UNCLASSIFIED

AD 285 811

Reproduced by the

ARMED SERVICES TECHNICAL INFORMATION AGENCY
ARLINGTON HALL STATION
ARLINGTON 12, VIRGINIA

UNCLASSIFIED

NOTICE: When government or other drawings, specifications or other data are used for any purpose other than in connection with a definitely related government procurement operation, the U. S. Government thereby incurs no responsibility, nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use or sell any patented invention that may in any way be related thereto.

285 811

ARMY RESEARCH OFFICE

285811
CATALOGED BY ASTIA
AS AD NO.

PROCEEDINGS OF THE

1962 ARMY SCIENCE CONFERENCE

United States Military Academy, West Point, N.Y.

20-22 JUNE 1962

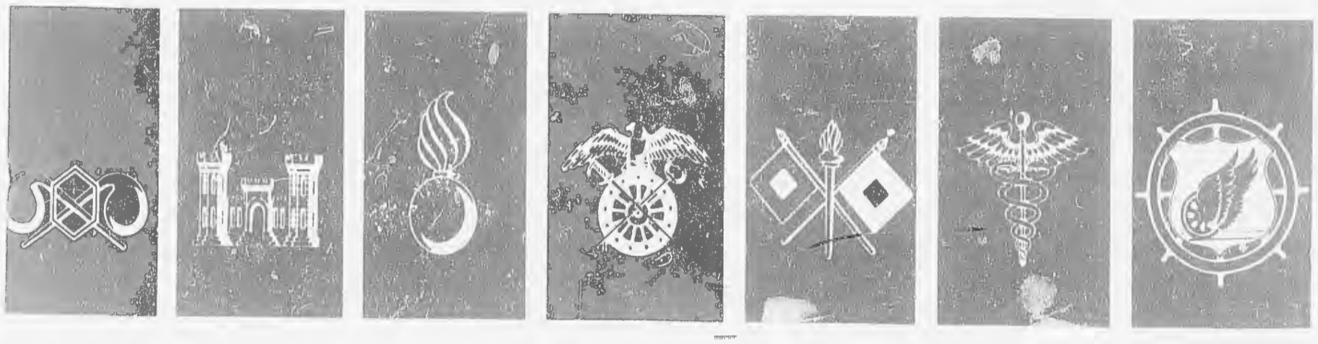
VOLUME I

(Principal Authors "A" thru "K")

OFFICE OF THE CHIEF OF RESEARCH AND DEVELOPMENT

☆ ☆ ☆ ☆ ☆ ☆ ☆

HEADQUARTERS
DEPARTMENT OF THE ARMY
Washington 25, D.C.



AVAILABILITY NOTICE

Qualified requestors may obtain copies of individual papers contained in this report from ASTIA. Requests for complete volumes should be sent to the address in the following paragraph.

Government agencies not authorized to utilize ASTIA may obtain copies of individual papers or the complete report from the Office, Chief of Research and Development, Department of the Army, Washington 25, D. C., Attention: Research Support Division.

Non - government agencies not accredited to ASTIA obtain copies of individual papers from the Office of Technical Services, Department of Commerce, Washington 25, D. C.

HEADQUARTERS
DEPARTMENT OF THE ARMY
OFFICE OF THE ADJUTANT GENERAL
WASHINGTON 25, D. C.

IN REPLY REFER TO
AGAM-P (M) 337 (5 Sep 62) CRD

19 September 1962

SUBJECT: Proceedings of Army Science Conference, 1962

TO: See Distribution

1. Inclosed for your information and use is Volume One of the "Proceedings of Army Science Conference, 1962," which is comprised of the unclassified papers which were read at the conference.

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

By Order of the Secretary of the Army:

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

1 Incl
Vol 1, Proceedings of Army
Science Conference, 1962

DISTRIBUTION:
As Indicated in Distribution
List in Volume

P R O C E E D I N G S

O F T H E

1962 ARMY SCIENCE CONFERENCE

UNITED STATES MILITARY ACADEMY, WEST POINT, N. Y.

20-22 JUNE 1962

V O L U M E I

(Principal Authors "A" thru "K")

INTRODUCTION

1. The Proceedings of the 1962 Army Science Conference are a compilation of all papers delivered at the conference. Its purposes are to:

   a. Provide a historical record of the conference.

   b. Serve as source material for scientists engaged in research for the Army.

   c. Enable conference participants to review papers in which they have a special interest.

2. The report consists of three volumes, as follows:

   Volume 1 - Unclassified papers
   Volume 2 - Unclassified papers
   Volume 3 - Classified papers

DWIGHT E. BEACH
Lieutenant General, GS
Chief of Research and Development

TABLE OF CONTENTS

PROCEEDINGS OF 1962 ARMY SCIENCE CONFERENCE

AUTHOR	TITLE	VOL	PAGE
Abbe, E. H. Hatch, H. P. Korytoski, R. D.	Studies of Nondestructive Testing Techniques to Control Properties of Case Hardened Steel Components	1	393
Abel, J. E. Kaufman, J. V. R. Reitzner, B.	Thermal Decomposition of Lead Azide in Various Environments	2	223
Abrahamson, E. P. II Alexander, J. A.	The Role of Electron Configuration on Properties in Dilute Solid Solution Alloys	1	1
Alexander, J. A.	See Abrahamson, E. P. II	1	1
Allen, A. O. Castorina, T. C. Johnson, E. R.	The Radiation Chemistry of Liquid Dinitrogen Tetroxide	1	103
Allen, F. J.	Some New Calculational Techniques in Neutron Shielding	3	1
Allen, F. J.	Weapon Debris as a High Altitude Lethal Mechanism	3	15
Allen, H. C. Howard, C. D. Huskins, C. W.	Synthesis of Carboxy Terminated Polybutadiene	3	29
Asptein, M. Comyn, R. H.	Nuclear Timing	1	135
Armendt, B. F. Hoffman, A. J.	Blast Effects of Simultaneously Detonated Multiple Explosive Charges	3	159
Assur, A.	Surfacing Submarines Through Ice	1	11
Ayers, O. E. Howard, C. D. Huskins, C. W.	Plastisol-Nitrocellulose Propellants Plasticized with N-F Oxidizers	3	37

AUTHOR	TITLE	VOL	PAGE
Bailey, S. D. Hasselstrom, T. Reese, E. T.	Regeneration of Food Flavors Through Enzymatic Action	1	285
Baker, Maj E. M. Sauberlich, H. E. Wallace, W. T. Wolfskill, S. J.	Tracer Studies of Vitamin C Utilization in Men; Metabolism of D-Glucurono-lactone-6-C14, D-Glucuronic-6-C14 Acid and L-Ascorbic-1 C14 Acid	1	21
Ball, D. H. Long, L., Jr. Mitra, A. K.	A Novel Carbohydrate Reaction	2	47
Barila, Lt Col T. Mon, G. Nunn, Capt D. E. Straub, H. H. Woodward, K. E.	A Fluid Amplifier Heart Pump	2	423
Barr, T. A., Jr. Cason, C. M. Ehrlich, J. J. Perkins, J. F.	Interim Report on the Design and Construction of the AOMC 8,000-kw Plasma Facility	1	221
Benson, R. C. Godwin, 2/Lt R. O. Mirarchi, M. R.	A Single-Pulse Ruby Laser for Ranging Applications	3	51
Berard, Capt C. W. Geever, E. F. Levenson, S. M. Rosen, H.	Wound Healing, Collagen, and Humoral Growth-Promoting Agents	2	251
Boatright, A. Metle, H.	Fast Response Solid State PME Detector for Laser Signals.	1	31
Boot, R. A. Cash, C. H.	COCODE Radar	3	65
Brand, F. A. Jacobs, H. Meindl, J. D. Weitz, S.	The Interaction of Electromagnetic Radiation and Semiconductors	1	339

AUTHOR	TITLE	VOL	PAGE
Brosious, Capt G. D.	Generation and Control an Outdoor Electromagnetic Environment Through Automatic Means	1	47
Callahan, Lt Col L. G. Jr.	Optimum Linear Filtering for Line Scan Imaging Systems	1	61
Carman, C. M.	Application of the Principles of Fracture Mechanics to Recoilless Rifle Problems	1	75
Carts, S. L. Jr.	Feasibility of Non-Metallic Mine Detection by Earth's Magnetic Field Distortion	1	89
Cash, C. H.	See Boot, R. A.	3	65
Cason, C. M.	See Barr, T. A. Jr.	1	221
Castorina, T. C.	See Allen, A. O.	1	103
Charlton, H. Neblett, J. H.	An Operations Research Model of Motor Truck Transport Derived from Nuclear Transport Theory	1	115
Christian, J. R. Goubau, G.	Investigations on a Beam Waveguide for Optical Frequencies	1	273
Coberly, R. H.	Launcher Dynamics (Summary)	3	79
Cohen, A. Strehlow, R. A. Wires, R. A.	Exothermic Reactions Behind Reflected Shocks	1	125
Comyn, R. H.	See Apstein, M.	1	135
Curchack, H. D. Tevelow, F. L.	Microwave Diagnostics in the Shock Tube	2	349
Davidson, W. Henry, M. C. Krebs, A. Noltes, J. G.	Organometallic Research Using Group IV Elements	1	309

AUTHOR	TITLE	VOL	PAGE
Davis, Capt T. E.	Peritoneal Dialysis in the Treatment of Renal Failure: I. The Postoperative and Post-Traumatic Group of Patients	1	153
Davis, T. R. A. Joy, Maj R. J. T.	Cold Injury and Cold Acclimatization in Man	1	173
Debey, A. L. Richard, V. W.	The DOPLOC Dark Satellite Tracking System	1	199
Dunn, R. L.	Polarization Sensitivity as a Discriminant for Classes of Radar Targets	3	91
Dziemian, A. J. Sperrazza, J.	Incapacitating Potential of any Projectile as Based on Some Simple Procedures	3	367
Edmondson, Col C. C. Murdoch, Col F. J. Jr.	Field Experimentation as an Aid to Operations Research	2	151
Egan, C. W. Purtell, J. P.	Development of Repetitive Kiloton Thrusts of Milliseconds Duration for Simulating the Firing and Determining the Reliability of Heavy Weapon Components	1	211
Ehrlich, J. J.	See Barr, T. A. Jr.	1	221
Eigelsbach, H. T. McCrumb, F. R. Jr. Saslaw, S. Tigertt, Col W. D.	Live and Killed Tularemia Vaccine: Evaluation in Animals and Man	1	235
Fast, 1/Lt R. W.	Linear Accelerator Research	3	105
Felkner, Sp4 I. C.	DNA Isolation by an Improved Procedure for Transformation of Bacillus SP	3	111
Fishbein, W. Rittenbach, O. E.	Combat Surveillance Correlation Radar	3	119

