cryogenic "Radio Frequency Tuner." DA-36-039-AMC-03321(E).
- (3) G. K. Gaulé, L. C. Dathe, "Switching System for Superconducting Circuit," invention disclosure, U. S. Army Electronics Patent Activity, Docket #14,238.
- (4) T. W. Bremer, "Superconducting Devices," McGraw-Hill Book Co., New York, 1962.
- (5) G. K. Gaulé, M. W. Woodruff, R. S. Logan, "Remotely Controlled Superconducting Switches and Actuators," invention disclosure, U. S. Army Electronics Patent Activity, Docket #14,713.
- (6) G. K. Gaulé, "Switching Circuits for the Charge and Discharge of Superconducting Solenoids," invention disclosure, U. S. Army Electronics Patent Activity, Docket #14,718.
- (7) R. Heim, Electric Contacts Handbook, Springer Verlag, Berlin, 1958, page 128.
- (8) Isolde Dietrich, Zeitschr. Physik 132, p. 231 (1952); Zeitschr. Physik 133, p. 499 (1952).
- (9) B. D. Josephson, Phys. Letters, 1, 251 (1962).
- (10) R. J. Allen, "Microwave Memory Techniques Project, "Final Report, AF30(602)-2414.
- (11) B. S. Deaver, W. M. Fairbanks, Phys. Rev. Letters, 7, 43 (1961). R. Doll and M. Nabauer, Phys. Rev. Letters, 7, 51 (1961).
- (12) G. K. Gaulé, "Rules for the Occurrence of Superconductivity among the Elements, Alloys and Compounds," USAELRDL Technical Report 2329.

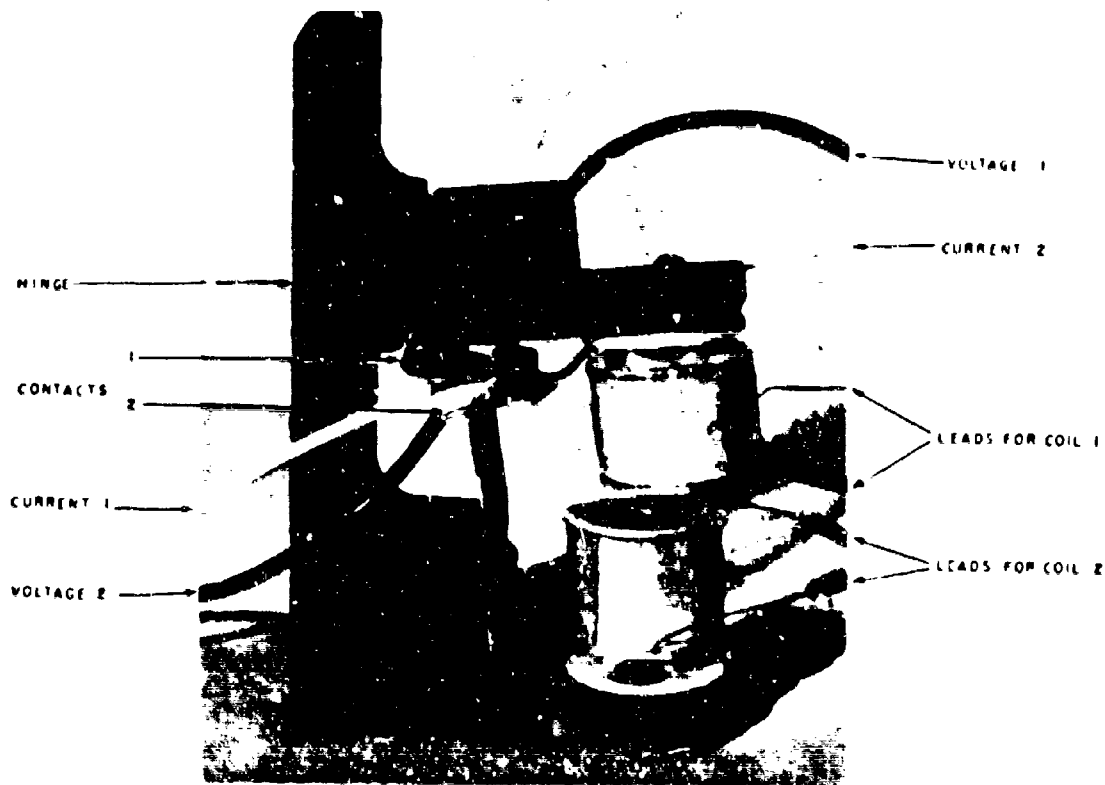


Fig. 1

Experimental superconducting relay switch. Coils have approximately 1.6 cm diameter and 1300 turns of Mo-Re wire, 0.25 mm thick. Critical current of contacts, $I_{cr} \approx 14$ A. Voltage leads on contacts are used to determine I_{cr} .

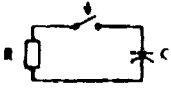

ONE TO ONE CORRESPONDING ENTITIES	CORRESPONDENCE BETWEEN LEAKAGE FREE CAPACITOR AND (SUPERCONDUCTING INDUCTOR)	
QUANTIZED CHARGE (FLUX) CHARGING PROCESS	$Q_0 = e = 1.6 \cdot 10^{-19}$ A SEC $Q = \int I dt = UC$ C IN FARAD = A SEC/V	$\phi_0 = hc/2e = 2.1 \cdot 10^{-15}$ V SEC $\phi = n \phi_0 = \int U dt = IL$ L IN HENRY = V SEC/A
LIMITATION OF CHARGE (FLUX)	$U < U_{MAX}$ DIELECTRIC BREAKDOWN	$I < I_{MAX}$ BREAKDOWN OF SUPERCONDUCTIVITY
TOTAL ENERGY STORED (IN JOULE=V A SEC) ENERGY STORED PER UNIT VOLUME (IN JOULE/cm ³)	$W = U^2 C / 2 = QU / 2$ $w = \frac{1}{2} \epsilon E^2 / 2$ $w = 0.04$ FOR $E = 10^6$ V/cm	$W = I^2 L / 2 = I \phi / 2$ $w = \frac{1}{2} \mu_0 H^2 / 2$ $w = 60$ FOR $H = 10^3$ A/cm
DISCHARGE THROUGH OHMIC LOAD, S, SWITCH CLOSSES (OPENS) AT t=0, PRIMED QUANTITIES, U', I' REFER TO LOAD		
FOR t < 0	I = I' = 0 U = U' = CONST U' = 0	U = U' = 0 I = I' = CONST I' = 0
FOR t > 0	I = I' = -U/R = -Q/RC U = U' = U_0 EXP (-t/RC)	-U = U' = -UR' = -\phi/IL I = I' = I_0 EXP (-tR'/L)

Fig. 2
 Analogies between the storage of electrical charge in a hypothetical
 leakage free capacitor (center) and the storage of magnetic flux in
 a superconducting inductor (right).

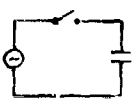
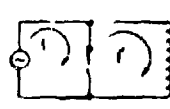
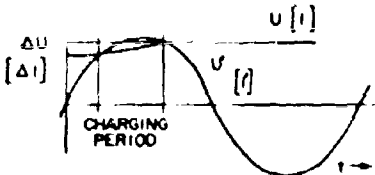
CORRESPONDING ENTITIES	LEAKAGE FREE CAPACITOR	SUPERCONDUCTING INDUCTOR
BASIC AC CHARGING CIRCUIT, WITH CONSTANT VOLTAGE (CONSTANT CURRENT) SOURCE		
		
CHARGING PERIOD BEGINS, SWITCH CLOSSES (OPENS) WHEN	U' = U + \Delta U	I' = I + \Delta I
DURING CHARGING PERIOD	I = I' > 0	U = U' > 0
CHARGING PERIOD ENDED, SWITCH OPENED (CLOSED) WHEN	I = I' = 0	U = U' = 0

Fig. 3
 Corresponding a.c. charging circuits, as in Fig. 2.

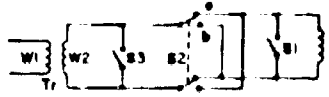
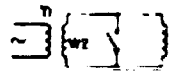
CHARGING CIRCUIT SIMILAR TO ABOVE, BUT WITH AUXILIARY TRANSFORMER T2 WINDING, W2 IS SUPERCONDUCTING

THIS CIRCUIT CAN ALSO BE USED FOR DISCHARGE, WHEN CONNECTED TO RESONANT LOAD CIRCUIT

CIRCUIT WITH REVERSAL SWITCH EQUIVALENT TO VIBRATOR CIRCUIT WITH CAPACITORS

W2 IS SUPERCONDUCTING AND STORES CURRENT I', FLUX Φ' UPON REVERSAL AND OPENING OF S2, FLUX IN T2 CHANGES TO Φ'' , PRODUCING VOLTAGE IN W1

SWITCHING SEQUENCE FOR DISCHARGE SEQUENCE FOR CHARGING ANALOGOUS SEQUENCE ALTERNATES WITH S2 IN POSITION 'a' AND S2 IN POSITION 'b'



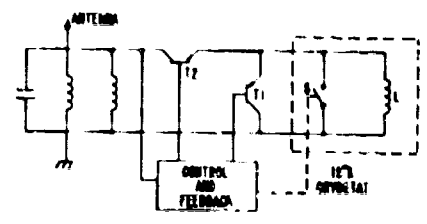
	S1	S2	S3
STANDBY a	1	0	0
DISCHARGE	0	1	0
	1	1	0
REVERSAL	1	1	1
STANDBY b	1	0	1
CONTINUED AS ABOVE			b

Fig. 4 Charging circuits for superconducting solenoids using superconducting switches and transformers. Primary transformer winding (W1) couples through wall of cryostat. Heavy current leads into the cryostat are thus unnecessary.

EMERGENCY TRANSMITTER FOR HIGH POWER PULSES

LOSSLESS STORAGE IS PROVIDED BY SHUNTING THE SUPERCONDUCTING INDUCTOR L WITH SUPERCONDUCTING SWITCH S. IN THE STANDBY POSITION CURRENT IS TRANSFERRED OUT OF THE CRYOSTAT AND GOES THROUGH SWITCHING TRANSISTOR T1 WITH $R \sim 10^{-3}$ OHMS FOR THE EFFECTIVE RESISTANCE OF T1. DECAY IN THE STANDBY POSITION IS DESCRIBED BY A TIME CONSTANT $\tau = L/R \sim 10^4 \dots 10^6$ SEC

OSCILLATION THROUGH OPENING AND CLOSING OF T1 AND T2 WITH RESONANCE FREQUENCY IN CONVENTIONAL MANNER



	S	T1	T2
LOSSLESS STORAGE	1	1	0
STANDBY	0	1	0
START OF OSCILLATION	0	0	1
OSCILLATION WITH RESONANCE FREQUENCY T1, T2 CONTROLLED BY FEEDBACK NETWORK	0	1	0
	1	0	1
	0	1	0
LOSSLESS STORAGE	1	1	0

Fig. 5 Example of potential future application of energy storage in superconducting solenoid. Lossless storage is given with superconducting switch S. Energy is lost only during the standby and operative periods, when solenoid current is controlled by switching transistors T1 and T2.

GAUL, HERSLIN, ROSS and
 LOGAN, 2d Lt, USAR

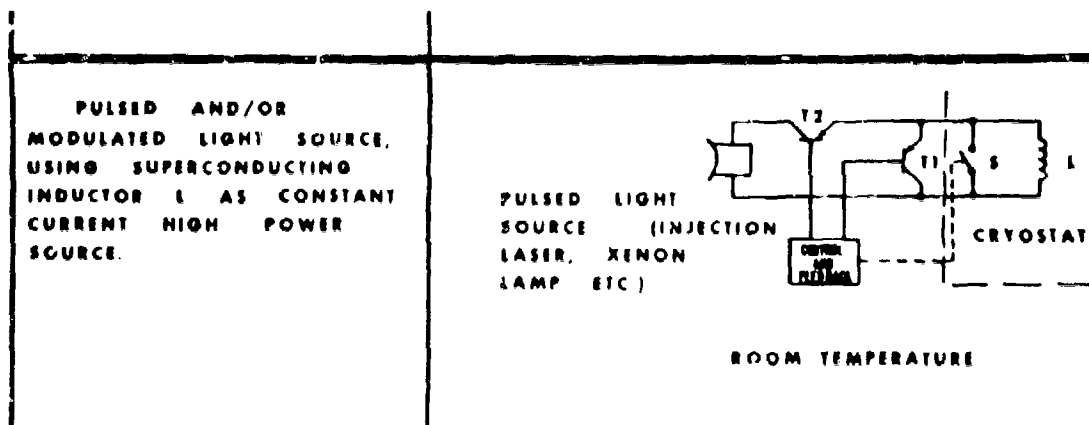


Fig. 6

Operation of pulsed light source from superconducting solenoid, similar to Fig. 4. In contrast to a capacitor bank, a superconducting solenoid is essentially a constant current source which is more suitable for many pulsed light sources.

RUBY LASER WITH VIBRATING REFLECTOR

E. A. GERBER and E. R. AHLSTROM*
U. S. ARMY ELECTRONICS RESEARCH
AND DEVELOPMENT LABORATORIES
FORT MONMOUTH, NEW JERSEY

INTRODUCTION

It is well known that conventional solid state lasers exhibit a temporal behavior of the output power which is completely random. Figure 1 shows a typical trace of a ruby laser where the time scale on the abscissa is 50 μ sec/cm. It is thought that this type of laser output is caused by the many electromagnetic modes existing in the laser cavity. According to Tang, Statz, and de Mars (1), multi-moding in solid state lasers can be explained as due to slow spatial cross-relaxation. Figure 2 illustrates this thought. Assume that just one axial mode is excited with nodal planes half a wavelength apart parallel to the mirrors. The induced emission produced by this standing wave is zero at the nodal planes and maximum in between. Under the assumption that no excitation can diffuse spatially, one obtains an inverted population that has a maximum at the nodal planes and minima in between. This inverted population distribution is not favorable for the first oscillating mode since in such places where the electric field is maximum there is a minimum density of inverted states. If the second oscillating mode is considered which has one half-wavelength more between the mirrors, it is easily recognized that the distribution of excited states is more favorable to this mode. In the middle of the crystal in particular, the maximum of the electric field and the maximum in the density of excited states coincide. This mode will thus have more gain than the one already oscillating and, therefore, will also go into oscillation. The above arguments can be continued for a third, fourth, etc., mode and thus multi-moding in solid state lasers can be explained.

GERBER and AHLSTROM

Tang, Stutz, and de Mars have further shown that multimoding caused by the described nonuniform spatial distribution in the inverted population can be avoided by using a traveling wave laser. Traveling waves de-excite the chromium ions in a spatially uniform manner precluding the simultaneous existence of many axial modes. In addition, single mode operation is expected to result in a regular spike pattern with exponential decay, a fact which has been experimentally verified by Tang, Stutz, and de Mars (1).

PURPOSE OF INVESTIGATION AND INSTRUMENTATION

It was thought that an effect similar to that caused by the traveling wave laser may be obtained if at least one reflector of the Fabry-Perot cavity is vibrated with sufficient amplitude and frequency. This measure could at least partially smear out a nonuniform distribution of excited states. To prove this out, experiments have been performed with a ruby laser with detached reflectors. Figure 3 shows the arrangement of the entire apparatus. The ruby rod was held between two brass sleeves and was pumped by a General Electric FT 524 helical flash lamp. One reflector was a stationary dielectric mirror with a reflectivity of 0.99, the other reflector provided the frequency sweeping mechanism: it consisted of an X-cut quartz vibrator. It was held at its circumference at three points and its silver coating served as electrode and mirror, with a reflectivity of 0.95. Vibrators with a resonance frequency of 1, 2, or 3 mc were used. They had a diameter of 38 mm and were driven in a thickness extensional mode of vibration. A solar cell with a parallel resistor to increase its frequency range, in combination with an oscilloscope, was used to record the power output of the laser. The total energy output of the laser was measured by observing the loss of charge of a capacitor through a phototube illuminated by the laser beam, as described by Schiel (2)

FIRST RESULTS, DETERMINATION OF VIBRATOR AMPLITUDE

Figure 4 displays a recording of the laser output with the mirror stationary, lower trace, and the mirror vibrating, upper trace. It can be seen that a regular spike pattern with exponential decay appears when the mirror is vibrating. These results are typical for many measurements of the spike structure which were

GERBER and AHLSTROM

made as a function of frequency and amplitude of the vibrating mirror, pumping power, and misalignment of the dielectric mirror. The amplitude of vibration was determined from the known properties of the vibrator and the applied HF current. Since the crystal was always driven at its series resonance frequency, a simple relationship between the amplitude ξ and the current i through the crystal holds (3), (4):

$$\xi = \frac{\pi^2}{4} \frac{\epsilon^S}{\omega C_0} i \quad (1)$$

where: ϵ^S is the dielectric permittivity at constant strain,
 e the piezoelectric stress constant,
 C_0 the static capacitance, and
 i the current through the crystal.

For the X-cut used, the value for ϵ^S is $4.5 \cdot 8.854 \cdot 10^{-12}$ farad meter and e is 0.173 coulomb meter⁻². The factor $\frac{\pi^2}{4}$ accounts for a sinusoidal distribution of the amplitude over the surface of the crystal which can be assumed to be present due to the boundary conditions employed (5).

EXPLANATION OF REGULAR SPIKING

Regular spiking was observable with an amplitude as small as .08 microns equal approximately to 1/10 of the ruby wavelength. Since this very small displacement will probably not be sufficient to cause any change in the axial distribution of excited states, unlike the experiment with a moving ruby rod recently reported (6), the following explanation for the reported observation is suggested. The quartz crystal represents, during one period of vibration, a plane, convex and concave mirror and thus, according to detailed theoretical calculations by Fox and Li, changes the losses for the oscillating modes in fast sequence (7). Since the losses are connected with the spot size on the mirrors, it is concluded that a rapid change of the spot size and with it the diameter of the oscillating modes is the cause for a smoothing-out process of nonuniform spatial distributions in the inverted population of the off-axis modes mainly. Obviously, if the amplitude of vibration attains greater values, axial modes will also be affected.

Pure optical effects due to the curvature of the vibrating mirrors are not sufficient to explain the regular spiking. The use of bimorph ceramic bender which did not vibrate but permitted the change of the radius of curvature of the mirror under static conditions from 20 m to infinity did not effect any change in the irregular spike output.

PROPERTIES OF SPIKE PATTERN AS FUNCTION OF PUMP POWER

Figure 5 shows a plot of the decay time T_d and the spike distance T_s of the regular spike pattern of a vibrating mirror laser as a function of the normalized pump power. The solid curves are obtained from a modified (to account for the ruby three-level system) Statz-de Mars theory pertaining to single axial mode operation (8). This theory is contained in two differential equations which connect the time derivatives of the population inversion and of the number of stimulated emission photons with physical and geometrical properties of the laser. The linearized solution of the two equations results in an expression for the number of stimulated emission photons which in turn leads to the equations represented by the two curves for T_d and T_s , taking into account the difference in the volume of cavity and active medium:

$$T_d = 2\tau^2 k l c \left[N_o L_r (W/W_t - 1) \right]^{-1} \quad (2)$$

$$T_s = 2\pi \left[\frac{N_o L_r}{\tau^2 k L_c} (W/W_t - 1) \right]^{-1/2} \quad (3)$$

where: $k = 4\pi^2 \nu \Delta\nu / c^3$

$\Delta\nu$ = width of fluorescent line = 300 GC

τ = lifetime of metastable state = $3 \cdot 10^{-3}$ sec

l = loss rate per pass = 0.03

(reflectivity of the mirrors 0.95 and 0.99, respectively)

N_o = No. of active ions per cc = $2.5 \cdot 10^{19}$

L_r = length of ruby = 3.3 cm

L_c = length of cavity = 51 cm

W/W_t = Pump power/Threshold pump power

Considering the various uncertainties in the experiments such as the influence of the imperfections of the ruby crystals, the time variation of the pump light, or the active volume of the cavity, the agreement between theory and experiment must be considered very good.

INFLUENCE OF MISALIGNMENT AND PUMP ENERGY ON ENERGY OUTPUT

All the previous measurements were performed with the two mirrors aligned for parallelism (without the ruby inserted). Since optical inhomogeneities of the ruby may offset the parallel alignment of the mirrors, measurements of the spike pattern and the output energy were performed with deliberate misalignment of the dielectric mirror within an arbitrarily chosen plane. Figure 6 shows the result. The optimum energy output occurs at $\pm 10'$ misalignment. The spike structure was always irregular in the case of the stationary mirror. In the case of the mirror vibrating, the decay was not completely exponential from approximately 0 to ± 20 minutes of arc. Outside of this region, completely exponential decay of the spike train was obtained. However, equidistance of the spikes was always observed in the case of the vibrating mirror. It is to be noted that an energy increase of 37% was obtained by vibrating the mirror.

Figure 7 which displays the output energy and the output energy increase as a function of the pump energy, reveals an even higher energy increase almost up to 100% due to the vibrating mirror. Again, in the case of the stationary mirror, the spiking remains completely irregular from threshold to the highest pump energy applied. In the case of the vibrating mirror, the spiking is regular up to a pump energy of 2200 Joules. Above this value, the decay shows some irregularities which, however, are not comparable with the completely random spiking obtained with the stationary mirror. The dependence of the character of the spike train and of the output energy from the amplitude of vibration of the mirror is less pronounced: At a certain threshold displacement (approximately .08 microns as mentioned previously in the paper) regular spiking occurs and the character of the spiking and the energy output is not influenced much more by the vibrational amplitude.

CONCLUSIONS

Considering all the observations of the character of the temporal output of the ruby laser and of the energy output influenced by the vibrating mirror, the following conclusions can possibly be drawn: The Stutz-de Mars theory quoted previously is based on the

assumption of the existence of a single mode in the cavity and predicts equally spaced spikes with exponential decay. It would seem, therefore, that the vibrating mirror provides a means for obtaining single mode operation in a solid state laser. Recent results, however, show that regular spike patterns do not necessarily indicate single-mode operation (9). But in the referenced case, the spiking is completely random at lower pumping levels and attains regularity at high pumping levels, whereas in this case the opposite is true. As described in the discussion of the results shown in Figure 7, spiking is completely regular at lower pumping levels and loses some of its regularity at higher levels. This fits very well into the single-mode picture, since higher pumping energy causes the appearance of more modes in the cavity, a fact which distorts the clean exponential decay of the spike train. In line with this picture, the following explanation for the surprising energy increase caused by the vibrating mirror is proposed: As is known (7), the dominant axial mode has the lowest losses in the Fabry-Perot cavity, and higher order longitudinal and transverse modes have increasing losses. If, therefore, higher order modes are suppressed by the vibrating mirror, more energy will be dumped into the low-loss dominant mode and, consequently, the efficiency of the laser will increase.

The authors wish to thank Drs. G. J. Goubau and H. Jacobs, and Mr. F. A. Brand of these Laboratories for helpful discussions and also Mr. E. Elder for his assistance in taking measurements.

REFERENCES

- (1) C. L. Tang, H. Statz, and G. de Mars, *Appl. Phys. Letters*, 2, 222 (1963).
- (2) E. Schiel, *Proc. IEEE*, 51, 365 (1963).
- (3) R. Bechmann, *Hochfrequenztechnik und Elektroakustik*, 56, 14 (1940).
- (4) E. A. Gerber, *Electronics*, 24, 142 (1951).
- (5) G. Sauerbrey, *Proc. 17th Annual Symp. on Frequency Control*, Fort Monmouth, N. J., p. 28 (1963).
- (6) J. Free and A. Korpel, *Proc. IEEE*, 52, 90 (1964)
- (7) A. G. Fox and Tingye Li, *Proc. IEEE*, 51, 80 (1963)
- (8) M. Birnbaum, T. Stocker, S. J. Welles, *Proc. IEEE*, 51, 854 (1963).
- (9) H. Statz, C. L. Tang, and G. de Mars, *Annual Meeting of the Am. Phys. Soc.*, New York, January 22-25, 1964.

FIGURE CAPTIONS

- Figure 1 Temporal behavior of the output of a ruby laser. The time scale on the abscissa is 50 μ sec/cm. The ordinate shows the power in arbitrary units.
- Figure 2 Diagram showing the field distribution for the first and second unstable longitudinal modes and inverted population density for the case in which only the first mode oscillates. (Courtesy of Dr. H. Statz).
- Figure 3 Experimental apparatus. The distance L_c of the mirrors was variable between 400 and 550 mm; the length L_r of the two ruby crystals employed 33 and 70 mm, their optical axis was parallel to the axis of the rod.
- Figure 4 Temporal output of a ruby laser. Lower trace: mirror stationary. Upper trace: mirror vibrating with 0.135 micron amplitude. Frequency: 3 mc. Abscissa: 20 μ sec/cm. Ordinate: power in arbitrary units. Pump energy: 1.5 times threshold, $L_c = 446$ mm, $L_r = 70$ mm.
- Figure 5 Spike distance and decay of spike envelope as a function of the ratio pump power/threshold pump power. Amplitude of vibrating mirror: 0.25 micron. Frequency: 1 mc. $L_c = 510$ mm, $L_r = 33$ mm.
- Figure 6 Output energy as a function of misalignment of dielectric mirror for stationary and vibrating quartz mirror. Pump energy: 2450 Joules equal to 1.35 times threshold. Amplitude of vibrating mirror: 0.17 microns. Frequency: 1 mc. $L_c = 446$ mm, $L_r = 70$ mm.
- Figure 7 Output energy and output energy increase as a function of pump energy for stationary and vibrating mirror. Amplitude of vibrating mirror: 0.145 microns. Frequency: 3 mc. Misalignments of dielectric mirror: $\pm 6^\circ$. $L_c = 446$ mm, $L_r = 70$ mm.

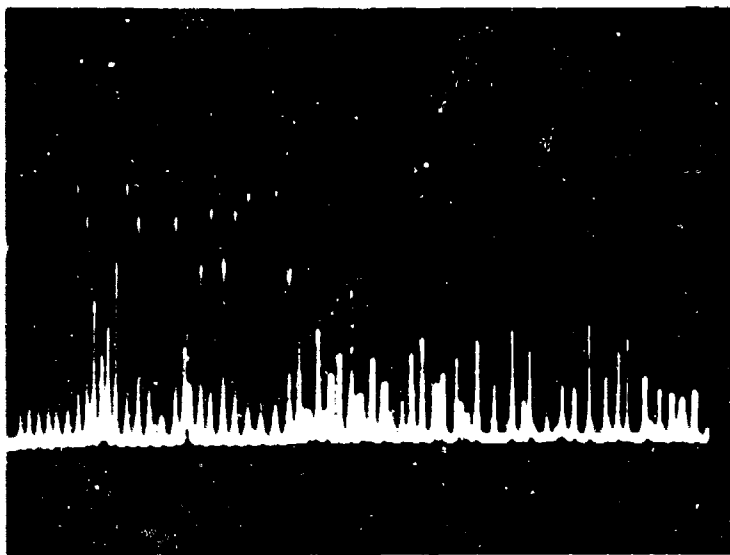


Fig. 1
Temporal behavior of the output of a ruby laser
The time scale on the abscissa is 50 μ sec/cm.
The ordinate shows the power in arbitrary units.

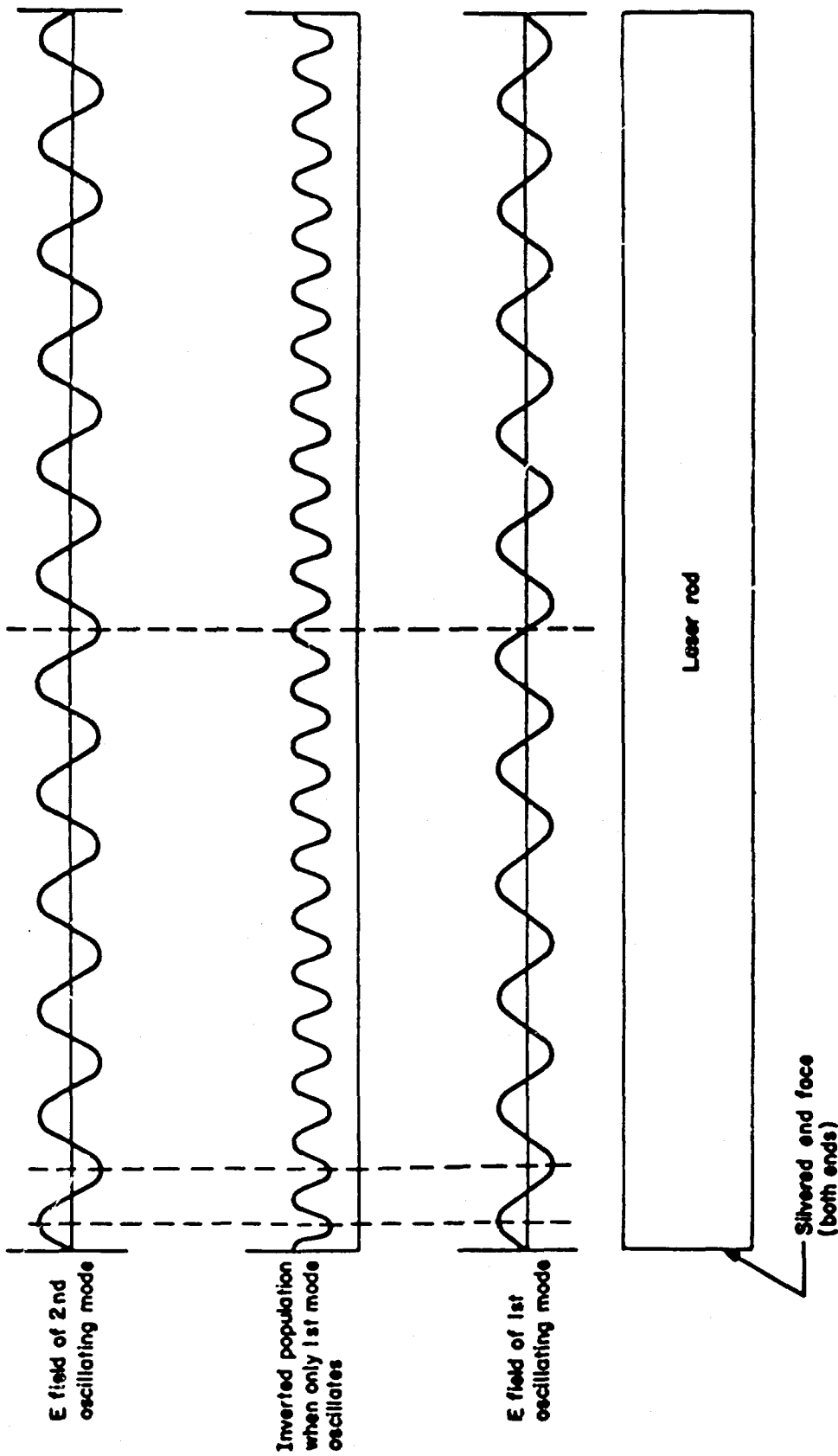


Fig. 2
Diagram showing the field distribution for the first and second unsteady longitudinal modes and inverted population density for the case in which only the first mode oscillates. (Courtesy of Dr. H. Stats)

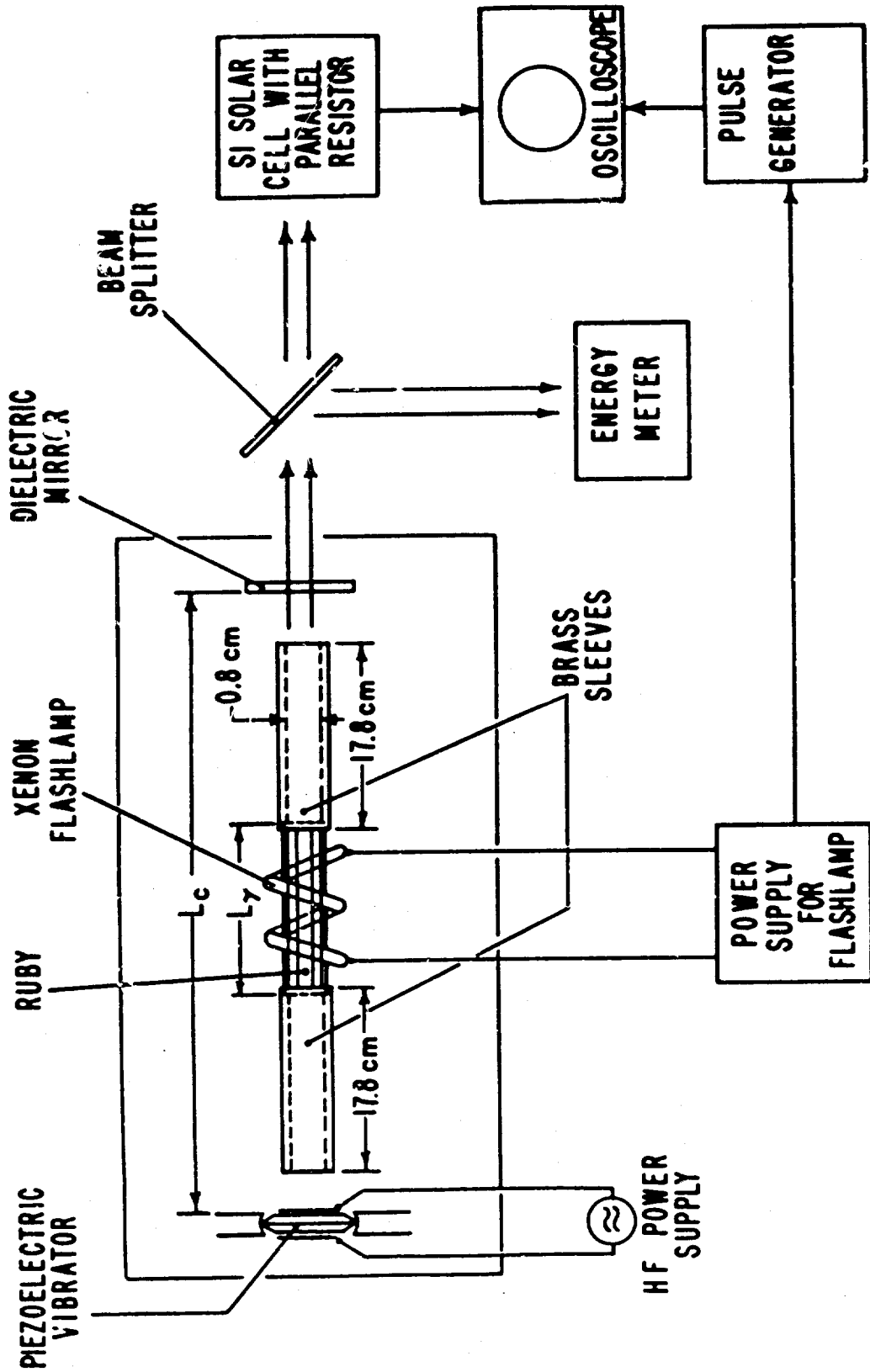


Fig. 3
 Experimental apparatus. The distance L_c of the mirrors was variable between 400 and 550 mm; the length L_r of the two ruby crystals employed 33 and 70 mm, their optical axis was parallel to the axis of the rod.

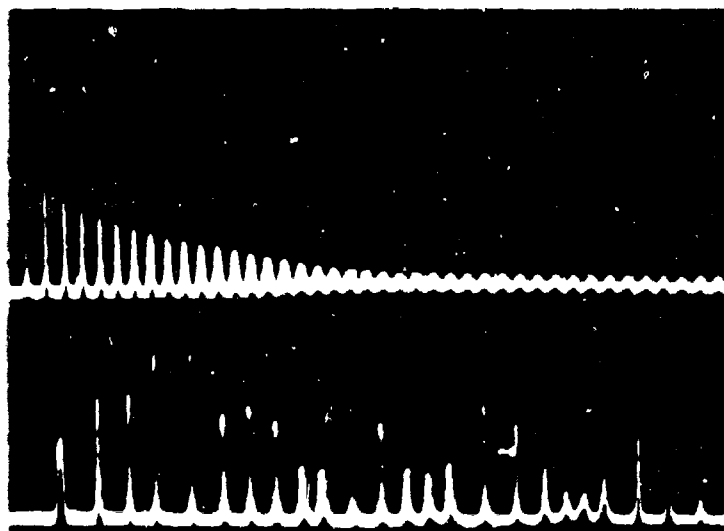


Fig. 4
Temporal output of a ruby laser. Lower trace: mirror stationary. Upper trace: mirror vibrating with 0.135 micron amplitude. Frequency: 3 mc. Abscissa: 20 μ sec/cm. Ordinate: power in arbitrary units. Pump energy: 1.5 times threshold. $L_C = 446$ mm, $L_T = 70$ mm.

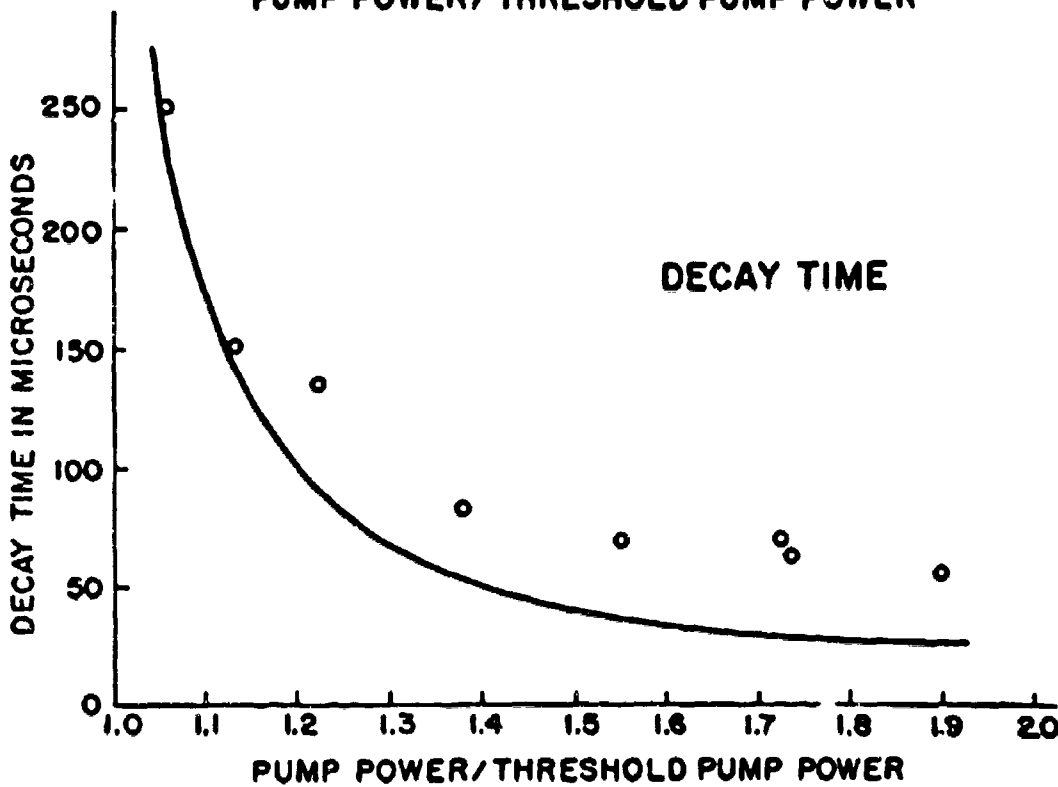
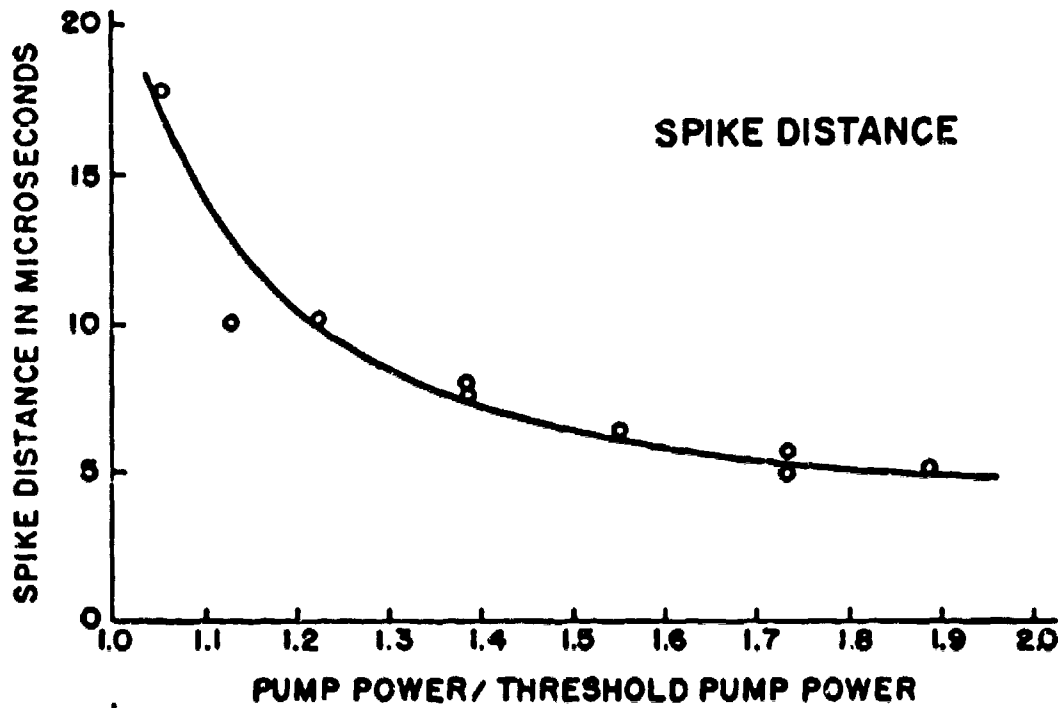


Fig. 5 Spike distance and decay of spike envelope as a function of the ratio pump power/threshold pump power. Amplitude of vibrating mirror: 0.25 micron. Frequency: 1mc. $L_c=510$ mm, $L_r=33$ mm.