AUTHOR	TITLE	VOL	PAGE
Fromm, B. W. Kramer, D. N. Poziomek, E. J.	Chemical Approaches to the Development of a Personal Expendable Detector for Chemical Agents	3	253
Geever, E. F.	See Berard, Capt C. W.	2	251
Gericke, O. R.	Determination of the Geometry of Hidden Defects by Ultrasonic Pulse Analysis Testing	1	249
Giardini, A. A. Kohn, J. A.	Ultra-Pressure High Temperature Research: A Frontier in the Quest for New Electronic Materials	1	259
Gisser, H. Portnoy, S.	Partially Fluorinated Aliphatic Compounds by Reductive Desulfurization of Substituted Thiophene	2	205
Godwin, 2/Lt R. O.	See Benson, R. C.	3	51
Goldberg, B.	Applications of Statistical Communication Theory	3	133
Golub, A.	User-Scientist Team Selects Optimum Weapon Systems	3	149
Goodman, Sp5 E. P. Ketcham, Capt J. S.	Human Responses to Four Anticholinergic Delirients	3	173
Goubau, G.	See Christian, J. R.	1	273
Groff, W. Sternberger, L. A. Sim, V. M.	Anticholinesterases (Nerve Agents): A Method for Continuous Human Bioassay of Agent Effect and Therapeutic Effects of Oximes	3	335
Haines, B. W. Hodges, 1/Lt D. R. Jones, W. I. Klein, F. Lincoln, R. E. Mahlandt, B. G.	Evidence that Rhesus Monkeys Die from Anthrax Toxin	1	367
Hasselstrom, T.	See Bailey, S. D.	1	285

AUTHOR	TITLE	VOL	PAGE
Hatch, H. P.	See Abbe, E. H.	1	393
Hatkin, L.	Fast Wave Radiators	1	295
Henry, M. C.	See Davidson, W.	1	309
Hodges, 1/Lt D.	See Haines, B. W.	1	367
Hoffman, A. J.	See Armendt, B. F.	3	159
Hoffman, F. A. Merola, A. Miles, T. D.	Replacing THPC in the APO-THPC Flame-Resistant Finish for Textiles	2	103
Howard, B. A. Jr.	Behavior of Non-Newtonian Fluids Under Conditions of High Pressure, Rapid Acceleration and High Velocity	1	317
Howard, C. D.	See Ayers, O. E.	3	37
	See Allen, H. C.	3	29
Huskins, C. W.	See Ayers, O. E.	3	37
	See Allen, H. C.	3	29
Ingram, L. F.	Air Blast Phenomena in an Arctic Climate	1	330
Jacobs, H.	See Brand, F. A.	1	339
Johnson, E. R.	See Allen, A. O.	1	103
Jones, W. I.	See Haines, B. W.	1	367
Joy, Maj R. J. T.	See Davis, T. R. A.	1	173
Kaufman, F. Kelso, J. R.	Laboratory Studies of Upper Atmosphere Chemical Reactions	1	355
Kaufman, J. V. R. Levy, P. W. Mapes, M. E.	Effects of Gamma-Ray Irradiation on Five Plastic Bonded HMX and DATB Compositions	3	213
Kaufman, J. V. R.	See Abel, J. E.	2	223
Kelso, J. R.	See Kaufman, F.	1	355
Ketcham, Capt J. S.	See Goodman, Sp5 E. P.	3	173

AUTHOR	TITLE	VOL	PAGE
Klein, F.	See Haines, B. W.	1	367
Kohn, J. A.	See Giardini, A. A.	1	259
Kolb, C. R.	Classification of Landscape Geometry for Military Purposes	1	379
Korytoski, R. D.	See Abbe, E. H.	1	393
Kramer, D. N.	See Fromm, B. W.	3	253
Krebs, A.	See Davidson, W.	1	309
Lamanna, C.	Problems and Progress in the Study of Oral Toxicity of Bacterial Toxins	2	1
Larson, F. R.	Static Fatigue Fracturing of High Pressure Gas Vessels	2	17
Leonard, F. Nelson, J. Nielson, C. A.	Materials for Biomechanical Applications	2	33
Levenson, S. M.	See Berard, Capt C. W.	2	251
Levine, S. Webb, H. C. Sr.	The Cover System	3	191
Levy, P. W.	See Kaufman, J. V. R.	3	213
Lincoln, R. E.	See Haines, B. W.	1	367
Liss, F. T. Salerno, J. Weiss, F.	The Krait Fuze System	3	305
Long, L. Jr.	See Ball, D. H.	2	47
Lowenthal, M. J.	Winds and Atmospheric Turbulence at Very High Altitudes	2	55

AUTHOR	TITLE	VOL	PAGE
McAfee, W. S.	The Attenuation of Radio-waves Propagated Through the Region of a High Altitude Nuclear Explosion	3	225
McClaran, O. B. Treglown, E. J. Shook, T. E.	Some Problems Associated with the Development of a Process for a Biological Agent	3	325
McClurg, G. H.	A Basis for the Long Range Planning of Army Research	2	67
McCrumb, F. R. Jr.	See Eigelsbach, H. T.	1	235
Macchia, R. P.	Radar Mapping Systems	3	203
Mahlandt, B. G.	See Haines, B. W.	1	367
Mapes, J. E.	See Kaufman, J. V. R.	3	213
Meindl, J. D.	See Brand, F. A.	1	339
Mendelson, Maj J. A.	Effect of Mafenide (SulfamylonR) and Penicillin in Topical Therapy of C.Perfringens-Contaminated Experimental Wounds	2	77
Merola, A.	See Hoffman, F. A.	2	103
Merrill, H. J.	Laser Progress & Applications	2	89
Metts, H.	See Boatright, A.	1	31
Miles, T. D.	See Hoffman, F. A.	2	103
Mirarchi, M. R.	See Benson, R. C.	3	51
Mitra, A. K.	See Ball, D. H.	2	47
Molella, D. J.	Evaluation of Hypersonic Probe Materials in a Water-Stabilized Electric Arc at Temperatures to 3200° (5795°F) for 60 Seconds	2	115

AUTHOR	TITLE	VOL	PAGE
Mon, G.	See Barila, Lt Col T.	2	423
Moore, L. F.	Studies of Rifle Effectiveness	3	233
Mow, C. C. Sadowsky, M. A.	Determination of the Critical Torque Inducing Buckling in a Twisted Spherical Shell Subject to Internal or External Pressure	2	129
Mueller, H. J. Singer, G. D. Tiller, H. D.	Processes of Photon and Exo-Electron Emission from Alkali Azides	2	141
Murdoch, Col F. J. Jr.	See Edmondson, Col C. C.	2	151
Neblett, M. J.	See Charlton, H.	1	115
Nelson, J.	See Leonard, F.	2	33
Nielson, C. A.	See Leonard, F.	2	33
Noltes, J. G.	See Davidson, W.	1	309
Nunn, Capt D.	See Barila, Lt Col T.	2	423
Ohmstede, W. D.	Solution of Non-Steady Soil Moisture Transfer	2	159
Palmer, J. S.	The Ecological and Epidemiological Aspects of Biological Field-Testing at Dugway Proving Ground	3	247
Papirmeister, B.	Effects of Mustard Gas* on T2 Bacteriophage and DNA* Synthesis	2	171
Pearce, C.	Effect of Ionizing Radiation on Pyridine	2	191
Perkins, J. F.	See Barr, T. A. Jr.	1	221
Portnoy, S.	See Gisser, H.	2	205
Poziomek, E. J.	See Fromm, B. W.	3	253

AUTHOR	TITLE	VOL	PAGE
Pradko, F.	Dynamic Simulator	2	209
Purtell, J. P.	See Egan, C. W.	1	211
Ramm, W. J.	Magnetic & Telluric Detection of High-Altitude Nuclear Explosions	3	263
Ravitsky, C.	An Infrared Antimissile Feasibility Study	3	277
Reese, E. T.	See Bailey, S. D.	1	285
Reitzner, B.	See Abel, J. E.	2	223
Richard, V. W.	See Debey, A. L.	1	199
Rinker, J. N.	Aerial Detection of Snow Surface and Undersnow Targets	3	291
Rittenbach, O. E.	See Fishbein, W.	3	119
Roberts, T. G.	The Significance of Spectroscopically Determined Plasma Jet Temperatures	2	235
Rosen, H.	See Berard, Capt C. W.	2	251
Sadacca, R.	Techniques for Optimizing Image Interpreter Performance	2	265
Sadowsky, M. A.	See Mow, C. C.	2	129
Salerno, J.	See Liss, F. T.	3	305
Saslaw, S.	See Eigelsbach, H. T.	1	235
Sauberlich, H. E.	See Baker, Maj E. M.	1	21
Savage, R. O. Jr. Tauber, A.	Growth of Hexagonal Ferrite Single Crystals for Application at Microwave and Millimeter-Wave Frequencies	2	333
Schmidt, S. J. Waite, A. H.	Exploring the Antarctic with Ice-Sounding Radar	2	383

AUTHOR	TITLE	VOL	PAGE
Schmidt, Th. W.	The Error Field Associated with Instrumentation for Position Determination	2	279
Sharp, C. E.	Artificial Ionospheres	3	317
Shook, T. E.	See McClaran, O. B.	3	325
Sim, V. M.	See Groff, W.	3	335
Singer, G. D.	See Mueller, H. J.	2	141
Smith, M. G.	Helicopter Armament	3	353
Spears, O. S.	A Model for Evaluating Surface-to-Surface Firepower	2	291
Sperrazza, J.	See Dziemian, A. J.	3	367
Steinbach, K. H. Varnum, F. B.	Recognition of Subsoil Targets by Means of Microwaves	3	377
Sternberger, L. A.	See Groff, W.	3	335
Strange, J. N.	The Attenuating Effects of a Bubble Screen on Underwater Shock	3	391
Straub, H. H.	See Barila, Lt Col T.	2	423
Straub, H. W.	Effects of Air Turbulence Upon Propagation of Light	2	309
Strehlow, R. A.	See Cohen, A.	1	125
Sulzberger, M. B.	Progress and Prospects in Idiophylaxis (Built-In Individual Self-Protection of the Combat Soldier)	2	315
Tauber, A.	See Savage, R. O. Jr.	2	333
Taylor, G. R.	Nuclear Altimeter	3	409
Tevelow, F. L.	See Curchack, H. D.	2	349

AUTHOR	TITLE	VOL	PAGE
Thiele, O. W.	Mesospheric Density Variability Based on Recent Meteorological Rocket Measurements	2	363
Tigertt, Col W. D.	See Eigelsbach, H. T.	1	235
Tiller, H. D.	See Mueller, H. J.	2	141
Treglown, E. J.	See McClaran, O. B.	3	325
Varnum, F. B.	See Steinbach, K. H.	3	377
Wachter, R. F.	Purification and Concentration Methods in the Development of Viral and Rickettsial Agents	3	419
Waite, A. H.	See Schmidt, S. J.	2	383
Wallace, W. T.	See Baker, Maj E. M.	1	21
Walter, E. L.	A Mathematical Model for a Ballistic Rocket	2	401
Webb, H. C. Sr.	See Levine, S.	3	191
Weiss, F.	See Liss, F. T.	3	305
Weitz, S.	See Brand, F. A.	1	339
Wires, R. A.	See Cohen, A.	1	125
Wogsland, N. C.	An Apparatur for Measuring the Bulk Modulus of Solid Propellants	2	411
Wolfskill, S. J.	See Baker, Maj E. M.	1	21
Woodward, K. E.	See Barila, Lt Col T.	2	423
Young, D. R.	Effect of Food Intake on Performance Capabilities	2	437

THE ROLE OF ELECTRON CONFIGURATION ON PROPERTIES IN DILUTE SOLID SOLUTION ALLOYS

E. P. ABRAHAMSON AND J. A. ALEXANDER
WATERTOWN ARSENAL LABORATORIES
WATERTOWN, MASSACHUSETTS

INTRODUCTION

In 1958 Abrahamson and Grant [1] reported a periodicity in the effect of dilute binary transition metal solute additions upon the brittle-to-ductile transition of chromium. The periodicity observed was not of the Mendeleev type but rather was based on the numbers of outer d and s electrons of the solute. A correlation was observed between the absolute value of the rate of change of transition temperature with solute addition and the number of ground state [2] outer d and s shell electrons of the solute atom. Dilute additions of elements with the same outer electron configuration had the same effect upon the properties of the base metal. It was inferred that this correlation resulted from some proportionality between the outer electronic structure of a solute atom and its contribution to bonding strength. From this work it was deduced that similar correlations should be found upon the systematic study of other parameters sensitive to atomic bond strength.

OBSERVATIONS

Recently, Abrahamson et al [3-9] have completed a series of systematic studies on the effect of dilute alloying additions upon the "one hour start-recrystallization temperature." The transition metals Fe, Cb, V, Ti, Zr and Ni were used as solvents. Most other transition metals were used as binary solid solution solutes in concentrations well within the region of solid solubility [10].

Consistently, the initial addition of solute linearly changed the recrystallization temperature. Figure 1 is a schematic of the forms of curves which have been obtained. As the concentration increases, the initial linear portion of the curve is terminated by a break in the curve generally to a lesser slope.

The effect of varying the recrystallization annealing time upon the form of the temperature-composition curve is shown in Figure 2. The initial slope is invariant while substantial change occurs in the slope beyond the break. The break itself appears to remain constant for the longer annealing times; apparent shift at the shortest annealing time may be attributed to the increased effect of the time required to attain the annealing temperature.

The most important observation is that the absolute slope of the initial linear portion of the curves has been shown to correlate with the free atom ground state outer electron configuration of the solute. This is typified by the results for vanadium as a solvent, which is shown in Figure 3, in a manner analogous to that reported by Abrahamson and Grant [1]. Figure 4 presents the same data plotted versus the Goldschmidt atomic radii, demonstrating the insufficiency of size factor considerations in explaining these effects. The absolute slope is used in Figure 3 since decreases in recrystallization temperature were observed for dilute additions to three base metals. Negative deviations in recrystallization temperatures were also noted by Semchysen with molybdenum as the solvent [11]. These systems are summarized in Figure 5. The break concentration, or limit of initial linearity, has also been found to correlate with the number of outer d electrons of the solute, Figure 6. The correlation curves are in the form of parallel sets of V's, all the elements on any one V having a specific number of s electrons.

In addition to the transition temperature and recrystallization temperature work of Abrahamson et al, the rate of grain growth, self diffusion coefficient, coefficient of thermal expansion and the lattice parameter are also grossly affected by dilute solute additions. Systematic studies on these parameters are still forthcoming and the universality of the electron configuration correlation is still subject to confirmation. However, it is possible at this time to propose some new ideas regarding these effects, and to question existing postulations.

DISCUSSION

The effect of dilute alloying elements upon the properties of pure metals has been shown to be extraordinary. Both solute impurity scavenging and segregation to grain boundaries or other imperfections have been postulated to explain this effect.

For this presentation the position will be taken that the cause of these dilute alloying effects is more homogeneous in nature and should be attributed to bulk solute-solvent interactions. Particularly important is the linearity of the initial portion of the property versus composition curves and the lattice parameter work of Cotta and Gazzara [12].

The linear portion of the property versus composition curve indicates that the addition of each solute atom accomplishes a constant change in the property being investigated. Beyond the break, added solute atoms no longer contribute to the same degree. Both the slope and the break concentration have been observed to correlate with the number of outer d shell electrons of the solute. These correlations have been observed for two material properties and in seven base metal binary transition metal systems.

On the basis that these correlations exist, their apparent generality and the form of the property versus composition curves, it is concluded that the initial linearity is the result of the solute-solvent atom interactions. The limit of linearity then represents the concentration at which solute atoms are unable, for any of several reasons, to contribute to the constant change in property required for extending the linear portions of the curves.

For the recrystallization process there have been two postulations with regard to the retardation of grain boundary migration. Lücke and Detert [13], have presented a theory which attributes retardation to impurity atom segregation at the moving boundary. Criani [14] defines a different kind of boundary, but otherwise is in general agreement with Lücke and Detert. Leslie et al [15] dispute these postulations and instead contend that grain boundary retardation is the result of solute segregation to "cell" boundaries in the deformed matrix ahead of the moving boundary.

Recent precision lattice parameter measurements by Cotta and Gazzara [12] on dilute iron-ruthenium alloys raise considerable doubt as to the validity of such segregation mechanisms. Cotta and Gazzara have reported a break in the curve of lattice parameter versus atomic percent at 0.17 atomic percent ruthenium. Figure 7 is a comparison of the lattice parameter and recrystallization temperature results for dilute iron-ruthenium alloys and is presented to demonstrate excellent agreement as to the location of the break.

It is observed that alloying element atoms initially accomplish a given amount of expansion of the lattice; subsequently a concentration is reached at which only a lesser degree of lattice expansion is accomplished. This observation would seem to rule out segregation to grain boundaries, since more economical accomodation of impurity atoms is considered to be the basis for such segregation.

Similarly, solute-impurity scavenging appears to offer little help in explaining such dilute alloying effects. The linear nature of the initial portion of the curve would indicate that only one species of impurity was being scavenged and the break would represent the completion of that scavenging.

The results for Ti and Zr base alloys, Figures 8 and 9, show that for two base metal systems with different residual impurity contents, the slope and the break composition are almost identical for the same added solutes. Further, the empirical relationships presented in a previous paper [5] for describing the correlation curves could scarcely be expected to hold if scavenging were the cause of this dilute alloying effect. The residual impurities in the various base metals were different in kind and in amount. The existence of empirical relationships between the changes in properties and the atomic configuration of the solutes and solvents would seem to be the strongest evidence for the conclusion that these are solid solution effects.

The interpretation of these data is that in the range of compositions studied, atoms are being added primarily to the bulk of each metal grain. The sites in grain boundaries which will accomodate impurity atoms more economically are essentially exhausted at a very low concentration. Thereafter the large majority of atoms will be accomodated in the grain bulk.

The rapid changes in properties are postulated to occur as added solute atoms influence isolated volumes of matrix atoms. This rapid change occurs only as long as each additional solute atom can exert its influence over isolated volumes of matrix atoms. Finally a concentration is reached at which the effect of each additional solute atom can no longer influence its characteristic volume of the matrix metal. Thereafter, additional alloying causes either interaction of the perturbed volumes or solute-solute clustering; both are generally accompanied by a reduction in the magnitude of the effect of the alloying element upon the property involved.

In this development the magnitude of the effect on properties is dependent upon the strength of the solute-solvent bonding; and the break concentration is dependent on the extent of the perturbation caused. These dependencies are naturally related and both have been found to correlate with the number of d shell electrons of the solute.

The nature of the configuration curves appears to be a function of the base metal being studied. There is a regularity to the apex position such that the d electron configuration at the apex is described by the following equation:

$$d_B = d_A + 1 + (s_A - 2)(s_B - 2) \qquad \text{(Equation 1)}$$

where d_B = solute d shell configuration at the apex
d_A = number of d shell electrons of the solvent
s_A = number of s shell electrons of the solvent
s_B = number of s shell electrons of the solute
all values being integral numbers equal to the free atom ground state configuration.