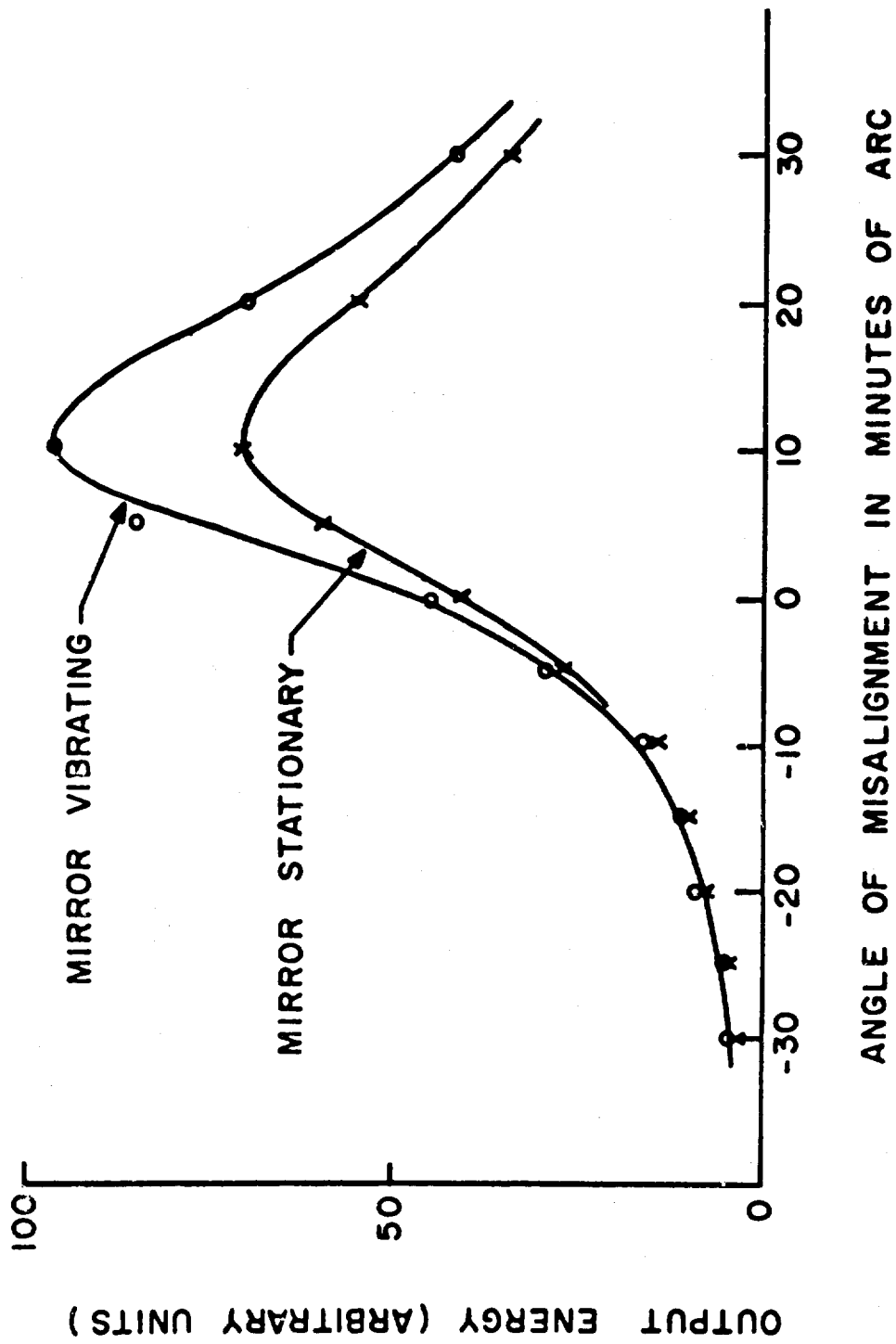


Fig. 6. Output energy as a function of misalignment of dielectric mirror for stationary and vibrating quartz mirror. Pump energy: 2450 Joules equal to 1.35 times threshold. Amplitude of vibrating mirror: 0.17 microns. Frequency: 1 mc. $L_c = 446$ mm, $L_T = 70$ mm.

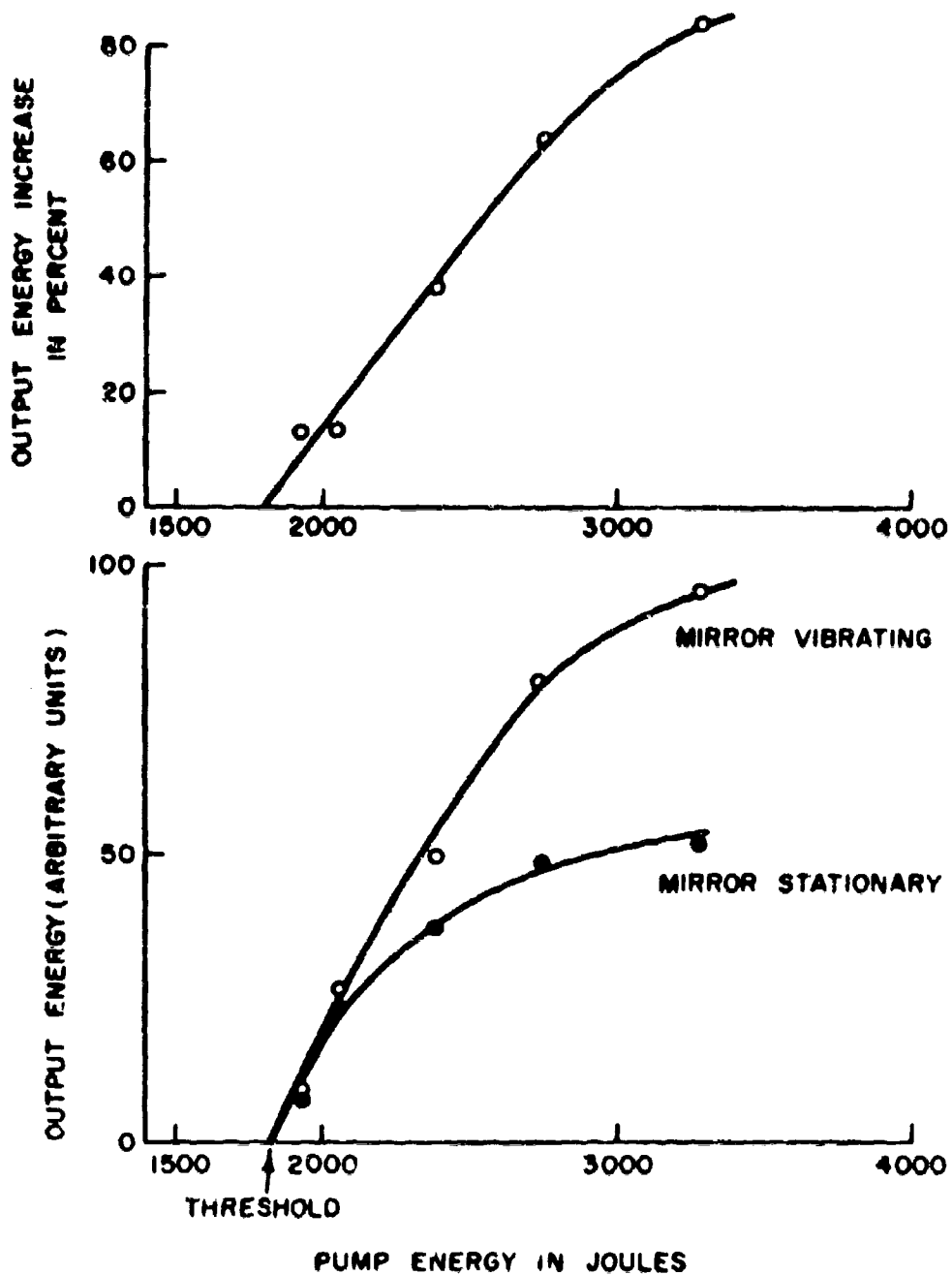


Fig. 7 Output energy and output energy increase as a function of pump energy for stationary and vibrating mirror. Amplitude of vibration of mirror: 0.145 microns. Frequency: 3 mc. Misalignments of dielectric mirror: +6'. $L_c = 446$ mm, $L_r = 70$ mm.

EVAPORATION OF ORGANIC COMPOUNDS
FROM METAL SURFACES AT HIGH VACUUM

H. GLISSER AND S. SADJIAN

FRANKFORD ARSENAL
PHILADELPHIA, PA.

The extensive space research now in progress has led to considerable activity in materials areas required for successful space operation. Ambient pressures encountered in space range from approximately 10^{-6} Torr at 125 miles to 10^{-12} Torr at 1400 miles, pressures at higher altitudes being considerably lower (1). There is a dearth of data on lubrication in high vacuum and consequently considerable work is being done in this area. In the interaction of high vacuum with lubrication, two factors must be considered (for fluids or semi-fluids lubricants). First, a lubricant film may evaporate from the metal surface at the low pressure and second, the absence of atmospheric gases may make unavailable those chemical reactions on which surface active additives depend in order to reduce friction or support heavy loads. Several investigators measured evaporation rates of oils and greases of low volatility in vacuum as a guide to establishing relative merit for space use (2,3). Empirical studies were conducted on the lubrication of high speed ball bearings at 10^{-5} to 10^{-6} Torr with petroleum base, silicone base, and ester base oils and greases of low volatility and successful operation was obtained for several thousand hours (2,4).

It is apparent that vapor pressure alone is an inadequate index of availability of a compound for lubrication at extreme low pressures. It has been established by a number of investigators that the coefficient of friction in the presence of a single molecular layer is the same as that of many layers (5,6,7,8). (Of course, the availability of a reservoir of material provides for continued availability of monolayers during continuous rubbing). It is to be expected that evaporation of the last monolayer would not be subject to the same physical laws that evaporation of the bulk of a compound is subject to. Considerable work has been done on the adsorption of monolayers of polar organic compounds on metal surfaces, and there is considerable evidence for compound formation (9,10), in some instances and in general where adsorption energy is in the neighborhood of the chemisorption range. Lubrication, i.e., friction

reduction and load carrying capacity, is presumed to be provided by the chemisorbed layers or compounds formed. It has also been shown that a single monolayer may have considerable durability (8,11). It follows that the effects of evaporation of lubricants on lubrication must be considered in the light of evaporation of the last monolayer. This is further borne out by a number of empirical studies in which it was shown that the performance of lubricants in high vacuum was considerably better than that expected from the bulk evaporation rate (12).

Friction reduction by monolayers applies to non-polar compounds (hydrocarbons of sufficient chain length) as well as to polar compounds. Further, although the adsorption energies of non-polar compounds on metal surfaces are not in the chemisorption range, it is to be expected that non-polar compounds would be adsorbed on metal surfaces by van der Waals forces and here too one would expect that vapor pressure would not reflect removal of the last monolayer. Thus even with petroleum lubricants better behavior in high vacuum lubrication is to be expected than that predicted from volatility alone. Indeed the empirical work published and unpublished in this area indicates that in general with both polar and non-polar lubricants, better lubrication is being achieved in high vacuum than is to be expected on the basis of evaporation rate data alone.

The foregoing consideration led to our investigation of the evaporation of the last monolayers of hydrocarbons from metal surfaces. Study of the adsorption of non-polar compounds on metal surfaces has been essentially limited to gases. Similarly, little work has been done on the coefficient of friction of metal surfaces coated with less than a close packed monomolecular layer. (Recent work on the adsorption of low molecular weight compounds in the vapor phase showed that coefficient of friction decreased to a minimum with increasing amount adsorbed as the monomolecular layer was reached (13)). We therefore decided to study the removal of the last monolayer of non-polar hydrocarbon on metal surfaces, using coefficient of friction measurement as an indication of amount of material remaining. This paper is a preliminary report of the finding in this work and is concerned primarily with the experimental techniques, establishment of data on coefficient of friction of surfaces that are incompletely covered with friction reducing films, and relationship of pressure to fraction of monolayer remaining on metal surfaces. Data were obtained on both polar and non-polar compounds and bear out the conclusions that are to be expected from earlier work.

EXPERIMENTAL

Materials. n-Dodecane, n-hexadecane (ASTM grade), and n-octadecane were Humphrey Wilkinson, high purity grade. n-Dodecylamine and dodecanol-1 were Eastman Kodak White Label. These compounds were fractionated under reduced pressure through a 36 in. column filled with glass helices with a 4:1 reflux ratio. The compounds were collected at the following temperatures and pressures: n-octadecane,

GLISSEY AND SADJIAN

150-150.5°C., 3.3 mm.; n-hexadecane, 120.4°C., 2.3 mm.; n-dodecane, 73-74°C., 3.8 mm.; n-dodecylamine, 93.5°C., 1.8 mm.; dodecanol-1, 105.5°C., 1.5 mm. To remove surface active impurities, the hydrocarbons were passed through a 12 in. adsorption column filled with 4 in. sections of activated Florisil, alumina, and Fuller's earth (top to bottom). Octadecanol-1, Eastman White Label, was used without further purification. Benzene was ACS grade and was passed through the Florisil, alumina, and Fuller's earth column to remove surface active impurities.

Friction Measurements. Friction measurements were run with 52100 steel riders on 1020 steel. Steel disks (1-1/8 in. dia. x 1/4 in. thick) were ground to a surface finish of 8-10 microinches rms. A final surface finish of 3-5 microinches rms (Profilometer measurements) were obtained by hand polishing with progressively finer emery cloth (2/0, 3/0, 4/0). Disks were degreased in hot benzene immediately after abrasion and again degreased in a Soxhlet extractor with benzene. Disks were then stored in a desiccator at $25^{\circ} \pm 1^{\circ}\text{C.}$, 20% R.H. for 24 hours prior to use. The riders used were 52100 steel balls (1/8 in. radius). The balls were washed in hot benzene and again degreased in a Soxhlet extractor prior to use.

Friction measurements were made at room temperature using a modified Bowden-Leben type apparatus(14). Normal and friction forces between the rider and disk were measured by 2 sets of four strain gages mounted on a 24ST aluminum ring. All measurements were made on new areas (disk and rider). Measurements were made with a 100g load at a sliding speed of 2.4 cm/min. Each result reported was the average of at least 3 separate traverses.

Deposition of Material. The hydrocarbons were deposited on the metal surface as a thick layer (excess) or as a monomolecular layer from benzene. When deposited as a thick layer, 0.05 ml of the material was pipetted onto the surface of the disk and allowed to spread until the surface was completely covered. Excess material was allowed to drain off for 5 minutes. When a monomolecular layer was deposited from benzene, the concentration of the solution was such that a monomolecular layer of the compound remained when 0.5 ml of the solution was pipetted onto the surface of the disk and the solvent permitted to evaporate.

Polar compounds were deposited as a condensed monolayer adsorbed on the steel disk by the retraction method using a melt of the compound (15).

Evaporation. The evaporation chamber was a 3 in. diameter glass tube approximately 7 in. long attached to a hot cathode ionization gage. A glass tube (end closed) of approximately 1 in. square section, 3 in. long was mounted inside the evaporation chamber and attached to a small Dewar mounted outside and on top of the evaporation chamber. A sheet of oxygen-free high conductivity copper was shaped into a square section tube and placed around the glass tube with a close fit. By filling the glass tube with liquid nitrogen through the externally

GISSER AND SADJIAN

mounted Dewar, a cold copper surface was maintained parallel to and approximately 1 cm above the test surface. New (out gassed) copper sheets were used for each determination. The purpose of the liquid nitrogen cooled surface was to prevent redeposition of organic compounds on the test surface after evaporation (and to condense any oil which back migrated from the diffusion pump). At a suitable level below the liquid nitrogen well was placed a copper platform (on which the test disks were placed). A cylindrical heater (100 watt, connected through a variable transformer) was close fitted into the thick section of the copper platform. A section of the evaporation chamber is shown in figure 1. A hole for use as a thermocouple well was machined in a disk parallel to and approximately 1/16 in. (surface to edge) below the surface. A disk so prepared was used uncoated (see later) in each experiment. Heater and thermocouple leads were brought out through the wall of the vacuum chamber in a conventional way. Two pumping systems were attached to the evaporation chamber. A rough system which consisted of a mechanical pump and liquid nitrogen trap and valve was used to evaporate excess material (where this was present) from the test surface. When this was completed, the rough vacuum system was valved off. The other pumping system consisted of a 2 in. oil diffusion pump, mechanical fore pump, liquid nitrogen trap, valve and variable leak. The latter was fed with dry nitrogen and used to maintain the pressure at a desired level.

The coefficient of friction (μ_k) of a conditioned uncoated disk was measured and the organic compound under study was deposited onto the surface of the disk either in an excess or as a monolayer and μ_k was again measured. The coated specimen and an uncoated specimen (with measured μ_k , referred to later as the standard) were placed in the test chamber, a thermocouple being placed in the well of the uncoated disk. The rough system was turned on and the reservoir of the condensing surface was filled with liquid nitrogen. During this time the "clean" vacuum system was being evacuated. After 1 hour (visible evidence of excess material disappeared within 5-10 minutes) the rough system was valved off and the clean system engaged. Test specimens were kept in the evaporation chamber for a period of 2 hours at the designated pressure and temperature. Temperature at the test surface during the test for room temperature runs was $25^\circ \pm 1^\circ\text{C}$. For experiments at elevated temperature, heating was started when the pressure reached approximately 1×10^{-5} Torr (clean vacuum system on). Temperature was maintained to $\pm 1^\circ\text{C}$. Specimens were then removed from the test chamber, kept in a desiccator over anhydrous calcium sulfate for one hour and μ_k measured on both disks. The coefficient of friction of the uncoated disk was always within 0.01 of its original value, in runs at room temperature with hydrocarbons indicating that there was no deposition of organic material originating from the pumping system or other source onto the steel surface during an experiment. In runs with the polar compounds, there was always some decrease in coefficient of friction of the uncoated disk (values are in the data tables). In experiments in which uncoated disks were heated in the vacuum chamber without a

coated disk being present, the coefficient of friction did not change more than 0.01, indicating that during the evaporation experiments, some contamination of the uncoated disk occurred. The extent of contamination was low and estimates (using the surface calibration data) will be given later.

RESULTS AND DISCUSSION

The coefficient of friction data obtained on the hydrocarbons in evaporation experiments at 25°C. are shown in Table I together with vapor pressures (from the literature). Evidently the bulk of material is removed at relatively high pressures. While removal of the last monolayer starts at relatively high pressures, the hydrocarbon is still present on the surface at relatively low pressures. Thus, the coefficients of friction of these surfaces lie in between that of the uncoated surface and that of the surface coated with a complete monolayer. To determine the significance of these values disks were coated with fractional monolayers of hexadecane and octadecane and coefficient of friction determined as a function of a monolayer present. The quantities of material used were based on the assumption that the surface was perfectly flat and the expression "fraction of a monolayer" used means fraction of amount of material required for one monomolecular layer assuming a perfectly flat surface. The quantities of material required for the theoretical monolayers were calculated from data in the literature (16). The material was deposited on the conditioned surface from benzene solution. The data are shown in figure 2. (Plots relating coefficient of friction to fraction of surface covered with organic compound will be referred to as surface calibration curves). These data were used to calculate the fraction of a monolayer remaining in the evaporation studies and the results are shown in figure 3. It is evident that the last layer of dodecane is removed at 7×10^{-3} Torr or higher, while with hexadecane and octadecane, the last layers have not been completely removed at 5×10^{-6} Torr.

Evaporation experiments on dodecanol-1, n-dodecylamine and octadecanol-1 at 25°C. indicated that substantially none of the last monolayer was removed at pressures in the neighborhood of 10^{-6} Torr. Experiments on dodecanol-1 and n-dodecylamine were therefore run at elevated temperatures. It was found (Table II) that at $2-5 \times 10^{-6}$ Torr, the last monolayer was progressively removed as the temperature increased. Removal is not complete at 80°C., however. Surface calibration curves were prepared (Figure 4) for fractional monolayers of dodecanol-1 and n-dodecylamine using a procedure similar to that described earlier for the hydrocarbons. The data were used to calculate the fraction of a monolayer of the polar compound remaining in the evaporation experiments and the results are in figure 5. Changes in coefficient of friction on the standard disks (Table II) indicate minor contamination. From the surface calibration data of figure 4 it may be estimated that the contamination of the standard surfaces was approximately 1/100 monolayer, except with dodecanol at

GISSER AND SADJIAN

TABLE I. Evaporation of Hydrocarbons at 25°C.
Coefficient of Friction on 1020 Steel
with 52100 Steel Rider

Hydrocarbon	μ_k Uncoated	μ_k Before Vacuum Exposure	Pressure (Torr)	μ_k After Vacuum Exposure
n-Octadecane (v.p. = 9.4×10^{-4} Torr at 25°C.)	0.31	0.16	6×10^{-3}	0.18
	0.31	0.16	5.5×10^{-6}	0.23
	0.32	0.16	1.6×10^{-6}	0.26
n-Hexadecane (v.p. = 9.9×10^{-3} Torr at 25°C.)	0.31	0.16	1×10^{-2}	0.25
	0.30	0.16	7×10^{-3}	0.25
	0.31	0.16	3.3×10^{-5}	0.27
	0.29	0.13	5×10^{-6}	0.28
n-Dodecane (v.p. = 0.27 Torr at 25°C.)	0.30	0.17	7×10^{-3}	0.31

80° in which case the standard was contaminated with approximately 1/60 monolayer.

The ease of removal of the hydrocarbon layers in contrast to the alcohols and the amine reflects the difference in energy of adsorption. Evidently in the case of the non-polar compounds, the adsorption energies, probably in van der Waals range, are such that the last layer does not persist completely in the pressure range of these experiments, but the amount remaining is a function of the pressure. To ascertain whether a time dependent phenomenon was involved, data were obtained after one, two and three hours at the low pressures, and no differences in the coefficient of friction were observed. The retention by the surface of partial monolayers is of significance even with polar compounds. Although it has been shown that only one complete monolayer is chemisorbed, further layers may be adsorbed with van der Waals energies (17). Thus if a polar compound is deposited in excess one would expect one monolayer to be chemisorbed and further material to be physically adsorbed. Whether the material above the first monolayer, after exposure to high vacuum would be a fraction of a single layer or several layers would depend

Table II. Evaporation of Polar Compounds at Elevated Temperatures.
Coefficient of Friction on 1020 Steel with 52100
Steel Rider.

Compound	Vapor Pressure ^a mm./°C	μ_k Uncoated	μ_k Coated Before		Exposure Temp, °C	μ_k After		Standard																	
			Vacuum Exposure	Vacuum		Vacuum Exposure	Vacuum	Before Exposure	After Exposure																
dodecanol-1	9.4x10 ⁻³ /25	0.31	0.12	0.12	25	0.14	0.31	0.31	0.26																
										5.3x10 ⁻² /45	0.33	0.13	0.13	45	0.20	0.33	0.30								
																		2.3x10 ⁻¹ /65	0.33	0.13	0.13	65	0.23	0.31	0.29
n-dodecyl- amine	2.1x10 ⁻² /25	0.31	0.12	0.12	25	0.11	0.31	0.31	0.29																
										4.1x10 ⁻¹ /65	0.31	0.11	0.11	65	0.16	0.33	0.24								
																		1.0 /80	0.31	0.12	0.12	80	0.28	0.32	0.29

^a Calculated from the Antoine equation (19).

GISSEK AND SADJIAN

on the magnitude of the adsorption energies involved in addition to the experimental conditions. There are insufficient data available at present to calculate quantities of polar compounds that may remain at specific pressures. It would appear however, that this would depend on molecular weight of the compound (or its chain length) and pressure in addition to degree of polarity. The polar compounds, on heating in vacuo, yield fractional monolayers on steel surfaces. This could indicate that at the temperatures in question, the forces holding the polar molecules to the surfaces are approximately the same as these holding the non-polar compounds at lower temperatures. It does not follow, however, that the configuration of the polar compounds in less than a monolayer on the metal surface is similar to that of the non-polar compounds, since selective adsorption at the polar group is still to be expected.

The distribution of compounds in less than a single monolayer is of interest. In depositing such fractional monolayers from solvent, there exists a possibility of non-uniform deposition because evaporation of the solvent may leave islands of material. The probability is not too remote that the entire fractional layer may consist of a series of islands of complete monolayers on a surface free from the compound in question. In this case, the term fraction of a monolayer may well mean fraction of the surface coated with islands of complete monolayer. To some extent the experimental data indicate the limit of size of these islands if they do exist. The coefficient of friction measurements were run at a rubbing speed of approximately 0.04 cm/sec. and apparatus response time was probably less than 0.01 sec. so that the length of travel to which the instrument responds, assuming point contact, is 0.0004 cm. However, the calculated diameter of the contact area was 0.003 cm., from which it follows that an island greater than 0.003 cm. dia. would be reflected in the friction measurements (assuming a flat surface, or average over many asperities). The variation in coefficient of friction measurements with fractional monolayers deposited from solvent was ± 0.015 for the non-polar compounds and ± 0.02 for the polar compounds within a single trace, which would indicate the absence of microscopic islands within the limits indicated. However, in general, a greater variation of results was observed with surfaces coated with fractional monolayers from solvent than those observed after evaporation at low pressure. It is reasonable to conclude that the observed coefficient of friction is the result of averaging by the apparatus over the non-uniformities of surface concentration of compound and also that there is a degree of non-uniformity greater when material is deposited from solvent than that obtained when the material is evaporated from a surface at low pressure. The non-uniformities may indeed be islands of condensed monolayers in a sea of uncoated surface, however, this does not influence the use of the data for estimating surface coverage by coefficient of friction measurement because of the averaging property of the apparatus. In measurements on the partial layers remaining after evaporation, the friction variation within a single

trace was somewhat greater than that obtained with either complete monolayers, excess material, or clean surfaces. The precise reason for this is obscure. Although it may be related to the uniformity of material remaining on the surface, no information is available which indicates that this may be the case. As the amount of material on the metal surface is progressively reduced below a complete monolayer, it is to be expected that the molecule would become progressively more parallel to the metal surface and for this reason significant reduction of coefficient of friction from that of an uncoated surface would pertain even at very small fractions of a monolayer. For this reason one would not expect a linear relation between fraction of monolayer remaining and change in coefficient of friction and indeed no such linear relation was found. Further work is necessary however to determine the distribution and configuration of molecules of fairly long chain length on metal surfaces. It has been observed earlier (18) that static coefficient of friction decreases with increasing number of carbon atoms, and that this reaches a minimum at approximately 17-18 carbon atoms. The data on fractional molecular layers supplements the earlier observation in that it shows that there is a distinct decrease in kinetic coefficient of friction from hexadecane to octadecane for corresponding fraction of the surface covered (fig. 2). Further study of fractional layers of organic compounds on metal surfaces may help to clarify the mechanism of friction reduction.

Finally a discussion of the nature of the surface on which measurements were made is in order. The metal surfaces before coating with the organic compound contained atmospheric gases, water vapor, and metal oxide. Conditioning of the surface before use served primarily to equilibrate the water vapor in the surface. What is significant here is that these variables had no effect on the capability of using coefficient of friction measurements to determine amount of organic material remaining on the surface. The cleaned steel surface after conditioning has μ_k 0.31-0.32. After evaporation of benzene (from a previously conditioned surface) μ_k was 0.31. These data indicate that there was no contribution of any of these variables in reducing the coefficient of friction of the metal surface below that of the conditioned metal surface on which no organic compounds were deposited. The major effect of conditioning was to increase uniformity in measurements from disk to disk. This uniformity probably resulted both from the more uniform and reproducible oxide layer on which measurements were made, than would have been the case had the surfaces not been conditioned.

GLISSER AND SADJIAN

LITERATURE CITED

1. Jaffe, L. D., and Rittenhouse, J. B., "Behavior of Materials in Space Environment", Technical Report No. 32-150 (Nov. 1, 1961).
2. Freundlich, M. E., and Hannon, O. H., *Lubrication Eng.* 17, 72-77 (1961).
3. Jagodowski, S. S., and Freundlich, M. E., "Investigation of Properties of Lubricants in High Vacuum", Third Progress Report - WADD Report No. 7556-1-3. Contract AF-33(616)-6845 (Oct 1960).
4. Clauss, F. J., "Lubricants and Self-Lubricating Materials for Spacecraft Mechanisms", Technical Report Lockheed Missiles and Space Div. 894812, Sunnyvale, Cal., 1961.
5. Langmuir, I., *Trans. Faraday Soc.* 15, 62 (1920).
6. Langmuir, I., *J. Franklin Inst.* 218, 143 (1934).
7. Bowden, F. P., and Tabor, D., "The Friction and Lubrication of Solids", Oxford University Press, 1950, pp 184-191.
8. Levine, O. and Zisman, W. A., *J. Phys. Chem.* 61, 1068-1188 (1957).
9. Rowe, J. W., "Vapor Lubrication and Friction of Clean Surfaces", Proc. Conf. on Lubrication and Wear. Institution of Mechanical Engineers, London, 1957.
10. Bowden, F. P., and Yound, J. E., *Proc. Roy. Soc.* A208, 311 (1951).
11. Cottingham, R. L., Shafrin, E. G., Zisman, W.A., *J. Phys. Chem.* 62, 513 (1958).
12. Coit, R. A., and Sorem, S. S., *Lubrication Engineering* 18, 438 (1962).
13. Fraioli, A.V., Healey, F.H., Zettlemyer, A.C., and Chesnok, J.J. Abstract of Papers, 130th Meeting, A.C.S., Atlantic City, N.J., Sept. 1956.
14. Bowden, F.P., and Tabor, D., "The Friction and Lubrication of Solids", Oxford University Press, 1950, p. 74-75.
15. Bigelow, W.C., Glass, E., and Zisman, W.A., *J. Colloid Science* 2, 563 (1947).
16. Adam, N.K., "The Physics and Chemistry of Surfaces", Oxford University Press, p. 46-55.
17. Trapnell, B.M.W., "Chemisorption", Butterworths Scientific Publications, London, 1955, p.4.
18. Bowden, F.P., and Tabor, P., "The Friction and Lubrication of Solids", Oxford University Press, 1950, p. 178-181.
19. Dreisbach, R.R., "Pressure-Volume-Temperature Relationships of Organic Compounds", Handbook Publishers, Inc., Sandusky, Ohio, 1952.

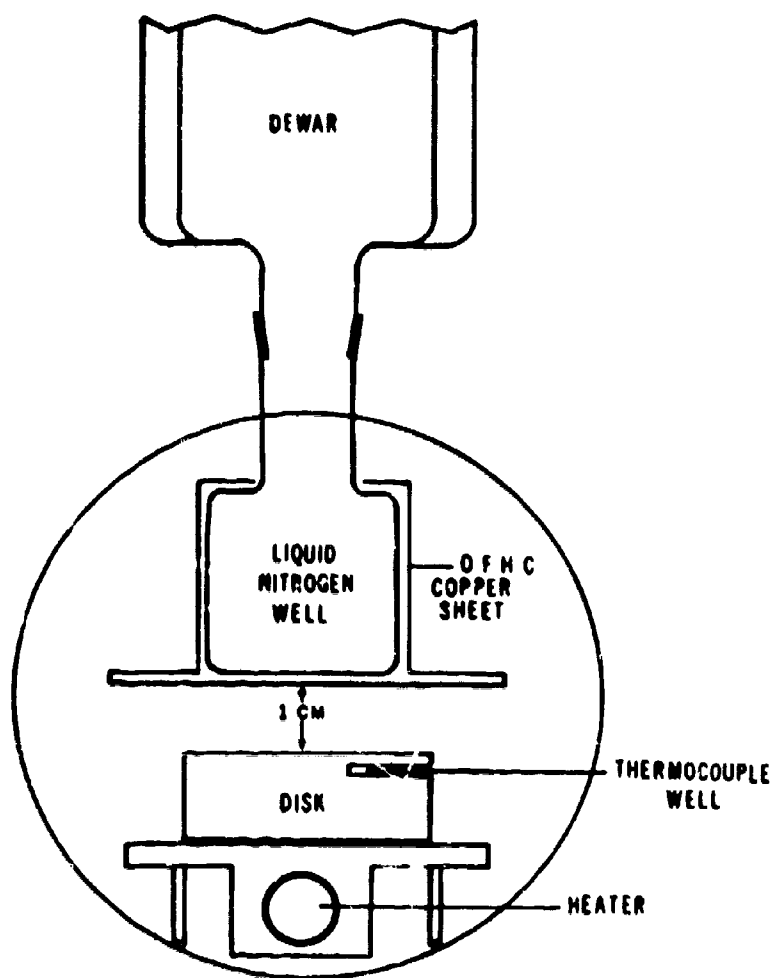


Figure 1. Evaporation Chamber

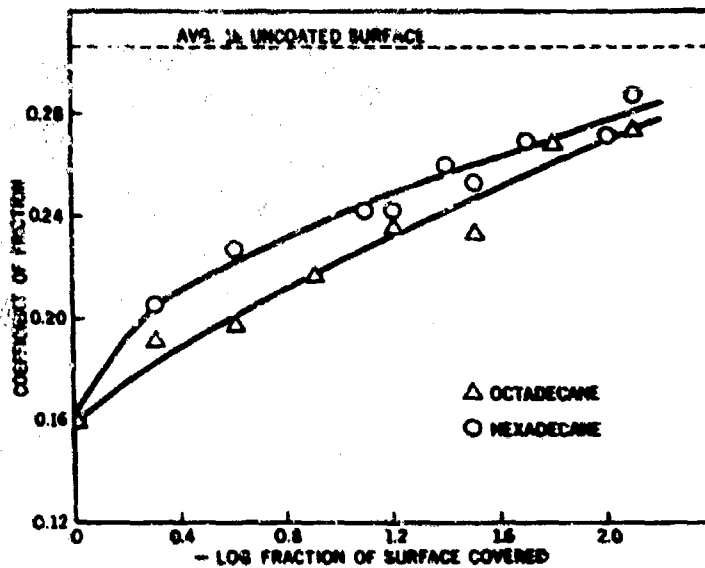


Figure 2. Coefficient of friction of steel surfaces coated with fractions of a monolayer of hydrocarbon.

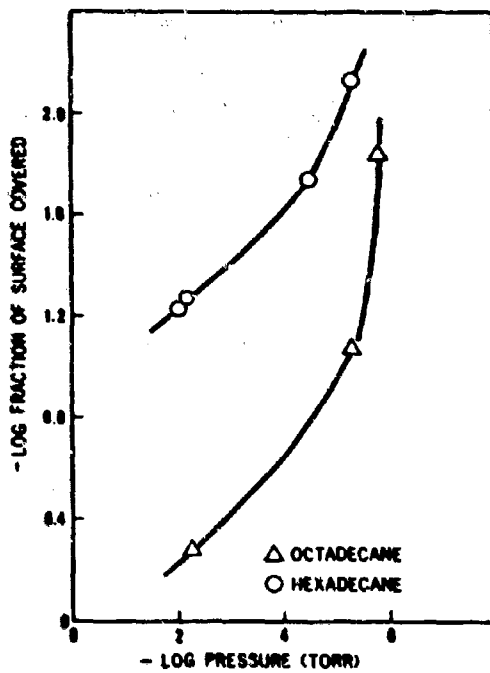


Figure 3. Evaporation of hydrocarbons from steel surfaces at low pressure.

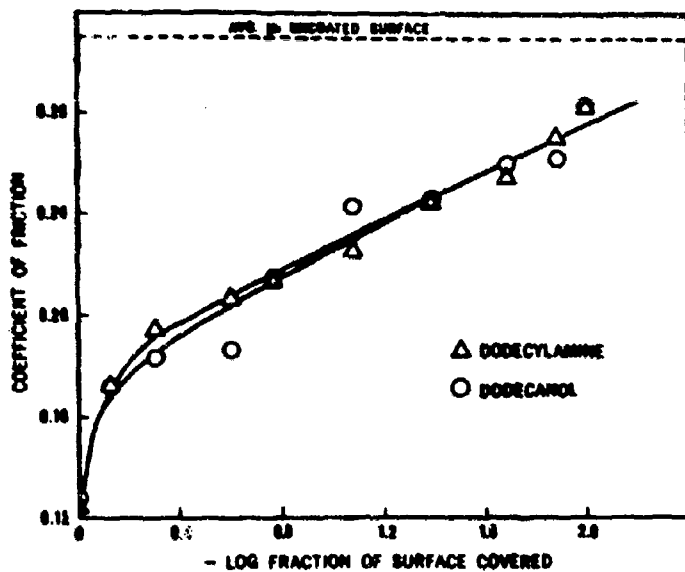


Figure 4. Coefficient of friction of steel surfaces coated with fractions of a monolayer of polar compounds.

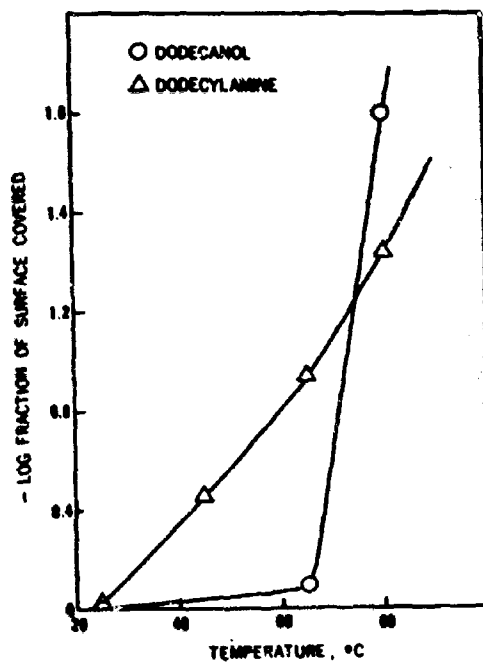


Figure 5. Evaporation of polar compounds from steel surface at $2-5 \times 10^{-6}$ torr.

THE EFFECT OF HYPERTHERMIA
ON PROTEIN TURNOVER IN INFECTION

IRVING GRAY, COLONEL, MSC
and PAUL V. HILDEBRANDT, CAPTAIN, VC
U. S. ARMY MEDICAL UNIT, FORT DETRICK, MARYLAND

The general interrelationship between infection and nutrition has been well established and recent reviews have summarized our present knowledge (1-4). The effect of infection on nitrogen metabolism has been investigated (5,6) and has been shown (7) to cause a shift toward negative nitrogen balance in virus infections (8). However, despite the fact that fever normally accompanies infections, little has been done to determine the relative contribution of the hyperthermia and the microorganism to the altered protein metabolism. The hypermetabolic state of individuals during febrile periods has long been recognized (9). It has been shown that for each °F rise in body temperature, the metabolic rate increases about 7% above basal (9). In addition, there are marked changes in electrolyte and carbohydrate metabolism. Protein metabolism is likewise affected primarily toward an increased breakdown of body protein (10,11). In our laboratories preliminary experiments have shown that rabbits infected with Pasteurella tularensis have an increased uptake of methionine-S-³⁵ during the early course of this febrile disease (12). Furthermore, we have observed and reported a marked increase in protein synthesis in a cell-free system derived from the brains of mice infected with the virus of Venezuelan equine encephalomyelitis (VEE) (13). In the studies to be reported here, we have tried to delineate the contribution of the hyperthermia and of the microorganism to the observed changes. To this end, we have studied protein turnover during controlled fever alone as well as during virus infection plus controlled fever.

MATERIALS AND METHODS

Three young chimpanzees, weighing approximately 17 kilograms, were used. Food was withheld for 24 hours prior to anesthesia. Each chimpanzee was placed under light surgical anesthesia with pentobarbital sodium. A superficial vein of an arm or leg was catheterized with polyethylene tubing for the administration of

GRAY and HILDEBRANDT

fluids, anesthetic agents, or the taking of blood specimens. A catheter was inserted to collect urine samples and measure urine output.

When the animal had been prepared, it was placed between plastic mats through which thermostatically controlled water circulated. A rectal thermister attached to the water-regulating apparatus (Gorman-Rupp Co. K-Thermia) automatically controlled the temperature. The temperature of the chimpanzee was held at 98-99°F as control; 105-106°F as hyperthermic. The experimental period lasted 24 hours from the time the temperature of the animal reached and stabilized at the selected temperature. Normal physiological saline in 5% glucose was administered as a slow drip via the intravenous catheter. Approximately 800 ml were administered during periods of euthermia and 1,000 ml during periods of hyperthermia.

Following zero time blood and urine samples, 14 mg of L-methionine- S^{35} (Me- S^{35}), containing 500 μ c, dissolved in 2 ml isotonic NaCl were injected intravenously into the femoral vein opposite the catheterized limb. Blood samples (5 ml) were collected in vials containing potassium oxalate at 1, 2, 4, 8, 16 and 24 hours after the administration of the methionine. Each time a blood sample was drawn, the urine volume excreted subsequent to the previous sample was noted and a portion saved for radioactivity assay and routine analysis.

The plasma was removed from the blood and 1 ml was fractionated into globulin and albumin fractions by the method of Pillimer-Hutchison (14). The soluble albumin was precipitated with trichloroacetic acid (TCA). Each fraction was washed with 5% TCA and dissolved in 1 ml of 0.5N NaOH. 100 μ l of each fraction were analyzed for nitrogen content by the Biuret method (15).

All planchets were counted in a gas flow counter to an error of less than 3%. The appropriate standard of the injected Me- S^{35} was counted and data reported as a fraction of the injected dose.

Protein turnover was followed under 3 different conditions in each animal. The first was at 98-99°F (control). The second was identical to the first except that the rectal temperature was raised to 106°F (fever). The third was identical to the second except the chimpanzee had been inoculated with 10^8 MIPLD₅₀'s (mouse intraperitoneal lethal dose 50) of VEE virus 24 hours prior to protein turnover studies. In chimpanzees, the virus of VEE produced its initial febrile response approximately 12 to 24 hours post inoculation. An additional 1 ml of blood was drawn at each sampling time for viremia determinations.

RESULTS

In the infected animals, the concentration of virus in the blood was $10^{3.75}$ MIPLD₅₀'s per ml when protein turnover studies were initiated. Baseline serology indicated that none of the chimpanzees had had previous experience with this virus, and in the infected animals, serology following the experiment indicated hemagglutination inhibition (HI) and complement fixation (CF) titers as summarized in Table I.

CHIMP NO.	HI	CF
1	5,120	128
2	≤10,240	512
3	≤10,210	256

TABLE I. Hemagglutination Inhibition (HI) and Complement Fixation (CF) Titers in the Serum of Chimpanzees Challenged with VEE.

In Figures 1a, b, c we see the marked effect on protein turnover resulting from fever. The decrease in the uptake of the labeled methionine into the plasma globulin is quite definite in two of the chimpanzees (Sam, Denise) but still apparent in the third (Ralph). No data are reported concerning the albumin fraction inasmuch as there was no observable change under the conditions of the experiment. However, when the same degree of hyperthermia is maintained in the presence of the virus, there is an obvious increase in protein turnover. In all 3 experimental situations, control, hyperthermic, hyperthermic with virus, the significant uptake of the Me-S³⁵ has occurred in the first 4 hours following the administration of the tracer. If we plot the ratio of each of the experimental curves to its own control from the 4-hour point on, we obtain the curves in Figure 2. It appears that in each of these experimental situations, the rate of loss of the S³⁵, the measure of the rate of protein breakdown, is the same but increased over the control rate. The increased rate of breakdown in the presence of a virus infection or fever could, if continued, contribute toward an increased nitrogen excretion. That virus infection does cause increased nitrogen excretion has recently been clearly demonstrated (7,8). This effect, related specifically to amino acids in bacterial infections, has been demonstrated in these laboratories (16). Furthermore, fever itself has been shown to

GRAY and HILDEBRANDT

have a similar effect on the breakdown of protein and nitrogen excretion (10,11). Thus, it is not unreasonable to conclude that a principal cause of increased nitrogen loss during acute infectious disease results from the fever rather than the microorganism.

Finally, if we consider that the breakdown rate of the globulin during fever alone is the same as fever with virus, then the increased uptake of the Me-S³⁵ in the presence of the virus must be due to an increased synthetic or uptake rate resulting from the direct effect of the microorganism. Our recent report (13) demonstrating the stimulation of protein synthesis by microsomes obtained from animals infected with VEE lends further support to this latter conclusion.

SUMMARY

Protein metabolism has been followed in control, hyperthermia and hyperthermia plus virus.

Breakdown is increased in both experimental groups over that in the control group. It is concluded that this increase, caused primarily by the fever, contributed to the previously observed increased nitrogen excretion.

Evidence is presented to support the concept that protein synthesis is stimulated by the presence of virus at a time simultaneous with that of the observed protein breakdown.

REFERENCES

1. Platt, B.S., Protein Malnutrition and Infection, Am. Jour. Trop. Med. Hyg. 6, 773, 1957.
2. Dubos, R.J., and Schaedler, R.W., Nutrition and Infection, J. Pediat., 55, 1, 1959.
3. Schneider, H.A., Nutritional Factors in Host Resistance, Bact. Rev. 24, 186, 1960.
4. Schneider, H.A., Genetics and Nutrition of the Host in Relation to Susceptibility to Infection, Ann. New York Acad. Sc. 55, 25, 1956.
5. Scrimshaw, N.S., Taylor, C.E., and Gordon, J.E., Interactions of Nutrition and Infection, Am. J. Med. Sc. 237, 367, 1959.
6. Reiss, Eric, Protein Metabolism in Infection, Metabolism 8, 151, 1959.

GRAY and HILDEBRANDT

7. Wilson, D., Bressani, R., Scrimshaw, N.S., Infection and Nutritional Status. I. The Effect of Chicken Fox on Nitrogen Metabolism in Children, *Am. Jour. Clin. Nutr.* 9, 154, 1961.
8. Gandra, Y.R., and Scrimshaw, N.S., Infection and Nutrition. II. Effect of Mild Virus Infection Induced by 17-D Yellow Fever Vaccine on Nitrogen Metabolism in Children, *ibid* 9, 159, 1961.
9. Best, C.H. and Taylor, N.B., *The Physiological Basis of Medical Practice*, 7th Ed. p. 895, Williams and Wilkins Co., Baltimore, Md., 1961.
10. Voit, F., *Sitz. d. Gesel. f. Morph. u. Physiol.* 11, 120, 1895.
11. Cajori, F.A., *et al*, *Jour. Biol. Chem.* 57, 217, 1923.
12. Gray, I. and Meislin, A., unpublished experiments, United States Army Medical Unit, Fort Detrick, Maryland.
13. Gray, Irving, *Biochem. and Biophys. Res. Comm.* 8, 247, 1962.
14. Pillemer, L. and Hutchison, M.C., The Determination of the Albumin and Globulin Content of Human Serum by Methanol Precipitation, *Jour. Biol. Chem.* 158, 299, 1945.
15. Wolfson, W.Q., Cohn, C., Calvary, E., and Ichiba, F., *Studies in Serum Proteins*, V. *Am. Jour. Clin. Path.* 18, 723, 1948.
16. Beisel, W.R., U. S. Army Medical Unit Report to Commission on Epidemiologic Survey of The Armed Forces Epidemiology Board, 1963.

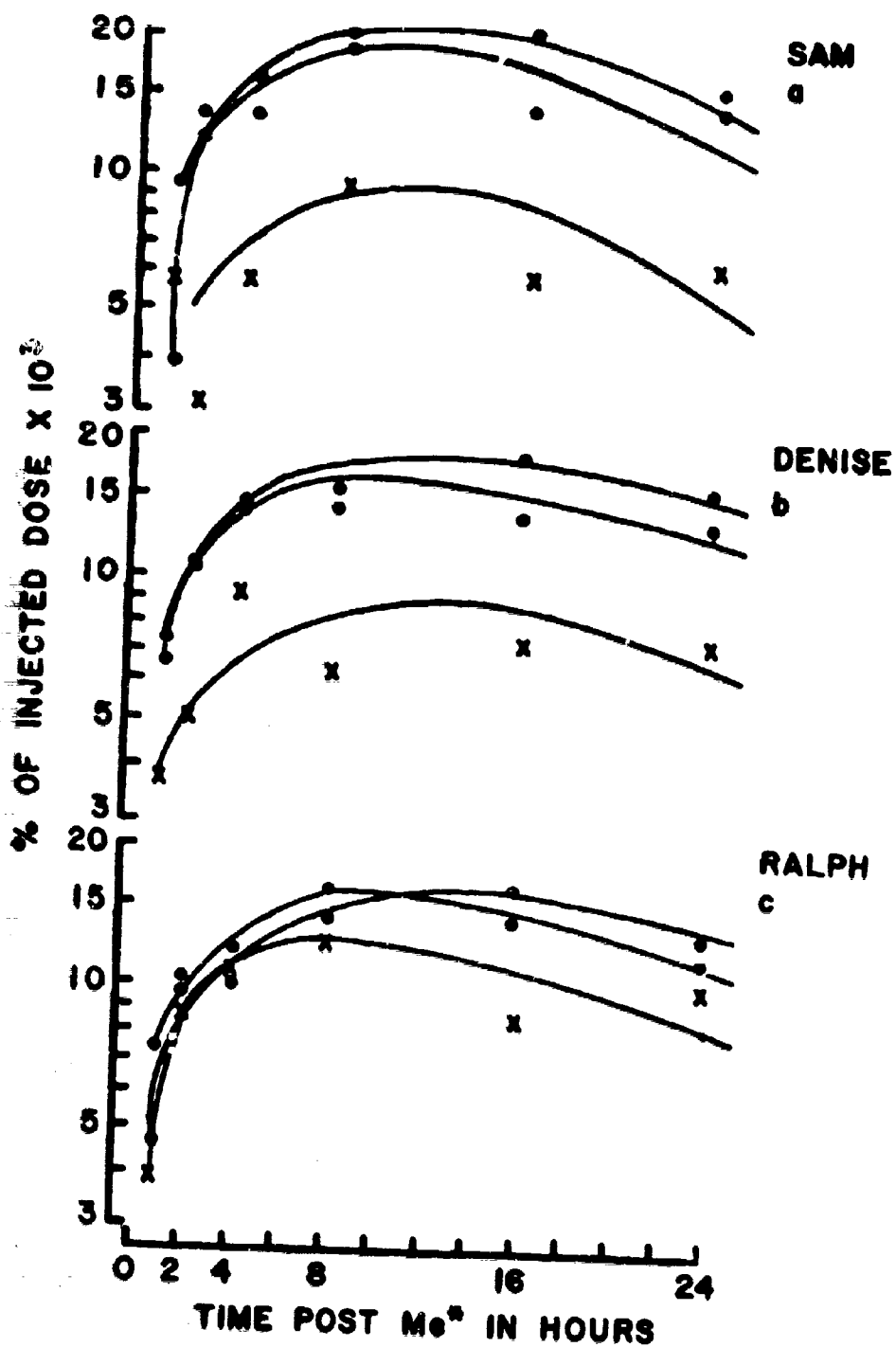


Figure 1: Plasma Globulin Turnover in Hyperthermic and Infected Chimpanzees.
 ● - Control, X = Hyperthermia, O = Hyperthermia + Virus

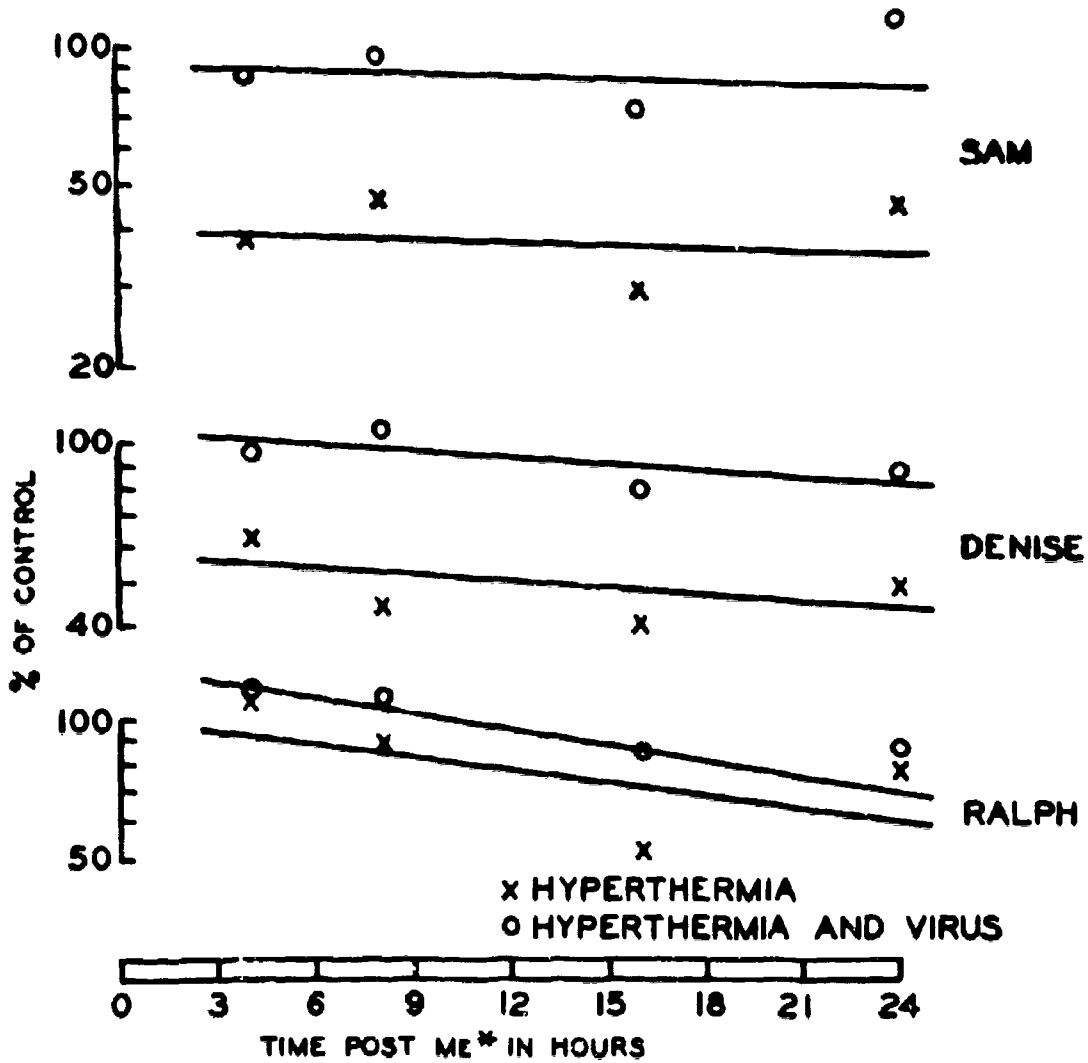


Figure 2: Plasma Globulin Breakdown in Hyperthermic and Infected Chimpanzees.

MOLECULAR MECHANISMS OF ANTIMICROBIAL ACTION

FRED E. HAHN
WALTER REED ARMY INSTITUTE OF RESEARCH
WASHINGTON, D.C.

Introduction. Infectious and other communicable diseases have at all times played an important and often decisive part in the outcome of military operations. The loss of efficiency or actual loss of life owing to communicable diseases has frequently rivaled or even surpassed losses from enemy action.

Napoleon went into the Russian campaign with an army of more than 500,000; when he began to retreat from Moscow, there were only 80,000 men left who were fit for duty. Most of the missing 420,000 had succumbed to disease. Of the other army of 500,000 which he assembled after the Russian disaster and led into the decisive battle of Leipzig, 220,000 had been lost for service by disease before the battle was joined. In the Civil War, 93,000 Federal soldiers died from enemy action but 186,000, exactly twice as many, from diseases. In World War I, the typhus epidemic in Serbia in 1915 killed over 150,000 and rendered Serbia an area virtually forbidden to military operations. In the Russian Civil War, 1917-1921, there were more than 25,000,000 cases of typhus with three million dead. In World War II, some 500,000 American servicemen acquired malaria with a total of 6.6 millions of days lost, and in the Korean War, 1950-1953, the total number of cases of communicable diseases was 85,000 with a loss of 2 millions of days.

Military Medicine must, therefore, recognize communicable diseases as a continuing and most formidable menace to the efficiency of the Army, particularly during special operations and in war and is dealing with these problems through the application of preventive and curative measures. In this communication I shall be concerned with scientific aspects of curative medicine.

The specific curative treatment of communicable diseases is known as chemotherapy. The scientific possibility of chemotherapy, first envisaged by Paul Ehrlich around the turn of the Century, was

developed by him and his associates into a medical actuality through the discovery of salvarsan in 1912. Since that time, massive efforts have continued at synthesizing or discovering chemical compounds which possess selective toxicity for disease-producing microorganisms and, hence, curative properties in communicable diseases. Contrary to the public belief in the universal efficacy of "wonder drugs," however, important categories of communicable diseases, for example the virus diseases, have remained refractory to chemotherapy. Furthermore, drug resistance is progressively emerging in diseases that were no longer considered to be serious threats to military health. For these reasons, continued and intensive research in chemotherapy is necessary.

From the molecular nature of chemotherapeutic drugs it follows logically that chemotherapy has a molecular basis. Not only are the chemical structures of all such drugs well known but in most instances they have been laboriously developed or modified in order to achieve optimal curative effects with minimal toxicity to patients. Theoretical efforts to establish relationships between chemical structure and chemotherapeutic potency are as old as the science of chemotherapy.

That the molecular nature of chemotherapeutic drugs must have correlates in the molecular organization of susceptible microorganisms was intuitively recognized by Ehrlich and expressed in his chemoreceptor hypothesis. Only during the past 25 years, however, has microbiological chemistry advanced to resolve the biochemical basis of drug action, and only during the past decade has the molecular biology of drug action begun to emerge as a field of research that elucidates the mechanisms of action of drugs at the primary, i.e., molecular, level of organization and promises to recognize principles that can perhaps be applied to a logical and premeditated design of new drugs. The long-range objective of this research is, therefore, to develop chemotherapy from an empirical into a predictive science. The present paper is concerned with perspectives and experimental studies in the molecular biology of drug action.

1. Inhibitors of Cell-Wall Synthesis. Bacteria, as well as most other microorganisms are enclosed by cell-walls which provide mechanical support and protection. The contents of the cell exert a significant internal pressure upon the walls, and damage to the wall or its deterioration will result in rupture and lysis of the cells unless special experimental precautions are taken. The classical example of a drug, active upon cell-walls, is penicillin.

When penicillin was added to cultures of Escherichia coli, growing exponentially in defined liquid media, and the turbidity of such cultures was measured at time intervals, the density of the cultures decreased precipitously after a certain lag (Fig. 1), owing to lysis of the organisms. Without mechanical agitation and upon adding sucrose for osmotic protection, lysis could be prevented (1).

Under the phase-contrast microscope a sequence of events was observed in these sucrose-containing cultures which is depicted in Fig. 2. The bacterial rods produced central or terminal globular extrusions that increased in size while the bacterial cell-walls became progressively empty of cytoplasm. Later the globes either separated from the cell-walls or retained parts of the walls attached. Finally the globular structures underwent partial vacuolization, releasing their entire content, and the only structural elements remaining were circular "ghosts" that represented the empty cytoplasmic membranes (1,2).

Independent biochemical evidence (summarized in 3) proves that penicillin leads to an accumulation of building blocks for cell-wall synthesis by interfering with the polymerization of this material into a basal constituent of the wall. Thus, penicillin interferes with the growth and maintenance of the intact cell-wall. Under conditions where the cytoplasmic content, particularly the proteins of the cell, continue to be synthesized, the wall finally ruptures and the cell content, inclosed in a cytoplasmic membrane extrudes (Fig. 2).

Examination of the chemical structures of penicillin and of the antibiotic, cycloserine (oxamycin), suggested to us a certain similarity between the two compounds (Fig. 3) and, hence, the possibility that they might have similar modes of action. This turned out, indeed, to be the case (4). Cycloserine also caused lysis of bacteria, it produced the same structural changes of bacterial cells as does penicillin, and it led to the accumulation of similar biochemical intermediates of cell-wall synthesis (4). When we referred our basic findings to Strominger for detailed enzymological study, he demonstrated that cycloserine blocks the formation of low-molecular building blocks for cell-wall synthesis at two well-defined and successive enzymatic steps (3).

Additional antibiotics which have been found to inhibit the biosynthesis of cell-walls in a manner similar to that of penicillin are bacitracin and novobiocin (3). In sum, it can be stated, therefore, that there exists one category of antimicrobial drugs which produce their effect by interfering with the growth and maintenance of cell-walls. The prototype of this category of drugs is penicillin which is perhaps the most important chemotherapeutic drug in current medical use. Since the constituent cells of mammalian tissues and organs do not possess the anatomical and biochemical equivalents of bacterial cell-walls, penicillin is generally non-toxic for man while highly toxic for bacteria.

2. Inhibitors of Protein Synthesis. Enzymes are of vital importance because either singly, or organized into cooperatively functioning systems, they catalyze the chemical reactions of the cell and thus constitute the metabolic machinery which is responsible for growth and multiplication. Any general inhibition of protein synthe-

HARD

sis amounts specifically to a condition in which microorganisms are rendered unable to synthesize enzymes. General inhibition of protein synthesis is produced by several important antibiotic drugs, for example, chloramphenicol, the tetracyclines, streptomycin, erythromycin, and the less important antibiotics, puromycin, streptogramin, and lincomycin.

A convenient way of studying protein synthesis uncomplicated by cellular growth and multiplication is afforded by the phenomenon of induced enzyme synthesis. Bacteria possess the inherited capacity of synthesizing numerous enzymes that are not normally constituents of these cells but whose formation can be induced by supplying their respective substrates to the organisms.

For example, when lactose, maltose, arabinose, or acetate was added to suspensions of *E. coli* in media that did not support growth, the cells failed at first to metabolize these various substrates. However, after 30 minutes, oxygen consumption began slowly and increased exponentially, indicating the progressive synthesis of induced enzymes, necessary for the metabolism of the initially added inducers (Fig. 4). In contrast, when chloramphenicol or chlortetracycline were present in such bacterial suspensions, no induced enzyme synthesis was observed (5). In bacteria which had synthesized the enzymes prior to the addition of the drugs, high enzyme activity was immediately evident and was not influenced by chloramphenicol (5). The drug, thus inhibited the induced synthesis of enzymes but not the catalytic activity of the enzymes once they had been synthesized. Similar results have been obtained with other antimicrobial drugs, and in all instances, such findings are indicative of an inhibition of protein synthesis.

An observed inhibition of protein synthesis can either be a specific effect, or it can result from a blockade of any of a number of essential supportive processes. We have devoted much effort to distinguishing among these possibilities (6,7) and conclude that the effects of chemotherapeutic drugs upon protein synthesis are specific (8).

Fig. 5 illustrates the specificity of inhibition of protein synthesis for streptomycin as an example (9). When this drug was supplied to a culture of *E. coli*, growing exponentially, bacterial density estimated spectrophotometrically at 420 m μ ceased to increase. Thus, streptomycin inhibited strongly an increase in bacterial mass and, hence, in protein. This was borne out in detail by chemical analysis (9). Synthesis of the other essential bacterial polymers, deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) continued as is shown by the continued increase in light absorption at 260 m μ . Even cell division, i.e., the actual formation of new cells, inclosed in new cell-walls, continued to some extent; this increase was measured in an automatic cell counter, verified by microscopic observation and is indicated in Fig. 5 as an increase with time in "total count."

Similar diagrams have been obtained for chloramphenicol, tetracyclines, erythromycin, and puromycin (summarized in 10). They indicate that the essential energy-generating and energy-transferring reactions, the vast majority of reactions of the intermediary metabolism, and many biosynthetic processes are not primarily influenced by the drugs and that the effects on protein synthesis are, indeed, highly specific.

Thus, a number of antimicrobial drugs which possess different and unrelated chemical structures exhibit in terms of microbial physiology the same mode of action, *viz.*, inhibition of protein synthesis. This seeming paradox is being resolved by investigation of the different primary mechanisms of action of the different drugs at the molecular level.

In considering inhibitions of protein biosynthesis, the center of our attention (Fig. 6) is occupied by the terminal mechanism through which amino acids for protein synthesis are assembled correctly into their genetically determined sequences and condensed through the establishment of peptide bonds. These processes are mediated through *ad hoc* assemblages made up of ribosomal particles and work copies of the genetic code, known as messenger-RNA. As the mechanistic details of the final events are still uncertain, the protein synthesizing system is symbolized in Fig. 6 as a "black box" for which only the input and output of material are indicated in detail.

The individual cistrons ("genes") in DNA contain the original information for the establishment of the correct linear amino acid sequences in the corresponding proteins to be synthesized. This information is laid down in the form of specific sequences of trinucleotides; each such triplet is a coding group for one amino acid, and the sequence of these "codons" is thought to determine the sequence of amino acids that must be established in protein synthesis (summarized, for example, in 11).

DNA is transcribed into a category of RNA which in base composition and, by inference, in the sequence of coding signals corresponds to the cistron from which it is copied. It is assumed that each RNA transcript of individual cistrons serves as a genetic messenger which informs and assembles a ribosomal system for protein synthesis (summarized in 12). In Fig. 6, this idea is summarized by the arrow which indicates the input of messenger-RNA into the "black box."

The lower right of Fig. 6 shows how amino acids and adenosine triphosphate (ATP) enter into the synthetic system. These amino acids are enzymatically activated and in one integrated reaction esterified with the terminal adenosine residue of a second class of RNA, known as "transfer-RNA." Each transfer-RNA molecule represents an adaptor for the correct placement of its attached amino acid alongside the messenger-RNA molecule in such a manner that complementary coding groups in the two types of RNA form specific hydrogen bonds. The

HAHN

process of combining amino acids and transfer-RNA is, therefore, summarized in Fig. 6 by the term "adaptation" and constitutes essentially the encoding of amino acids in preparation for protein synthesis.

The encoded amino acyl-transfer-RNA units do not reach the ribosomal system in the black box spontaneously and directly in such a manner that protein synthesis results. There are interposed two further enzymatic reaction steps. Uncertainty about the details of these reactions is indicated by a question mark behind the term "transfer" (Fig. 6). Irrespective of this uncertainty the amino acid code is read by the synthetic system in a stepwise progressing linear fashion by which the amino acids are assembled and condensed, beginning at the N-terminal member of the chain.

The individual transfer-RNAs are being recycled and re-used as the synthetic process progresses. Once the protein chain is completed, it is released from the site of its assemblage. The messenger-RNA serves several times over in individual acts of protein synthesis but is eventually degraded; its building blocks are recycled into a new transcription process as suggested by the term "turnover" (Fig. 6).

Our own work has been much concerned with the action of the antibiotic, chloramphenicol. Evidence has been summarized previously (13) that this drug does not act upon the activation and adaptation steps (Fig. 6). We have found, however, that chloramphenicol causes an accumulation of unused messenger-RNA.

When E. coli was exposed to the drug for several duplication times and chloramphenicol was then removed, much of the RNA that had accumulated during the inhibition of protein synthesis was enzymatically degraded by the organism (14). Such metabolic instability is characteristic for messenger-RNA (12).

The most characteristic property of messenger-RNA, however, is its base composition. This RNA, as stated above, is derived from DNA by a direct transcription process and, hence, possesses a base composition which corresponds to that of the DNA from which it is copied. If it could be shown that the labile portion of chloramphenicol-RNA had a DNA-like base composition it would suggest that the material were messenger-RNA. This hypothesis was tested with Bacillus cereus as a test organism (15) because the base compositions of its DNA and its global RNA differ markedly, a circumstance which should lend significance to any analytical findings on chloramphenicol-RNA.

The left half of Fig. 7 shows the accumulation of RNA in B. cereus during 90 minutes of exposure to chloramphenicol. The bacteria were then freed from chloramphenicol, and the right side of Fig. 7 shows the subsequent loss from the organism of a large portion

of the RNA which had accumulated during inhibition of protein synthesis by chloramphenicol. This lost material appeared in the form of free purine and pyrimidine bases in the suspending medium; hypoxanthine was derived from adenine and uracil from cytosine. The bases resulting from degradation of chloramphenicol-RNA were separated by ion-exchange chromatography and quantitated spectrophotometrically; the results of these analyses are tabulated in Table 1.

Table 1

Base Composition of Nucleic Acids of Bacillus cereus in Mole per Cent

	Guanine	Adenine (+Hypoxanthine)	Cytosine	Thymine or Uracil	Sum of Pyrimidines
DNA	17.8	32.2	17.5	32.6	50.1
Chloramphenicol-RNA Fragments	20.8	29.8	5	44	49.3
RNA	31.0	25.5	20.1	23.4	43.5

Examination of the data in Table 1 shows that the relative abundance of guanine, adenine (+hypoxanthine), and of the sum of the pyrimidine bases in the degradation products of chloramphenicol-RNA was quite similar to the base composition of the DNA of B. cereus but markedly different from the composition of the global RNA of the organism. These findings (15) are in accord with the hypothesis that the RNA which breaks down after removal of chloramphenicol is messenger-RNA which originally accumulates because it can not be utilized for protein synthesis in the presence of the drug. Some preliminary information exists to the effect that labile, DNA-like RNA which accumulates under such conditions possesses messenger-RNA activity when tested in cell-free model systems of protein biosynthesis.

Such a cell-free system (16), containing ribosomes, some form of messenger, radioactive amino acids, essential enzymes, transfer-RNA, and sources of energy has been used widely during the past two years in studies on protein synthesis. All drugs cited above inhibit the incorporation of radioactive amino acids in this system (reviewed in 10). We have specifically established this fact for the antibiotic, erythromycin (17). Results of this study are listed in Table 2.

Table 2
Inhibition by Erythromycin of Polyphenylalanine Synthesis in a
Cell-free System

Composition of the System	Counts per min per total ppt.
Complete System	33,649
Energy Source (ATP) omitted	2,762
Poly U (messenger) omitted	764
Erythromycin (1.0 μ moles) added	14,815
Erythromycin (1.5 μ moles) added	6,765

The availability of cell-free model systems of protein synthesis which are inhibited by chemotherapeutic drugs affords experimental approaches to the solution of problems of mechanisms of action at the molecular level. One set of such mechanisms under current consideration would require that the drug molecules interact directly with the ribosomal particles that mediate protein synthesis. We have shown such a direct interaction of E. coli ribosomes and radioactive chloramphenicol.

To graded amounts of ribosomal suspensions were added identical quantities of chloramphenicol alone or in mixture with erythromycin. After collecting the ribosomes by high-speed centrifugation they were quantitated by chemical analysis for their ribose moiety and counted for radioactive chloramphenicol. The diagram, Fig. 8, shows that the amount of chloramphenicol bound to ribosomal particles was proportional to the quantities of these particles and that a given concentration of erythromycin inhibited the interaction between ribosomes and chloramphenicol by a constant percentage, irrespective of the amounts of ribosomal material present. Such a situation is indicative of reversible inhibition and suggests that two inhibitors of protein synthesis, chloramphenicol and erythromycin, competed for the same site of attachment on the ribosomes. Since streptomycin and the tetracyclines did not interfere with the interaction between ribosomes and chloramphenicol, it is inferred that they possess different mechanisms of action.

3. Inhibitors of RNA Synthesis. The vital importance of messenger-RNA has become apparent during the preceding discussion of genetically controlled protein synthesis. Data in Table 2, for example, show that a cell-free system was virtually non-active with respect to amino acid incorporation when the messenger was withheld. As messenger-RNA must be continuously reassembled in microorganisms

(12) any substance that inhibits RNA synthesis should be a very effective antimicrobial compound.

So far, only one family of antibiotics, exemplified by the prototype, actinomycin D, is known to have this mechanism of action (18). Actinomycin D combines with DNA; the immediate consequence is a blockade of the transcription of all RNA, including messenger-RNA. Soon thereafter, when the last molecule of pre-existing messenger-RNA has been consumed, protein synthesis comes to a standstill, and later DNA formation also ceases.

Actinomycin itself is not of practical clinical importance. However, in terms of general principles of action, the drug suggests that a search for substances that inhibit RNA synthesis at the level of the genome by interfering with the transcription mechanism would hold the promise of yielding useful drugs.

4. Inhibitors of DNA Synthesis. It is self-evident that an inhibition of DNA formation, i.e., in biological terms an induced inability of a microorganism to duplicate its genetic material, will lead to radical antimicrobial effects. The antibiotic, mitomycin C is a natural substance which has this effect (19). The compound inhibits DNA synthesis; RNA and protein synthesis remain relatively unaffected for a time but the viability of test bacteria is markedly decreased. The mechanism of action of mitomycin C is still unknown. Current hypotheses assume that the antibiotic produces in vivo cross-links between the two companion helices of DNA; such a mechanism of action, once proved, would tend to characterize mitomycin C as a natural radiomimetic compound.

As an anticancer agent, the antibiotic has disappointed. However, mitomycin C exemplifies a principle of action, viz., inhibition of gene replication, that may conceivably be exploited in chemotherapy research. There exists, however, an element of doubt whether any agent acting in that manner could be rendered sufficiently specific for microbial DNA so as to eliminate antimitotic or other cytotoxic effects in man.

Summary. Antibiotics have been found to inhibit (1) cell-wall synthesis (penicillin), (2) protein synthesis (chloramphenicol, tetracyclines, streptomycin, erythromycin), (3) RNA synthesis, and (4) DNA synthesis. Mechanisms by which these effects are produced have been studied and typical results of this research are reported. Chloramphenicol and erythromycin appear to inhibit protein biosynthesis by direct interaction with ribosomes. In all instances, the antimicrobial actions occurred in the general area of macromolecular biosynthesis and assemblage.

REFERENCES

1. Hahn, F.E., and J. Ciak. *Science* 125, 119 (1956).
2. Duguid, J.P. *Edinburgh Med. J.* 53, 402 (1946).
3. Strominger, J.L. *Fed. Proc.* 21, 134 (1962).
4. Ciak, J., and F.E. Hahn, *Antibiotics & Chemotherapie* 9, 47 (1959).
5. Hahn, F.E., and C.L. Wisseman. *Proc. Soc. Exptl. Biol. Med.* 76, 533 (1951).
6. Hahn, F.E., C.L. Wisseman, and H.E. Hopps. *J. Bact.* 67, 674 (1954).
7. Hahn, F.E., C.L. Wisseman, and H.E. Hopps. *J. Bact.* 69, 215 (1955).
8. Hahn, F.E. *Proc. Fourth Internat. Congr. Biochem.* 7, 104 (1959).
9. Hahn, F.E., J. Ciak, A.D. Wolfe, R.E. Hartman, J.L. Allison, and R.S. Hartman. *Biochim. Biophys. Acta.* 61, 741 (1962).
10. Hahn, F.E. *Proc. Third Internat. Congr. Chemotherapy, in the press* (1964).
11. Crick, F.H.C. *Progr. Nucleic Acid Res.* 1, 164 (1963).
12. Jacob, F., and J. Monod. *J. Molecular Biol.* 3, 318 (1961).
13. Hahn, F.E. *Antimicrobial Agents Annual 1960*, 310 (1961).
14. Hahn, F.E., M. Schaechter, W.S. Ceglowski, H.E. Hopps, and J. Ciak. *Biochim. Biophys. Acta.* 26, 469 (1957).
15. Hahn, F.E., and A.D. Wolfe. *Biochem. Biophys. Res. Comm.* 6, 464 (1962).
16. Wrenberg, M.W., and J.H. Matthaei. *Proc. Nat. Acad. Sci.* 47, 1588 (1961).
17. Wolfe, A.D., and F.E. Hahn. *Science, in the press* (1964).
18. Hurwitz, J., J.J. Furth, M. Malamy, and M. Alexander. *Proc. Nat. Acad. Sci.* 48, 1222 (1962).
19. Shiba, S., A. Terawaki, T. Taguchi, and J. Kawamata. *Nature* 183, 1056 (1959).

HAHN

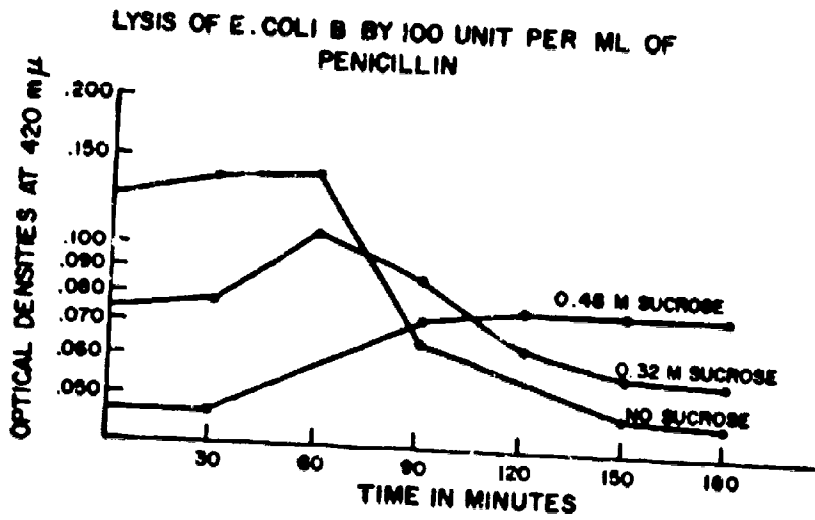
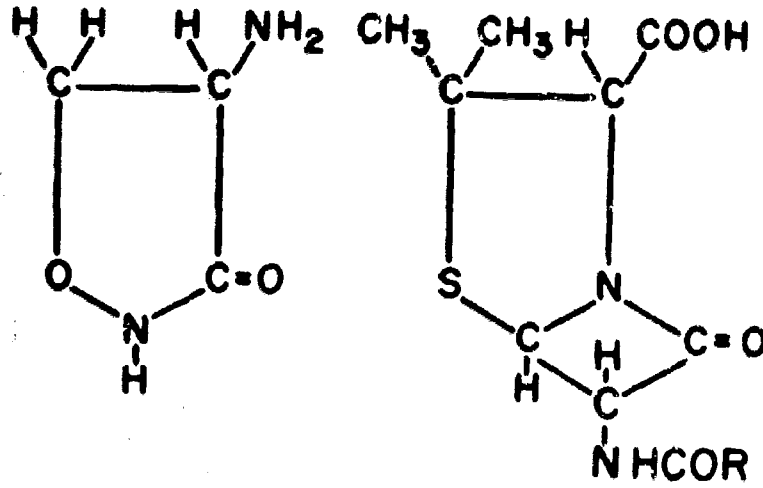


Figure 1.

Figure 2. Development of Spheroplasts by Penicillin-Exposed Escherichia coli.



HAIN



D-OXAMYCIN

PENICILLIN

Figure 3.