Much could be said in descriptive terms regarding the correlation curves in Figures 8 and 9. Rather than to elaborate on such observations, only those worthy of further study and consideration will be enumerated:

1. The negative changes in recrystallization temperature.
2. The changes in the form of the correlations, i.e., inverse to upright V's with increasing numbers of d electrons in the solvent.
3. The consistent displacement of the apex of the V in accord with Equation 1, i.e. dependent upon both the number of solvent d electrons and the number of s electrons in the solute and solvent.

The determination of lattice parameters in dilute alloy systems will continue in an effort to achieve better understanding of the electron configuration correlations through the use of this more readily interpretable parameter.

Obviously the electron configuration of an atom in the solid state is not that of the free atom; however, the fact that various properties consistently plot in accordance with the free atom configuration indicates that their state in dilute solution is, in some manner, proportional to the ground state configuration. It should be expected that any property sensitive to electron configuration would, in this dilute alloying range, be found to correlate in a similar fashion.

While it is obvious that there are many unanswered questions regarding various facets of the subject of dilute alloying effects, it is now obvious that the electronic structures of solutes and solvents are most important. Many postulations have been made regarding atomic diameter correlations and much thinking has been devoted to the generation of size effect arguments. It has been shown that the simpler concept of number of ground state outer electrons provides a much more reliable basis for correlation and, it is hoped, that this presentation will generate thinking in these terms.

CONCLUSIONS

1. Consideration of the electron configuration of solutes and solvents is important to the understanding of dilute solid solution phenomena.

2. Initial linear changes in recrystallization temperature, brittle-to-ductile transition temperature and lattice parameter are the result of solute-solvent interactions in the bulk of the grains. Such effects are the result of individual solute atoms influencing a characteristic volume of solvent.

3. The break in the property versus composition curve results from a saturation characterized by the inability of each subsequent solute atom to influence its characteristic volume.

REFERENCES

1. E. P. Abranamson, II and N. J. Grant; Trans ASM, 1958, Vol. 50, pp. 705-721.

2. C. Kittel, Introduction of Solid State Physics; John Wiley and Son, Inc., New York, 1956, p. 66.

3. E. P. Abrahamson, II and B. S. Blakeney, Jr., Trans Met Soc, AIME, 1960, Vol. 218, pp. 1101-1104.

4. E. P. Abrahamson, II, The Recrystallization of Iron as a Function of Binary Solute Addition and Time, WAL TR 830.3/2, January 1961.

5. E. P. Abrahamson, II, Trans Met Soc AIME, 1961, Vol. 221, pp. 1193-1196.

6. E. P. Abrahamson, II, Trans Met Soc AIME, Vol. 221,pp. 1196-1198.

7. E. P. Abrahamson, II, Trans Met Soc AIME, Vol. 221, pp. 1199-1201.

8. E. P. Abrahamson, II, The Role of Dilute Binary Transition Element Additions on the Recrystallization of Titanium, WAL TR 830.3/6, July 1961, to be published, Trans Met Soc AIME.

9. E. P. Abrahamson, II, The Role of Dilute Binary Transition Element Additions on the Recrystallization of Nickel, WAL TR 830.3/7, November 1961, to be published Trans Met Soc AIME.

10. M. Hansen, Constitution of Binary Alloys, McGraw-Hill Book Co., New York 1958.

11. M. Semchysen, The Metal Molybdenum, ASM, Cleveland, 1958, pp. 281-329.

12. L. J. Cotta and C. P. Gazza a, The Lattice Parameters of Iron-Ruthenium Solid Solution Alloys, Proceedings, Denver Research Institute X-Ray Diffraction Conference, 1961, to be published.

13. K. Lücke and K. Detert, Acta Met 1957, Vol. 5, pp. 628-637.

14. R. A. Oriani, Acta Met, 1959, Vol. 7, pp. 62-63.

15. W. C. Leslie, J. T. Michelak and F. W. Aul, The Annealing of Gold Worked Iron, to be published Trans Met Soc AIME.

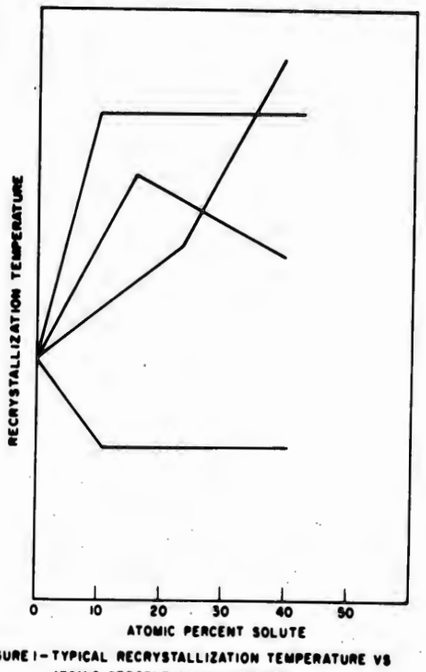


FIGURE 1 – TYPICAL RECRYSTALLIZATION TEMPERATURE VS ATOMIC PERCENT SOLUTE CURVES

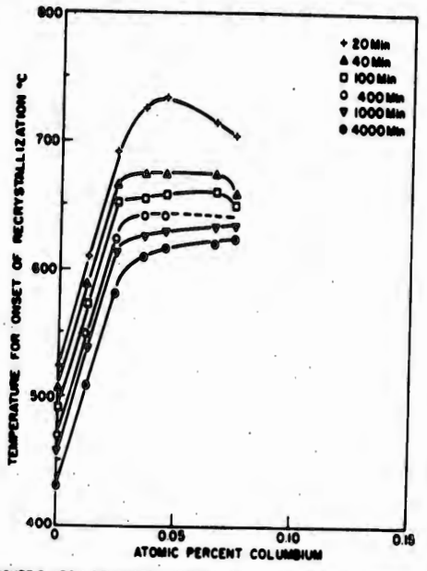


FIGURE 2 – IRON RECRYSTALLIZATION TEMPERATURE AS AFFECTED BY COLUMBIUM FOR A VARIETY OF TIMES, COLD WORK CONSTANT AT 75% AFTER ABRAHAMSON (4)

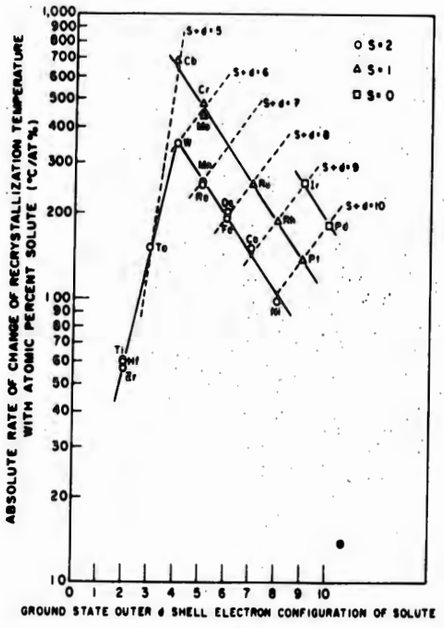


FIGURE 3 – INITIAL RATE OF CHANGE OF RECRYSTALLIZATION TEMPERATURE WITH ATOMIC PERCENT VERSUS NUMBER OF OUTER d SHELL ELECTRONS OF ADDED ELEMENT AFTER ABRAHAMSON (5)

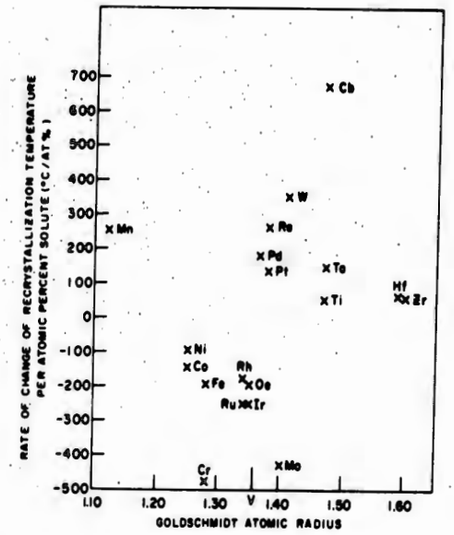


FIGURE 4 – COMPARISON OF RATE OF CHANGE OF RECRYSTALLIZATION TEMPERATURE PER ATOMIC PERCENT SOLUTE VERSUS SOLUTE ATOMIC RADIUS.

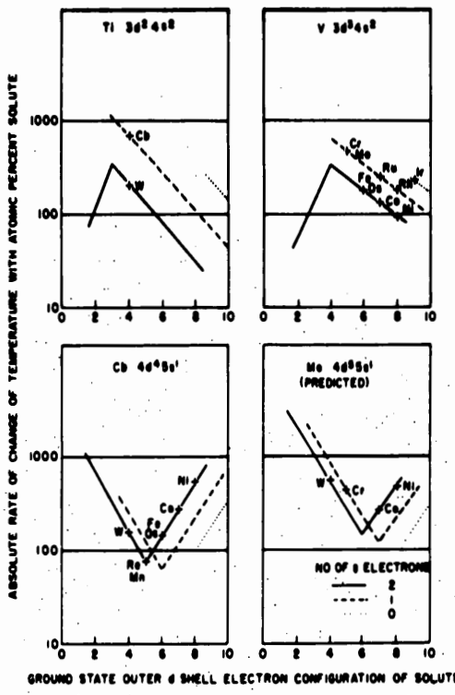


FIGURE 5-ELECTRON CONFIGURATION CORRELATIONS SHOWING THOSE ELEMENTS WHICH CAUSE NEGATIVE DEVIATIONS IN RECRYSTALLIZATION TEMPERATURE