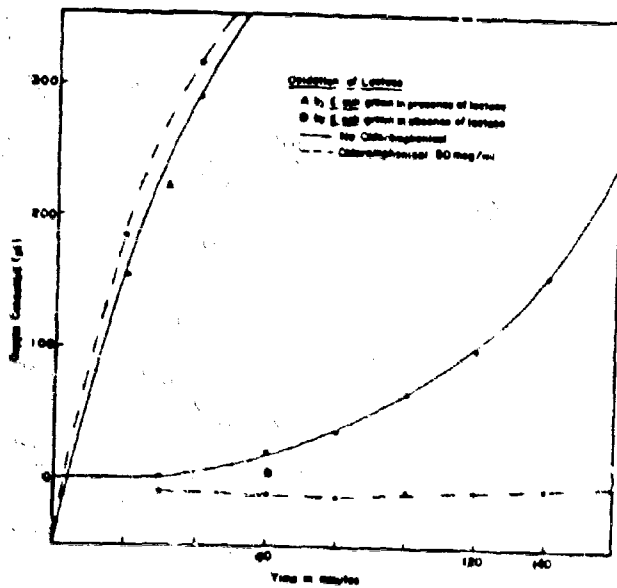


Figure 4.

Inhibition of Induced Enzyme Synthesis

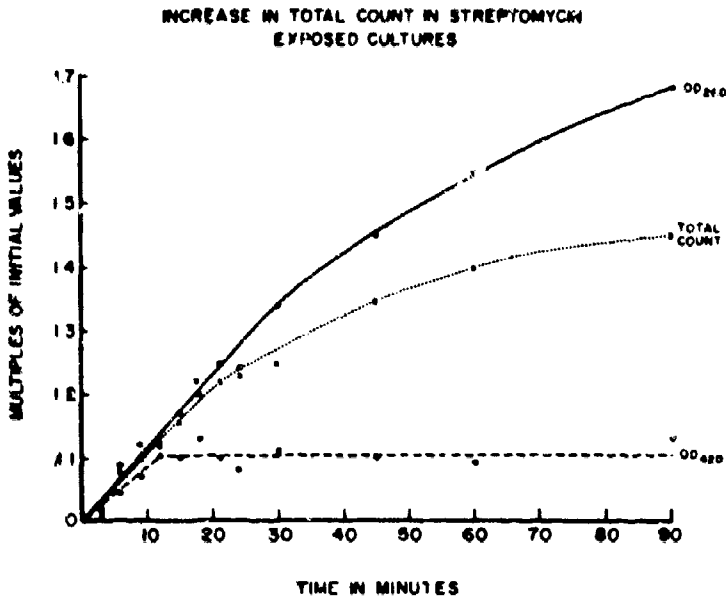


Figure 5.

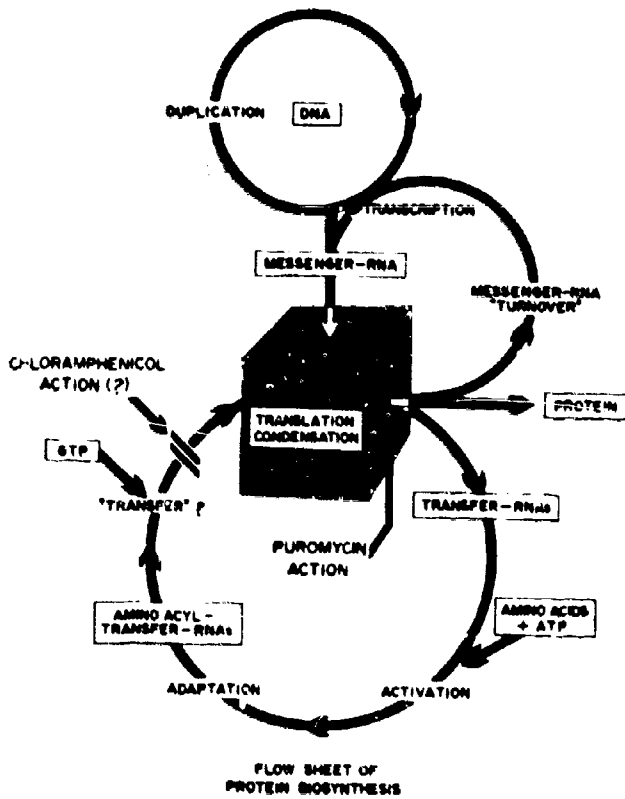


Figure 6.

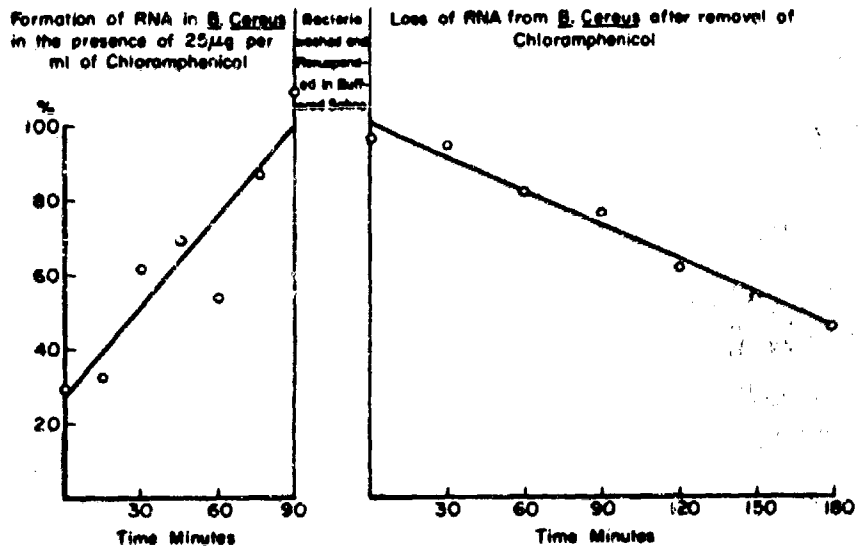


Figure 7.

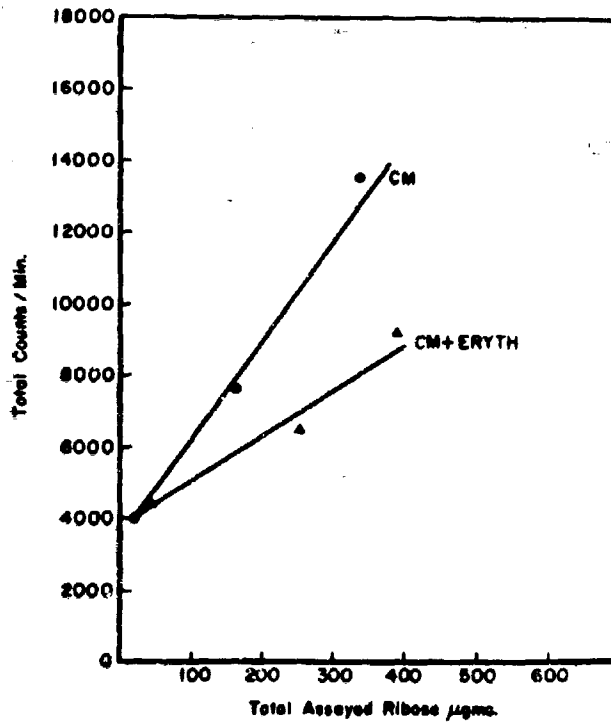


Figure 8.

Binding of Chloramphenicol- C^{14} by Ribosomes

INFLUENCE OF TRAUMA AND HEMOLYSIS
ON HEMORRHAGIC SHOCK IN DOGS

ROBERT M. HARDAWAY, III, COLONEL, MC
WALTER REED ARMY INSTITUTE OF RESEARCH
WASHINGTON, D. C.

Intravascular hemolysis has in general been considered to be detrimental. Hemolytic transfusion reaction, hemolytic crises, old bank blood, etc., have all been associated with high morbidity and mortality. During investigations attempting to assess the role of trauma in irreversible hemorrhagic shock in dogs, it became apparent that hemolysis was a prominent feature of trauma. If dogs (under Nembutal anesthesia) were subjected to 100 blows on one thigh with a padded mallet, it was noted that while hemoglobinemia was not prominent immediately, it was very prominent after two days (Fig. 1A). This amount of hemoglobinemia was roughly equivalent to that produced by the injection of 20 ml. of hemolyzed autogenous blood (Fig. 1A). Because traumatized dogs were much more susceptible to hemorrhagic shock than non-traumatized dogs, it was thought that perhaps mortality was correlated with hemolysis. A series of experiments was therefore set up to assess the role of hemolysis in hemorrhagic shock. The present paper reports the results of these experiments.

Materials and Methods

The principles of laboratory animal care as promulgated by the National Society for Medical Research were observed. One hundred twenty-three mongrel dogs of either sex weighing between 12 and 18 kg. were sedated with morphine. Under local anesthesia and with sterile conditions a polyethylene catheter was inserted into a femoral artery and used to record blood pressure on a Sanborn recorder. Another polyethylene catheter was inserted into a femoral vein and used to administer electrolyte solutions, medication and blood. The second femoral artery was exposed but

* These experiments will be reported elsewhere.

HARDWAY

not catheterized. When blood was to be drawn for sampling or for hemorrhaging, a short polyethylene catheter was inserted into the artery and immediately withdrawn and discarded after use. Artery and tubing were allowed to flush before a sample was collected. The animals were divided into 10 groups as follows:

Group 1. Fifteen dogs were bled through a short polyethylene catheter directly into a clinical type of blood donor bag (Fenwal) containing ACD solution. Blood was allowed to flow into the bag until it contained 400 ml., care being taken to mix the blood well. The tube was withdrawn and tied close to the bag. A second clinical donor bag (Fenwal), this time containing heparin as an anticoagulant, was then used in a similar manner, blood being withdrawn until the mean aortic pressure was 40 mm. Hg. The animal was allowed to stabilize for 30 minutes during which time the aortic pressure usually rose somewhat. At this time additional blood was withdrawn using a new short length of polyethylene tubing to bring the pressure again to 40 mm. Hg. This blood was used for tests or discarded after measurement. For the next 3½ hours the mean arterial pressure was kept at 40 mm. Hg. either by withdrawing small amounts of blood through a new segment of plastic tubing or by administering small amounts of ACD blood intravenously. No heparinized blood was given. At the end of the 4-hour shock period, a blood sample was again taken through a new short plastic tube after allowing the artery to flush itself. Meanwhile all bled blood had been measured, recorded and discarded. After the final sample, all blood remaining in the ACD bag and all blood in the heparin bag was returned to the animal intravenously through a recipient set filter. The central venous pressure was measured during and after the retransfusion using the intravenous catheter. The blood used for samples or discarded during the shock period was replaced with normal saline. After the retransfusion, the catheters were withdrawn and the wounds sutured under sterile conditions. The animals were returned to their cages where food and water were available but no other treatment given. All animals were observed for 48 hours or until death. Animals alive at 48 hours were considered survivors. Blood samples were obtained before start of hemorrhage, 30 minutes after hemorrhage, and just before the retransfusion after four hours of shock. Samples were tested as follows: 1) Lee White clotting time using siliconized tubes kept in a constant temperature water bath; 2) hematocrit by the microtechnique; 3) prothrombin time; 4) fibrinogen level; (14) 5) total protein; 6) platelet count; 7) endogenous heparin level using azure-A dye precipitation with colorimetry. (19,20)

Group 2. Thirty-four dogs were treated in the same manner as in Group 1 except that before bleeding, 20 ml. of blood was withdrawn, frozen in a dry ice bath, thawed, and immediately returned to the animal.

HARDAWAY

Group 3. Thirteen dogs were treated in the same manner as in Group 2 except that 100,000 units of fibrinolysin (Thrombolylin, Merck, Sharpe and Dohme) was given intravenously according to the following schedule:

- 12,500 units after initial blood sample.
- 12,500 units after administration of hemolyzed blood.
- 6,250 units every 30 minutes during shock period.

The remainder of the total dose was added to the retransfusion at the end of the shock period.

All of these animals were paired with 13 animals of Group 2 and done on the same table at the same time. Selection as to which animal was in Group 2 or Group 3 was accomplished just before the procedure by the toss of a coin.

Group 4. Thirteen dogs were placed on the operating table and, after preparation, 100 ml. of blood was withdrawn, frozen by dry ice, and reinjected into the animal. Samples were taken as in the other groups but no hemorrhagic shock was induced.

Group 5. Two dogs were treated as in Group 4 except that only 50 ml. of blood was hemolyzed.

Group 6. Two dogs were treated as in Group 4 except that only 20 ml. of blood was hemolyzed.

Group 7. Thirteen dogs were treated as in Group 2 (hemorrhagic shock and 20 ml. of hemolyzed blood) except that instead of 20 ml. of hemolyzed blood, an equivalent amount of hemoglobin was given. An attempt was made to purify this hemoglobin by extracting hemolyzed cells with chloroform and using the remaining hemoglobin after filtration.

Group 8. Fourteen dogs were given fibrinolysin in the same dosage and schedule as those dogs in Group 3 but were not subjected to hemorrhagic shock nor injected with hemolyzed blood.

Group 9. Eight dogs were subjected to hemorrhagic shock and 20 ml. of hemolyzed blood as in Group 2 but in addition were given heparin according to the following dose schedule: 1,000 units or 10 mg. per kg., 80% of which was given before hemorrhage and 20% one hour later. This probably is an inadequate dosage to prevent DIC in view of the hypercoagulability of dog blood, particularly as influenced by shock and hemolyzed blood.

Group 10. Nine dogs were given fibrinolysin as in Group 8, (no hemorrhagic shock) but in addition were given 100 ml. of hemolyzed blood after the first dose of fibrinolysin.

HARDAWAY

Results

Mortality.

	<u>No. of Dogs</u>	<u>Died</u>	<u>% Mortality</u>
Group 1 Shock alone	15	2	13
2 Shock and hemolyzed blood (20 ml.)	34	31	91
3 Shock, hemolyzed blood (20 ml.) and fibrinolysin	13	5	38
4 Hemolyzed blood alone (100 ml.)	13	0	0
5 Hemolyzed blood alone (50 ml.)	2	0	0
6 Hemolyzed blood alone (20 ml.)	2	0	0
7 Shock and purified hemoglobin	13	11	84
8 Fibrinolysin alone	14	0	0
9 Shock, hemolyzed blood (20 ml.) and heparin	8	8	100
10 Hemolyzed blood (100 ml.) and fibrinolysin	9	0	0

Hypotension alone at 40 mm. Hg. mean aortic pressure for four hours resulted in a mortality of 13%. However, the addition of 20 ml. of hemolyzed autogenous blood increased this to 91%. This is highly significant ($P < .001$). When fibrinolysin is administered to animals with an otherwise 91% fatal shock, however, the mortality is reduced to 38%. This was done in paired experiments making possible the method of sequential sampling (Fig. 2). By following the chart, (see Fig. 2) it was possible to stop the experiment when statistical validity was attained ($P < .05$). This is statistically valid evidence that fibrinolysin is effective in treatment of this type of shock in the dog. Groups 4, 5, and 6 had only hemolyzed blood given without any shock procedure. It is evident that even 100 ml. of hemolyzed blood (compared with 20 ml. administered in Groups 2 and 3) is innocuous by itself.

HARDAWAY

In Group 7 an attempt to purify the hemoglobin from the hemolyzed blood was not helpful in reducing mortality. It was somewhat helpful in other parameters (see below) (Fig. 5). This may indicate that the toxic agent is not hemoglobin but is closely associated with it.

In Group 8 it is shown that fibrinolysin in the dosage used produces no mortality. In Group 9 an attempt to protect the animals with heparin was unsuccessful although as seen below, it was quite successful in preventing the fibrinogen fall. The reason for this is unknown but may possibly be due to a specific toxic combination of heparin with hemolyzed blood or possibly to an inappropriate heparin dosage. Fibrinolysin in combination with hemolyzed blood is innocuous (Group 10).

Central Venous Pressure. In no case was there any elevation of central venous pressure in any group at any time during the experiment including readings taken after the retransfusion.

Silicone Clotting Times. A prolonged silicone clotting time after shock (above 60 minutes) has previously been shown to be associated with irreversible hemorrhagic shock. (9) The present experiments confirm this. In Group 1, shock alone causes a lengthening of clotting time but only into the border zone of 55 minutes. This shock is usually not fatal (13%). However, Group 2 with a mortality of 91% shows a clotting time at the end of the shock period of 141 minutes. Hemolyzed blood alone will cause a transient but marked prolongation (in fact may never clot) but this rapidly returns to normal within four hours and is not associated with mortality. This happens only if large amounts (100 ml.) are given. Small amounts (20 ml.) as were given to shocked animals have no effect. The cause for this prolonged silicone clotting time is obscure. It correlates to some extent with endogenous heparin (see below) but this is not the only factor. Silicone clotting times show little correlation with glass clotting times; in fact the glass time is frequently normal when the blood will never clot in silicone. The blood can be made to clot with the addition of almost anything (calcium, particulate matter, glass).

Prothrombin Times. These roughly parallel the silicone clotting times in that those groups with greatly prolonged silicone clotting times also show significantly ($P < .01$) prolonged prothrombin times. There is a significant ($P < .01$) difference between the shocked animals given hemolyzed blood (Group 2) and those given a purified hemoglobin (Group 7). In this parameter the purification process seemed to attenuate the effect of hemolyzed blood separating the toxic substance from hemoglobin whereas it was not significant in mortality (see above). A significant ($P < .01$) elevation in prothrombin time occurs 30 minutes after 100 ml. of hemolyzed blood alone but this returns to normal within four hours.

HARDAWAY

This parallels similar changes in the silicone clotting time and the heparin level.

Hematocrit. The hematocrit level fell in all shocked animals and can be used as a measure of hemodilution. However, it is felt that total protein is a better index of dilution.

Total Protein. The protein level fell in all shocked animals, denoting, at least in large part, a hemodilution. The figures are not listed here but were used, for lack of a better method, for correcting the values of fibrinogen and platelets after shock (see below).

Fibrinogen. (Fig. 3) Fibrinogen levels were corrected for changes in dilution of the blood by using the changes in total protein as an index of dilution. The hematocrit and total protein both fall precipitously during hemorrhage and then rise slightly during the shock period. This is due no doubt at least in part to dilution of the blood by extravascular fluid. Fibrinogen values are corrected by the following formula:

$$\frac{\text{gm. total protein before hemorrhage} \times \text{mg\% fibrinogen after shock}}{\text{gm. total protein after shock}} =$$

mg% fibrinogen after shock corrected value.

Fibrinogen decreases only moderately (5%) but significantly ($P < .05$) in shock alone, (Fig. 3) but decreases much more (17%) ($P < .01$) when 20 ml. of hemolyzed blood is injected. This points toward the activation of the clotting process by hemolyzed blood which is easily counteracted when shock is not present. Fibrinogen decreases markedly ($P < .01$) and afibrinogenemia frequently results if fibrinolysin is given during either shock or after hemolyzed blood or both, whereas the same dosage of fibrinolysin alone has significantly less effect. This may be explained as follows: Disseminated intravascular coagulation (DIC) causes a moderate fall in fibrinogen but never afibrinogenemia. However, some fibrinogen is converted to fibrin which is acted on by fibrinolysin. This keeps clearing one side of the fibrinogen \rightarrow fibrin reaction enabling the reaction to go to completion. Fibrinolysin in moderate concentration has no effect on fibrinogen. (11) Fibrinogen actually rises in the presence of shock with hemolyzed blood if heparin is added to prevent intravascular clotting (Fig. 4). This is likely due to the rapid mobilization or manufacture of fibrinogen under the influence of the stress of shock. (12) The fall in fibrinogen (which takes place without heparin) takes place in the face of this mobilization or manufacture. In the present case this manufacture probably exceeds 100 mg% (25 mg% per hour) during the shock period if its prevention by heparin is taken as a base line (Fig. 4). In some cases it may be as great as 50 mg% per hour. (13) Other experiments (18) have shown that after shock, dogs elevate their

HARDAWAY

fibrinogen at the rate of 9 mg% per hour on the average for 24 hours. It keeps rising for 48 to 36 hours or longer. (12,13)

Hemolyzed blood alone has only a transient effect even in doses of 100 ml. (Fig. 3). Fibrinogen falls transiently but is essentially normal within the 4-hour period. Purified hemoglobin, while still toxic in the present experiments, seems less so than whole hemolyzed blood. In fact, the purified product as produced for these experiments produced only a slight fall in fibrinogen (Fig. 5).

Fibrinogen values may fall due to one or more of the following mechanisms:

- 1) Blood dilution. This is corrected as above.
- 2) Failure of fibrinogen manufacture due to liver damage.

This does not start to occur for six hours even in hepatectomized dogs. (15)

3) Destruction of fibrinogen by fibrinolysin. This occurs only in high levels of fibrinolysin which were not present in these experiments except in those groups in which it was administered. Even this high dosage is relatively ineffective under normal circumstances when there is no intravascular coagulation.

4) Consumption of fibrinogen in an episode of disseminated intravascular coagulation (DIC). This apparently occurs in irreversible shock. (3-7)

Endogenous Heparin. (Fig. 6). The appearance of endogenous heparin is quite definite and significant in the fatal shock of Group 2. It is almost completely absent in controls which are not shocked (Group 8). It appears transiently after large doses of hemolyzed blood alone but quickly disappears. It seems to appear as a response to DIC and probably as a protective mechanism to prevent DIC. Endogenous heparin appears synchronously with prolonged silicone clotting times and prothrombin times (see above). However, it does not account for the changes in these determinations except in part.

Platelets. (Fig. 7). All groups showed a decrease in platelets. This was particularly significant in groups 2, 3 and 4. The group to show the least fall was Group 8, the only group to have neither hemolyzed blood nor shock. Thus it would appear that both hemolyzed blood and shock cause a decrease in platelets. Nearly all of the post shock samples showed marked clumping of the platelets probably reflecting an increased platelet stickiness. This is possibly a stage in the formation of platelet thrombi which occurs in DIC. Platelet thrombi are seen after incompatible blood transfusion. (16) Platelets completely disappear after administration of endotoxin. (17). A decrease in platelets is a hallmark of DIC. (3)

DISCUSSION

Red cells contain an agent which incites blood

HARDWAY

coagulation (1). However, rather large scale hemolysis is harmless under otherwise normal conditions. The intravenous injection of distilled water is harmless to normal humans. (2) The only effects are a transient hemoglobinemia and hemoglobinuria. As was demonstrated in the present experiments 100 ml. of autogenous hemolyzed blood produced no harmful results other than a transient, but dramatic, prolongation in silicone clotting time, a transient endogenous rise in heparin, and a mild and transient fall in fibrinogen. However, in the presence of otherwise nonfatal hemorrhagic shock, even 20 ml. of autogenous hemolyzed blood is fatal in over 90% of animals. This fatal effect is correlated with the production of an episode of disseminated intravascular coagulation (DIC). (3) It is postulated that the stagnant capillary circulation of hemorrhagic shock (4-6) plus an increase in coagulability of the blood (7,8) causes clotting of the blood in the microcirculation. Hemolysis markedly assists this coagulation due to the clotting factor in red cells. If these clots persist long enough, microscopic focal tissue necrosis due to infarction results in such vital organs as the liver, kidneys, pancreas, heart and other organs. (9) The evidence for the DIC episode is typical. (3) This includes the coagulation changes described in the present experiments, namely, a greatly prolonged silicone clotting time, the appearance of an endogenous anticoagulant (heparin) and the decrease of all clotting elements. Those measured in the present experiments include fibrinogen, prothrombin, and platelets. Other experiments have also shown decreases in factors V and VII. (17) Fall of fibrinogen level is due to its being consumed in a process of DIC and in the face of a tremendous production of fibrinogen which occurs as a result of the stress of hemorrhage and shock, which may exceed the rate of 50 mg% of fibrinogen an hour. (12,13) A significant fibrinogen drop (in the absence of large concentrations of fibrinolysin) is associated with irreversibility and death. These results are explained as follows: Coagulation uses up the clotting elements and stimulates the production of an anticoagulant in an attempt to stop the process. That the toxic effects of hemolyzed blood are due to its stimulation of DIC, is supported by the protective effect that treatment with fibrinolysin has on these animals as shown in the present experiments. Heparin appeared not to be effective in protecting the animals from death but was very effective in preventing the using up of fibrinogen. The reason for this is not known but may be due to some specific reaction of heparin and the hemolyzed blood or due to inappropriate heparin dosage. Fibrinolysin in large doses destroys fibrinogen but in small dosages does not. (10,11) However, fibrinolysin, even in small dosages, causes the almost complete destruction of fibrinogen in the presence of shock. (11) This occurred in the present experiments and is evident when one compares the levels of even a fairly large fibrinolysin dosage by itself with the results of the same dosage in the presence of shock. This indicates that while low levels of fibrinolysin will not attack fibrinogen, that in shock there is the conversion of fibrinogen all or part of the way to fibrin (possibly only to monomer fibrin) in

HARDAMAY

an episode of DIC and that the fibrinolysin then attacks the fibrin enabling the fibrinogen → fibrin equation to go to completion with the almost complete destruction of fibrinogen. This fall in fibrinogen is more dramatic and for a different reason than without the fibrinolysin. It is coordinated with survival of the animal whereas a lesser but significant fall in fibrinogen without fibrinolysin is coordinated with death in irreversible shock.

An attempt was made to isolate the toxic factor in hemolyzed blood. It is believed not to be hemoglobin. Administration of an equivalent quantity of autogenous hemoglobin (to 20 cc. of hemolyzed blood) was much less effective in producing the changes in the blood clotting mechanism and also was less lethal. The toxic factor seemed to be located in the chloroform extracted fraction and is probably a lipoprotein. This problem is being worked on further in this laboratory.

SUMMARY AND CONCLUSIONS

- 1) A small amount (20 ml.) of endogenous hemolyzed blood causes an increase in mortality of dogs in hemorrhagic shock from 13% to 91%.
- 2) This lethal effect is due to the stimulation by hemolyzed blood of disseminated intravascular coagulation (DIC), (3) which is also stimulated by hemorrhage and shock. The lethal factor is a clotting factor in the red cell.
- 3) Large amounts (100 ml.) of hemolyzed blood alone is harmless, but even a small amount (20 ml.) is lethal in the presence of shock.
- 4) Hemolyzed blood causes the conversion of fibrinogen to fibrin.
- 5) Fibrinolysin will prevent death in irreversible hemorrhagic shock both as produced by hemorrhage alone (10) or as influenced by hemolyzed blood.
- 6) Endogenous heparin is stimulated by DIC which is produced either by hemorrhage alone, hemolyzed blood alone, or by a combination of both. It is probably a protective mechanism.
- 7) A markedly prolonged silicone clotting time occurring during hemorrhagic shock is an accurate prognostication of irreversibility.
- 8) Irreversibility is correlated with a fibrinogen fall, when the fibrinogen fall is due to consumption in DIC.
- 9) Central venous pressure is not elevated even though the shock is proven to be irreversible.
- 10) Irreversibility is due to an episode of DIC. (3)

References

1. Quick, A. J., et al.: The Clotting Activity of Human Erythrocytes. Theoretical and Clinical Implications. Am. J. M. Sc., 228: 207-213, 1954.
2. Crosby, W. H.: Studies of Hemoglobinuria. Proc. VII Internat. Cong. of Internat. Soc. Hematol., Rome 1958. Vol. II, p. 424-

HARDAWAY

- 425, Grune & Stratton, New York, 1960.
3. Hardaway, R. M., et al.: The Syndromes of Disseminated Intravascular Coagulation. Review of Surgery. In press.
 4. Hardaway, R. M., et al.: A New Theory of Shock. *Mil. Med.*, 128: 198-208, 1963.
 5. Hardaway, R. M.: The Role of Intravascular Coagulation in the Etiology of Shock. *Ann. Surg.*, 155: 325-338, 1962.
 6. Hardaway, R. M., et al.: Vascular Spasm and Disseminated Intravascular Coagulation. Influence of the Phenomena One on the Other. *AMA Arch. Surg.*, 83: 173-180, 1961.
 7. Hardaway, R. M., et al.: Studies on the Role of Intravascular Coagulation in Irreversible Hemorrhagic Shock. *Ann. Surg.*, 155: 241-250, 1962.
 8. Hardaway, R. M., et al.: Role of Norepinephrine in Irreversible Hemorrhagic Shock. *Ann. Surg.*, 156: 57-60, 1962.
 9. Hardaway, R. M., et al.: Evidence of Disseminated Intravascular Coagulation in Clinical Shock. In preparation.
 10. Hardaway, R. M., et al.: Prevention of "Irreversible" Hemorrhagic Shock with Fibrinolysin. *Ann. Surg.*, 157: 39-47, 1963.
 11. Hardaway, R. M., et al.: Mechanism of Action of Fibrinolysin in the Prevention of Irreversible Hemorrhagic Shock. *Ann. Surg.*, 157: 305-309, 1963.
 12. Hardaway, R. M., et al.: Influence of Stress on Fibrinogen. In preparation.
 13. Hardaway, R. M., et al.: Studies on the Fibrinogen Replacement Rate in Dogs. In preparation.
 14. Quick, A. J.: Quantitative Estimation of Fibrinogen in Hemorrhagic Diseases. p. 436-439. Lea and Febinger, 1957.
 15. Johnson, D., et al.: Unpublished data.
 16. McKay, D. G., et al.: Alterations in Blood Coagulation Mechanism After Incompatible Blood Transfusion. *Am. J. Surg.*, 89: 583-592, 1955.
 17. Hardaway, R. M., et al.: Blood Clotting Mechanism in Endotoxin Shock. *Arch. Int. Med.* In press.
 18. Hardaway, R. M., et al.: Influence of Extracorporeal Handling of Blood During Hemorrhagic Shock in Dogs. In preparation.
 19. Houchin, D.: A Colorimetric Method for Estimation of Heparin Levels in Blood and Biological Solutions. In prep.
 20. Bassiouni, M.: The Estimation of Heparin in Blood. *Brit. J. Clin. Path.*, 6: 39, 1953.
 21. Villavicencio, J. L., et al.: Clinical Determination of Fibrinolytic Activity. *Arch. Surg.*, 78: 639-646, 1959.



Fig. 1A. On the left are three blood samples taken from a dog subjected to 100 blows of a padded mallet on one thigh. The first tube was drawn before the trauma, the second tube 30 minutes following the trauma, and the third tube 48 hours after the trauma. Note that hemolysis which was only mild right after the trauma had increased markedly by the end of two days. This is thought to be due to absorption of hematoma and ecchymosis probably via lymphatics. If no hemolyzed blood had been absorbed, the mild hemolysis shown in the second tube would have completely disappeared by 48 hours. On the right are two tubes drawn before and after the administration of 20 ml. of autogenous blood which had been frozen and thawed. Note that the amount of hemolysis 48 hours after trauma and immediately after the administration of 20 ml. of hemolyzed blood is approximately the same.

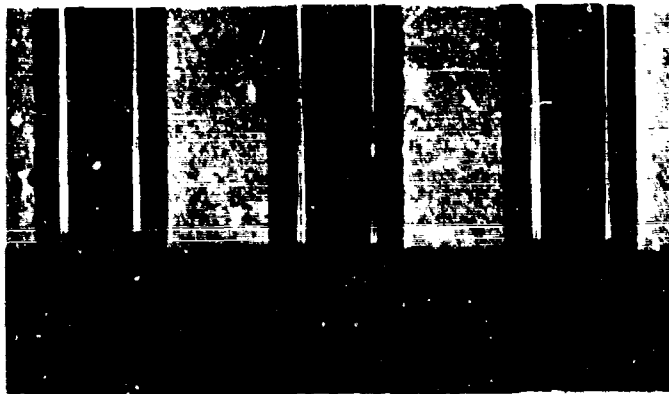
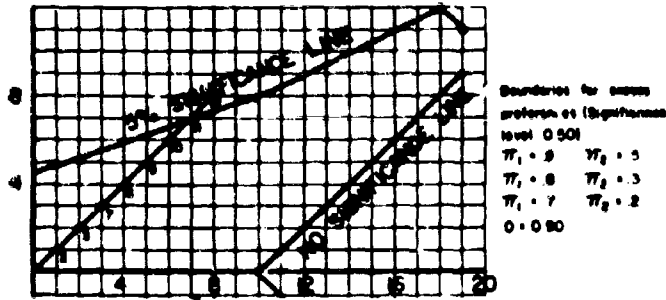


Fig. 1B. Three specimens of blood taken from a Group 4 dog given hemolyzed blood alone without any shock procedure. The first tube was drawn before administration of hemolyzed blood, the second tube was drawn 30 minutes after administration of hemolyzed blood, and the third tube four hours after administration of hemolyzed blood. Note that there is dramatic clearing of the hemoglobinemia during the 4-hour period. This clearing corresponds to the recovery towards normal of the corresponding changes in fibrinogen, clotting time and endogenous heparin.

HARDAWAY

FIBRINOLYSIN PROTECTION
SIGNIFICANCE CHART FOR 50% ADVANTAGE



and either up or down according to the results. If the treated dog survived and the untreated one died, the line was up and to the right one square. If the reverse happened the line was drawn to the right and down one square. If both dogs had the same outcome no entry was made. If the line thus produced crosses the 5% significance line, the result is significant at the 5% level. The numbers adjacent to the individual entries refer to the numbers of pairs of dogs required to attain that position. The omitted numbers were those showing no difference in their outcome.

Fig. 2. Statistical treatment of survival of paired Gp. 2 and Gp. 3 dogs. One of each pair (Gp. 3) was given fibrinolysin whereas its mate (Gp. 2) had none. A line is started at the left of the heavy horizontal line. If there was a difference in the outcome between the two dogs a line was drawn one square to the right

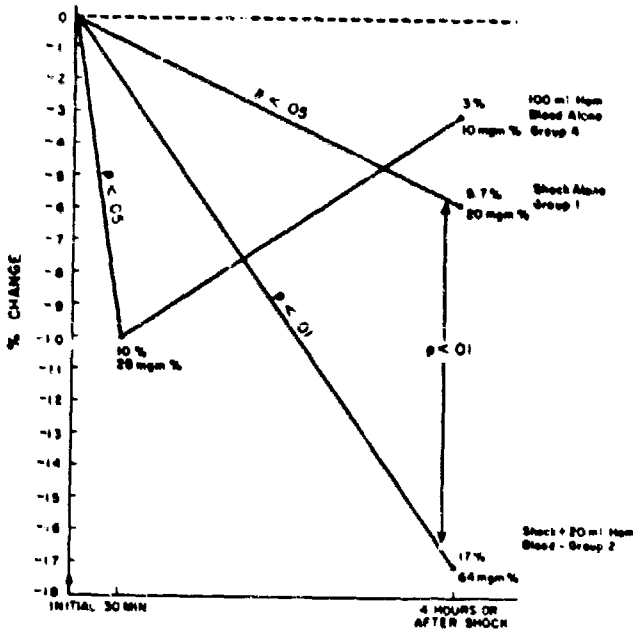


Fig. 3. % fibrinogen change. All values have been corrected for blood dilution using the total protein method as described in the text. There is a mild but significant fall in fibrinogen in the group subjected to shock alone (Gp. 1). This shock is in most cases reversible. However, there is a much more dramatic fall in fibrinogen in those dogs given hemolyzed blood and subjected to shock (Gp. 2), and which procedure resulted in death in most cases. There is a significant difference in this fall between groups 1 and 2.

100 ml. of hemolyzed blood alone without any shock procedure (Gp. 4) resulted in an immediate significant decrease in fibrinogen but this had returned to normal levels by the end of a 4-hour period. This reflects an immediate conversion of fibrinogen to fibrin by the hemolyzed blood but the deficiency is quickly made up because no further intravascular clotting was taking place.

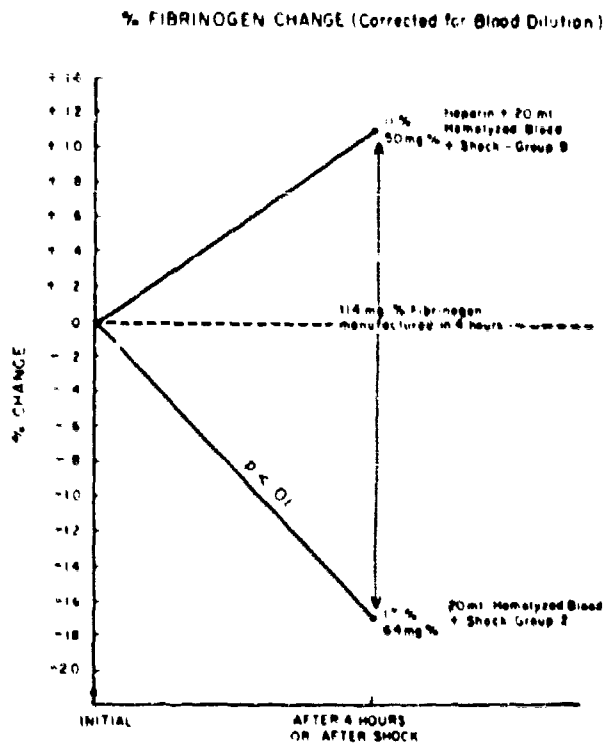


Fig. 4. Fibrinogen levels of some of the groups. Dogs subjected to hemorrhagic shock after administration of 20 ml. of hemolyzed blood (Gp. 2) show a moderate and significant decrease in fibrinogen at the end of the shock period. If heparin is administered to dogs subjected to this same procedure (Gp. 9) there is an actual increase in fibrinogen over the period of shock. If one assumes that the heparin prevents in part the conversion of fibrinogen to fibrin, then by subtraction one can say that approximately 114 mg% of fibrinogen were converted to fibrin in the nonheparinized group. The increase in fibrinogen in the heparinized group is accounted for by the manufacture or mobilization of new fibrinogen.

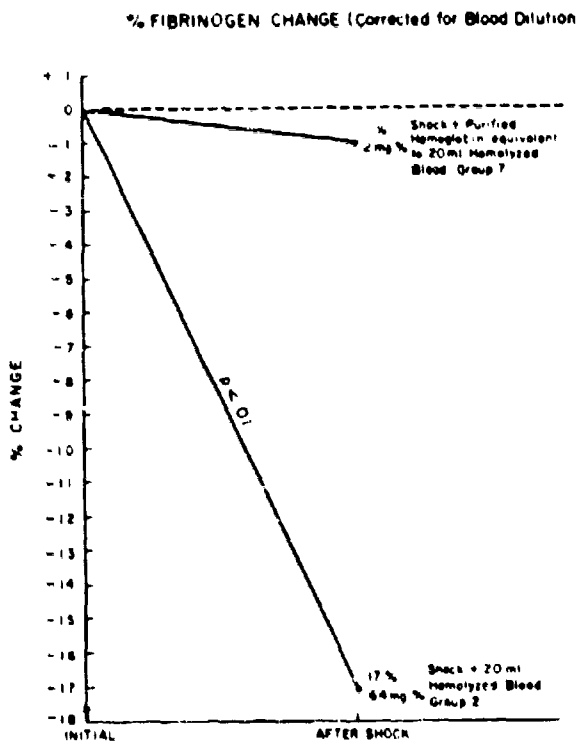


Fig. 5. Fibrinogen levels of some of the groups. Hemorrhagic shock after the administration of 20 ml. of hemolyzed (Gp. 2) blood causes a moderate but significant fall in fibrinogen. However, if instead of whole hemolyzed blood, an equivalent quantity of hemoglobin purified from autogenous blood is administered there is only a slight and nonsignificant decrease in fibrinogen as a result of hemorrhagic shock.

ENDOGENOUS HEPARIN

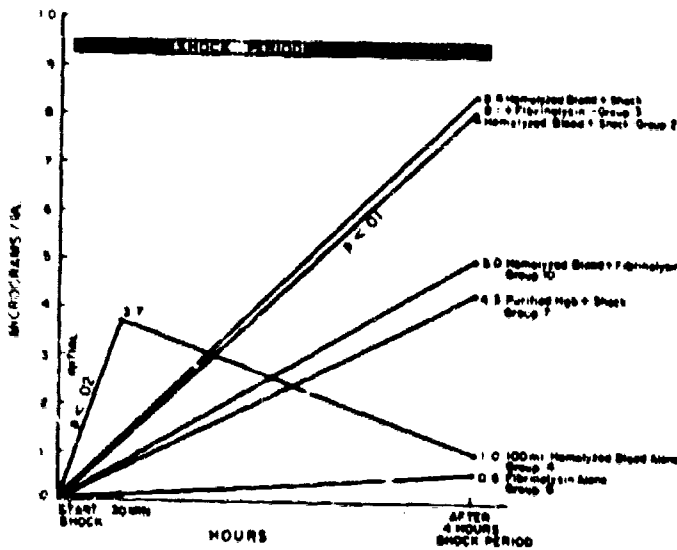


Fig. 6. Endogenous heparin. Endogenous heparin appeared in all groups subjected to hemorrhagic shock. There was a significant level of endogenous heparin produced particularly in those dogs subjected to hemorrhagic shock after administration of 20 cc. of hemolyzed blood (Gp. 2). Dogs given fibrinolysin alone (Gp. 8) showed no endogenous heparin. The administration of 100 ml. of hemolyzed blood alone (Gp. 4) without shock caused the immediate development of significant quantities of endogenous heparin at

30 minutes but this had disappeared by the end of four hours. This parallels the changes in fibrinogen and silicone clotting time in this same group.

PLATELETS

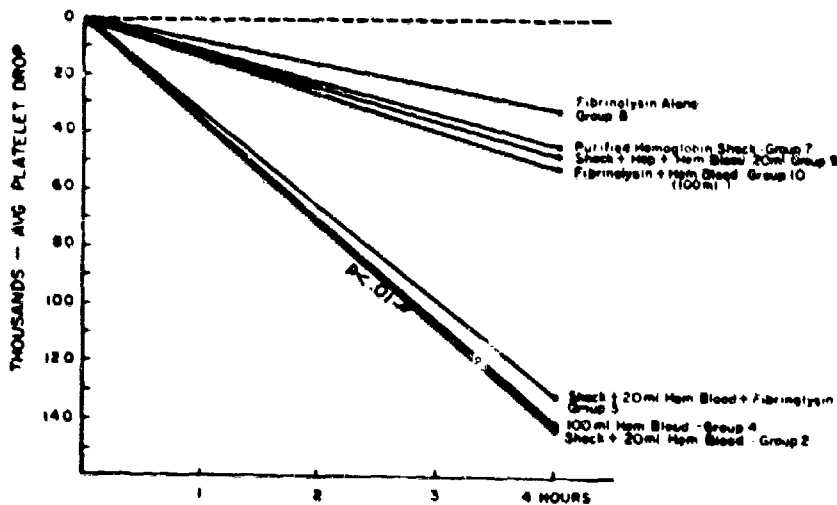


Fig. 7. Platelet changes. All groups showed some decrease in platelets. The group subjected to fibrinolysin alone (Gp. 8) showed only a slight and insignificant fall. However, the group subjected to irreversible hemorrhagic shock (Gp. 2) showed a marked

and significant decrease. All dogs subjected to hemorrhagic shock showed marked clumping of the platelets at the end of the shock period possibly reflecting an increased platelet stickiness.

FRICION HYDRO PNEUMATIC SUSPENSION SYSTEM

MAJOR WILLIAM J. HART JR, ORDNANCE CORPS
ROBERT J. OTTO and ALEX H. SINCLAIR
U. S. ARMY TANK-AUTOMOTIVE CENTER
WARREN, MICHIGAN

BACKGROUND

Mobility has been a key factor in many historic victories and defeats throughout the history of warfare. The defeat of France in 1940 can be attributed primarily to the rapid employment of German armored and mechanized forces that spearheaded the attack and achieved surprise. The excellent roadnets of Western Europe and favorable terrain enabled these units to move rapidly to their objectives and maintain an effective supply line. However, similarly equipped units were defeated decisively in Russia, where deep mud hindered tactical movement and adequate roads were either nonexistent or overloaded. Even the brilliant penetrations of General Patton's Third Army were dependent principally on the existence of good roadnets, and at the end of World War II, the mobility of armored and mechanized units still depended to a great extent on the availability and the condition of roads. During the post-war period, greater emphasis was placed on the achievement of a higher degree of ground mobility. This objective is particularly important when it is considered that there are vast land areas of the world where few roads exist and destruction resulting from the use of nuclear weapons may make the use of existing roads impossible.

Military requirements generated by the combat arms generally reflect the need for tactical vehicles that are capable of achieving good cross-country speeds over all types of terrain, as well as other mobility criteria such as swimming, deep-water fording, and air transportability. The U. S. Army Tank Automotive Center, through its research and engineering facilities, and in coordination with other research and development agencies, is working toward the solution of many technical problems which must be overcome in order to equip combat forces with vehicles which will provide a high degree of mobility under a wide variety of terrain conditions.

HART, OTTO, and SINCLAIR

One vehicle component that significantly limits the achievement of good cross-country speeds is the suspension system. Even today, with the many technological advances that have been made in the design of automotive components, the speed of a tank generally does not exceed 5-7 miles per hour over extremely rough terrain because of undesirable vehicle vibration characteristics at higher speeds.

Under the direction of the Components Research and Development Laboratories, U. S. Army Tank Automotive Center, an extensive engineering investigation of suspension systems was initiated in order to more clearly define the dynamic relationships between a vehicle and the terrain over which it must travel, and to establish parameters which can be applied directly to vehicle design toward achieving a significant improvement in vehicle cross-country mobility.

DESIGN PARAMETERS

Current research in the field of vehicle dynamics indicates that the ability of a vehicle to traverse rough terrain rapidly is dependent primarily on vehicle vibration characteristics which, in turn, are related to spring rate, damping, vertical wheel travel, and contour of the ground wave. There are other parameters such as angle of approach, angle of departure, weight distribution and track configuration which influence obstacle and soft-soil performance.

Effective operation of a tank is limited significantly by oscillation about the transverse axis (pitch axis) located directly below the center of gravity of the vehicle at the level of the wheel hub. Vehicle pitching, induced by terrain irregularities, imposes severe limitations on the ability of the crew to perform their duties effectively, and particularly limits the use of optical instruments for observation and target acquisition. Also, vehicle pitching interferes with efficient performance of other operations such as driving, and gun loading, and may result in personnel injuries and premature failure of vehicle components. Cross-country speeds, therefore, have been limited generally to 10 miles per hour or less in order to retain a reasonable degree of combat effectiveness. To remove or significantly diminish these limitations, it is necessary to obtain, through proper design and selection of components, better vehicle vibration characteristics where the pitch angle decreases as speed increases, and a level of vertical and angular acceleration is maintained as low as possible when traversing a given piece of terrain. This can be accomplished to some extent by decreasing the vehicle natural frequency, thereby decreasing the resonance speed over a given ground wave.

Because the natural frequency of pitch (f_0) is directly proportional to the spring rate (k) (Figure 1a), it is apparent that a reduction in the spring rate will result in a corresponding

reduction in the natural pitch frequency. Since a similar relationship exists between the natural frequency of bounce (f_y) and the spring rate (k) (Figure 1b), the natural frequency of bounce will be reduced also, if the spring rate is reduced. The selection of a soft spring has some limitations, however, and these limitations must be taken into account, or the advantages gained by lowering the natural frequency will be counteracted by the creation of conditions which are undesirable.

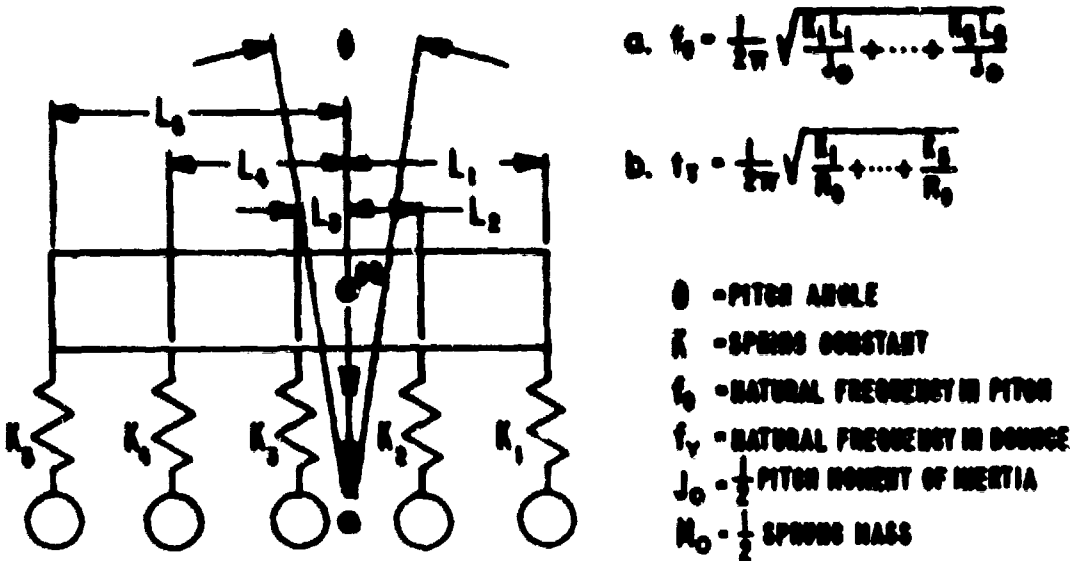


FIGURE 1 - SIMPLE VIBRATION SYSTEM

If a low spring rate is used without proper damping, a single impact, such as firing the main armament or passing a single obstacle, may cause prolonged oscillation. Although a soft spring will allow the wheels to move freely over minor surface irregularities without disturbing the vehicle hull significantly, it may not be capable of absorbing the kinetic energy transmitted by high-impact loads without bumpout. In order to achieve a significant measure of improvement in vibration characteristics, it is necessary to determine the optimum combination of springing and damping characteristics for maximum performance. Either a variable spring rate, variable damping, or both appear to be desirable in order to provide a suspension system sufficiently versatile so that its vibration characteristics can be optimized to suit a particular vehicle application and operational environment.

Analysis of vehicle ride dynamics indicated that a suspension system incorporating a low spring rate and a vertical wheel travel of approximately 16 inches would enable a 40-45 ton tank to traverse rough terrain at speeds of 24-27 miles per hour. Initial investigations were limited to the use of metallic springs which,

HART, OTTO and SINCLAIR

when designed to provide the necessary characteristics, required an excessive amount of critical hull space. The use of external shock absorbers created additional hull drag which would degrade vehicle performance in deep mud.

The aim of ATAC's current development program was to provide an integrated, multiple function, compact, light-weight suspension system. The resulting design characteristics included internal damping, a low spring rate, variable ground clearance, suspension lockout and a 14-18 inch wheel travel. In order to accomplish these objectives, non-metallic springs were included in the analysis. As a result, the hydropneumatic spring emerged as the media which offered the greatest potential within the present state of the art because of its inherent high-energy storage capacity and nonlinear spring rate.

Since the hydropneumatic spring adheres to the Gas Laws and closely approximates an adiabatic process ($PV^\gamma = \text{Constant}$), gas pressure increases exponentially as increasing wheel forces are transmitted to the spring:

$$(1) \quad P_2 = P_1 \left[\frac{V_1}{V_2} \right]^\gamma$$

Wheel deflection can be related directly to a change in gas volume by use of a linkage between the suspension arm and the spring; therefore, the spring rate (k) is proportional to the static pressure (P_1) and the change in gas volume (ΔV).

$$(2) \quad k \approx P_1 \left[\frac{V_1}{V_2} \right]^\gamma \quad \text{OR} \quad k \approx P_1 \left[\frac{V_1}{V_1 - \Delta V} \right]^\gamma$$

In order to transmit wheel forces to the spring and provide a means of damping, a three-vaned hydraulic rotary actuator offered several advantages. The torque capacity of a vane-type rotary actuator is proportional to the total vane area, the mean radius of the vanes, the number of vanes and the hydraulic pressure.

$$(3) \quad T = PAR_m N$$

A three-vaned actuator provides high torque capacity in a compact unit without developing excessive hydraulic pressures. Further increase in the number of vanes would not provide sufficient angular

deflection, and consequently would limit vertical wheel travel. The use of a hydraulic rotary actuator enables the spring characteristics to be varied by rotating the roadarms, thereby varying the vehicle ground clearance and the static pressure in the spring.

DEVELOPMENT

The development of a tank suspension system based upon the rotary hydropneumatic machinery was initiated in 1960. The design began with an 8100 pound static wheel load and dynamic loads of over 100,000 pounds. Four automatically controlled operating heights were to be provided, and a vertical wheel travel of 14-18 inches was required, using a 16-inch roadarm and 33-inch diameter roadwheels.

Allowing the necessary space for porting and for the physical size of the vanes, it was determined that 73° of rotation could be obtained from a three-vaned hydraulic rotary actuator without physical contact between movable and fixed vanes. An angular displacement of 73° resulted in a total vertical wheel travel of 18.4 inches (Figure 2). Four operating positions were selected within the 73° limit in order to provide some variation of wheel travel and vehicle ground clearance. Neglecting friction, and

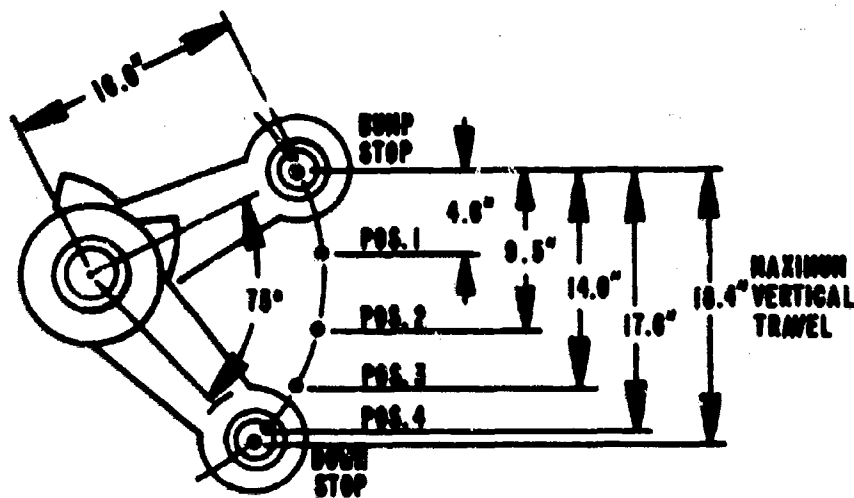


FIGURE 2 - VERTICAL WHEEL TRAVEL

using a rotary actuator 9 inches in diameter, calculations revealed that a torque of 64,000 lb.-in./psi. could be developed, and that a total of 81.5 cu. in. of oil would be displaced if the total angular

HART, OTTO and SINCLAIR

displacement of 73° were used. Assuming a 95% mechanical efficiency for the rotary actuator, the maximum torque required to lift the vehicle occurs in position number 1 (Figure 2) where the roadarm is almost horizontal and creates the longest lever arm. In this position 135,800 lb.-in. are required to lift the vehicle, resulting in a requirement for a 2110 psi hydraulic pressure. Position number 1 can be maintained with only 1915 psi, since friction helps to support the vehicle. A 3000 psi pressure-compensated hydraulic pump with a maximum flow of 28.5 gallons per minute at full flow was selected to supply the system. Since the maximum oil capacity of a single actuator is 81.5 cu. in., and the maximum lift pressure is 2110 psi, the pump is capable of raising the vehicle rapidly. To raise the vehicle from the upstop to position number 4 (69°) requires 77 cu. in. of oil per actuator. The necessary oil can be supplied by the pump in less than 7 seconds, with the pump running at 2700 rpm.

Design of the hydropneumatic spring was governed by several considerations:

1. The oil capacity required is 81.5 cu. in.
2. The initial gas precharge should be a minimum.
3. Maximum dynamic pressure should be limited to about 6000 psi.
4. The free piston type accumulator should, if possible, fit inside the arm shaft in order that the complete suspension unit might be constructed as a package.
5. Adiabatic gas compression would be assumed.

Calculations were made for all wheel positions using the following relationships and using a maximum gas volume of 132 cu. in. and assumed precharges (Figure 3):

$$(4) \quad P_3 = P_2 \left[\frac{P_0 V_0}{P_0 V_0 - P_2 \Delta V_0} \right]^{1.4} \quad \text{WHICH IS DERIVED FROM THE RELATION: } P_3 = P_2 \left[\frac{V_2}{V_3} \right]^{1.4}$$

By calculation, computer simulation and laboratory test it was determined that a gas precharge of 1350 psi would be suitable, since it resulted in a maximum pressure of 6620 psi, which is close to the specifications established. Lower peak pressures can be obtained but higher precharges would be required. Using the data already

HART, OTTO and SINCLAIR

established, the spring rate curves for each operating position were calculated and plotted (Figure 4).

Damping in rebound is achieved by passing oil through a fixed orifice damping valve. The valve chosen was selected primarily from experience gained in other hydraulic applications, and its adequacy is to be determined through tests of the complete system. Calculated spring rates in the vicinity of the static position were on the order of 350-900 lb/in, depending upon the position of the roadarm, and calculated vehicle natural frequencies were on the order of 39-62 cycles per minute.

The components described above were fabricated, subjected to laboratory tests and installed in a test vehicle. Early prototypes used external accumulators, but redesign in early stages of development resulted in the configuration shown in Figure 5. The need for a stable firing platform for the main armament and a means of preventing immobilization of the vehicle due to hydraulic failure precipitated the development of a friction brake system. This system provides an additional capacity for absorbing great dynamic loads imposed on front and rear wheels when traversing rough terrain at high speeds.

The first redesign resulted in the addition of friction disc brake units to front and rear suspension units as shown in Figure 6. During the final 7° of upward travel of the roadarm, a cam on the arm shaft causes the brake piston to displace approximately 3 cu. cm. of oil into an annular cavity behind a ring-type piston, thereby raising the oil pressure to approximately 1400 psi and engaging the friction discs. The brake unit is capable of absorbing 700,000 lb-in of torque, thereby reducing the kinetic energy transmitted to the hull. Approximately one half of the bumpout energy is absorbed in the brake unit, and the remainder is divided between the hydro-pneumatic spring and the solid metal bump stop. An important function of the brake unit is to provide a means of locking the suspension in a selected position to provide a rigid firing platform for the vehicle weapons system. Actuation in this mode is accomplished by applying oil pressure directly to the brake piston by means of an external control valve. Pressure is maintained by an accumulator during periods in which the hydraulic pump is inoperative.

Oil entering or leaving the hydropneumatic spring passes through a fixed orifice damping valve which is designed to provide damping in rebound only. Wheel movement induced by terrain roughness results in the transfer of oil from the actuator to the hydro-pneumatic spring (through the damping valve in which the poppet is unseated) allowing relatively unrestricted flow. In rebound, the poppet is seated and oil returning to the actuator is forced through a series of small orifices to provide the necessary damping.

EVALUATION AND TESTING

Computer simulations conducted at the U. S. Army Tank Automotive Center enabled a preliminary technical evaluation to be made prior to actual on-vehicle tests. Using an analog computer and terrain traces of the Perryman Cross-country Course at Aberdeen Proving Ground, a constant-speed evaluation was made of two equally-weighted vehicles. One vehicle was equipped with the Friction Hydro Pneumatic Suspension and the other with the current production torsion bar suspension. The results confirm the fact that a significant reduction in pitch and bounce vibration can be expected at speeds of 25 miles per hour using the Friction Hydro Pneumatic Suspension. A sample of the results is shown in Figure 7. Since it was necessary to make numerous assumptions in order to program the computer, it was considered essential to determine the suspension characteristics by on-vehicle testing.

In order that test conditions could be repeated, a test course was constructed using a prepared roadway and rigid obstacles which varied in height from 6 inches to 12 inches. Obstacles were spaced 20 feet apart, since this was considered to be the worst condition for the vehicles being evaluated. Two tanks sprung with torsion bars and one tank sprung with the Friction Hydro Pneumatic Suspension were used. The objective of the evaluation was to determine the maximum speed that each vehicle could achieve over a series of 10 obstacles of a given height, spaced 20 feet apart. Instrumentation was installed in order to record vertical and pitch accelerations, pitch angles and vehicle speed. Only experienced drivers were used and the limit of each vehicle was determined by the ability of the driver to maintain control of the vehicle. The three test vehicles were run over courses consisting of 6-inch, 8-inch and 12-inch obstacles in speed increments of 5 miles per hour.

In order to provide some criteria against which the vibration characteristics could be measured, it was necessary to establish a level of pitch amplitude and vertical acceleration which would be considered acceptable for effective operation of the vehicle. A review of the research in the field of vehicle dynamics conducted by Lehr in Germany⁽¹⁾ revealed that vertical acceleration in excess of 0.4g (12.9 ft/sec²) would interfere with efficient use of optical instruments of the tank's fire control system and that pitch amplitude should be reduced to approximately ± 20 using the proper arrangement of springing and damping. The criteria of 0.4g (12.9 ft/sec²) falls within the "discomfort zone" (10-14 ft/sec²).⁽²⁾

The criteria established for this suspension evaluation was established as 10 ft/sec² or less for effective tank operation since it coincides very closely to the criteria established by Lehr and is also the lower limit of discomfort. A total pitch angle of 40 was arbitrarily selected as an acceptable limit, however, it is anticipated that this criteria will be more accurately determined by

a more detailed analysis of the human factors involved, as well as the physical capabilities of the optical instruments used in the tank.

Although evaluation of the Friction Hydro Pneumatic Suspension System has not been completed, the results achieved to date are significant. Maximum speed of the control vehicles sprung by standard production torsion bars was found to be limited to approximately 11 miles per hour over a course consisting of 10 six-inch obstacles, spaced 20 feet apart. At this speed pitching was so violent that drivers were unable to maintain control of the vehicles and experienced difficulty in remaining seated even with the aid of seat belts. The vehicle sprung with the Friction Hydro Pneumatic Suspension System negotiated the course at the maximum speed that could be achieved by its power plant, 22 miles per hour, without causing the driver to experience difficulty in steering or control. Testing was repeated using 8-inch and 12-inch obstacles. Although maximum speeds were reduced as the obstacle size was increased, the differential in speed capability remained approximately the same (Figure 8), even though the gross power-to-weight ratio of one of the control vehicles was 50% greater than the Friction Hydro Pneumatic test vehicle. The second control vehicle had an equal power-to-weight ratio. An advantage in power-to-weight ratio did not seem to affect the results significantly.

Analysis of the data recorded revealed that over the six-inch obstacles, pitch angles remained at a low level ($\leq 3^\circ$) on the Hydro Pneumatic test vehicle throughout the speed range (0-22 miles per hour). Pitch acceleration also remained at a low level (≤ 5 rad/sec²). The best performance achieved by the control vehicles resulted in pitch angles and pitch accelerations so high (16.6° and 11 rad/sec²) that the driver was unable to pass the resonance speed (Figure 9). Vertical accelerations recorded show the same general characteristic (Figure 10). If the criteria for effective operation of the vehicle is established at 12.9 ft/sec², as suggested by Lehr, the Friction Hydro Pneumatic test vehicle remains effective throughout the speed range (0-22 mph) while the control vehicles become increasingly ineffective at approximately 10 miles per hour.

It is difficult to measure a vehicle's total effectiveness, however, because of the many human and mechanical variables involved. In order to present an objective evaluation, no definite cutoff is shown in Figure 10 and only a gradual decrease in effectiveness is shown for accelerations above 10 ft/sec². Vibrations increased in magnitude as the size of the obstacles was increased, but in all cases the Friction Hydro Pneumatic test vehicle was capable of negotiating the course at higher speeds than the control vehicles.

The evaluation is continuing in order to determine the performance limits when wave length and amplitude of the ground wave

HART, OTTO and SINCLAIR

are not constant. Data collected during this evaluation will be applied to the development of other vehicle suspension systems where high cross-country speed is a required characteristic.

CONCLUSIONS

1. It is possible to design a suspension system which provides a 41-ton tank with approximately 100% increase in cross-country speed without increasing the power-to-weight ratio of the vehicle.

2. A suspension system incorporating a low spring rate and adequate damping will provide a significant reduction in pitch and bounce vibrations, thereby increasing vehicle cross-country speed capability without degrading the effectiveness of the crew or the vehicle weapons system.

3. It is feasible to design cross-country vehicles with a speed capability of 30-35 miles per hour within the present state of the art.

4. A major advancement in mobility machinery has been achieved by means of a theoretical analysis of suspension design characteristics and performance parameters.

An extensive amount of research and engineering work must still be accomplished in order to develop suspension systems capable of providing efficient operation at high speeds over all types of terrain, particularly in the fields of springing and damping media, suspension geometry, terrain characteristics and terrain sensing devices.

The promising results achieved to date have already had a significant impact on the design of suspension systems for military vehicles. Perhaps even more important is the impetus given to the initiation of companion research and development programs which will contribute toward improving the effectiveness of future military vehicles under a wide variety of terrain conditions.

REFERENCES

- (1) Lehr, E. "The Springing and Damping of the Suspension of Armored Vehicles", Berlin, January 28, 1944 (Translation of presentation lecture to the Ministry for Armaments in War Production).
- (2) Ordnance Proof Manual, Volume II, Aberdeen Proving Ground, Maryland, OPM 60-305, "Human Engineering", 20 November 1957, Page 5.

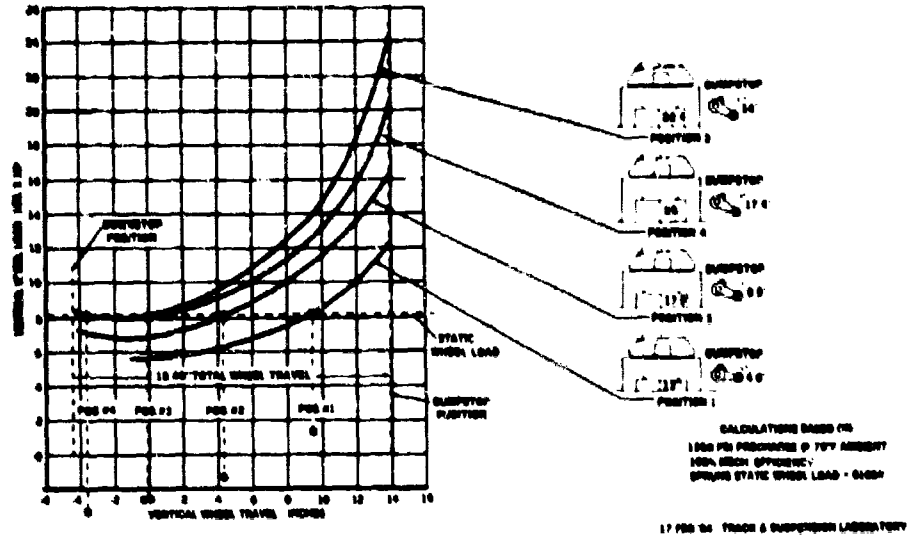


FIGURE 3 CALCULATED SPRING CHARACTERISTICS

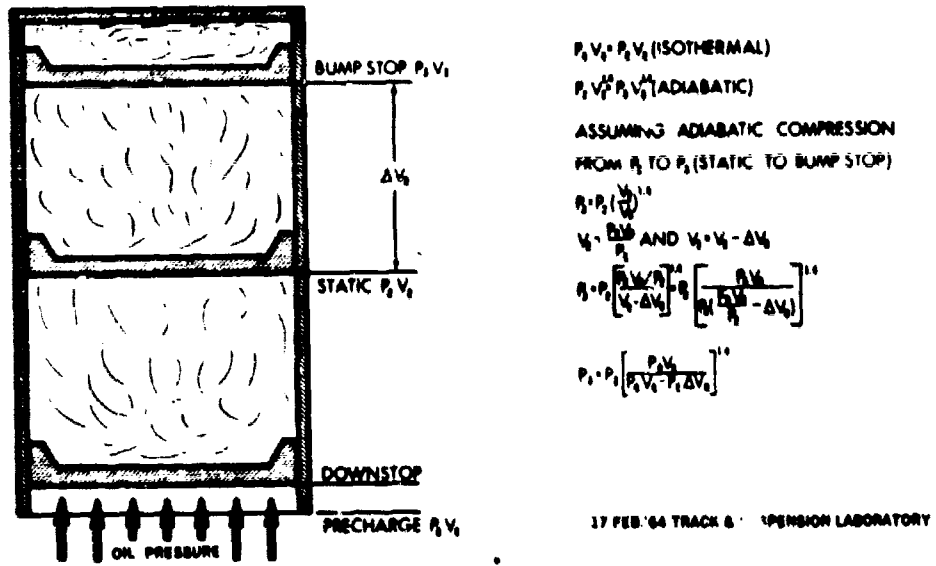


FIGURE 4 SPRING CALCULATION:

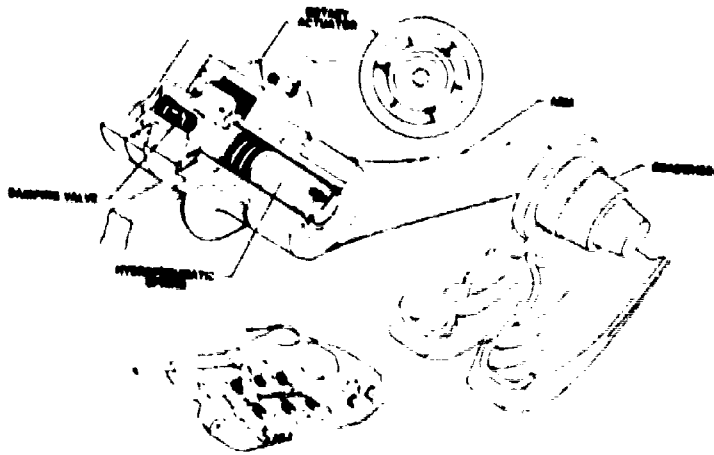


FIGURE 5 INTERMEDIATE UNIT

17 FEB 64 TRACK & SUSPENSION LABORATORY

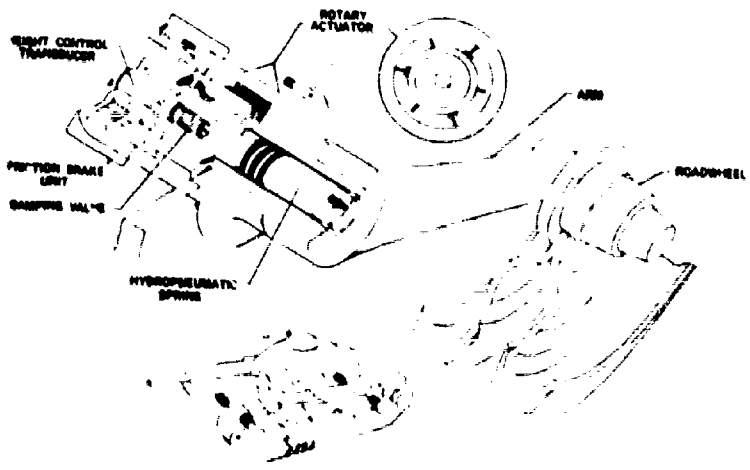


FIGURE 6 FRONT & REAR UNIT

17 FEB 64 TRACK & SUSPENSION LABORATORY

Computer Simulation

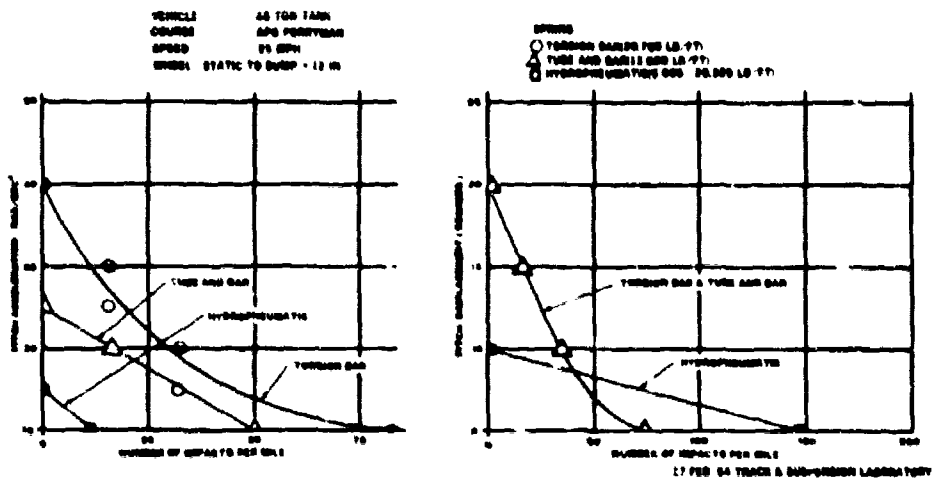


FIGURE 7 WHEEL TRAVEL RESEARCH STUDY

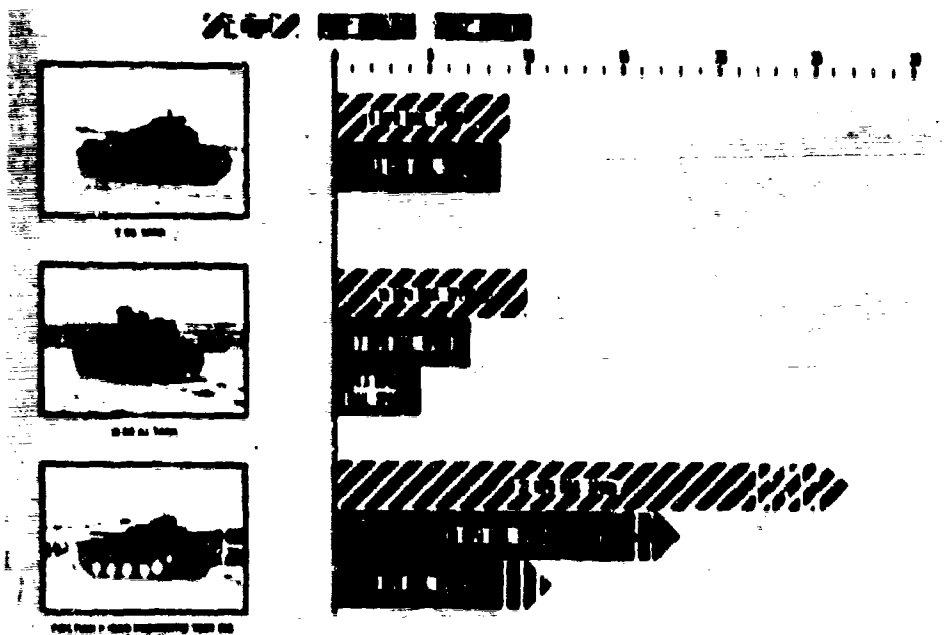


FIGURE 8 MAXIMUM SAFE SPEED OVER SERIES OF TEN OBSTACLES SPACED 20 FEET APART

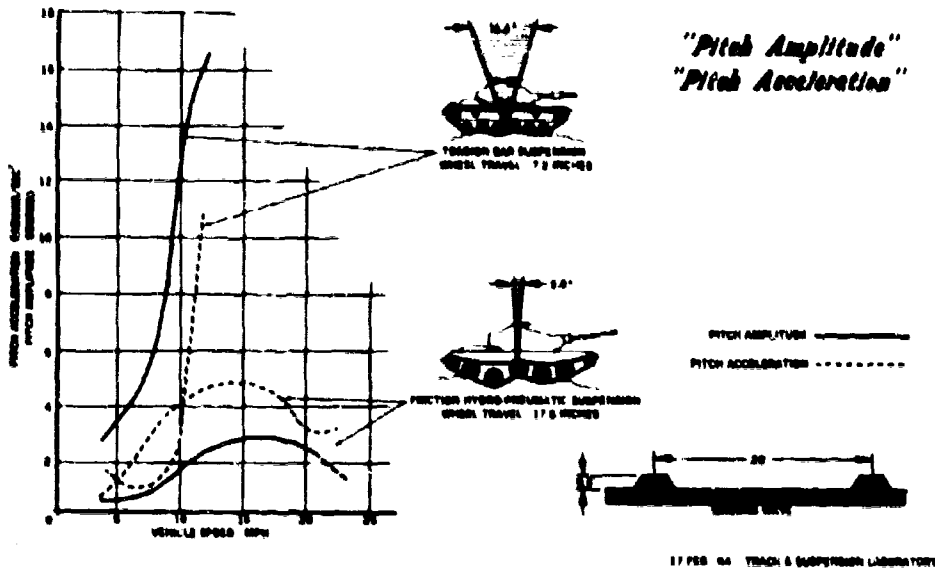


FIGURE 9 VIBRATION CHARACTERISTICS

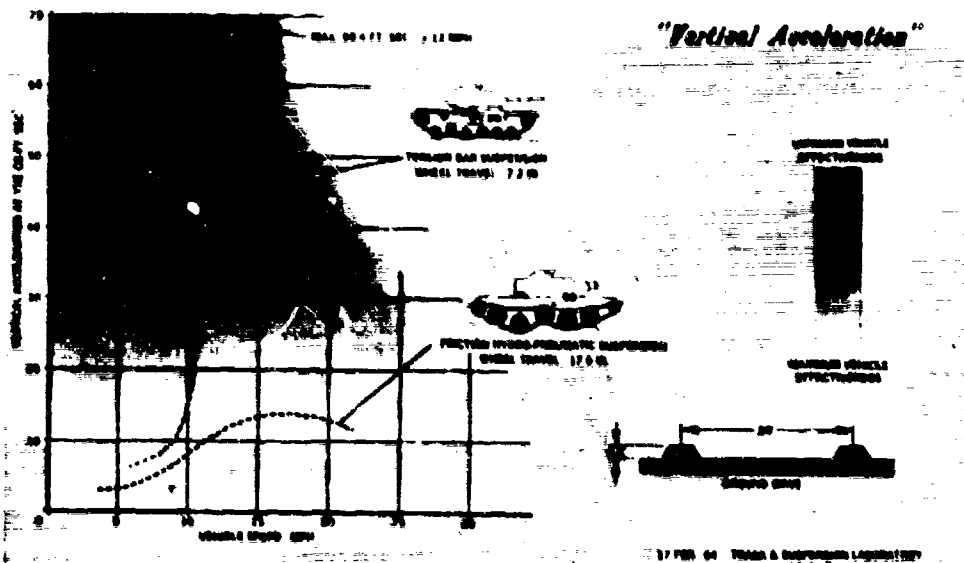


FIGURE 10 VIBRATION CHARACTERISTICS

THE INFLUENCE OF STORAGE, AEROSOLIZATION, AND REHYDRATION
ON THE PERMEABILITY OF PASTEURELLA TULARENSIS

DORA K. HAYES and VICTOR J. CABELLI
DUGWAY PROVING GROUND
DUGWAY, UTAH

During the development of a collecting fluid for aerosolized *Pasteurella tularensis*, Cabelli (1962) observed that the number of culturable cells recovered was greatly increased by the presence of the carbohydrates, melezitose, trehalose, and sucrose in the collecting solution. The polyhydric alcohol, inositol, also improved these recoveries. Table 1 summarizes these findings. All recoveries are related to that obtained in 0.075M sucrose solution which also contained the additive and which has been arbitrarily set at 1.00. Since it is advantageous to hold collected samples for an eighteen to twenty-four hour period before assay, recoveries both before and after holding are shown.

TABLE 1

Solution Composition	Initial Recovery Relative to 0.075M Sucrose-0.075M Inositol	24-Hour Recovery Relative to 24-Hour Recovery in Sucrose- Inositol
Inositol (0.075M)		
Melezitose (0.075M)	2.93	1.48
Inositol (0.075M)		
Trehalose (0.075M)	1.65	1.52

Cabelli (1962) also observed that the presence of orthophosphate in the collecting solution improved recoveries. He suggested that the phosphate present outside the cell provided a concentration gradient more favorable to retention of this ion by the stressed cell. The carbohydrate additives probably prevent diffusion of other essential metabolites outward after desiccation and collection. These additives may also inhibit the entry of deleterious molecules by means of solvent drag during rehydration.

The complete collecting fluid developed by Cabelli (1962) contained phosphate, sucrose, inositol, l-cysteine.HCl, thiourea and spermidine phosphate. The incorporation of cysteine is dictated by

the relatively high requirement for cysteine in growth media of this organism. The findings of Mager (1955) suggest that the polyamines, such as spermine and spermidine, stabilize osmotically fragile organisms. Melesitose and trehalose were not incorporated in the collecting fluid, since neither sugar is readily available in large quantities at a reasonable price.

The analytical method for phosphate ion is relatively simply and rapid. By examining the behavior of this ion under a variety of situations on which the bacterial cell was stressed, we hoped to increase our understanding on the processes which occur within the bacterial cell and at the interfaces between cell and environment during collection. If the reasons for phosphate requirement were elucidated, a sound basis could be provided for preventing deleterious effects of collection and of holding.