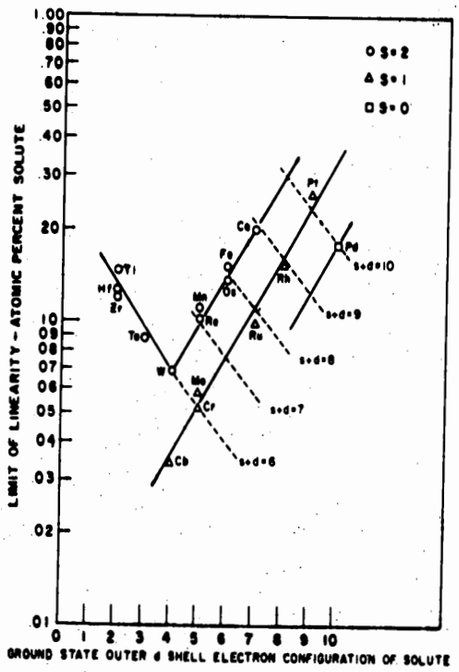


FIGURE 6 – LIMIT OF LINEARITY IN ATOMIC PERCENT SOLUTE VERSUS NUMBER OF OUTER d SHELL ELECTRONS OF ADDED ELEMENT AFTER ABRAHAMSON (5)

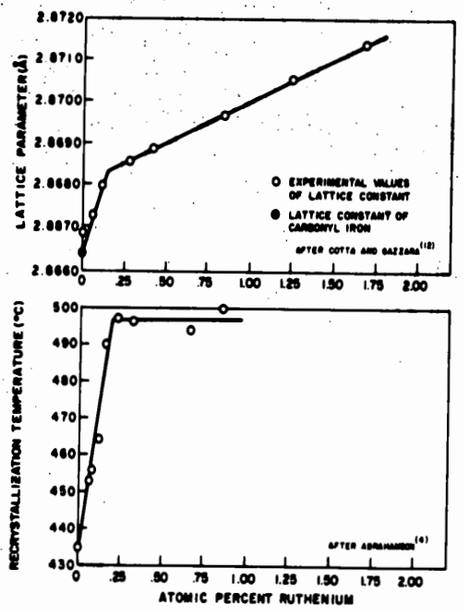


FIGURE 7-THE ROLE OF RUTHENIUM ON THE RECRYSTALLIZATION TEMPERATURE AND LATTICE PARAMETER OF IRON.

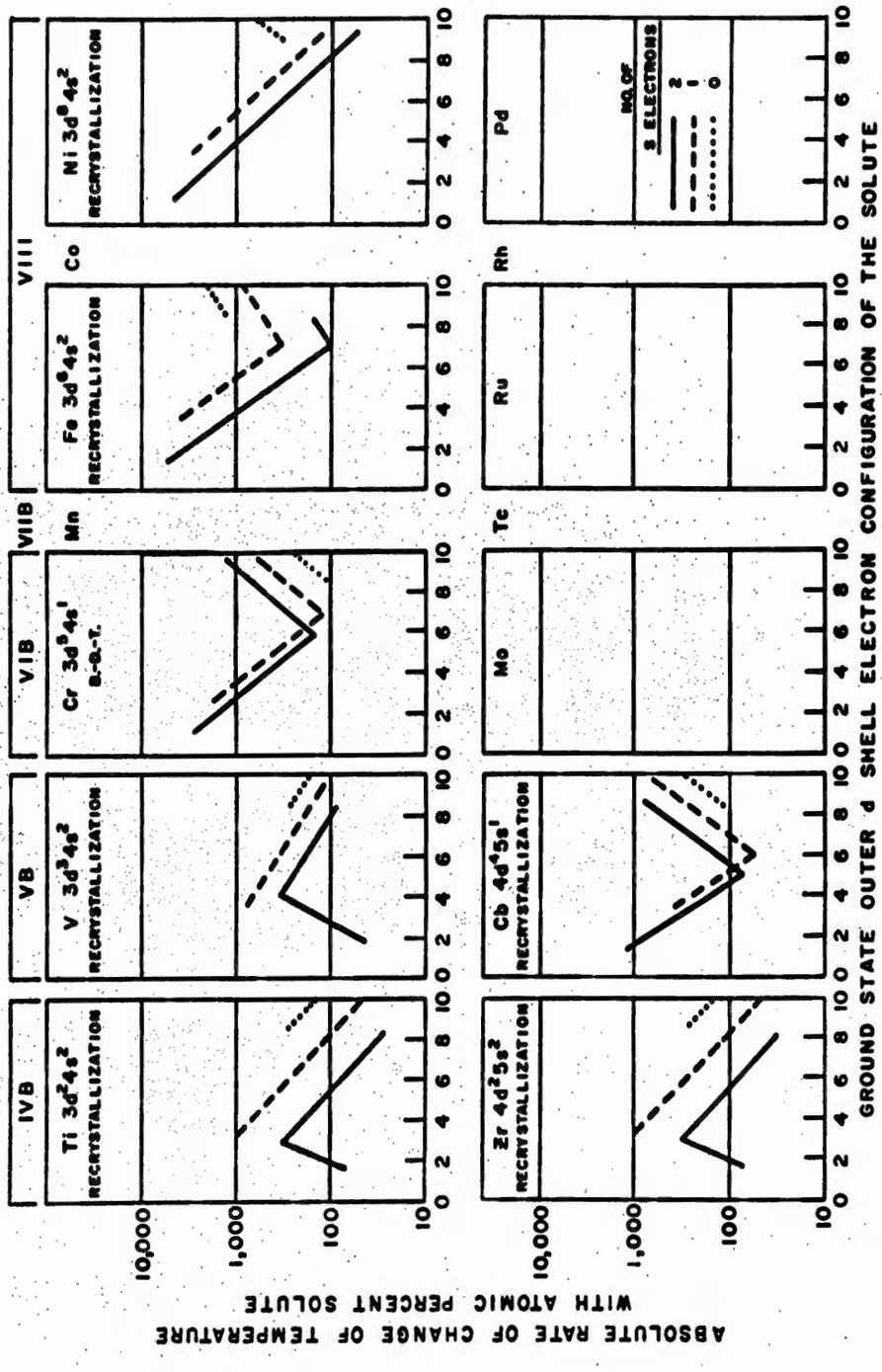


FIGURE 8-THE ROLE OF OUTER ELECTRON CONFIGURATION ON SEVERAL PROPERTIES OF DILUTE BINARY TRANSITION ALLOYS.

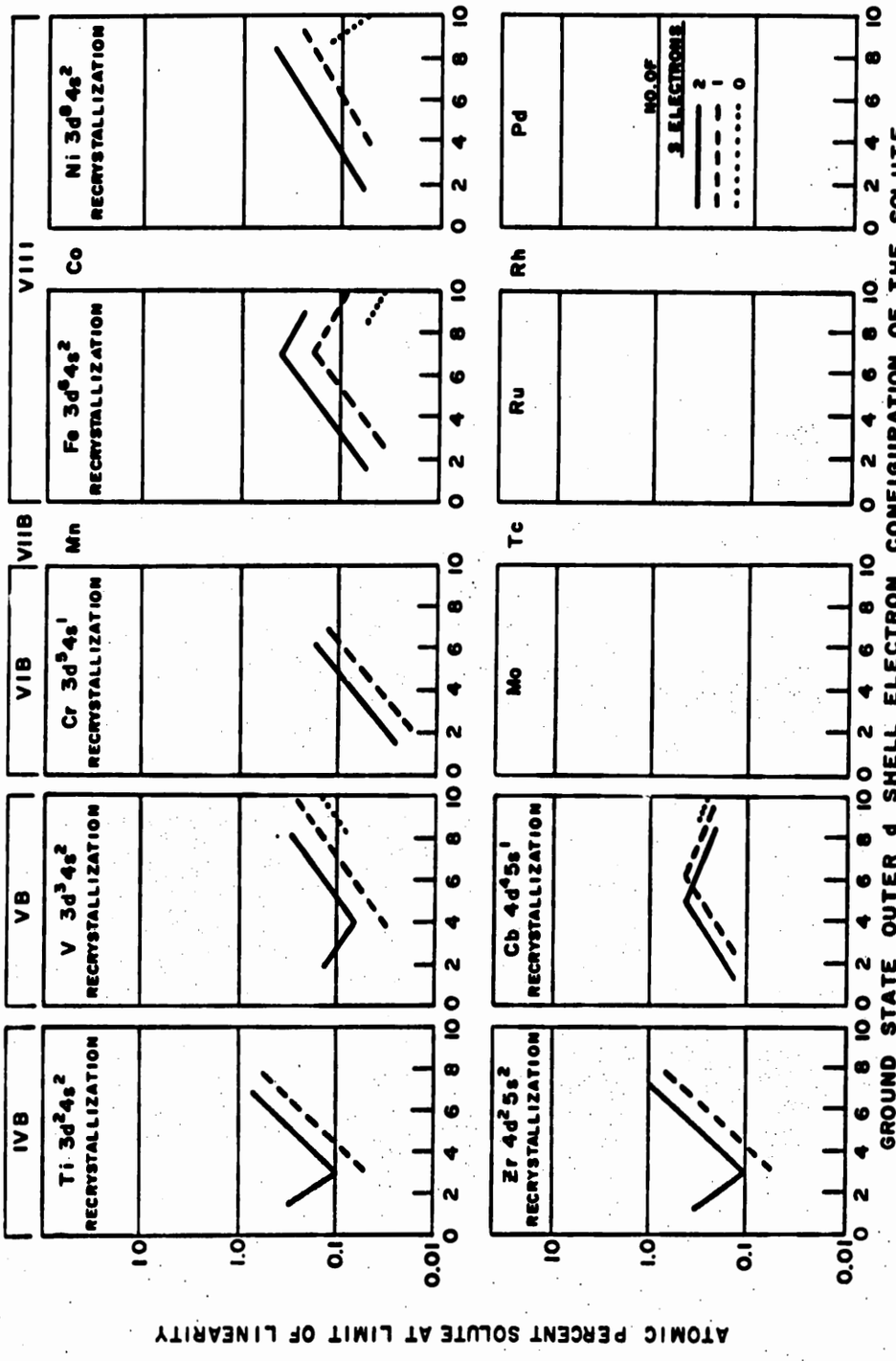


FIGURE 9 - THE ROLE OF OUTER ELECTRON CONFIGURATION ON THE LIMIT OF LINEARITY.

Surfacing Submarines Through Ice

A. Assur
USA, Cold Regions Research and Engineering Laboratory
Hanover, N. H.

The operation of submarines under ice and penetration through ice sheets have become a reality, although the process of breaking through ice was largely based upon trial and error.