A number of workers have observed that phosphate metabolism is altered by stress and during aging. Throughout the aging process, phosphatase activity in animal cells may show an increase. Often there is uncoupling of oxidation from phosphorylation (Bourne, 1962). Acidosis caused by carbon dioxide intoxication can result in altered body fluid phosphate levels in invertebrates and mammals (Brown, 1957; Hayes, 1962). Phosphate metabolism in bacterial cells may be affected by stresses in one or more of the following ways:

- (1) The ability of the cell to take up phosphate may be decreased because enzymes on or near the surface of the cell which are necessary for phosphate transport have been destroyed.
- (2) A specific phosphate transport site may be blocked, deformed or destroyed.
- (3) Cell-environment interphases may be "stretched" to permit leaching of phosphate ion.
- (4) Available intracellular space may be increased by injury.
- (5) Esterases which hydrolyze organic phosphates may be activated. An increase in intracellular phosphate concentration will enhance passive diffusion outward, especially under conditions which are unfavorable for phosphate retention.

Materials and Methods: In all of the experiments, the organism studied was *P. tularensis*. To study movements of materials in or out of cells, suspensions of the bacteria were washed in sucrose (5%)-inositol (2.7%) solution and then resuspended in the test solutions. Dynamic aerosols were produced using a University of Chicago Technical Laboratory atomizer. Collections were made for ten minutes, using the all glass impinger with a fill of 18.5 ml and an air flow rate of 6 liters per minute.

Collected cells were concentrated as shown in Figure 1. Results are expressed either as micrograms of phosphate per equivalent volume of impinger fluid or as micrograms of phosphate per μg . of bacterial nitrogen.

Trichloroacetic acid soluble phosphorus was determined by the method of Fiske and Subarow (1925) and total phosphorus was

determined by assay of sulfuric acid digests of material containing organic phosphates. Nucleic acids were extracted and fractionated according to the Borenblum-Ghain-Healy procedure, using the technique of Schmidt-Thannhauser as described by Schmidt (1957). The number of washings with organic solvents were decreased, since a very small amount of cellular material was available. Ultraviolet absorptions of cell extracts were determined using a Beckman DK-2A recording spectrophotometer.

Results and Discussion: The efficacy of the wash solution used was determined by washing in 62 day old cell suspension with sucrose-inositol (5-2.7%) and resuspending the cells in solutions which contained inositol (2.7%) and either sucrose at various concentrations or all the components of the synthetic collection fluid of Cabelli (1962). The cells were stressed by holding aliquots at 4°, 27°, and 37° C for twenty-four hours. Table II summarizes the results of this study.

TABLE II
EFFECT OF COMPOSITION OF SUSPENDING FLUID AND TEMPERATURE ON SURVIVAL OF PASTEURELLA TULARENSIS

Solution Composition	Per Cent Recovery		
	4°C	27°C	37°C
Sucrose (5%)	95	58	0.42
(15%)	72	50	0.19
(22%)	56	36	0.16
Synthetic fluid	75	88	19.0

*All solutions contained inositol (2.7%).

The increased sucrose concentrations did not improve the recovery of organisms held at 37°C. However, it is evident that the complex synthetic fluid was far superior at room temperature or at 37°C.

The phosphate concentrations of the cells and of the supernatant medium are summarized in Table III.

TABLE III
EFFECT OF COMPOSITION OF SUSPENDING FLUID AND TEMPERATURE ON INORGANIC PHOSPHATE OF PASTEURELLA TULARENSIS

Solution Composition*	Intracellular inorganic phosphate-microg. P in cells from one ml of suspension			Extracellular inorganic phosphate-microg. P. per ml supernant		
	4°	27°	37°C	4°C	27°C	37°C
Sucrose (5%)	57	51	37	7	17	35
(15%)	44	55	18	9	24	59
(22%)	50	45	14	8	20	51
Synthetic fluid	55	55	53	950	900	900

*All solutions contained inositol (2.7%). Initial concentration of cells 1.1×10^{11} cells per ml.

At low temperatures, a solution consisting of sucrose and inositol in fairly low concentrations maintains the cells adequately. As the temperature is increased, the requirement for a solution which contains phosphate, as well as a number of other components, is required. It can be observed from Table III that the synthetic fluid is conducive both to maintenance of culturability and to retention of intracellular phosphate. It may also be observed from Table III that the intra- and extracellular phosphate concentration added together totals more at 37° C than at 4° C. This suggests that the cell organic phosphate is probably breaking down to inorganic phosphate, which is then liberated to the supernatant solution.

That the inorganic phosphate does break down to liberate trichloroacetic acid soluble phosphate is indicated in Table IV. In this experiment, both inorganic phosphate and total phosphate were determined. It will be observed that both the inorganic phosphate and the total phosphate in the cells decrease. The inorganic phosphate and total phosphate concentrations are almost identical. This indicates that under the conditions of this experiment, almost all of the phosphate lost by the cells is found as inorganic phosphate in the supernatant. The quantity of total phosphate lost is indicative of the action of one or more intra- or extracellular phosphatases.

TABLE IV
EFFECT OF TEMPERATURE ON CELLULAR PHOSPHATE IN
PASTEURELLA TULARENSIS

Temperature °C	Microg. inorganic P in cells from 1 ml of suspension	Microg. total P in cells from 1 ml of suspension	Microg. inorganic P per ml supernant*	Microg. total P per ml. supernant
4	60	415	12	12.3
27	55	387	17	17.1
37	25	205	31	32.0

* Initial concentration of cells per ml suspension - 8×10^{10}

Cells stressed and prepared as outlined in Figure 1 were examined in order to determine the effects of such treatment. Three collecting fluids, the synthetic fluid (SYN), sucrose-inositol (5, 2.7%) and gelatin-saline (0.1, 0.85%) were chosen. Recoveries in sucrose-inositol both immediately and after holding were close to those obtained in SYN. Gelatin-saline is a poor holding solution for *P. tularensis*. Typical recoveries are illustrated in Figure 2. It may be observed that after 24 hours of holding, culturable cells recovered from gelatin-saline collecting solution are considerably decreased. Cells which are not stressed, but are merely held at refrigerator temperature for 24 hours do not show decreased recoveries as illustrated in Figure 3.

HAYES and CABELLI

The phosphate content of cells collected at two relative humidities and at different ages is illustrated in the next series of tables. The results are expressed as micrograms of inorganic phosphate found in the cells from 35 ml. of collecting solution. In a spray suspension that was 19 days old, there is a significant difference in the phosphate content of cells collected and held in SYN as compared to those collected and held in gelatin-saline. This difference becomes more pronounced at 32 day suspension after holding 18 hours. Sucrose-inositol appears to afford some protection to cells collected from the 32 day old spray suspension, but this protection appears to disappear at 47 days. Since the absolute amounts of cells collected during each run was different, it became necessary to compare results between collecting fluids used on a certain day, but not to compare absolute amounts of phosphate among days.

Table V contains results obtained at 50 per cent relative humidity and Table VI contains similar data obtained at 80 per cent relative humidity.

TABLE V

INORGANIC PHOSPHATE CONTENT* OF AEROSOLIZED P. TULARENSIS** CONCENTRATED FROM 35 ML OF COLLECTION FLUID					
Cell Age days	Holding Time hours	Microg. inorganic phosphorus of cells collected in			
		SYN	Gel-saline	Sucrose- inositol	
19	1	3.7 ± 0.43	1.9 ± 0.28	3.8 ± 0.04	
	18	2.6 ± .25	1.0 ± .00	-	
32	1	3.1 ± .69	2.5 ± .48	2.2 ± .17	
	18	3.3 ± .63	0.7 ± .17	1.8 ± .07	
47	1	1.8 ± .64	1.6 ± .18	1.5 ± .20	
	18	3.0 ± .39	1.8 ± .09	1.9 ± .27	

*Expressed ± standard errors.

**Relative humidity - 50%, Temperature - 50°C, 10 min. collection.

TABLE VI

INORGANIC PHOSPHATE CONTENT* OF AEROSOLIZED P. TULARENSIS** CONCENTRATED FROM 35 ML OF COLLECTION FLUID					
Cell Age days	Holding Time hours	Microg. inorganic phosphorus of cells collected in			
		SYN	Gel-saline	Sucrose- inositol	
20	1	2.4 ± 0.52	2.9 ± 0.09	1.3 ± 0.23	
	18	3.9 ± .20	1.2 ± .08	2.4 ± .05	
34	1	3.6 ± .24	2.7 ± .36	2.0 ± .08	
	18	-	1.6 ± .17	3.4 ± .10	
48	1	2.2 ± .39	1.5 ± .09	1.1 ± .27	
	18	3.5 ± .14	1.9 ± .16	2.2 ± .06	

*Expressed ± standard error.

**Relative humidity - 50%, Temperature - 50°C, 10 min. collection.

HAYES and CABELLI

A problem raised by this series of experiments was the observation that although the number of viable cells decreased by at least a factor of 10 in the gelatin saline, as illustrated in Figure 2, the inorganic phosphate did not decrease in the same fashion. Our expectation had been that a non-reproducing cell placed in environments low in phosphate would lose inorganic phosphate. Possibly there was a source for the inorganic phosphate in the cell. Table VII indicates that there was a greater decrease in total phosphate than in inorganic phosphate. Two trials were conducted at 50 per cent relative humidity and one trial was conducted at 80 per cent relative humidity. There is a suggestion that at the more favorable relative humidity, less phosphate is lost. There is even an indication of increase of total phosphate in SYN fluid.

Several lines of inquiry are suggested by these data. The first is the determination of the compounds which are broken down during stress. The second is the identification of the responsible enzyme(s). This identification might dictate the incorporation of an inhibitor in the collecting fluid. The third is the determination of the type of process occurring during the period of time in which phosphate is being lost to the fluid.

TABLE VII
CELLULAR INORGANIC AND TOTAL PHOSPHATE IN
AEROSOLIZED PASTEURELLA TULARENSIS AFTER EIGHTEEN
HOURS HOLDING IN IMPINGER FLUID

Collecting Fluid	Per cent change inorganic phosphate	Per cent change total phosphate		
		Trail I	Trail II	Trail III
		50% RH	50% RH	80% RH
SYN	-17	-10	+ 4.3	+ 23
Gel-saline	-29	-63	-64	- 55
Sucrose-inositol	-15	-30	-	+ 2.5

The third process was examined in cells stressed by heating to 37° C. Figure 4 illustrates the type of data obtained in the two collecting fluids which did not contain phosphate. The shape of the curve suggests that if a kinetic analysis of the data were performed, at least two processes would be found to contribute to the total process of phosphate loss. At low temperatures, the metabolic component is probably not contributing significantly to the total process. At 37°C, however, probably a sum of metabolic processes result in a greater organic phosphate breakdown and loss from the cell.

Attempts were made to utilize phosphorus-32 in kinetic studies of this process. However, the studies that were performed were those in which attempts were made to exchange $P^{32}O_4$ in supernatant solutions with inert phosphate in the cell. Conditions under which these exchanges occur in significant amounts are difficult to obtain with small numbers of stressed cells. Preliminary results indicate that the space available to phosphate increases in the stressed cell but much work remains to be done in this area in order

to obtain maximum uptake. Phosphate which is taken up by these cells under resting conditions is readily lost upon washing with phosphate-free solutions. Binding or retention by the membrane is not a prime factor in retaining the radioisotope, if, indeed it has been taken up within the cell.

No work has been done toward identification of specific enzymes responsible for the hydrolysis of organic phosphates. Bacterial ribonuclease and one or more phosphomono- and diesterases are likely candidates.

Identification of the phosphate compounds which are lost by the stressed cell has begun. Cells were treated as outlined in Figure 1. Control cells were treated similarly except that aerosolization was omitted. Absorption in the ultraviolet of compounds extracted by the trichloroacetic acid was determined. Ultraviolet absorbing materials are found both in the extract from stressed and unstressed cells. The relative amounts of material absorbing at 2600 Å and at 2800 Å differs in the two groups of cells. A greater percentage of material absorbing at 2600 Å is found in the extract from unstressed cells. This can be interpreted to indicate that during stress, mono- and di-nucleotides which would normally be found in the trichloroacetic acid extract are lost.

Table VIII summarizes the results obtained by examination of total nucleic acid phosphate and ribose-nucleic acid (RNA) phosphate of stressed and unstressed cells. These assays were performed on the trichloroacetic acid residue after lipid phosphates had been removed.

TABLE VIII
EFFECT OF SUSPENDING FLUID ON NUCLEIC ACID PHOSPHORUS
IN AEROSOLIZED AND UNAEROSOLIZED CELLS* AFTER A
TWENTY-FOUR HOUR HOLDING PERIOD

Collecting Fluid	Per cent decrease total nucleic acid phosphorus		Per cent decrease RNA phosphorus	
	Aerosol	Control	Aerosol	Control
SYN	24	0	40	0
Gel-saline	67	36	69	42
Sucrose- inositol	31	11	50	8

*Cell suspension - age 18 days.

SYN fluid appears to protect the stressed cell against loss of total nucleic acid phosphate and RNA phosphate during the holding period. Greatest loss of these phosphates is apparent in gelatin saline solution. The sucrose inositol solution appears to be intermediate between the other two collecting fluids. Control samples held in gelatin saline lose a greater proportion of the nucleic acid fraction of organophosphorus compounds than do samples held in the other two fluids. Assay of the phosphates extracted with the lipid fraction of the cell did not reveal any striking differences among

HAYES and CABELLI

any of the collecting fluids.

It is postulated that the phosphate present in the complex collecting fluid, SYK, prevents breakdown of organic phosphates in stressed cells. Partial protection is also obtained with sucrose-inositol solutions. The ability to protect the cell from phosphate loss is not so apparent when the cell to be protected has been aerosolized. Figure 5 summarizes the theoretical equilibria existing between the phosphate compounds of the bacterial cell and its environment. Further experimentation may reveal that the postulated "stable fraction" is subject to hydrolysis under suitable conditions of stress.

Conclusions:

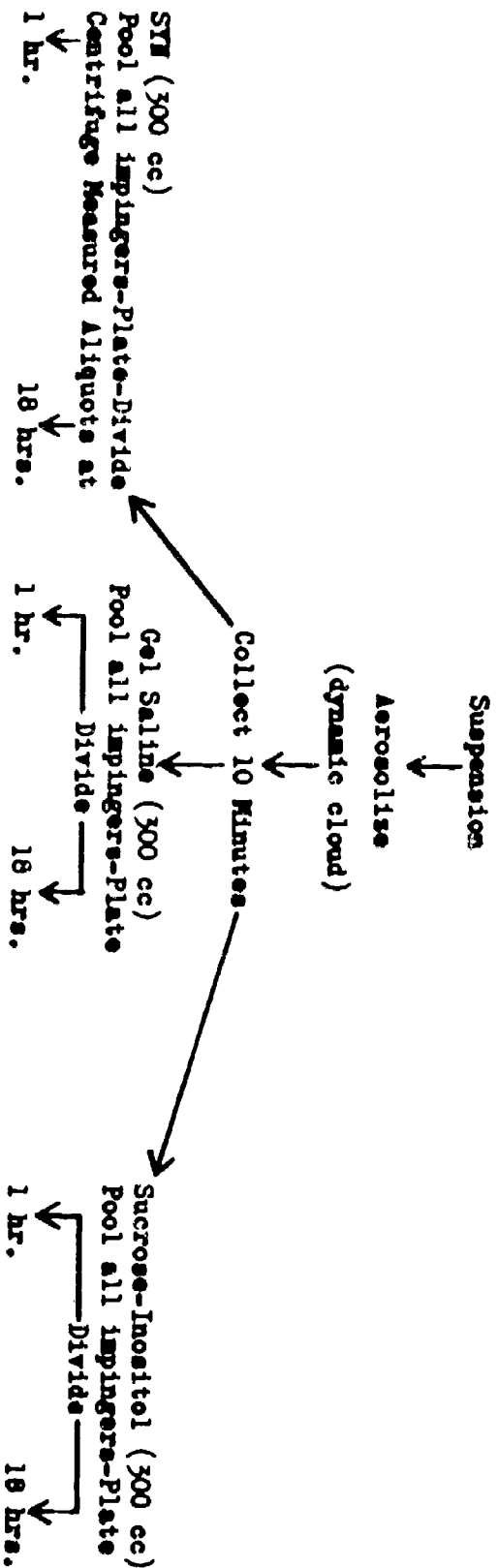
1. Aerosolized cells collected in phosphate-free solutions lose the ability to maintain phosphates against a concentration gradient. Factors which are age dependent are involved in this ability. Other stresses also result in a loss of intracellular phosphate, and the rate of loss appears to be temperature dependent.
2. Organic phosphate of the cell appears to be the source of inorganic phosphate lost by the cell. A fraction of cell protein isolated as the nucleoprotein fraction appears to be a source of this phosphate.
3. Nucleotides or breakdown products of nucleotides and proteins can be observed in supernatants and extraction solutions. The amounts of both nucleotides and proteins lost appear to increase as a function of temperature.

BIBLIOGRAPHY

1. Bourne, G.H., 1962. Cited in Chem. and Eng. News 40, Feb. 19, 108.
2. Brown, S.H. and Prasad, A.S., 1957. Am. J. Physiol. 190, 462.
3. Cabelli, V.J., 1962. Technical Report DPGR 314, Feb.
4. Fiske, C.H. and Subarrow, Y., 1925. J. Biol. Chem. 63, 461.
5. Hayes, D.K., Singer, L. and Armstrong, W.D., 1962. Proc. Soc. Exptl. Biol. Med. 109, 126.
6. Mager, J., 1955. Nature, 176, 933.
7. Schmidt, G., 1957. Methods in Enzymology, Vol. III, p. 671, eds. S. P. Colowick and N.O. Kaplan, Academic Press, New York, New York.

FLOW SHEET - PREPARATION OF CELLS FOR INORGANIC PHOSPHATE ANALYSIS

Figure 1



Centrifuge - measure aliquots - wash cell twice - sucrose-inositol determine inorganic phosphate of washed cells

Effect of Impinger Fluid Composition On
Recovery of Aerosolized *Pasteurella tularensis*
50% RH

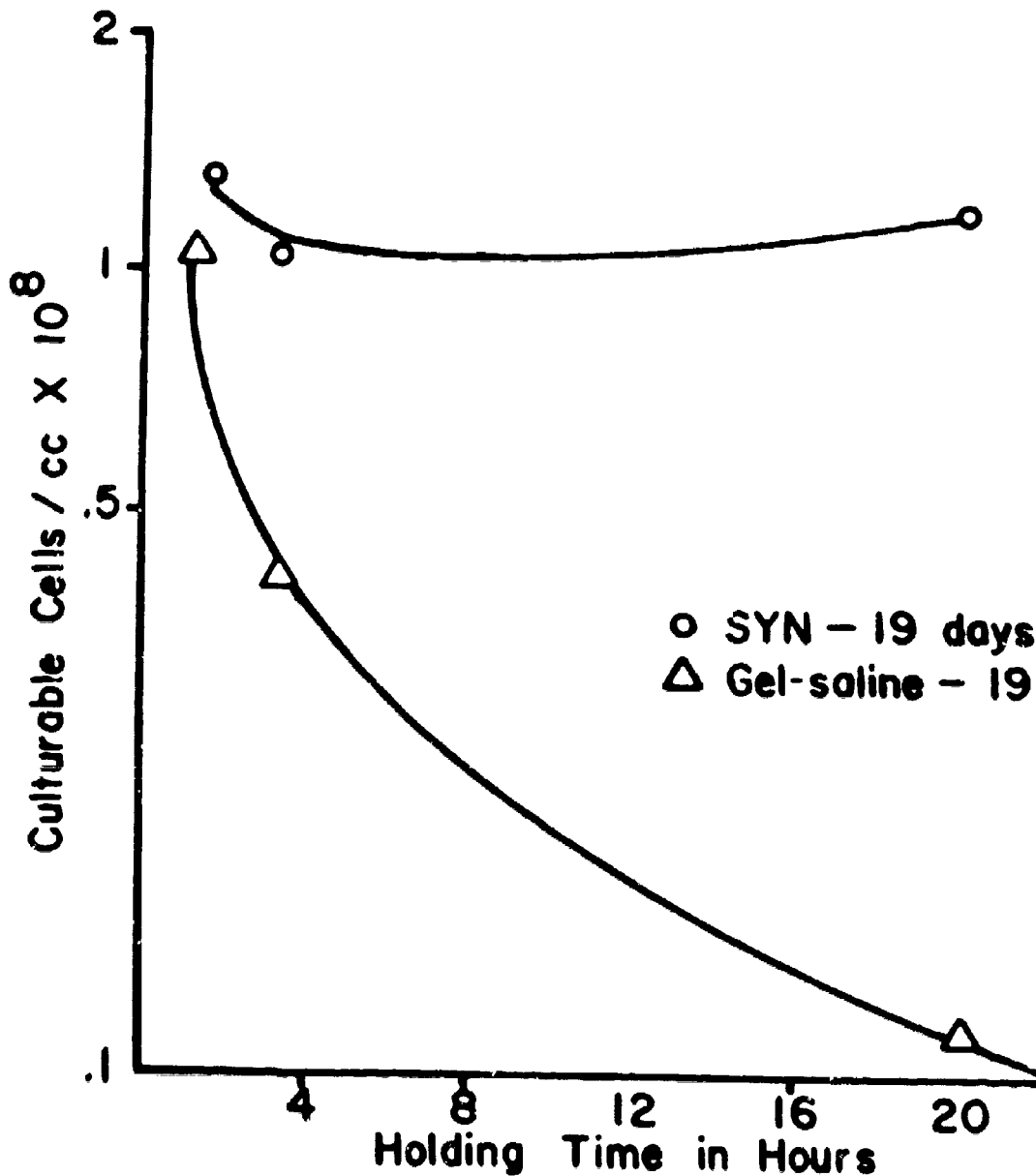


Figure 2

Effect of Composition of Suspending
Fluid on Viability of *P. tularensis*
Cell Age - 19 days

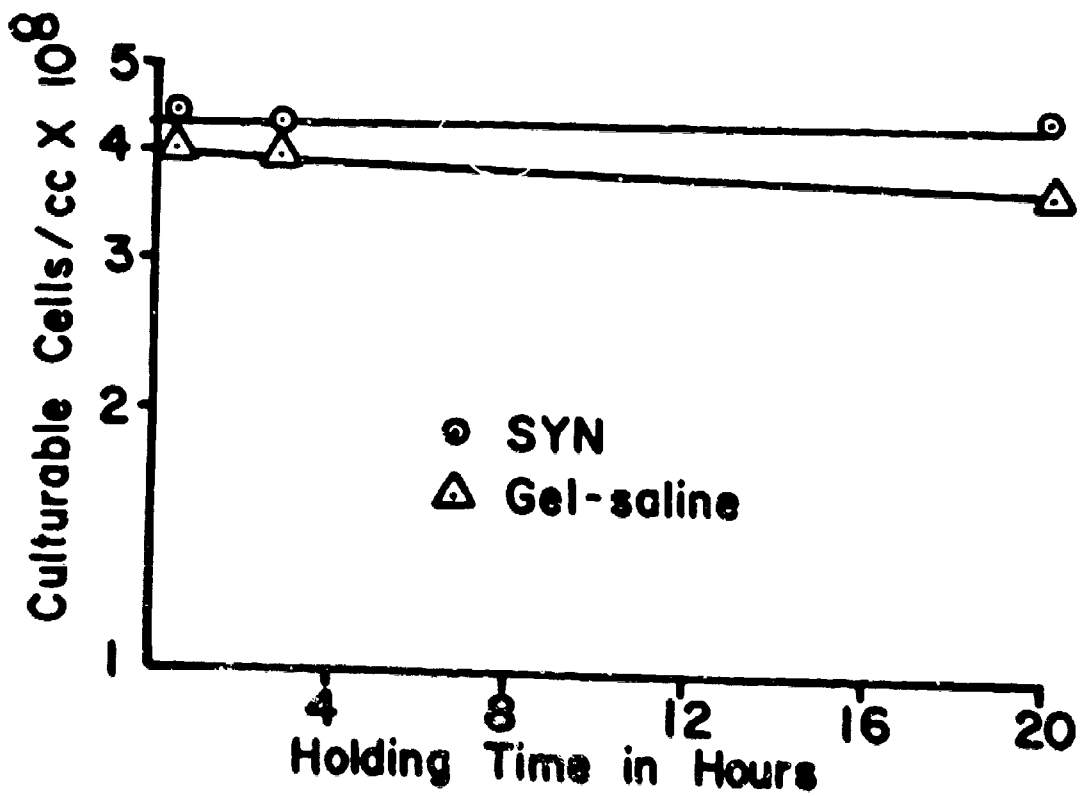


Figure 3

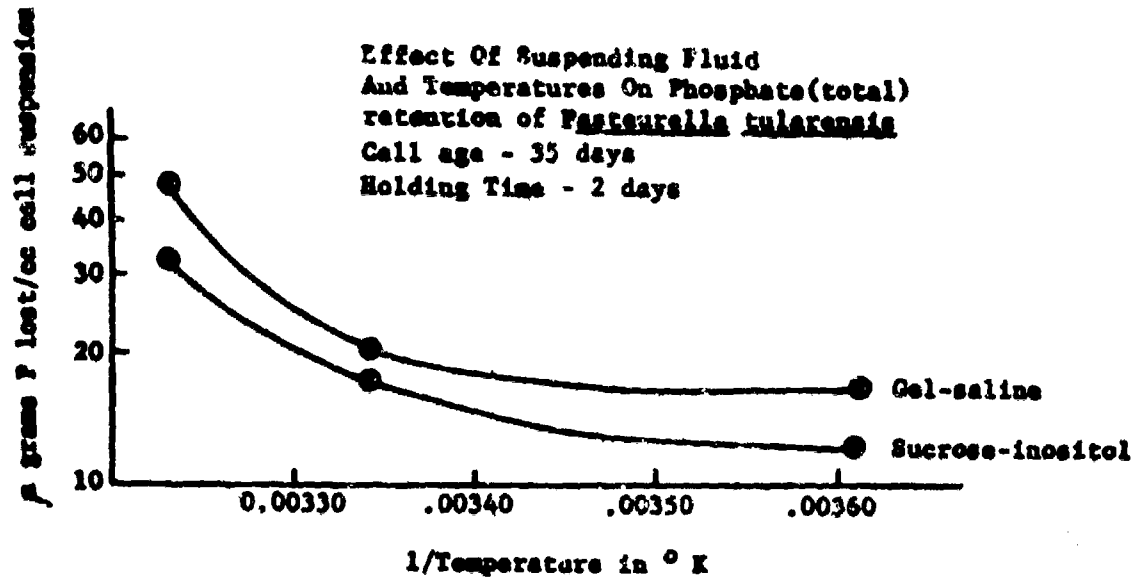


Figure 4

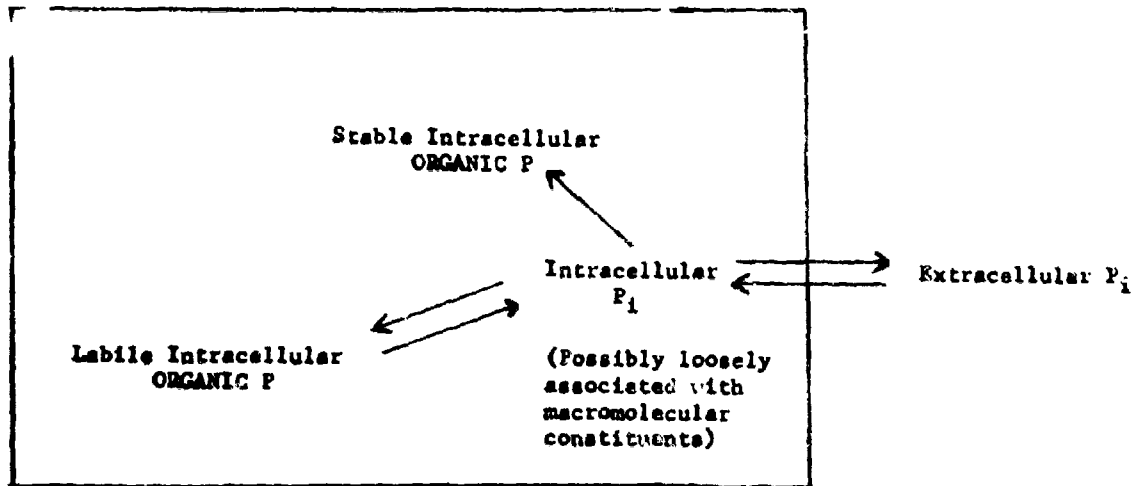


Figure 5

HENRY

PREPARATION AND USES OF SOME
NEW ORGANOLEAD CHEMICALS

MALCOLM C. HENRY(a)
U. S. ARMY NATICK LABORATORIES
NATICK, MASSACHUSETTS

We have been interested for some period of time in the properties of organolead compounds as they may be applied to micro-biological control, as biological toxicants, for rodent repellency, polymer stabilizers, lubricant additives, soil treating agents, herbicides and marine and anti-fouling paints. Although organolead compounds have been known for many years, only tetraethyllead and more recently tetramethyllead, have found commercial use. As an antiknock agent in internal combustion engines, tetraethyllead has an estimated output of 500 million pounds per year and is in fact the top commercial metal-organic. The demand, however, has not kept pace with increasing gasoline consumption in recent years owing to octane improvement techniques and to the increased use of compact cars which do not require premium grade gasoline. The effect of turbine engines in automobiles could have a disastrous effect on the consumption of tetraethyllead. Because of these problems in the lead industry, The International Lead Zinc Research Organization (ILZRO) was formed from a world-wide group of lead and zinc producers to find new applications for lead and zinc chemicals and to thus offset the possible decline in the use of tetraethyllead in automotive and aviation motor fuels.

In a cooperative venture involving mutual interests, (ILZRO) and the U. S. Army Natick Laboratories have been investigating new syntheses and new applications of organolead chemicals. New chemicals are prepared and characterized at the Natick Laboratories. They are then screened and tested for possible applications, both internally and by external elements. The results have been

(a) Substantial contribution by Dr. Kenneth Hills, Research Fellow at the U. S. Army, Natick Laboratories and sponsored by the International Lead Zinc Research Organization is herewith gratefully acknowledged.

HENRY

most encouraging. In cooperation with the Organic Chemical Institute, Utrecht, Holland, pilot plant quantities of numerous organolead chemicals are now being produced and distributed free of charge to interested industrial, government and academic institutions.

This paper will deal with the significant developments of three phases of this program, namely, syntheses, screening and testing, and the production of organolead chemicals on a large scale.

I. SYNTHESES

Chemical compounds containing metallic atoms bonded directly to carbon, traditionally called organometallic compounds, have widely different properties than those compounds containing only carbon-carbon bonds. Consideration of the properties of homologous compounds where carbon atoms are replaced by Group IV-B elements, Silicon, Germanium, Tin and Lead, has intrigued organic synthetic chemists for a number of years. Compounds containing lead-carbon bonds, known classically as organolead compounds are of particular interest because they represent the extreme case compared to carbon-carbon type molecules (1).

Initial synthetic efforts dealt with a study of some reactions of hexaphenyldilead, $(C_6H_5)_3PbPb(C_6H_5)_3$, (2) as well as the discovery of a new class of organolead chemicals called organolead arsonates (3), organic analogues of lead arsonate the common inorganic insecticide.

More recently, interest has been directed towards synthesizing organolead compounds which contain, in addition to lead-carbon bonds, at least one and sometimes two, lead-sulfur-carbon bonds. A number of routes were attempted in order to obtain these compounds (Slide #1). As a result of these experiments, it was discovered that organolead halides when reacted with mercaptans in the presence of an organic base such as triethylamine, gave an excellent general method for synthesizing this type of organolead compound (4).

(Slide #2) - describes some of the interesting types of compounds that have been made using this preparative technique. In contrast to the parent mercaptans, these compounds are odorless, white and sometimes yellow, low melting organic compounds, soluble in typical organic solvents and showing properties typical of organic molecules. They have low melting points and may be recrystallized from common organic solvents.

II. SCREENING AND TESTING PROGRAM

Preliminary screening of the newly synthesized organolead chemicals was carried out at several different laboratories. The results of this phase of the program have resulted in the discovery

HENRY

of several interesting properties that suggest a number of useful applications.

The Worcester Foundation for Experimental Biology in Shrewsbury, Massachusetts is a nonprofit educational and research institution dealing basically with physiological processes. These laboratories have been interested in organolead compounds as they affect or disturb otherwise normal physiological processes.

In screening organolead compounds it was found that thio-benzyl triphenyllead, $(C_6H_5)_3PbSCH_2C_6H_5$, and thiophenyl triphenyllead, $(C_6H_5)_3PbSC_6H_5$, showed excellent anti-inflammatory properties. Compounds with this type of activity are rather rare and are usually valuable in treatments of rheumatoid arthritis, allergies and asthma.

Another compound, thioacetyl triphenyllead $(C_6H_5)_3PbSC(O)CH_3$ showed sufficient anti-androgenic activity to make it a possible candidate for the treatment of acne, baldness or affecting beard growth inhibition in man.

The ramifications, screening methods and evaluation of these experiments are described in a new book edited by R. I. Dorfman (5).

(Slide #3) - illustrates the molecular configurations of these compounds and the possible uses assigned as a result of the screening experiments.

The Tropical Research Medical Laboratory in Puerto Rico, has also received samples of organolead compounds for screening. This laboratory, a member of the World Health Organization dealing with communicable diseases has been particularly concerned with the tropical disease bilharzia (schistosomiasis). This serious infection which attacks the urinary tract of man is carried by fresh water snails and is particularly prevalent in the Middle and Far East as well as in Africa.

Several organolead compounds have been effective in controlling these fresh water snails under controlled conditions. This work is continuing enthusiastically.

At the Organic Chemistry Institute, TNO, Utrecht, Holland, a number of organolead compounds were tested as possible fungicides, against a representative series of fungi.

(Slide #4) - shows how effective these compounds are as potential fungicides. With concentrations of as low as one part per million complete inhibition of botrytis allii was noted.

HENRY

The United States Department of Interior, Fish and Wildlife Service, have shown by their screening program that certain types of organolead compounds have excellent rodent repelling properties.

(Slide #5) - shows the kind of results which have been obtained for a variety of organolead compounds including the organolead-sulfur compounds. These results make these compounds potential candidates for impregnating agents of paper, fabric and other materials subject to attack by rodents.

As lubricant additives, organolead-sulfur compounds also show promise. Studies carried out at the Ethyl Corporation Laboratories in Detroit, Michigan, with a standard L_1 ball wear testing machine showed several organolead-sulfur compounds to be the best lubricant additives tested in those laboratories to-date. A patent application has been filed for the use of these compounds in this particular application.

Other suggested uses for these compounds have included applications in marine and anti-fouling paints, pesticides, stabilizers for chlorinated polymers, catalysts for polyurethane foams and radical producers in esterifications.

III. ORGANOLEAD CHEMICAL PRODUCTION

Under ILZRO sponsorship, the Organic Chemistry Institute, TNO, has instituted scaled-up syntheses of specific organolead compounds. These compounds are now available, free of charge to the world-wide scientific community as a further inducement to develop new uses for organolead chemicals (6).

Organolead-sulfur compounds have been included in this program because of their interesting properties and applications.

The future for organolead chemicals in general and organolead-sulfur compounds in particular seems assured as better synthetic routes become available. As adequate quantities of these chemicals become distributed, new uses and applications will assuredly follow.

The United States Army Laboratories, by effectively contributing to this pioneering area of chemistry make it possible to take advantage of the latest developments in research and to accomplish the military mission more effectively.

HENRY

REFERENCES

1. A. R. Anderson, "Abundance of Metal Organics", Chemical Week, Sept. 17, 1960.
2. Krebs, Adolph W. and Malcolm C. Henry, "Some Reactions of Hexaphenyldilead", J. Org. Chem. 28, 1911 (1963)
3. Malcolm C. Henry, "Synthesis of Organolead Arsonates and Arsinates", Inorganic Chemistry 1, 917 (1962)
4. Malcolm C. Henry and Adolf W. Krebs, "Synthesis of Organo-lead Sulfur Compounds", J. Org. Chem. 28, 225 (1961)
5. Ralph I. Dorfman, Experimental and Biological Test Methods, Academic Press, 1962
6. Chemical Week, "Organolead Compounds", Aug. 3, 1963

HENRY

SLIDE NUMBER ONE

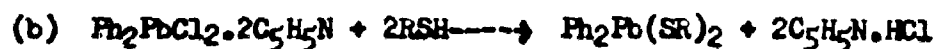
PREPARATION OF ORGANOLEAD SULFUR COMPOUNDS



1. Lead Mercaptide Method



2. Pyridine Method

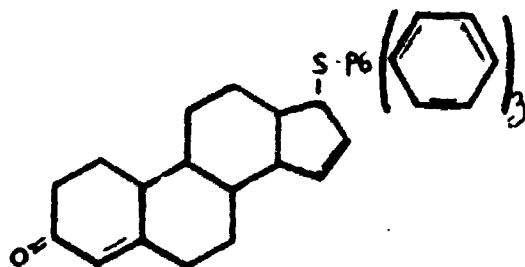
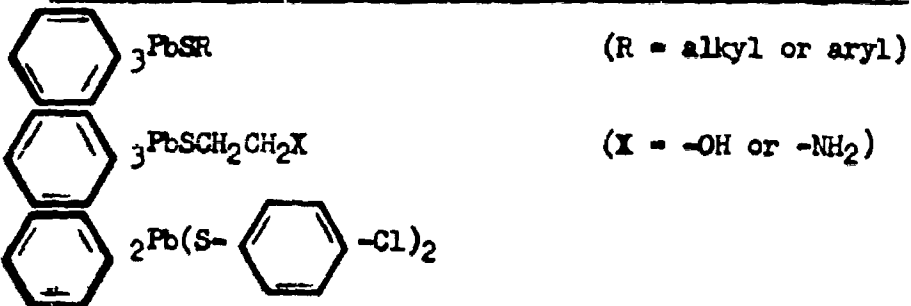


3. Triethylamine Method



SLIDE NUMBER TWO

SOME EXAMPLES OF ORGANOLEAD SULFUR COMPOUNDS PREPARED

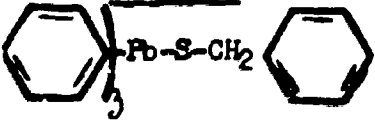
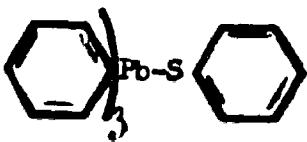
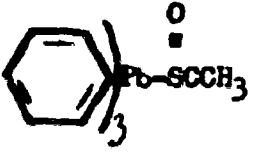


"17- . thiotestosterone" derivative

HENRY

SLIDE NUMBER THREE

PHYSIOLOGICAL AFFECTS OF ORGANOLEAD COMPOUNDS

<u>Compound</u>	<u>Potential Uses</u>
 thiobenzyl triphenyllead	<u>Anti-inflammatory agents</u> for rheumatoid arthritis allergies asthma
 thiophenyl triphenyllead	
 thioacetyl triphenyllead	<u>Anti-androgenic activity</u> for control of acne baldness or affecting beard growth in man

HENRY

SLIDE NUMBER FOUR

FUNGICIDAL PROPERTIES OF ORGANOLEAD COMPOUNDS

B = Botrytis allii; P = Pencillium niger; A = Aspergillus niger;
R = Rhizopus nigricans Activity indicated - minimal concentration
in parts per million causing complete inhibition of visible growth.



	$\frac{B}{100}$	$\frac{P}{100}$	$\frac{A}{100}$	$\frac{R}{100}$
Ph ₄ Pb	100	100	100	100
Ph ₃ PbPh ₃	500	500	500	500
Ph ₃ PbMe	200	500	500	500
Ph ₃ PbCl	0.2	0.5	0.5	1
(Ph ₃ Pb) ₂ S	100	100	50	100
Ph ₃ PbSMe	2	2	2	5
Ph ₃ PbSPr	1	2	2	2
Ph ₃ PbSCH ₂ Ph	1	5	2	2
Ph ₃ PbSOOPh	2	5	2	5

SLIDE NUMBER FIVE

RODENT REPELLENCY TEST

Laboratory: U. S. Fish & Wildlife Service, Denver, Colorado

Candidate Repellent	Repellency to House Mice No. repelled/No. tested
Lead benzylmercaptide	10/10
Triphenyllead chloride	10/10
Triphenylmethyllead	10/10
Thiomethyl triphenyllead	10/10
Thiophenyl triphenyllead	10/10
Diphenyllead dichloride	7/10
Bis-triphenyllead sulfide	7/10
Hexaphenyldilead	2/10

HIGGINS

X-BAND SEMICONDUCTOR SWITCHING AND
LIMITING USING WAVEGUIDE SERIES TREES

VINCENT J. HIGGINS

U. S. ARMY ELECTRONICS RESEARCH & DEVELOPMENT LABORATORIES
FORT MONMOUTH, NEW JERSEY

INTRODUCTION

The fact that the impedance level of a semiconductor diode can be varied by the application of a bias voltage has led to the use of semiconductor diodes in RF switching and limiting applications. The use of semiconductor diodes to control the level of 1-6, microwave power transmission has been detailed in the literature. 9-13, 15-18

Insertion of a diode in a waveguide or other transmission line results in attenuation and transmission of RF power incident upon the diode. Attenuation of the power incident upon the diode is achieved by reflection or absorption, or both. When RF power is transmitted past the diode with little loss, the ratio of power incident upon the diode to power transmitted past the diode is termed insertion loss. Similarly, when RF power is attenuated by the diode with little transmission, the same ratio is termed isolation.

The level of RF power transmission is controlled through the application of forward and reverse bias potentials to the diode terminals. For a particular diode, the bias requirements will depend upon the semiconductor material (i.e., silicon, germanium, or gallium arsenide), the frequency of operation, and the transmission line environment. For example, a semiconductor diode shunted simply across the center of a waveguide at 9000 Mc/s is usually biased in the forward direction to obtain a high impedance, allowing transmission past the diode with little loss, and biased in the reverse direction to obtain a low impedance, preventing transmission past the diode.^{3,4,6,13} At 1000 Mc/s this is not necessarily the case. With the same diode now shunted across a transmission line of lower characteristic impedance (coaxial or stripline), a forward bias results in a low impedance and a reverse bias results in a high impedance. The bias requirements for transmission and attenuation of incident RF power are now the opposite of the X-band case. This is primarily due to changes in the frequency dependent parasitic reactance

HIGGINS

attributed to diode lead inductance and package capacitance and to capacitive reactance changes in the transition region or barrier layer capacitance.

At frequencies in the X-band region (8.2 - 12.4 kMc/s), particularly 9300 Mc/s, the simplest form of the semiconductor switch consists of a diode shunted across a waveguide along the guide axis. 3,4,10,13 The operation of the diode in this simple form of switch can be explained qualitatively by referring to the assumed diode equivalent circuits of Fig. 1.

In the equivalent circuit the nonlinear capacitance of the diode, in this case a varactor, is attributed to transition region or barrier layer capacitance. This capacitance is predominant over any diffusion capacitance arising from minority carrier storage. 2,9,19 The barrier layer capacitance as a function of voltage is defined approximately as²

$$C_0(V) \approx \frac{C_0}{\left(1 - \frac{V}{\phi}\right)^n} \quad (1)$$

where C_0 is the zero bias carrier capacitance, ϕ is the contact or "built-in" voltage of the barrier and is a function of semiconductor doping with impurity atoms. For abrupt junctions, e.g., alloy junctions, point contact diodes, n is two; for graded junctions, e.g., diffused mesa types as in most varactors, n is three.

The application of a forward bias voltage greater than the contact or barrier potential ϕ will effectively short the barrier. For this bias condition, the diode equivalent circuit is an R-L circuit shunted by the package capacitance. In the absence of conductivity modulation, R is simply the spreading resistance R_s , and L_s is the lead inductance. The diode impedance is then:

$$Z_D = \frac{L_s/C_p - j R_s/\omega C_p}{R_s + j(\omega L_s - 1/\omega C_p)} \quad (2)$$

If the parameters L_s and C_p are such that antiresonance occurs, that is $\omega L_s = 1/\omega C_p$, the diode impedance Z is simply

$$Z_D = \frac{L_s}{R_s C_p} - j \frac{1}{\omega C_p} \quad (3)$$

For an antiresonant frequency of 9300 Mc/s, the ratio L_s/C_p is 30 kilohms squared and the diode impedance is very large for small values of R_s . Thus, transmission of incident microwave power is achieved with little loss, for if the diode impedance is much larger than the characteristic impedance of the standard X-band waveguide, power division between the matched waveguide load and the diode is small.

Consider now the semiconductor diode biased in the reverse

HIGGINS

direction. This negative bias results in a large barrier resistance which is shunted by the capacitive reactance of the barrier layer. The diode equivalent circuit is now a series R-L-C circuit shunted by the package capacitance. R_s is the spreading resistance, L_s is the lead inductance, and C_B is the barrier layer capacitance. The impedance of the negatively biased diode is

$$Z_D = \frac{R_s + j(\omega L_s - 1/\omega C_B(-v))}{R_s + j[\omega L_s - (1/\omega C_p + 1/\omega C_B(-v))]} \quad (4)$$

If the negative bias is such that the lead inductance L_s resonates with the barrier capacitance $C_B(-v)$, that is, $\omega L_s = 1/\omega C_B(-v)$ the expression for diode impedance (Z_D), reduces to:

$$Z_D = \frac{R_s(1 - j\omega C_p R_s)}{R_s^2 \ll X_p^2} \quad (5)$$

For a series resonant frequency of 9300 Mc/s and a lead inductance of 3 nanohenries, the zero bias capacitance of the diode must be of the order of 0.2 to 0.4 picofarads. Then from Equation (1), it is evident that a proper value of negative bias will reduce the zero bias capacitance to a value where series resonance occurs. In this resonant condition the diode impedance Z_D , Equation (5) is very small. Thus, microwave power incident upon the negatively biased diode is mostly reflected with little absorption, and high isolation is achieved.

It must be pointed out that while the simple theory of operation outlined above gives correlation between experimental and predicted results, the correlation is unique to diodes, such as the silver-bonded germanium varactors, whose parameters satisfy the resonant conditions specified (Table I). Other diodes, such as silicon diffused junction varactors (Table I), whose parameters do not satisfy the resonance requirements, have been used in this simple switching mode and have given good experimental results. However, an extension of the same analysis for nonresonant conditions in the package and junction fails to predict with any reasonable degree of accuracy the experimental results obtained. The reason for this discrepancy is as yet unknown, but it is believed due to a difference between the assumed and actual diode equivalent circuits as seen by the incident microwave energy.

The simple model of a diode shunted across a waveguide is an example of shunt mode switching one of the two basic modes of semiconductor switching operation; the other being appropriately the series mode. In the simple mode, as described, a diode is inserted across a transmission line of characteristic impedance Z_0 in parallel with matched load and generator impedances. In the simple series mode, the diode is inserted in the same transmission line in series with matched load and generator impedances (Fig. 2).

HIGGINS

The following will describe a three-element limiting and switching configuration, operating in a series mode at a center frequency of 9375 Mc/s.

THE SERIES SWITCH

A photograph of the three-element series switch is shown in Fig. 3. The switch consists basically of three series arms (tees), each arm containing a crystal mount terminated by a shorting plate exactly one-quarter of a guide wavelength behind the crystal seat. Each series arm is separated by quarter guide wavelengths along the main guide. This antiresonant spacing makes the isolation of each of the three diodes almost additive. For this particular configuration, the separation between each series arm, along the main guide, is $5\lambda_g/4$ where λ_g is 4.48 cm - or a frequency of 9375 Mc/s. In each series arm, the crystal seat is a distance of $\eta\lambda_g/2$, where η is any integer, from the junction of the main guide and series arm.

If it were possible to move the shorting plate to the crystal seat, a nominal halfwavelength from the waveguide junction, the plate would be translated to the wall of the main guide effectively shorting out the series arm. This would be almost as if no series arm existed and microwave energy would propagate down the main guide with little insertion loss.²⁰

Similarly, a perfect open circuit across the guide at the crystal seat would be translated to the wall of the main guide as an infinite impedance. This would have the effect of cutting off all the main guide lying to the right of the waveguide junction, providing almost infinite isolation. (This assumes incident power propagates left to right.) If at the crystal seat, an instantaneous change between open and short circuits were obtainable, an ideal switch would be realized.

The use of semiconductor diodes readily lends itself to the microwave circuit described above. It has been shown how the impedance level of a semiconductor diode in a waveguide environment at a frequency in the 9 kMc/s range will change with sudden changes in the applied bias. It has been reported that the time required for a diode to switch between high and low impedance states, as the bias is suddenly changed, is a few nanoseconds. 5,-16

The operation of the three-element series switch is the reverse of the simple shunt mode operation. In the series switch, diodes are inserted in each series arm and biased negatively to obtain transmission, and biased positively to prevent transmission.

Silver-bonded germanium varactors of Japanese manufacture, and silicon diffused junction varactors of the M450-type were tested for switching action. The characteristics of these diodes are listed in Table I. In this particular series configuration, the better switching performance was obtained using the silver-bonded

HIGGINS

germanium varactor diodes. The parameters of the silver-bonded diodes fully satisfy the resonant conditions described in the simple shunt mode operation, based on the assumed diode equivalent circuit. When forward biased, the diode equivalent circuit becomes an R-L circuit shunted by the package capacitance. Since $\omega L_s = 1/\omega C_p$, the diode is essentially a loaded tank circuit of high impedance which will, to a great degree, prevent transmission of microwave power. When reverse biased, the diode equivalent circuit is an R-L-C series circuit shunted by the package capacitance. The diode is usually biased negatively to the point where $1/\omega C_p(-V) = \omega L_s$. Then the diode equivalent circuit in this resonant condition is simply R_s , the spreading resistance, shunted by C_p , the package capacitance. The net result is a low impedance circuit which allows transmission of microwave power with little loss.

Figures 4 and 5 show isolation and insertion loss as a function of frequency for the diode series switch. The data of Fig. 4 illustrates typical switching performance obtained using GSR2 silver-bonded germanium varactors. Insertion loss lower than that of Fig. 4, of the order of 0.25 db, is obtainable at the expense of a decrease in isolation of a few db, and with a narrowing of 10 to 15 Mc/s of the 20 db and 30 db isolation bandwidths.

Figure 5 shows the series switching performance using selected silicon junction varactors. These units were selected by testing each diode individually for switching ratio, i.e., ratio of isolation to insertion loss. In the shunt case for best switching performance, diodes with low junction capacitance and high cutoff frequency were found to have best switching ratios. Any similar basis for choosing silicon units for acceptable switching action in the three-element configuration was unsuccessful and the empirical approach described above had to be used.

In Fig. 5, the peak isolation occurs at a frequency of 9450 Mc/s, 75 Mc/s greater than the design frequency of the three-element configuration. This shift is attributed to a susceptance introduced by the back plate in each series arm which, at the design frequency, is exactly a quarter of a guide wavelength behind the crystal axis. This susceptance varies rapidly with frequency. It is believed that this susceptance interacts with the diode admittance such that for the silicon units, with their higher package capacitance, peak isolation is attained at 9450 Mc/s rather than at the design frequency of 9375 Mc/s. This may also explain the narrower bandwidths obtained with the silicon varactors.

In the measurement of isolation and insertion loss versus frequency, the incident power level was 500 mw CW using the GSR2 diodes. The GSR2 is rated as capable of dissipating 500 mw. However, VSWR measurements indicated that in this particular mode of operation, 50 to 60 percent of incident power is absorbed so that as a safety factor, the incident power levels were restricted to

HIGGINS

upper levels of 500 mw CW.

The silicon units with 6-volt breakdown voltages are rated as capable of dissipating 250 mw and were tested at 250 mw CW incident power levels.

SERIES LIMITER

The three-diode series switch has been successfully operated as a passive microwave power limiter for CW input up to 650 mw. The use and operation of the three-element switch correspond with Garver's criteria "... that any diode switch providing high isolation with diode conduction will function passively as a limiter. Low RF power does not cause significant diode conduction, while high RF power results in conduction which changes the diode impedance, increasing the attenuation."¹⁵

By short-circuiting the diode biasing terminals of the three-element switch, a relatively flat limiting characteristic has been obtained. Using three silver-bonded diodes, the output power is limited to 1.8 mw for an incident power of 500 mw CW, as illustrated in Fig. 6. The output characteristic of Fig. 6 while not perfectly flat, shows an increase of only 0.7 mw in output for an increase in input power from 10 to 500 mw. For the same three diodes, Fig. 7 depicts the frequency dependence of the limiter showing a peak isolation of 24.6 db with a 20 db isolation bandwidth of 207 Mc/s. The low-level insertion loss at 250 microwatts is less than 0.6 db from 9.1 to 9.45 Gc/s rising to 1.2 db at 9.6 Gc/s. Other units have been tested which give a low-level insertion loss of less than 1.05 db over the whole band for loss of 1 db in isolation, and a slight narrowing of the 20 db bandwidth.

A technique, giving limited output power at levels lower than can be obtained with the three-diode limiter just described, is available. This technique utilizes the rectification properties of the silver-bonded diode which is reported to have a rectification ratio of 10^6 .¹⁹

The diode first seen by the incident microwave energy, diode A in Fig. 8 is inserted in its mount in a direction opposite to the direction of insertion of diodes B and C. The diode terminals are then connected to each other. Measurements indicate that most of the incident microwave energy is absorbed in diode A, giving rise to a substantial rectified current which at 500 mw of input power is as high as 10 ma. Since diodes B and C have been inserted in their respective mounts in a direction of "easy current flow," the current from diode A biases diodes B and C into forward conduction. In this state, each diode represents a high impedance and if driven deep enough into conduction to a point where resonance occurs, the diode impedance can be very large. Thus, at increased input power levels (200-600 mw) the large diode impedance tends to maintain a

HIGGINS

low output power level, but not a flat output characteristic.

Figure 9 depicts the output characteristic for this type of limiter with the output power limited to 1.1 mw for 650 mw of input power. The frequency dependence of the limiter is illustrated in Fig. 10. A peak isolation of 30 db with a 20 db isolation bandwidth of 245 Mc/s is shown. The low-level insertion loss is less than 1.0 db from 9.1 to 9.45 Gc/s, rising to 1.4 db at 9.6 Gc/s. This form of the three-diode limiter gives higher peak isolation and broader 20 db bandwidths at the expense of an increased insertion loss.

Because of the power limitations of the diodes and the physical dimensions of the three element waveguide structure, direct application of this particular limiter to military systems is limited. However, application of techniques derived from investigations of microwave semiconductor switching mechanisms of which the three element switch and limiter is only one result, is not limited. These investigations demonstrated the feasibility of using semiconductor limiters and switches as replacements for gas discharge type duplexers and receiver protectors in certain Army radars. This led to USAELRDL sponsorship and supervision of several programs to develop semiconductor receiver protectors for the manpacked personnel radars AN/PPS-4, 5, and 6.

A technique first used and developed in the three diode limiter was employed by the USAELRDL contractor in order to meet the technical requirement and delivery date. This was the technique, described previously, of using the rectified current of one diode to bias other diodes with all diodes mounted in the same physical configuration. Figure 11 is a cross sectional view of a 500 watt peak diode limiter for use as receiver protector in the AN/PPS-6 radar and is also the prototype for a 1 EW peak semiconductor duplexer for the AN/PPS-4. In this limiter a high power detector diode is recessed in the wall of the waveguide a short distance preceding the limiting diodes and is decoupled sufficiently from the main transmission line. Incoming RF power is sampled and detected by the rectifying diode and its output fed into the limiting diodes through insulated bias chokes. This effectively switches the diodes lowering the threshold of limiting, i.e., providing high isolation at lower power levels, as was the case with the three diode limiter when the rectified output of the first diode biased the other diodes. Fig. 12 is a photograph of the three diode limiter shown with a 1 KW X-band duplexer for the AN/PPS-4 (lower left) and a 500 watt X-band crystal protector for the AN/PPS-6 (lower right). In the left wall of each of the lower units in Fig. 12 the recessed rectifying crystal is visible.

SUMMARY AND CONCLUSIONS

The purpose of this report was to describe a technique for achieving series mode switching and limiting at X-band, using

HIGGINS

semiconductor diodes in a waveguide structure. Switching was achieved by the application of forward and reverse bias potentials to the diode terminals. When operated as a switch, the three-element configuration provided high isolation and low insertion loss, using either silver-bonded germanium varactors or silicon junction varactors. The better switching performance, i.e., higher isolation and lower insertion loss over wider bandwidths was obtained using the silver-bonded germanium diodes.

The three-element series configuration has also functioned passively as a microwave power limiter. Good limiting action was attained only when silver-bonded germanium units were used. This can be attributed primarily to the low voltage at which these diodes enter conduction (approximately 0.3 volt) and the diode parameters which allow resonant operation. The silicon units enter conduction at approximately 0.8 volt and have higher package capacitance, 0.4 μf , precluding resonant operation in the 9 to 10 Gc/s region.

One of the principal disadvantages of this series mode of operation is power absorption. Most of the incident microwave power is absorbed within the diodes. Thus, for reliable operation, incident power levels must be restricted to levels lower than the dissipation ratings of the particular type diode being used to prevent diode burnout.

The effect of harmonic generation has not been investigated. It is believed that this effect would only become serious in the case of the limiter at the higher power levels. Thus, if the harmonics were removed by a filter, the isolation at the fundamental frequency would be enhanced.

Because of size and power absorption the three-diode series limiter and switch has little direct system application. A two-diode series limiter and switch has been made which will reduce the size more than half and provide adequate receiver protection in moderate power (50 to 100 watts peak) radar systems. More important, however, are the techniques and information gleaned from investigations associated with the development of the three-diode series switch and limiter. This investigation determined the feasibility of further engineering effort and together with techniques incorporated in the diode series limiter made possible the successful development of semiconductor receiver protectors and duplexers, thereby increasing the capability and reliability of military radars.

The author is grateful to the following personnel of USAELRDL: Mr. G. E. Hambleton for guidance and many helpful discussions; Messrs. F. A. Brand and W. G. Mathei for their suggestions and interest; and to Mr. T. Sadd and Sage Laboratories for construction of the three-element configuration.

HIGGINS

REFERENCES:

1. D. J. Grace, "Some Applications of Crystal Rectifiers in Broadband Microwave Circuits," *Microwave Crystal Rectifier Symposium Record*, Fort Monmouth, N. J., p. 313, February, 1956.
2. A. Uhlir, Jr., "The Potential of Semiconductor Diodes in High-Frequency Communications," *Proc. IRE*, vol. 46, pp. 1099-1115, June, 1958.
3. M. R. Millet, "Microwave Switching by Crystal Diodes," *IRE Trans. on Microwave Theory and Techniques*, vol. MTT-6, pp. 284-290, July, 1958.
4. R. V. Garver, E. G. Spencer and M. A. Harper, "Microwave Switching Techniques," *IRE Trans. on Microwave Theory and Techniques*, vol. MTT 6, pp. 378-387, October, 1958.
5. R. V. Garver, "High-Speed Switching of Semiconductors-II," *IRE Trans. on Microwave Theory and Techniques*, vol. MTT-7, pp. 272-276, April, 1959.
6. R. Lucy, "Microwave High-Speed Switch," *Proc. 1959 Electronic Components Conference*, Philadelphia, Pa., pp. 12-15, May 6-8, 1959.
7. G. C. Messenger, "A Review of Parametric Diode Research," *Proc. 1959 Electronic Components Conference*, Philadelphia, Pa., pp. 220-222, May 6-8, 1959.
8. C. J. Spector, "A Design Theory for the High-Frequency P-N Junction Variable Capacitor," *IRE Trans. on Electron Devices*, vol. ED-6, pp. 347-351, July, 1959.
9. R. V. Garver, J. A. Rosado and E. F. Turner, "Theory of the Germanium Diode Microwave Switch," *IRE Trans. on Microwave Theory and Techniques*, vol. MTT-8, pp. 108-111, January, 1960.
10. M. Bloom, "Microwave Switching with Computer Diodes," *Electronics*, vol. 33, pp. 85-87, January 15, 1960.
11. A. F. Harvey, "Duplexing Systems at Microwave Frequencies," *IRE Trans. on Microwave Theory and Techniques*, vol. MTT 8, pp. 415-431, July, 1960.
12. J. Galejs, "Multidiode Switches," *IRE Trans. on Microwave Theory and Techniques*, (Correspondence), vol. MTT-8, pp. 566-569, September 1960.
13. D. W. Feldman and B. R. McAvoy, "A 100 db Microwave Semiconductor Switch," *Review of Scientific Instruments*, vol. 32,

HIGGINS

pp. 74-76, January, 1961.

14. S. Kita, J. Okajima and M. Chung, "Parametric Amplifier Using a Silver Bonded Diode," IRE Trans. on Electron Devices, vol. ED-8, pp. 105-109, March, 1961.

15. R. V. Garver and D. Y. Tseng, "X-Band Diode Limiting," IRE Trans. on Microwave Theory and Techniques, (Correspondence), vol. MTT-9, p. 202, March, 1961.

16. R. V. Garver, "Theory of TEM Diode Switching," IRE Trans. on Microwave Theory and Techniques, vol. MTT-9, pp. 224-238, May, 1961.

17. T. H. B. Baker, "Semiconductor Diode Waveguide Switch," Electronic Technology, vol. 38, pp. 300-304, August, 1961.

18. L. Gould, M. Bloom and R. Tenenholtz, "Semiconductor Microwave Control Devices," NEREM Record, 1st Edition, pp. 138-139, November, 1961.

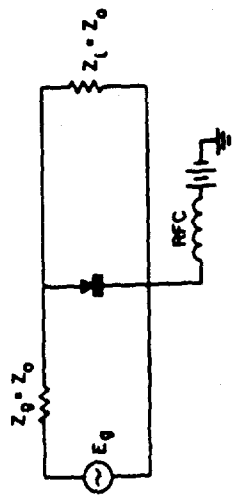
19. S. Kita, "Microwave Applications of the Silver Bonded Diode," Electronics, vol. 35, pp. 86-87, 11 May 1962.

20. G. C. Southworth, Principles and Applications of Waveguide Transmission, pp. 236-240, D. Van Nostrand Company, Inc., 1950.

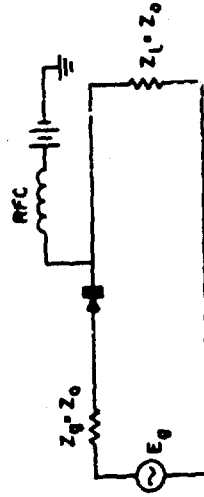
TABLE I

DIODE PARAMETERS

<u>Parameter</u>	<u>Symbol</u>	<u>Si Junction</u>		<u>Ge Silver-Bonded Varactor Diodes</u>		<u>Units</u>
		<u>MA450 Type</u>		<u>GSB1A</u>	<u>GSB2</u>	
Lead Inductance	L_B	2	3	3	3	nh
Package Capacitance	C_p	0.4	0.1	0.1	0.1	pf
Zero Bias Junction Capacitance	C_0	0.6-2.0	0.3	0.3	0.3	pf
Cutoff Frequency	F_c	90-120	60(min)	100(min)	40(min)	kMc/s
Breakdown Voltage	V_B	6	6	6	11	volts

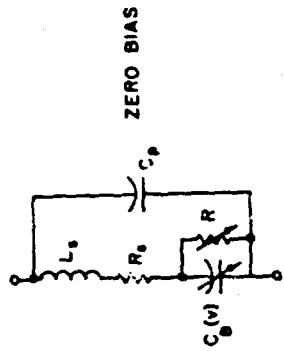


SIMPLE SHUNT MODE

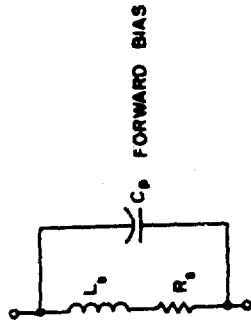


SIMPLE SERIES MODE

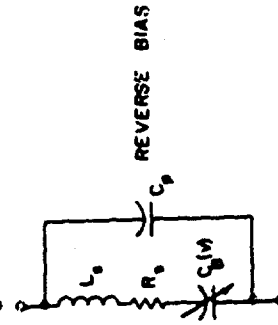
FIG. 2 SIMPLE SERIES AND SHUNT MODE



ZERO BIAS



FORWARD BIAS



REVERSE BIAS



FIG. 1 ASSUMED DIODE EQUIVALENT CIRCUITS

HIGGINS



FIG 3 PHOTOGRAPH-THREE-ELEMENT SWITCH

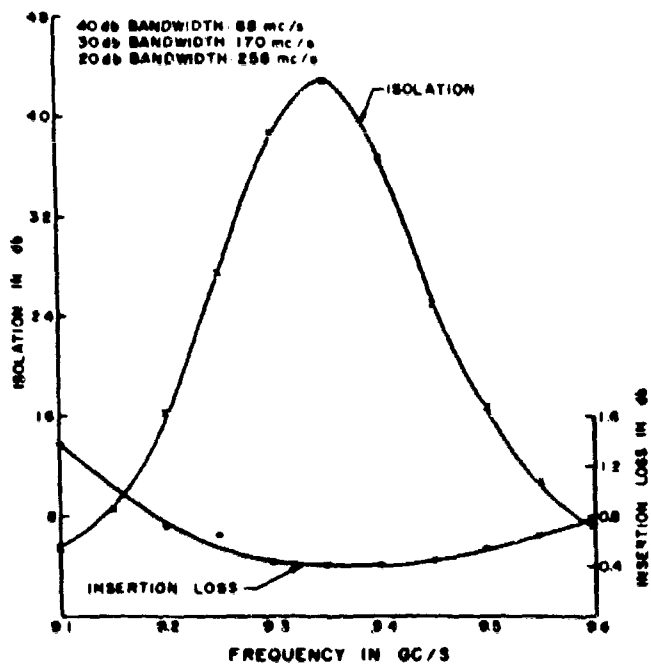


FIG. 4 TYPICAL SWITCHING PERFORMANCE OF GS82 TYPE SILVER-BONDED DIODES

HIGGINS

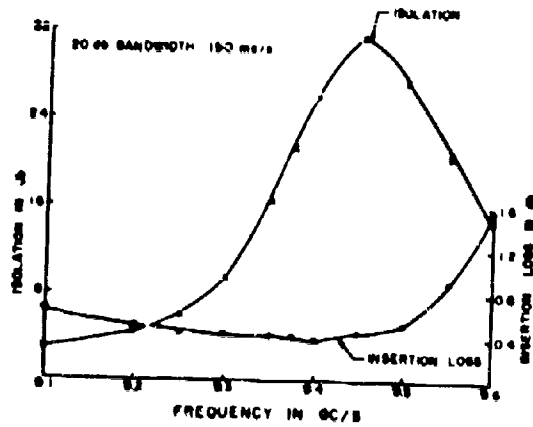


FIG. 5 TYPICAL SWITCHING PERFORMANCE OF SELECTED SILICON JUNCTION VARACTORS

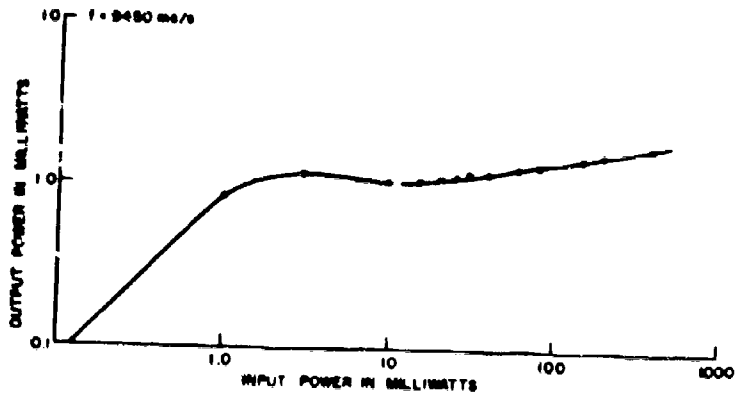


FIG. 6 OUTPUT CHARACTERISTIC OF THREE-ELEMENT LIMITER USING SILVER-BONDED DIODES WITH BC TERMINALS SHORTED

HIGGINS

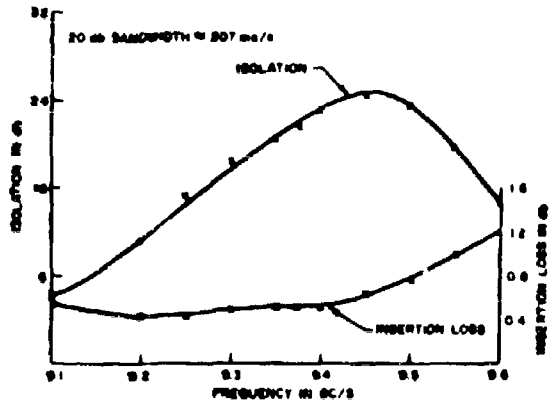


FIG. 7 FREQUENCY DEPENDENCE OF THREE-ELEMENT LIMITER USING SILVER-DIPPED DIODES WITH DC TERMINALS SHORTED

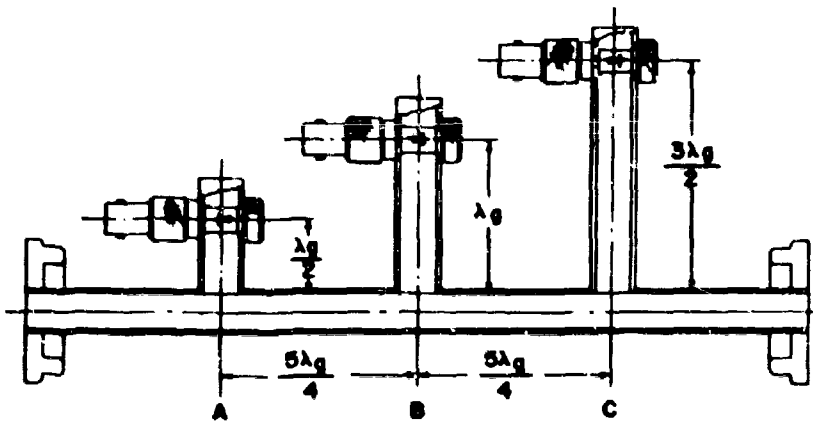


FIG. 8 THREE-ELEMENT LIMITER SHOWING POSITION OF DIODES

HIGGINS

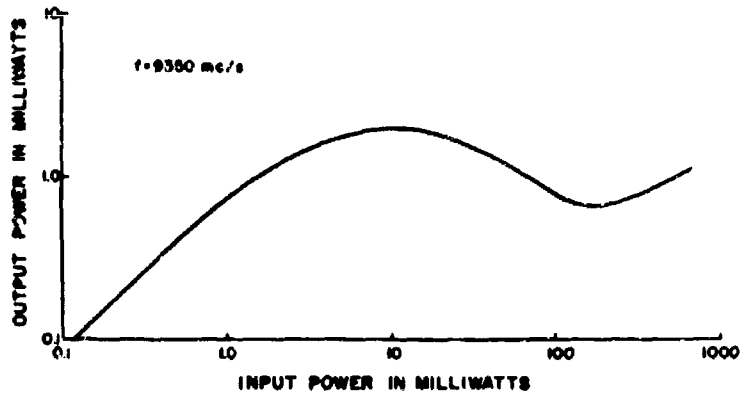


FIG. 9 OUTPUT CHARACTERISTIC OF THREE-ELEMENT LIMITER WITH FIRST DIODE 'DRIVING' OTHER TWO DIODES

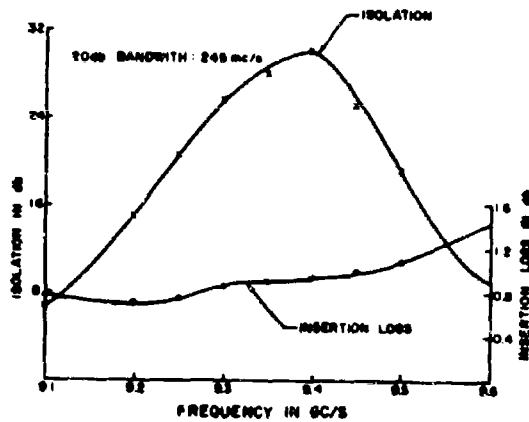


FIG. 10 FREQUENCY DEPENDENCE OF THREE-ELEMENT LIMITER WITH FIRST DIODE 'DRIVING' OTHER TWO DIODES

HIGGINS

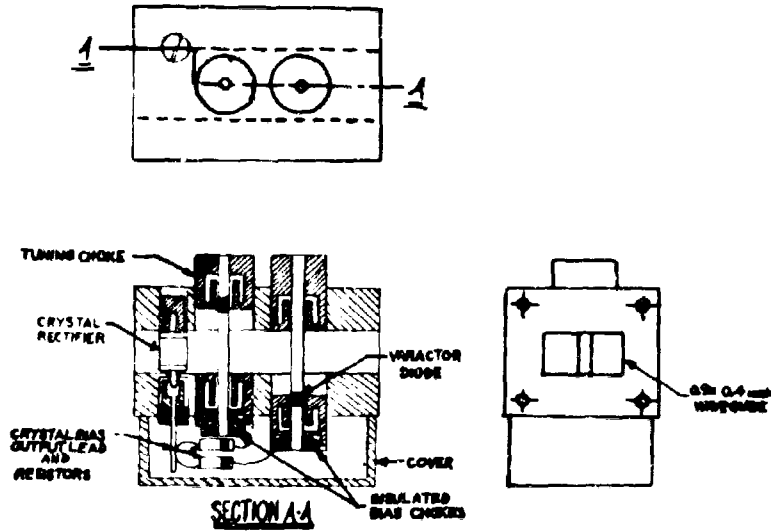


FIG. 11 CROSS SECTIONAL VIEW OF 500 WATT PEAK POWER X-BAND POWER LIMITER

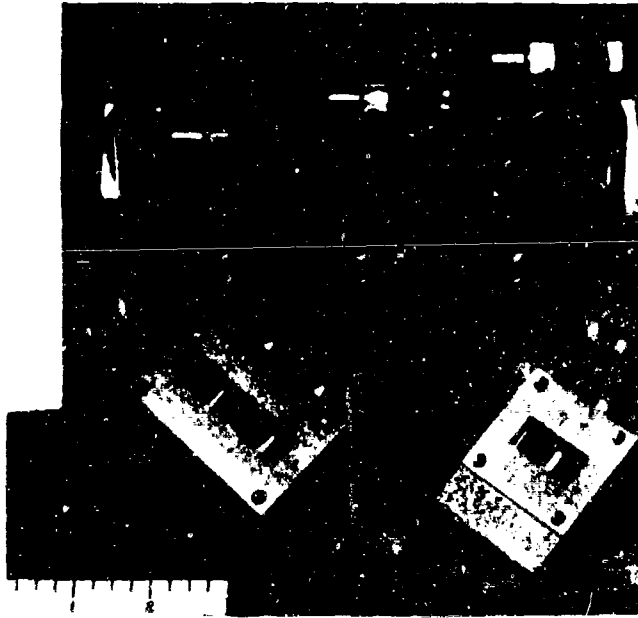


FIG. 12 THREE DIODE LIMITER AND SWITCH WITH 1 KW SEMICONDUCTOR DUPLEXER FOR AN/PPS-4 RADAR (LOWER LEFT) AND 500 WATT SEMICONDUCTOR RECEIVER FOR AN/PPS-6 (LOWER RIGHT)

MONOCYCLE POSITION MODULATION

WILLIAM A. HUBER
USAEIRD
FORT MONMOUTH, NEW JERSEY

INTRODUCTION

Information that has been encoded into a binary digital signal stream constitutes a large part of the traffic that must be transmitted over both military and commercial communication systems. It is therefore important to be able to transmit this type of signal in an efficient manner. Monocycle Position Modulation⁽¹⁾ (MPM) provides such a facility by furnishing a means for increasing the rate of transmission of binary digital data without extending the bandwidth. Monocycle Position Modulation is a particular form of time position modulation that features a signal bit structure that is easy to generate and simple to modulate. The MPM signal stream inherently contains synchronizing information and possesses a power spectrum amenable to efficient transmission, and a signal wave shape convenient for correlation detection. An analysis of the power spectrum is given in the Appendix.

This paper will describe the basic MPM concepts and how they are applied to encode a binary digital signal stream at the transmitter, and decode the resulting MPM signal stream at the receiver. Because the MPM system is synchronous and depends on phase identification for decoding at the receiver, a special synchronizing system has been developed. This system possesses unique characteristics in that it has both short time response under dynamic conditions and infinitely long time memory under static conditions. The design and operation of this synchronizing system will be discussed. An experimental model of the MPM system operating directly from a binary digital signal stream at 3000 bits/second has been built and the results of operational tests performed in the laboratory on this system will be given.

MPM CONCEPT

A salient feature of the conventional binary digital signal

HUBER

as its step-function transition between adjacent mark/space and space/mark data. This step-function characteristic affects the signaling rate, channel bandwidth, error rate and implementation necessary for its utilization as an information conveyor. The MPM system replaces the step-function with one of the form

$$g(t) = \begin{cases} (1 - \cos 4\pi t/T) & \text{if } 0 < t < T/2 \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

The binary and corresponding MPM signal bits are illustrated in Fig. 1.

Binary digital data is encoded into a mark/space or on/off signal format; the signals being established along a time scale at a periodic rate. This is illustrated in Fig. 2a. The corresponding MPM signal, Fig. 2b, consists of a serial stream of time quantized $(1 - \cos 4\pi t/T)$ signals. The time quantizing is exactly specified to any one of 8 distinct time locations once every T period. Each of these time positions within period T specifies a 3-bit code such as arbitrarily assigned in Table I. By using MPM techniques it is possible to transmit a 3-bit code in the same time required to transmit a 2-bit conventional binary digital code. Hence 50% more information is transmitted during equal time periods by using MPM.

The MPM system is designed to operate directly on a binary digital signal stream such as shown in Fig. 2a. In practice, the binary digital information is read synchronously out of store, such as a magnetic or punched paper tape, and converted into the MPM signal format, Fig. 2b. The MPM encoding is accomplished by reading simultaneously 3-bit binary digital code groups and converting these to the corresponding MPM time position code in accordance with Table I. Referring to Fig. 2 and Table I, the first 3-bit binary digital code group is (010) and the corresponding MPM code requires that the signal bit be placed in the 2nd time slot as indicated in Fig. 2b. The 2nd 3-bit binary digital code is (011) indicating that the corresponding MPM signal bit be placed in the 3rd time slot. The binary digital to MPM encoding proceeds continuously in like fashion where 3-bit binary digital code groups are converted to the equivalent MPM code. Referring to Fig. 2b, it will be noted that the MPM signal stream contains one signal bit for each time period T. This characteristic gives positive identification when the system is operating regardless of the code being transmitted. More important, this quasi-periodic characteristic of the MPM signal stream can be used to convey synchronizing information without the expenditure of additional power or transmission time.

It is convenient to consider each of the 8 time quantized positions within period T to be represented by a single stationary phasor. The angular positions that these phasors can occupy are shown in Fig. 3a, where dashed lines are used to indicate the 8 possible locations. As only one MPM signal bit is transmitted in any single time interval T, only one phasor location is occupied during

HUBER

that interval. It should also be noted that all phasor locations are fixed and equally spaced $\pi/4$ radians apart. This equal spacing between phasor locations allows the use of an 8 harmonic count of the basic MPM bit frequency $f = 2/T$ to accurately establish the 8 time quantized positions within each time interval.

The MPM signal stream Fig. 2b, can be expressed as a function of time as follows:

$$f(t) = \sum_{n=-\infty}^{\infty} \left[1 - \cos \frac{4\pi}{T} \left(t - \frac{kT}{16} - nT \right) \right] \left\{ u \left(t - \frac{kT}{16} - nT \right) - u \left[t - \frac{kT}{16} - T \left(n + \frac{1}{2} \right) \right] \right\} \quad (2)$$

where $k = 0, 1, 2, 3, 4, 5, 6$ or 7

$$u(t) \begin{cases} 0 & t < 0 \\ 1 & t \geq 0 \end{cases}$$

As can be seen from (2) the specific value of k determines the time quantized position occupied by the signal bit in the corresponding time period T . The factor k therefore contains the message content.

Decoding the MPM signal stream at the receiver consists of regenerating the original binary digital signal stream from the transmitted MPM signals. A prerequisite for accomplishing this is that the frequency $f = 2/T$ and its zero phase position be known. This is illustrated in Fig. 4 where (a) is the MPM signal stream corresponding to the transmission of periodic binary (000) and (b) is the required receiver generated reference signal. The reference signal is obtained from the local clock, which in turn is synchronized from the incoming signal stream.