This paper limits itself to a theoretical treatment of the problem which, combined with the tests already available, should serve as a basis for further experimentation and for the extension of operational capabilities beyond existing experience.

Two groups of criteria are derived below. One is for the condition when a submarine cannot possibly surface. This is when the buoyancy force is just sufficient to produce a first crack in the ice sheet. The other is the condition when complete break-through with a collapse of the ice sheet can be accomplished.

Another consideration is whether the submarine applies a force "topping" at rest against the ice or whether it approaches the ice sheet with a vertical velocity. In the first case the force required under static loading is computed, in the second case the energy required to penetrate under impact load. Rather quick loading in either case is necessary to maintain the stability of the submarine. Once the sail has penetrated, slow loading can be applied for complete surfacing. In this case a plastic failure of the ice sheet has to be considered.

One of the biggest problems in the mechanics of sea ice is the vertical variation in stress-strain relations. Corresponding equations based upon recent advances in the physics of sea ice conclude the paper.

Assur

The behavior of an ice sheet under cylindrical bending (line and strip loads) is governed by the differential eq. $d^4w/dx^4 = -4w/l_1^4$ (1) with the general solution

$$w = e^{x}(C_1 \cos x + C_2 \sin x) + e^{-x}(C_3 \cos x + C_4 \sin x) \quad (2)$$

Under circular bending the differential eq.

$$\frac{d^4w}{dr^4} + \frac{2}{r}\frac{d^3w}{dr^3} - \frac{1}{r^2}\frac{d^2w}{dr^2} + \frac{1}{r^3}\frac{dw}{dr} = -\frac{w}{l^4} \quad (3)$$

with the general solution in terms of Kelvin functions

$$w = C_1 \operatorname{ber}\rho_1 + C_2 \operatorname{bei}\rho_1 + C_3 \operatorname{ker}\rho_1 + C_4 \operatorname{kei}\rho_1 \quad (4)$$

has to be used. w - deflection, x and r distance from load, $\chi = x/l$, $\rho_1 = r/l$. $l_1 = \sqrt[4]{4D/\kappa} = \sqrt{2}\,l$ - the "action width", $l = \sqrt[4]{D/\kappa}$ - the "action radius", $D = \frac{Eh^3}{12(1-\mu^2)}$ - the "flexural rigidity", h - ice thickness, E and μ Young's modulus and Poisson's ratio for ice. κ - the "foundation modulus".

Both solutions have to be used in case of a submarine since we deal with a long but finite load. Even the sail is a long, roughly elliptical load.

The derivation of eqs (1) and (3) assumes that the reaction is proportional to the deflection of the ice sheet. κ is equal to the density of water if a load pushes from above, but it has the same value if a load pushes from below, since water together with the ice has to be lifted; otherwise, a vacuum would occur. In case of a cracked ice sheet, eqs. (1) and (3) hold, with κ equal to the density of ice until the ice sheet is lifted out of the water, but then complete collapse has occurred anyhow.

It is important whether we consider the load as flexible, readily following the shape of the deformed ice sheet, or as rigid.

In the case of a flexible, elliptical load, the author gives

$$\frac{M}{p} = \frac{(1+\mu)}{8}\frac{1}{\alpha}\frac{d\psi}{d\alpha} + \frac{1}{8\pi}\frac{(1-\varepsilon)}{(1+\varepsilon)}(1-\mu). \quad (5)$$

P - load producing the moment M leading to first cracking perpendicular to the long axis of the load (sail in case of submarines). $a = \alpha l$ "equivalent radius" of the ellipse, ε - ratio of short to long diameter.

$\psi = -\frac{4}{\pi} \kappa ei\rho_1$ (Ref 1).

We can consider the submarine also as a flexible line load.

$$\frac{M}{P} = -\frac{1}{8\lambda}\left[\sqrt{2} + \int_{-\infty}^{\lambda}\left(\mu \frac{d^2\psi}{d\rho_1^2} + \frac{1}{\rho_1}\frac{d\psi}{d\rho_1}\right)d\rho_1\right] \qquad (6)\text{(Ref 2)}$$

$\lambda = L/l$ - the relative length of the submarine in this case.

Considering the relative half-width of the submarine $\alpha_1 = a/l$

$$\frac{M}{P} = -\frac{1}{8\lambda}\left[2\sqrt{2}\,\frac{\sin\alpha_1}{\alpha_1 e^{-\alpha_1}} + \int_{-\infty}^{\lambda}\left(\mu \frac{d^2\psi}{d\rho_1^2} + \frac{1}{\rho}\frac{d\psi}{d\rho_1}\right)d\rho_1\right] \qquad (7)$$

This is still an equation for a flexible load leading to first cracking.

The other extreme is a rigid load which does not deform but acts as a stamp without producing a bending of the ice sheet along the axis of the sail or the submarine itself. This simplification allows us to obtain a series of solutions, the derivation of which otherwise would become unmanageable.

General case

Let us assume a finite rigid line load or a finite rigid strip load. It will produce cylindrical bending perpendicular to the load and semicircular deflection dishes at the ends with circular bending. Combined, these end dishes can be treated similarly to the hydrostatic reaction under a point load if we deal with a finite load or under a circular load of the radius in case of a strip load. The total relative length of a submarine is, therefore, $\lambda + 2\alpha$. Solution (4) leads to the center deflection $w(0) = P_0/8\kappa l^2$ (ref 1,2) with P_0 - the contribution of the end deflection dishes to the total reaction.

We define $R_0 = w/w(0)$ with w_0 corresponding to the equivalent center deflection of the end dishes depending upon loading geometry and state of the ice, even if cracked into wedges.

The deflection due to a finite line load of the intensity p is $w_1(0) = \frac{p}{2\sqrt{2}\kappa\ell} \cdot \frac{p-p_0}{2\sqrt{2}\kappa\ell^2\lambda}$. Again we define $R_1 = w_1/w_1(0)$ for the general case of a strip load. Our assumption of a rigid load leads to

$$w = w_0 = w_1 = \frac{P}{2\sqrt{2}\kappa\ell^2\lambda} R_1 - \frac{2\sqrt{2}}{\lambda} \frac{R_1}{R_0} w$$

or $w = \frac{P}{8\kappa\ell^2} \frac{2\sqrt{2}R_1}{\lambda+\varsigma}$ (8) with the abbreviation

$\varsigma = 2\sqrt{2}R_1/R_0$ or $\boxed{\omega = \frac{2\sqrt{2}R_1}{\lambda+\varsigma}}$ and $\frac{1}{\omega} = \frac{\lambda}{2\sqrt{2}R_1} + \frac{1}{R_0}$ (9)

for $\omega = w/w(0)$.

For a strip load, the moment developed under the load is $m_1 = p\ell_1 R_m/4$ where R_m gives the moment in the general case as a ratio to the moment $p\ell_1/4$ under a line load.

In this case we have

$$m_1 = \frac{PR_m}{2\sqrt{2}\lambda}\left(1 - \frac{\varsigma}{\lambda+\varsigma}\right)$$ and considering

$m_1 = \sigma h^2/6$ (σ flexural strength)

$$\boxed{P/h^2 = \frac{\sqrt{2}}{3} \frac{\lambda+\varsigma}{R_m} \sigma}$$ (10) for the "loading factor" in the elastic case. For plastic failure $m_1 = \sigma h^2/4$ and $P/h^2 = \frac{\sqrt{2}}{2}\left(\frac{\lambda+\varsigma}{R_m}\right)\sigma$ (11).

The maximum deflection at failure can be derived as $w_m = \frac{\sigma h^2}{6\kappa\ell^2} \frac{R_1}{R_m}$ (12). Since $P = \frac{8\kappa\ell^2(\lambda+\varsigma)}{2\sqrt{2}R_1} w$, the work (or kinetic energy) to achieve failure becomes

$$W = \int_0^{w_m} \frac{4\kappa\ell^2(\lambda+\varsigma)}{\sqrt{2}R_1} w\,dw = \frac{\sqrt{2}\kappa\ell^2(\lambda+\varsigma)}{R_1} \frac{\sigma^2 h^4}{36\kappa^2\ell^4} \frac{R_1^2}{R_m^2}$$

Substituting $\ell^2 = \ell_0^2 h^{3/2}$

$$\boxed{W = \frac{\sqrt{2}(\lambda+\varsigma)}{36\kappa\ell_0^2} \frac{R_1}{R_m^2} \sigma^2 h^{5/2}}$$ (13)

The kinetic energy required to break through by bouncing against the ice sheet increases rapidly with the ice thickness (h) and is proportional to the square of the flexural strength (σ). An operation which succeeded in the fall may fail, therefore, in the winter just because of the higher strength even with the same thickness. In the case of static loading the dependence is not so severe.

For plastic loading the factor 6 in eq (12) has to be changed to 4 and the factor 36 in eq (13) to 16.

Specific cases

For the initial longitudinal cracking under a finite line load the same eqs hold as in the general case since $R_0 = R_1 = R_m = 1$. It can be shown, however, that cracks at the end of the finite line load would originate first with resulting small wedges. Once small wedges form, the resulting stresses at the top of free wedges are

$$\sigma_0 = \frac{6\sqrt{2}}{(\lambda+2\sqrt{2})\pi} P/h^2 \qquad (14a)$$

or 27% higher than according to eq. (10). This theoretical conclusion, however, loses its significance in case of a submarine since considerable confined wedging of the ice should occur at the ends.

Complete collapse of the ice sheet occurs upon formation of longitudinal cracks at the distance $\frac{\sqrt{2}}{2}\pi\ell$ from the line load. In this case $R_0 = 1$ and $R_1 = 2$ since the load is distributed on two edges of the ice sheet already broken in the middle. $\varphi = 4\sqrt{2}$ and $\omega = 4\sqrt{2}/\lambda + 4\sqrt{2}$ satisfying the proper conditions for both $\lambda = 0$ and $\lambda = \infty$.

The maximum negative moment is

$$m_1 = -\frac{P\ell}{2} e^{-\pi/4} \qquad |R_m| = \sqrt{2} e^{-\pi/4}$$

and the loading factor $P/h^2 = \frac{1}{3} e^{\pi/4} (\lambda + 4\sqrt{2})\sigma$ (14b)

giving the solution for the ice thickness which can be penetrated.