From the reference signal Fig. 4b, it is possible to establish the four quadrature phase relationships represented by phasor diagram Fig. 3b. Here the four orthogonal phasors are shown as solid lines to indicate that all four signals are continuously available at the receiver. Decoding is accomplished by performing periodic correlations between the incoming MPM signal stream and the four quadrature reference signals; identification of the quantized time position occupied by the transmitted signal being established in relation to the reference phasor providing maximum correlation. Correlation⁽²⁾ techniques can be readily applied for signal decoding because of the orthogonal relationships between various time quantized positions used for encoding, the fixed duration and sinusoidal nature of the MPM signal bit, and the availability of 4 quadrature reference signals. The decoding procedure continuously identifies the time slot occupied by the MPM signal bit during each period T . When the time position is established simple logic transforms this information into the corresponding binary digital code in accordance with Table I.

SYNCHRONIZATION

While no special provision is made for the transmission of synchronizing data such information is inherently contained in the

HUBER

MPM signal stream by nature of the encoding technique. This means that the system does not require additional time or power to transmit synchronizing data. To extract synchronizing data from the transmitted signal requires a memory characteristic at the receiver. For conventional operation a short time memory such as supplied by a "fly-wheel" type of oscillator is sufficient. However, the system operation can be considerably enhanced by the use of a master clock that is frequency controlled by a special analog memory that will be described.

To visualize how synchronizing information is conveyed by the MPM signal stream it will be remembered that the signal bit can occur only one of 8 equally displaced time slots within the synchronous period T . Each successive time slot being displaced $T/16$ second. Now if each of the $(1 - \cos 4\pi t/T)$ signal bits is caused to generate a rectangular signal of duration much less than $T/16$ it would be possible to compare this pulse stream with a locally generated sine wave of period approximately $T/16$. Any frequency or phase difference between pulse train and sine wave can be used to generate an error voltage that in turn is applied to a voltage controlled oscillator thereby effecting a frequency change in the sine wave until it is locked with the pulse train. This technique of frequency and phase lock is an old one that is widely used. The only difference here is that the pulse train is not strictly periodic but does have a quasi-periodic character, in that successive pulses have random spacings that are integer multiples of $T/16$. This means that when these pulses are superimposed on a locally generated sine wave of the same period, coincidence will always occur at the same phase angle. This can be seen by reference to Fig. 5 where coincidence is shown at zero cross-over of the sine wave.

A simplified block diagram of the synchronizing system is shown in Fig. 6. The function of the phase comparer is to effect the superimposing of the waves shown in Fig. 5 and produce an output signal proportional to any phase angle change. This signal is then stored in the memory circuit for subsequent application to the voltage controlled oscillator used as the master clock. The conventional approach to the design of a store is to use an RC integrator for the memory that has a time constant much greater than the rate of data readin. The disadvantage of the RC and similar types of integrators is that when the transmitted signal is lost the memory discharges because of its finite time constant. Such a condition produces an undesirable change in the local clock frequency. What is desired is that when the transmitted signal is lost, the memory retains its last corrected value and therefore holds the local clock frequency accordingly as this is the best bit of synchronizing information that is available. This will assure minimum drift of the local clock improving the probability that synchronization will not be lost on short term signal fades.

An analog memory circuit has been developed which meets the

HUBER

above requirements, in that it possesses a memory with the ability for rapid change under dynamic conditions and infinite memory under static conditions. This memory uses a pair of tape-wound magnetic cores (3) made of square-loop magnetic material. Cores of this type have their domains oriented along the direction of the tape in either of two opposing directions with their flux density being essentially constant. The remanent state of the core can be defined by the net flux, which is simply the difference between the fluxes in the oppositely directed domains. It is the magnitude of the remanent flux that is used to store frequency and phase error information. To use the tape-wound magnetic core as a memory for frequency and phase control the value of the remanent flux is changed in accordance with the error signal developed as a result of frequency or phase differences between the comparing signals. A nondestructive readout of the flux state is provided by means of an rf carrier induced in the cores.

CONCLUSIONS

The MPM system is designed for the purpose of increasing the efficiency in transmission of binary digital data. To determine how successful the MPM technique is in accomplishing this, an experimental model of the system operating at 3000 bits/second, has been built and tested under laboratory conditions. The results of these tests indicate that the system is feasible, operates reliably and performs as predicted. The system was operated over a short wire line with a channel bandwidth of approximately 300 to 3500 cycles/second. A controllable source of Gaussian noise was added at the input to the line. With the transmission of random bits the system produced on an average of 1 error in 10^6 bits with a signal-to-noise ratio of 20 db. At 18 db signal-to-noise, the error rate was 1 in 10^4 bits. The operation of the synchronous phase lock system, Fig. 6, has been extremely stable. By using the tape-wound magnetic core memory as the source of control for the master clock no degradation in its frequency stability was noted.

The increase in the information density contained in the MPM signal over that of the original binary digital signal is the direct result of the encoding technique. The price that must be paid for this speed increase is a moderate increase in terminal equipment and a possible increase in the error rate for threshold signal-to-noise conditions. No additional bandwidth is required to handle the increased transmission speed. To substantiate this last statement, reference is made to the power spectrum analysis of the two signaling conditions, random MPM and binary digital, given in the Appendix.

The first part of the analysis is general and shows how the signal power is distributed as a function of frequency for a signal composed of a sequence of pulses whose positions, relative to their unmodulated positions, vary according to the modulation. This modulation is treated statistically, the position of any pulse being assumed to be statistically independent of any of the other pulses. The gen-

HUMER

eral expression derived for the spectral power distribution is given in (23). Equation (23) discloses that the power spectrum consists of two parts, a continuous term and a discrete term. In the absence of modulation, $K(f) = 1$ and the continuous term vanishes. This is the discrete spectrum. The continuous term in (23) is the result of random modulation, which modifies the discrete spectrum by a factor $K(f)$ and introduces a power term with a continuous spectrum.

When the general results are applied to the MPM signal the resulting power spectrum is given by (27) where as before the first term specifies the continuous spectrum and the second term, the discrete spectrum. A graphical plot of $\Phi(f)$, (27) is shown in Fig. 7 where the continuous term is represented by the shaded area indicating a "smearing" of the power as compared to the periodic condition where the power is concentrated in discrete spectral lines. The vertical arrows indicate the frequency location of the discrete power in the randomly modulated MPM signal. Referring to Fig. 7 it can be seen that most of the power is below frequency $3/T$. If the dc term is neglected, as is permissible in the MPM system, approximately 82% of the ac power is located in the frequency band between $1/4T$ and $3/T$. This is a favorable distribution with respect to channel utilization. Referring to Fig. 2b it will be noted that the MPM signaling frequency is $1/T$ and each signal is encoded with 3 bits. Hence a rate slightly better than a bit per cycle is being transmitted in the $1/4T$ to $3/T$ frequency band.

Equation (28) expresses the ac power distribution as a function of frequency for the binary digital signal as shown in Fig. 2a. This power function is plotted as a dashed line in Fig. 7. The power distribution of the binary digital signal being represented by the area under this curve. It can be seen that this power distribution is concentrated at the lower frequencies. Only approximately 50% of the ac power is located in the $1/4T$ to $3/T$ frequency band. When this is compared with approximately 82% of the total ac power being located in the same frequency band for the MPM system its superiority with respect to channel utilization is clearly indicated. It should also be noted that power spectral comparison is made for an MPM system transmitting 50% more information than the comparing binary digital system.

ACKNOWLEDGMENT

The author gratefully acknowledges the support and encouragement of Dr. Hans K. Ziegler, Chief Scientist, USAECOM and Mr. Robert S. Boykin, Director, Communications Department, USAELRDL. Indebtedness is also acknowledged to Mr. Richard E. Loisel for system implementation and operation, and to Mr. Alois J. Mattes for his contributions to circuit design and logic. The author would like to thank Dr. Richard O. Duda, Stanford Research Institute for his helpful discussion and mathematical analysis of randomly modulated MPM contained in the Appendix.

APPENDIX

The following is a general analysis of a signal composed of a sequence of pulses whose positions, relative to their unmodulated positions, vary according to the modulation. This modulation is treated statistically, the position of any pulse being assumed to be statistically independent of any of the other pulses. The results of the general analysis are applied to the MPM signal where the starting times t_n can be assumed any one of eight possible values with equal probability.

The method of analysis proceeds as follows: The MPM signal $f(t)$ Fig 2b is replaced by a signal $f_N(t)$ which is identical with $f(t)$ for $-NT < t < (N+1)T$ and is zero elsewhere. The Fourier transform $F_N(f)$ is found, and the average energy $\overline{|F_N(f)|^2}$ is determined. This is divided by $(2N+1)T$ to obtain the average power and the power spectrum of $f(t)$ is found as the limit.

$$\phi(f) = \lim_{N \rightarrow \infty} \frac{\sum |F_N(f)|^2}{(2N+1)T} \quad (3)$$

Let $g(t)$ Fig 1 be the pulse signal and let $f(t)$ Fig 2b be the MPM signal. Then

$$f(t) = \sum_{n=-\infty}^{\infty} g(t-nT-t_n) \quad (4)$$

$$f_N(t) = \sum_{n=-N}^N g(t-nT-t_n) \quad (5)$$

$$\begin{aligned} F_N(f) &= \int_{-\infty}^{\infty} f_N(t) e^{-j2\pi ft} dt \\ &= \sum_{n=-N}^N \int_{-\infty}^{\infty} g(t-nT-t_n) e^{-j2\pi ft} dt \\ &= \int_{-\infty}^{\infty} g(t) e^{-j2\pi ft} dt \sum_{n=-N}^N e^{-j2\pi fnT} e^{-j2\pi ft_n} \end{aligned} \quad (6)$$

or

$$F_N(f) = G(f) \sum_{n=-N}^N z^n e^{-j2\pi ft_n} \quad (7)$$

where

$$G(f) = \int_{-\infty}^{\infty} g(t) e^{-j2\pi ft} dt \quad (8)$$

$$z = e^{-j2\pi fT} \quad (9)$$

Then

$$\begin{aligned} |F_N(f)|^2 &= |G(f)|^2 \sum_{n=-N}^N \sum_{m=-N}^N z^n z^{*m} e^{-j2\pi f t_n} e^{j2\pi f t_m} \\ \overline{|F_N(f)|^2} &= |G(f)|^2 \sum_{n=-N}^N \sum_{m=-N}^N z^n z^{*m} \overline{e^{-j2\pi f t_n} e^{j2\pi f t_m}} \end{aligned} \quad (10)$$

Where we use a bar to denote ensemble (statistical) averages. Assuming that t_n and t_m are statistically independent, $n \neq m$ and assuming that the probability densities $p(t_n)$ are the same for all n , we have

$$\overline{e^{-j2\pi f t_n} e^{j2\pi f t_m}} = \begin{cases} 1 & \text{if } n = m \\ \left| \overline{e^{-j2\pi f t_n}} \right|^2 & \text{if } n \neq m \end{cases} \quad (11)$$

Let $K(f) = \left| \overline{e^{-j2\pi f t_n}} \right|^2$ (12)

and $\delta_{nm} = \begin{cases} 1 & \text{if } n = m \\ 0 & \text{if } n \neq m \end{cases}$ (13)

Then $\overline{e^{-j2\pi f t_n} e^{j2\pi f t_m}} = \delta_{nm} + (1 - \delta_{nm}) K(f) = \delta_{nm} (1 - K(f)) + K(f)$ (14)

Substituting (14) in (10) we obtain

$$\begin{aligned} \overline{|F_N(f)|^2} &= |G(f)|^2 \left[\sum_{n=-N}^N \sum_{m=-N}^N z^n z^{*m} \delta_{nm} (1 - K(f)) + \sum_{n=-N}^N \sum_{m=-N}^N z^n z^{*m} K(f) \right] \\ &= |G(f)|^2 \left[(2N+1)(1 - K(f)) + K(f) \left| \sum_{n=-N}^N z^n \right|^2 \right] \end{aligned} \quad (15)$$

Thus we obtain the following approximation to the power spectrum:

$$\phi_N(f) = \frac{\overline{|F_N(f)|^2}}{(2N+1)T} \quad (16)$$

$$= \frac{|G(f)|^2}{T} \left[(1 - K(f)) + K(f) \frac{\left| \sum_{n=-N}^N z^n \right|^2}{(2N+1)} \right] \quad (17)$$

Let

$$J_N(f) = \sum_{n=-N}^N z^n \quad (18)$$

Evaluating this series we obtain

$$J_N(f) = \cos N2\pi fT + \sin N2\pi fT \cot \pi fT \quad (19)$$

We are primarily interested in the limiting behavior of $|J_N(f)|^2 / (2N+1)$ as $N \rightarrow \infty$. For N large and ϵ small

$$\int_{-\epsilon}^{\epsilon} \frac{|J_N(f)|^2}{2N+1} df \cong \frac{1}{2N+1} \int_{-\epsilon}^{\epsilon} \sin^2 N2\pi fT \cot^2 \pi fT df \cong \frac{1}{T} \quad (20)$$

Where the approximation becomes exact (for $\epsilon \rightarrow 0$) as $N \rightarrow \infty$. Thus $|J_N(f)|^2 / (2N+1)$ approaches an impulse train, the period being $1/T$.

$$\lim_{N \rightarrow \infty} \frac{|J_N(f)|^2}{2N+1} = \frac{1}{T} \sum_{k=-\infty}^{\infty} \delta(f - \frac{k}{T}) \quad (21)$$

Using this result in connection with (17) and (19)

$$\phi(f) = \lim_{N \rightarrow \infty} \phi_N(f) \quad (22)$$

$$\phi(f) = \frac{|G(f)|^2}{T} \left[(1-K(f)) + K(f) \frac{1}{T} \sum_{k=-\infty}^{\infty} \delta(f - \frac{k}{T}) \right] \quad (23)$$

For the periodic condition where t_n is not random, $K(f)=1$

$$\phi(f) = \frac{|G(f)|^2}{T^2} \sum_{k=-\infty}^{\infty} \delta(f - \frac{k}{T}) \quad (23a)$$

The general results (23) are now applied to MPM signals (1) where starting times assume one of the values $kT/16$, $k=0, 1, 2, \dots, 7$, then

$$|G(f)|^2 = \left[\frac{\sin \frac{\pi f T}{2}}{\pi f \left[1 - \left(\frac{fT}{2} \right)^2 \right]} \right]^2 \quad (24)$$

The quantity $K(f)$ is found from the probability density function

$$P(t_n) = \frac{1}{8} \sum_{k=0}^7 \delta(t_n - \frac{kT}{16}) \quad (25)$$

and

$$K(f) = \left[\frac{\sin \frac{\pi}{2} fT}{8 \sin \frac{\pi}{16} fT} \right]^2 \quad (26)$$

Combining (23) (24) and (26) we obtain the power spectrum for randomly modulated MPM signal.

$$\begin{aligned} \phi(f) = & \frac{1}{T} \left[\frac{\sin \frac{\pi}{2} fT}{\pi f \left[1 - \left(\frac{fT}{2} \right)^2 \right]} \right]^2 \left[1 - \left(\frac{\sin \frac{\pi}{2} fT}{8 \sin \frac{\pi}{16} fT} \right)^2 \right] \\ & + \frac{1}{T^2} \left[\frac{\sin \frac{\pi}{2} fT}{\pi f \left[1 - \left(\frac{fT}{2} \right)^2 \right]} \right]^2 \left[\frac{\sin \frac{\pi}{2} fT}{8 \sin \frac{\pi}{16} fT} \right]^2 \sum_{k=-\infty}^{\infty} \delta \left(f - \frac{k}{T} \right) \end{aligned} \quad (27)$$

By the same general approach it can be shown that the power spectrum for the randomly modulated binary digital signal is

$$\phi_{bd}(f) = \frac{T}{2} \frac{\sin^2 \frac{\pi fT}{2}}{\left(\frac{\pi fT}{2} \right)^2} \quad (28)$$

REFERENCES

1. USAFPA, Patent Application Docket No. 12,518.
2. Y. W. LEE, et al, "Applications of Correlation Analysis to the Detection of Periodic Signals in Noise", Proc IRE Vol. 38, 1950.
3. HAROLD S. CRAFTS, "Design of a Magnetic Variable Component for Adaptive Network", 1963 WESCOM convention, San Francisco, California.

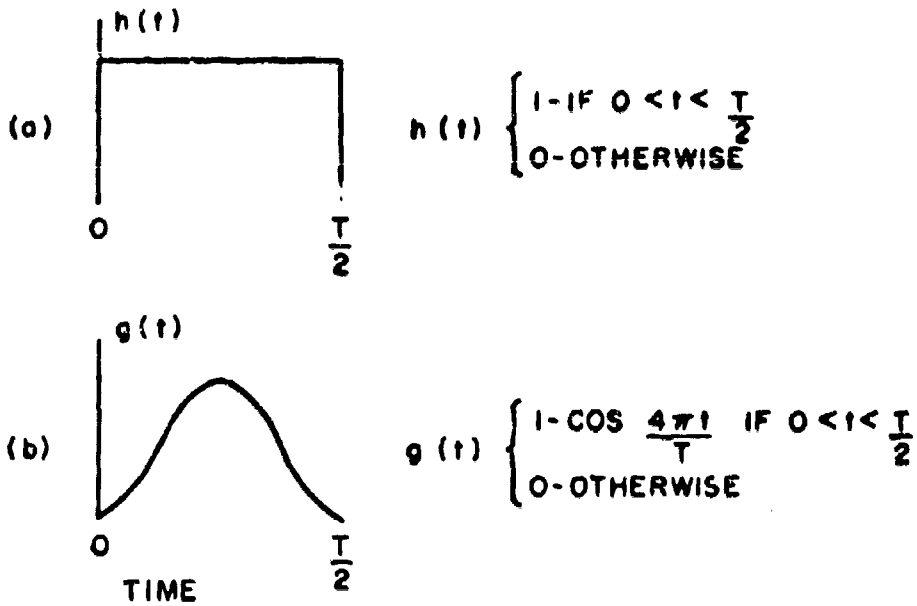


FIG. 1 (a) BINARY DIGITAL SIGNAL BIT
(b) MPM SIGNAL BIT

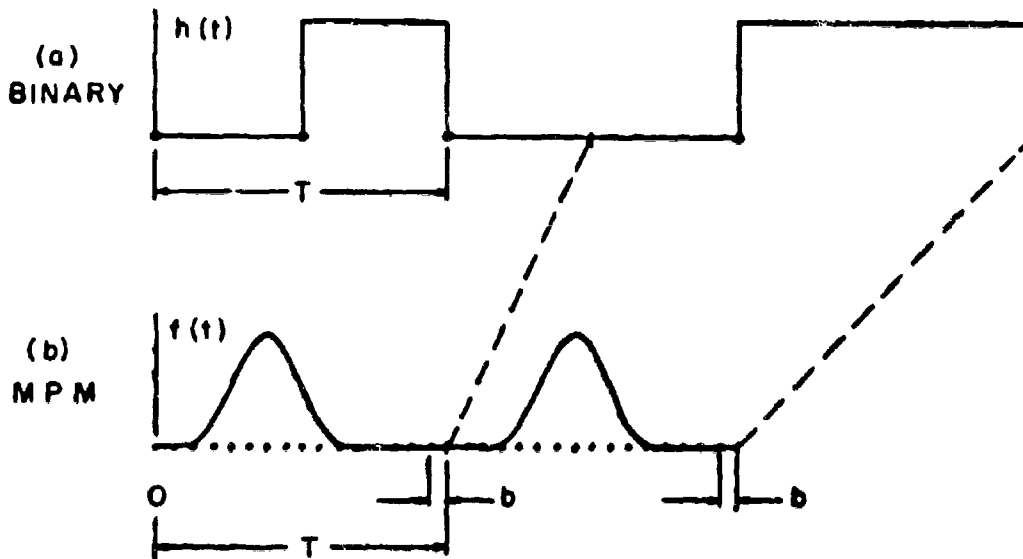


FIG. 2 COMPARISON OF BINARY AND MPM SIGNAL STREAMS

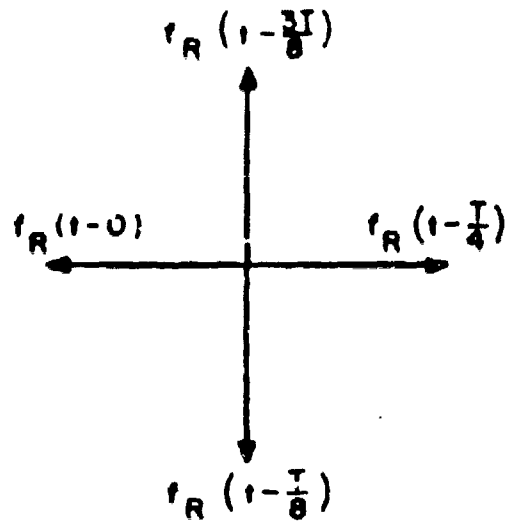
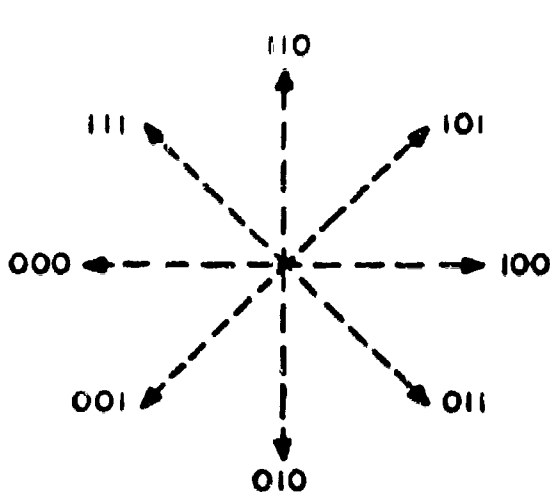


FIG. 3a PHASOR REPRESENTATIONS OF ENCODING POSITIONS

FIG. 3b PHASOR REPRESENTATIONS OF REFERENCE SIGNALS USED FOR CORRELATION DETECTION AT THE RECEIVER

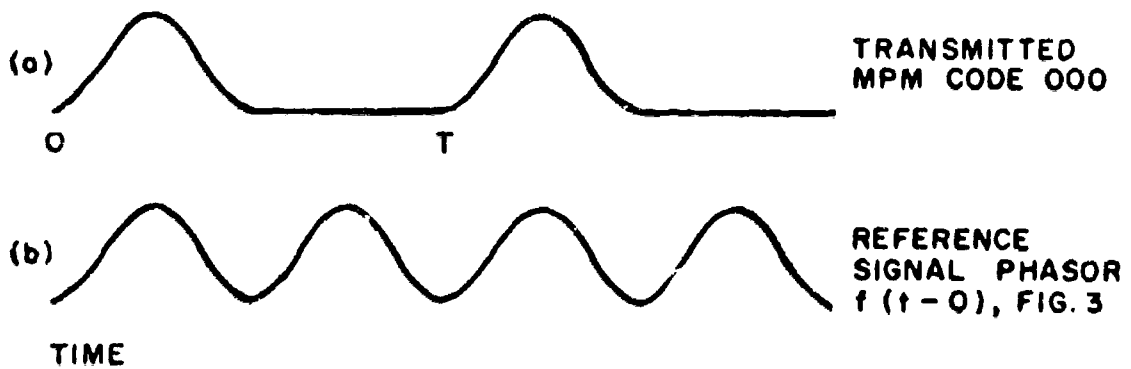


FIG. 4 ILLUSTRATING INITIAL ALIGNMENT BETWEEN TRANSMITTED MPM CODE 000 AND REFERENCE PHASOR $f(1-0)$

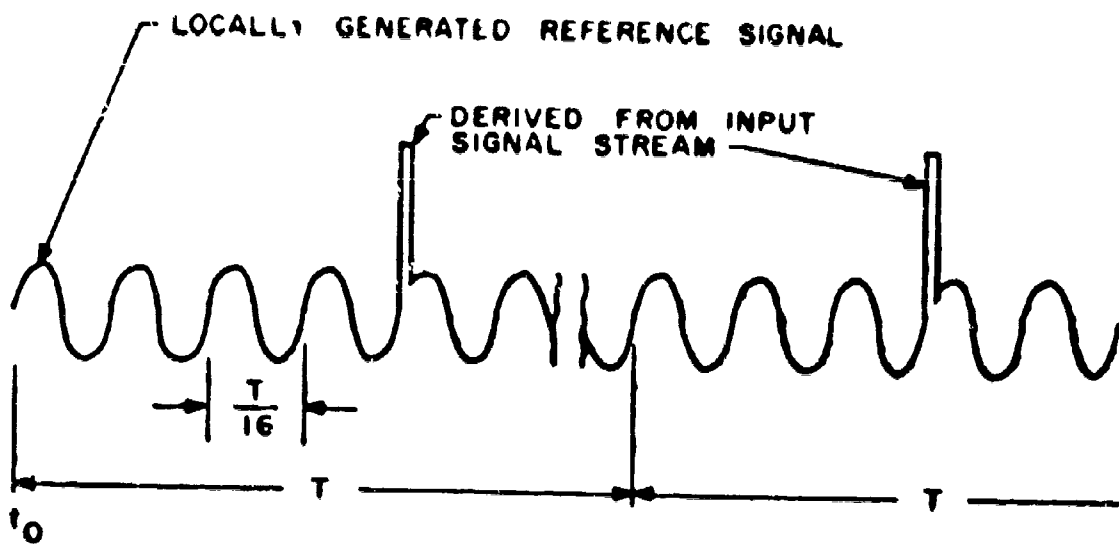


FIG. 5 SUPERIMPOSED DERIVED SYNCH AND LOCALLY GENERATED REFERENCE SIGNAL

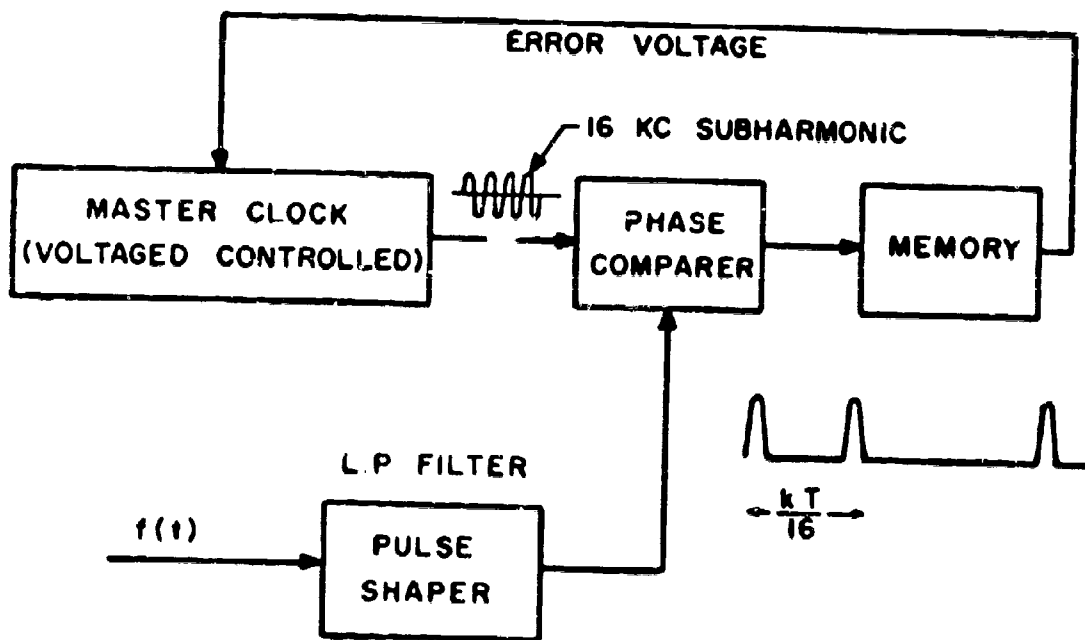


FIG. 6 PHASE LOCKED OSCILLATOR AT RECEIVER

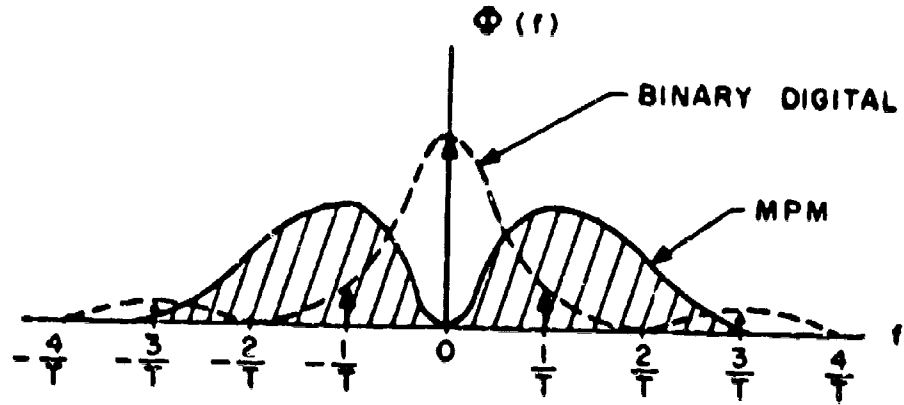


FIG. 7 POWER DISTRIBUTION FOR RANDOMLY MODULATED SIGNALS

TABLE I

(k)	<u>M P M</u>	<u>B I N A R Y</u>
	0	000
	1	001
	2	010
	3	011
	4	100
	5	101
	6	110
	7	111

DISTRIBUTION LIST

DEPARTMENT OF DEFENSE

Director of Defense, Research & Engineering Washington 25, D. C.	1
Defense Documentation Center Building 5 Cameron Station Alexandria, Virginia 22314	250
Chief Defense Atomic Support Agency Washington 25, D. C.	1
Director Weapons Systems Evaluation Group Department of Defense Washington 25, D. C.	1
Director Armed Forces Institute of Pathology Washington 25, D. C.	1
Commandant Industrial College of the Armed Forces Fort Lesley J. McNair Washington 25, D. C. Attn: Library	1
Commandant National War College Fort Lesley J. McNair Washington 25, D. C. Attn: Library	1
Commandant Armed Forces Staff College Norfolk, Virginia Attn: Library	1

DEPARTMENT OF THE ARMY

Assistant Secretary of the Army (R&D) Department of the Army Washington 25, D. C.	1
Chief of Staff United States Army Washington 25, D. C.	1
Chief of Research and Development United States Army Washington 25, D. C. Attn: Scientific and Technical Information Division	10

DEPARTMENT OF THE ARMY (Continued)

Commanding Officer Army Research Office (Durham) Box CM, Duke Station Durham, North Carolina	1
Chief US Army R&D Liaison Group (9983) (Europe) APO 757 New York, New York	1
Commanding Officer US Army R&D Group (Far East) (9984) APO 343 San Francisco, California	1
Chief US Army Security Agency Arlington Hall Station Arlington 12, Virginia	1
Deputy Chief of Staff for Logistics United States Army Washington 25, D. C.	1
Deputy Chief of Staff for Military Operations United States Army Washington 25, D. C.	1
Deputy Chief of Staff for Personnel United States Army Washington 25, D. C.	1
Assistant Chief of Staff for Intelligence United States Army Washington 25, D. C.	1
Commanding Officer US Army Personnel Research Office Washington 25, D. C.	1

Army Materiel Command

Commanding General U. S. Army Materiel Command Washington, D. C.	1
Commanding General U. S. Army Materiel Command Attn: AMCSA-Chief Scientist Washington, D. C.	1
Commanding General U. S. Army Materiel Command Attn: AMCMS Washington, D. C.	1
Commanding General U. S. Army Materiel Command Attn: AMCMU Washington, D. C.	1
Commanding General U. S. Army Materiel Command Attn: AMCAD Washington, D. C.	1
Commanding General U. S. Army Materiel Command Attn: AMGIS Washington, D. C.	1
Commanding General U. S. Army Materiel Command Attn: AMCCP-C Washington, D. C.	1
Commanding General U. S. Army Materiel Command Attn: AMCRD-S Washington, D. C.	1
Commanding General U. S. Army Materiel Command Attn: AMCRD-I Washington, D. C.	1
Commanding General U. S. Army Materiel Command Attn: AMCRD-DE Washington, D. C.	1
Commanding General U. S. Army Materiel Command Attn: AMCDS Washington, D. C.	1
Commanding General U. S. Army Materiel Command Attn: AMCRD-DW Washington, D. C.	1
Commanding General U. S. Army Missile Command Redstone Arsenal, Alabama	10

Army Materiel Command (continued)

Commanding General U. S. Army Missile Command Attn: Director of Missile Intelligence Redstone Arsenal, Alabama	1
Commanding General White Sands Missile Range Las Cruces, New Mexico	1
Commanding General U. S. Electronics Proving Ground Fort Huachuca, Arizona	1
Commanding General U. S. Army Satellite Communications Agency Fort Monmouth, New Jersey	1
Commanding General U. S. Army Tank-Automotive Center Warren, Michigan	1
Commanding General U. S. Army Electronics Command Fort Monmouth, New Jersey	1
Commanding General U. S. Army Mobility Command Warren, Michigan	1
Commanding General U. S. Army Munitions Dover, New Jersey	2
Commanding General U. S. Army Supply & Maintenance Command Washington, D. C.	1
Commanding General U. S. Army Test & Evaluation Command Aberdeen Proving Ground, Maryland	1
Commanding General U. S. Army Weapons Command Rock Island Arsenal Rock Island, Illinois	1
Commanding General U. S. Army Weapons Command Attn: R&D Directorate Rock Island Arsenal Rock Island, Illinois	2
Commanding General Natick Laboratories Natick, Massachusetts	2
Commanding Officer U. S. Army Foreign Science & Technology Center Munitions Building Washington, D. C.	5
Commanding Officer U. S. Army Ballistic Laboratories Aberdeen Proving Ground, Maryland	1

Army Materiel Command (continued)

Commanding Officer U. S. Army Human Engineering Laboratories Aberdeen Proving Ground, Maryland	1
Commanding Officer U. S. Army Coating & Chemical Laboratory Aberdeen Proving Ground, Maryland	1
Commanding Officer U. S. Army Cold Regions Research and Engineering Laboratory Hanover, New Hampshire	2
Commanding Officer U. S. Army Materiel Research Agency Watertown Arsenal Watertown, Massachusetts	1
Commanding Officer Harry Diamond Laboratories Washington, D. C.	2
Commanding Officer U. S. Army Nuclear Defense Laboratory Edgewood Arsenal, Maryland	1
Commanding Officer Aberdeen Proving Ground Aberdeen Proving Ground, Maryland	1
Commanding Officer Dugway Proving Ground Salt Lake City, Utah	1
Commanding Officer Edgewood Arsenal Edgewood Arsenal, Maryland	1
Commanding Officer Frankford Arsenal Philadelphia, Pennsylvania	2
Commanding Officer Picatinny Arsenal Dover, New Jersey	2
Commanding Officer Rock Island Arsenal Rock Island, Illinois	2
Commanding Officer Springfield Armory Springfield, Massachusetts	2
Commanding Officer Watervliet Arsenal Watervliet, New York	2
Commanding Officer U. S. Army Biological Laboratories Fort Detrick, Maryland	1
Commanding Officer U. S. Army Chemical Research & Development Laboratories Edgewood Arsenal, Maryland	1

Army Materiel Command (continued)

Commanding Officer 1
U. S. Army Electronics Research & Development Laboratory
P. O. Box 205
Mountain View, California

Commanding Officer 1
U. S. Army Engineer Research & Development Laboratories
Fort Belvoir, Virginia

Commanding Officer 1
U. S. Army Transportation Research Command
Fort Eustis, Virginia

Office of the Project Manager, CCIS-70 1
Army Materiel Command
Fort Belvoir, Virginia

Office, Chief of Engineers

Chief of Engineers 4
Department of Army
Washington, D. C. 20315
Attn: Chief Scientific Advisor

Commanding General 1
U. S. Army Engineer Division
Ohio River Division Laboratories
Cincinnati 27, Ohio

Director 2
U. S. Army Waterways Experiment Station
P. O. Box 631
Vicksburg, Mississippi

Director 1
U. S. Army Engineer Geodesy Intelligence
and Mapping R&D Agency
Fort Belvoir, Virginia

Office of The Surgeon General

The Surgeon General 1
United States Army
Washington 25, D. C.

Commanding General 1
U. S. Army Medical R&D Command
Washington 25, D. C.

Commanding Officer 1
U. S. Army Research Institute of Environmental Medicine
Natick, Massachusetts

Office of The Surgeon General (continued)

Commanding Officer	1
U. S. Army Medical Research Laboratory	
Fort Knox, Kentucky	
Commanding Officer	1
U. S. Army Medical Research & Nutrition Laboratory	
Fitzsimons Army Hospital	
Denver 30, Colorado	
Commanding Officer	1
U. S. Army Surgical Research Unit	
Brooke Army Medical Center	
Fort Sam Houston, Texas	
Commanding Officer	1
Medical Equipment Development Laboratory	
Fort Totten, New York	
Director	1
Walter Reed Army Institute of Research	
Walter Reed Army Medical Center	
Washington 12, D. C.	

Other Army Agencies

Army War College, Library-B-205	1
Carlisle Barracks	
Carlisle, Pennsylvania	
Command and General Staff College	1
Fort Leavenworth, Kansas	
Attn: Library	
Commanding General	1
U. S. Continental Army Command	
Fort Monroe, Virginia	
U. S. Continental Army Command	1
U. S. Combat Development Experimentation Center	
Fort Ord, California	
Office of Special Weapons Development	1
Fort Bliss, Texas	
Combat Operations Research Group	1
Headquarters, Continental Army Command	
Fort Monroe, Virginia	
President	1
U. S. Army Air Defense Board	
Fort Bliss, Texas	
Combat Developments Office	1
U. S. Army Infantry School	
Fort Benning, Georgia	
Superintendent	1
United States Military Academy	
West Point, New York	

DEPARTMENT OF THE NAVY

Chief of Naval Operations 1
Washington 25, D. C.
Chief of Naval Research 1
Office of Naval Research (Code 407)
Washington 25, D. C.
Director 1
U. S. Naval Research Laboratories
Washington 25, D. C.
Director 1
Weapons Systems Analysis Division
Bureau of Naval Weapons
Washington 25, D. C.

DEPARTMENT OF THE AIR FORCE

Headquarters U. S. Air Force (AFCOA) 1
Washington 25, D. C.
Commander 1
Air Force Systems Command
Andrews Air Force Base
Washington 25, D. C.
Commander 1
Office of Aero Space Research
Building T-D
Washington 25, D. C.
Director 1
Air University Library
Maxwell Air Force Base, Alabama
Attn: AUL-8641
Commanding Officer 1
Wright Patterson Air Force Base, Ohio

OTHER GOVERNMENT AGENCIES

Atomic Energy Commission 1
1901 Constitution Avenue, N. W.
Washington 25, D. C.
National Aeronautics and Space Agency 1
Washington 25, D. C.
National Bureau of Standards 1
Connecticut Avenue & Van Ness Streets, N. W.
Washington 25, D. C.
Human Performance Requirements 1
NASA, Ames Research Center
Moffett Field, California
Director 1
Science Information Exchange
Smithsonian Institution
Washington 25, D. C.
National Bureau of Standards 1
Boulder, Colorado

OTHER GOVERNMENT AGENCIES (continued)

National Institutes of Health Bethesda, Maryland Attn: Director	1
National Science Foundation Washington 25, D. C.	1
The Library of Congress Washington 25, D. C.	1

OTHER AGENCIES

British Joint Services Mission Thru: Foreign Liaison Office Assistant Chief of Staff, Intelligence Department of the Army Washington 25, D. C.	1
Canadian Joint Staff Thru: Foreign Liaison Office Assistant Chief of Staff, Intelligence Department of the Army Washington 25, D. C.	1
Human Resources Research Office The George Washington University 2013 G Street, N. W. Washington 7, D. C.	1
National Academy of Sciences 2101 Constitution Avenue, N. W. Washington, D. C.	1
Research Analysis Corporation 6935 Arlington Road Bethesda, Maryland	1
Special Operations Research Office American University 4501 Massachusetts Avenue, N. W. Washington 16, D. C.	1
Mathematics Research Center, U. S. Army University of Wisconsin Madison, Wisconsin	1