The kinetic energy required for a complete collapse of the ice sheet can be computed from

$$W = \frac{\sqrt{2}\, e^{\pi/2}}{36 k \ell_0^2} (\lambda + 4\sqrt{2}) \sigma^2 h^{5/2} \qquad (14c)$$

The work required to break completely through is almost 10 times the work to achieve initial failure for a short load and 5 times that for very long line loads.

A number of approximations have been derived for finite strip loads so that the width of the submarine can be accounted for.

The exact solution considers a rectangular finite load of the width 2a with two rounded ends of the radius a.

On the basis of solution (4), it can be shown that

$$R_0 = \frac{\rho}{\pi\alpha^2}\left(1+\alpha\,\mathrm{ker}'\alpha\right) \qquad (14b).$$

On the basis of (ref 3), we derive

$$R_1 = \frac{1-B}{\alpha} \quad \text{with} \quad B = e^{-\alpha_1}\cos\alpha_1 \quad \text{and} \quad \alpha_1 = a/\ell_1 = a/\sqrt{2}\,\ell .$$

Therefore, $$\varsigma = \frac{\pi}{2\sqrt{2}}\frac{(1-B)\alpha}{(1+\alpha\,\mathrm{ker}'\alpha)} \qquad (14)$$ and $$\omega = \frac{2\sqrt{2}\,(1-B)}{\alpha\left(\lambda+\frac{\pi}{2\sqrt{2}}\frac{(1-B)\alpha}{1+\alpha\,\mathrm{ker}'\alpha}\right)} \qquad (15)$$

Also $$R_m = \frac{\sin\alpha_1}{\alpha_1 e^{\alpha_1}}$$ so that

$$P/\lambda^2 = \frac{\sqrt{2}}{3}\left(\lambda+\frac{\pi}{2\sqrt{2}}\frac{(1-B)\alpha}{(1+\alpha\,\mathrm{ker}'\alpha)}\right)\frac{\alpha_1 e^{\alpha_1}}{\sin\alpha_1}\,\sigma \qquad (16)$$

and $$W = \frac{\sigma\lambda^2}{6\kappa\ell^2}\frac{(1-B)}{\sqrt{2}}\frac{e^{\alpha_1}}{\sin\alpha_1} \qquad (17)$$ define the conditions under which a submarine cannot possibly break through.

More complicated is the final collapse of the ice sheet under the action of the submarine considered as a finite strip load.

For the deflection of the ice wedges which form at the ends of the load, we can base our derivations upon (ref 4). We obtain

$$R_0 = \frac{4\pi}{(\Gamma\frac{3}{4})^4}\left(1-\frac{\alpha}{2}\right) \qquad \Gamma = 1.2254 \qquad (18)$$

The factor $(1-\alpha/2)$ is an approximation showing the effect of the width of the submarine upon the deflection of the end dishes. The exact solution is a complicated matrix.

R_1 as above for a strip load leads to

$$\rho = \frac{(\Gamma\frac{3}{4})^4(1-B)}{\sqrt{2}\,\pi}\frac{}{\alpha(1-\alpha/2)} \quad \text{and} \quad \frac{1}{\omega} = \frac{\lambda\alpha}{2\sqrt{2}(1-B)} + \frac{(\Gamma\frac{3}{4})^4(1-\alpha/2)}{4\pi} \qquad (19)$$

Space does not permit reproducing the lengthy derivation for the predicted location $x = \chi_0 l_1$ of the final crack under collapse defined by

$$tg\chi_0 = \frac{\cos\alpha_1 (e^{2\alpha_1} - 1) + \sin\alpha_1 (e^{2\alpha_1} + 1)}{\cos\alpha_1 (e^{2\alpha_1} - 1) - \sin\alpha_1 (e^{2\alpha_1} - 3)} \tag{20}$$

as well as for the moment

$$|m_u| = \frac{p l_1}{4\alpha_1} \left\{ tg\chi_0 / \sqrt{2} \left[\cos\alpha_1 (e^{2\alpha_1} + 1) - 2\sin\alpha_1 \right] - \left[\sin\alpha_1 (e^{2\alpha_1} + 5) - 2\cos\alpha_1 \right] \right\} = \tag{21}$$

$$= \frac{p l_1}{4} R_m(\alpha)$$

The required loading against the ice sheet is

$$P/h^2 = \frac{\sqrt{2}}{3} \frac{(\lambda + \frac{(\Gamma\frac{3}{4})^4 (1-B)}{\sqrt{2}\pi\alpha (1 - \alpha/2)})}{R_m(\alpha)} \sigma \tag{22}$$

and the work

$$W = \frac{\sqrt{2}}{36 k l_0^2} \left[\delta + \frac{(\Gamma\frac{3}{4})^4 (1-B)}{\sqrt{2}\pi\alpha (1 - \alpha/2)} \frac{\sigma^2 h^{5/2}}{R_m^2(\alpha)} \right] \tag{23}$$

If the submarine can maintain its stability under slow loading, then it may be desirable to take advantage of plastic flow. No exact solutions exist. Approximate solutions in ref (5) have been criticized in ref (6). However, expedient temporary solutions can be given. For the case where the submarine cannot break through, we assume R_0 as in eq (14b) $R_1 = 2\sqrt{2}/3.3$

$$\delta = \frac{\pi\alpha^2}{3.3} (1 + \alpha \, ker'\alpha) \tag{24}$$

$$\frac{1}{\omega} = 0.4125\lambda + \frac{\pi\alpha^2}{8(1 + \alpha \, ker'\alpha)} \tag{25}$$

$R_m = \frac{2\sqrt{2}}{3.64} (1 - 0.405\alpha)$ so that the loading

$$P/h^2 = 0.910 \frac{[\lambda + \frac{\pi\alpha^2}{3.3} (1 + \alpha \, ker'\alpha)]}{(1 - 0.405\alpha)} \tag{26}$$

is not going to accomplish its purpose w/o initial failure.

More important is the complete plastic collapse of a broken ice sheet.

Assuming simply $R_0 = 1$, $R_1 = 2\sqrt{2}/3.3$

and deriving $R_h = \frac{2\sqrt{2}}{2.72}(1-0.682\alpha)$ from (ref 5), we have

$$\frac{1}{\omega} = 1 + 0.4125\lambda \tag{27}$$

and the loading factor

$$P/h^2 = 0.680\,\frac{\lambda + 2.43}{(1-0.682\alpha)}\,\sigma \tag{28}$$

will achieve complete break-through, with initial failure.

All these solutions are not readily usable so that nomograms are being prepared giving the required results in a minimum of time. For a given submarine a simple curve will produce the desired information.

Stress-strain relations

The application of plate mechanics to sea ice had been handicapped severely in the past by lack of information about strength and stress-strain relations as a function of temperature and salinity. A sea ice sheet does not resemble the classical homogeneous plate. It has a hard crust at the top and is almost mush near the bottom. Recent progress in the physics of sea ice has considerably clarified these problems. It is now possible, in principle, to predict the flexural strength (σ) in both directions of bending with tensile stresses developed in the top fibers as well as in the bottom layers, as well as the effective Young's modulus E, flexural rigidity D, action width l_1, and action radius l, all of which enter the theoretical relations given above. Temperature and snow conditions at the top (ref 7) have to be known or can be guessed based upon the season of the year as well as the salinity of the ice sheet, which can be estimated. (Ref 8) has shown how the tensile strength itself is related to brine volume, which is determined by temperature and salinity on the basis of phase relations and the geometry of brine pockets. Later work has refined and corrected these relations. The simplest possible relation is

$$\sigma_t = \sigma_0\left(1 - \frac{\nu^2}{\nu_0}\right) \tag{29}$$

$\nu = \frac{\gamma_i}{\gamma_b}\frac{S}{S_b}$ - the brine volume of sea ice. γ_i - theoretical density of sea ice, γ_b - density of brine, S - salinity of sea ice, S_b - the salinity of brine in phase

equilibrium with the temperature Θ (temperature, C, negative sign omitted)

$\frac{1}{S_t} = 1 - \frac{54.11}{\Theta}$ holds up to-7.1 C and only approximately below because of precipitation of salt. (S - absolute ratio).

Similar to (29) a simple relation $E = E_0 (1 - \frac{\nu}{\nu_0})$ can be assumed which holds very well according to recent experiments.

Abbreviating, we write $\nu = \frac{ca\,S}{\Theta}$

At some distance from the bottom the ice does not resist strain at the temperature $\Theta_0 = ca\,S/\nu_0$

Assuming a linear temperature profile with the gradient Θ'

$E/E_0 = 1 - \frac{1}{1 + \nu_0 \frac{z}{z_1}}$ (30) with $z_1 = ca\,S/\Theta'$.

The neutral axis shifts from its normal position in the middle of the plate (classical theory) towards the top. The condition

$$\int_0^h \left(1 - \frac{1}{1+\nu_0 z/z_1}\right) z\,dz \Big/ \int_0^h \left(1 - \frac{1}{1+\nu_0 z/z_1}\right) dz$$

gives the solution

$$\eta = \frac{1}{2\left(1 + \frac{\beta}{1-\beta} \ln \beta\right)} - \frac{\beta}{1-\beta} \qquad (31)$$

for η - the relative position of the neutral axis, counting from the bottom

$$\beta = \frac{\gamma_i}{\gamma_t} \frac{54.11}{1 - S_t} \frac{S}{\nu_0 \Theta_1} \qquad (32)$$

Θ_1 - temperature at the top of the ice sheet which can be estimated if air temperature and thickness of snow cover are given (ref. 7).

The reduction (G) in the moment of inertia because of decreasing Young's modulus with depth can be computed from a modified cross section, assuming a fictitious width proportional to E/E_0 with unity at the top and zero at the bottom. From the condition

$$G = \int_0^h \frac{E}{E_0} z^2 dz - (\eta h)^2 \int_0^h \frac{E}{E_0} dz$$

we obtain, after lengthy derivations,

$$G = \left(\frac{1}{3} - \eta^2\right) - \left[\frac{1}{2(1-\beta)^2} + \eta^2 \ln\beta\right]\frac{\beta}{1-\beta} + \left[\frac{2}{1-\beta}\right]\left(\frac{\beta}{1-\beta}\right)^2 - \left[\frac{3}{2} - \ln\beta\right]\left(\frac{\beta}{1-\beta}\right)^3 \qquad (33)$$

$G/1-\eta$ gives the reduction in section modulus, $\frac{\sigma_t G}{1-\eta}$ the reduction in flexural strength σ compared with the tensile strength σ_t at the top which in turn is strongly influenced by temperature and salinity (ref 8).

The derivation of the flexural strength with tension at the bottom is more complicated since partial failure or plastic yielding may occur in the weak bottom layer.

Practical use of these theoretical relations can be made upon preparation of nomograms.

References

(1) Assur, A. (1962), Traffic over frozen or crusted surfaces, 1st International Conference on the Mechanics of soil-vehicle systems, Torino

(2) Assur, A. (1959), Criteria for landing bomber and fighter aircraft on floating ice sheets (U), SIPRE Technical Report 58, also Army Science Conference 1957

(3) Hetenyi, M. (1947), Beams on elastic foundation

(4) Nevel, D. E. (1961), The narrow free infinite wedge on an elastic foundation, CRREL Research Report 79

(5) Meyerhof, G. G. (1960), Bearing capacity of floating ice sheets, Proc. ASCE 2627, EM 5

(6) Assur, A. (1961), Discussion: Bearing capacity of floating ice sheets, Proc. ASCE 2842, EM 3

(7) Assur, A. (1956), Airfields on floating ice sheets with supplements, SIPRE Report 36

(8) Assur, A. (1960), Composition of sea ice and its tensile strength, CRREL Research Report 44, also Arctic Sea Ice Conference, NAS-NRC Publ. 598 (1958)

TRACER STUDIES OF VITAMIN C UTILIZATION IN MEN; METABOLISM OF D-GLUCURONOLACTONE-6-C^{14}, D-GLUCURONIC-6-C^{14} ACID AND L-ASCORBIC-1-C^{14} ACID

E. M. BAKER, MAJOR, MSC, H. E. SAUBERLICH, S. J. WOLFSKILL and W. T. WALLACE
U.S. ARMY MEDICAL RESEARCH & NUTRITION LAB, FITZSIMONS GEN. HOSPITAL DENVER, COLORADO

Recent studies (1) indicated that D-glucuronolactone caused increased blood ascorbic acid levels as well as increased urinary excretion of ascorbic acid in men, whereas D-glucuronic acid did not do this. To check the possible conversion of D-glucuronolactone to L-ascorbic acid, it was decided to study the metabolism of the lactone in two ways. One was to give D-glucuronolactone-6-C^{14} orally and then isolate urinary ascorbic acid to determine if any of the labeled lactone had been converted to L-ascorbic acid. The other was first to label the total body ascorbic acid pool with L-ascorbic-1-C^{14} acid and then test with various loads of D-glucuronolactone to see if any changes would take place in the specific activity and rate of excretion of ascorbic acid. Further, an attempt was made to see if the total body ascorbic acid and its rate of utilization were related to the fat-free body weight.

Experimental. The D-glucuronolactone-6-C^{14} and D-glucuronic-6-C^{14} acid were obtained from the National Bureau of Standards, Washington, D. C. The L-ascorbic-1-C^{14} acid was obtained from the California Corporation for Biochemical Research, Los Angeles, California. These compounds were checked for purity prior to use by melting point measurement and by paper chromatography. The activity of the above compounds was checked by radioassay. D-glucuronolactone-6-C^{14}, 4.11 μc/mg, or D-glucuronic-6-C^{14} acid, 3.72 μc/mg, were freshly dissolved in distilled water and swallowed immediately by healthy men in quantities of 20 μc of each appropriate compound for each man together with a cold carrier load of 2 g of D-glucuronolactone or D-glucuronic acid. For the men receiving L-ascorbic-1-C^{14} acid, each swallowed 20 μc of L-ascorbic-1-C^{14} acid having a specific activity of 1.35 mc/mM. No cold carrier was given to these subjects. Following this, each man swallowed 2 g of freshly dissolved D-glucuronolactone on the 3rd or 9th day.

Total daily urine was collected and measured from all subjects; these samples were refrigerated and 2.0 ml of each were taken for counting.

Immediately after receiving the tracer quantity of D-glucurononolactone, D-glucuronic acid or L-ascorbic acid, the men were made to expire directly through a $CaCl_2$ drying train into a 5-liter Cary-Tolbert ionization chamber connected to a vibrating reed electrometer. The $C^{14}O_2$ activity, total CO_2, and the volume of flow were recorded automatically on a 6-channel recorder.

The total activity of each urine sample was determined by use of a liquid scintillator counter using P-dioxane-toluene. Oxalate in selected samples was isolated as calcium oxalate, recrystallized 4 times, and dissolved in 1 N hydrochloric acid for counting in the liquid scintillator. Quantitative determination of the total oxalate was done with the Archer method (2). Efficiencies for all liquid scintillator counting samples were determined individually by use of added standard C^{14} samples.

Urinary ascorbic acid levels were chemically determined by the Schaffert method (3). The urinary ascorbic acid was then isolated by the method described by Jackel et al. (4). After the dinitrophenylhydrazone (DNPH) derivatives were recrystallized, these were dissolved in P-dioxane and applied to weighed planchets and counted in a gas flow counter. All DNPH derivatives were recrystallized to constant activity which usually required 4 to 6 recrystallizations.

Total body tissue volume (V) was estimated in duplicate tests using a body volumeter based on displacement of water (5). From body weight (M) and V, fat (F) in kg was calculated according to an equation developed in this laboratory: $F = 4.834\ V - 4.366\ M$.

Results. Figure 1 shows the rate of $C^{14}O_2$ expiration in 3 men, as measured with an ionization chamber electrometer immediately following the swallowing of 20 μc of D-glucuronolactone-6-C^{14} or D-glucuronic-6-C^{14} acid. The cumulative excretion of $C^{14}O_2$ after ingestion of D-glucuronolactone was 44% of the total carbon-14 dose in 2 experiments, whereas after ingestion of D-glucuronic acid it was 68%. The peak rate of excretion occurred within 25 minutes in the lactone experiments, whereas in the glucuronic acid experiments it did not occur until after the 4th hour following ingestion. No $C^{14}O_2$ activity could be detected in the expired air of the subjects receiving L-ascorbic-1-C^{14} acid even using a 15-liter ionization chamber for greater sensitivities. A fraction of 1% oxidation to CO_2 during the first 8 hours could have been easily detected by this technique.

The urinary ascorbic acid which was isolated as the DNPH derivative had no radioactivity in the labeled D-glucuronic acid

experiments on 2 men. However, the DNPH derivatives isolated from the urinary ascorbic acid of the 3 men who had received D-glucurono-lactone-6-C^{14}, the specific activities ranged from 2.1 x 10^{-3} to 3.6 x 10^{-3} μc/mg ascorbic acid. This, of course, was dependent upon the amount of ascorbic acid excreted daily which ranged from 26.4 to 39.0 mg.

Further, an experiment was made with 20 μc of D-glucurono-lactone-6-C^{14} added to freshly voided urine together with cold carrier D-glucuronolactone and L-ascorbic acid. From this mixture the ascorbic acid was isolated as the DNPH derivative, recrystallized at least 4 times, and counted to determine the degree of contamination, if any, by the C^{14} labeled lactone. Less than 0.04% of the initial 20 μc could be detected in the ascorbate derivative, indicating that in the experiments on ingestion of D-glucuronolactone-6-C^{14} that less than 10% of the ascorbate derivative activity could have been from contamination under the most unauspicious set of conditions.

Figure 2 shows the percent of total activity excreted in the urine as oxalate from one of 2 men who had received 20 μc of L-ascorbic-1-C^{14} acid. This was done to see if the oxalate excretion arising from the C^{14} ascorbate was constant and, further, to study the effect of cold D-glucuronolactone, D-glucuronic acid, and L-ascorbic acid on the oxalate and ascorbic acid excretion. The oxalate activity was, indeed, remarkably constant, remaining at a level of 50 to 52% of the total activity excreted and changed only when the body ascorbate pool was overloaded or otherwise affected. Both subjects, who received cold D-glucuronolactone on the 3rd or 9th day following the L-ascorbic-1-C^{14} acid, excreted less total urinary oxalate activity whereas the urinary ascorbate total activity increased and was of a lowered specific activity. On the following day, the excretion of oxalate radioactivity returned to the 50% level. On the 2nd day following overloading with the cold lactone, the proportion of the total radioactivity excreted in the form of oxalate decreased, for unknown reasons, and did not return to the 50% level until after a further 2 days. Also shown (Figure 2) is a similar effect exhibited by overloading with 0.5 g of cold L-ascorbate. Further, note that 2 g of cold D-glucuronic acid, if anything, caused increased oxalate excretion. The above effects, at the least, indicate a marked dilution of the body ascorbate pool presumably by ascorbate arising from glucuronolactone.

The half-life of the L-ascorbic acid was 16 days in the 3 subjects receiving labeled ascorbate. This was obtained by plotting the percent of the initial dose remaining in the body as a logarithmic function of time. An average biological half-life of 16 days for ascorbic acid has been reported to occur in 3 diseased patients (6).

Table I shows the body pool size of ascorbic acid in 5 men. To calculate this, the specific activity of urinary ascorbic acid

(Column 1) was divided into the total activity remaining in the body at the end of the collection period (Column 2) to obtain the total body ascorbic acid (Column 3). In the last 3 tests the pool size was estimated by assuming that one-fourth of the C^{14} labeled D-glucuronolactone supplied ascorbic acid which, in turn, was incorporated into the ascorbic acid pool. On Subject No. 2 the estimated pool size was 1.65 g compared with 1.64 g when this subject was measured using L-ascorbic-1-C^{14} acid, thus showing that, indeed, one-fourth of the lactone went to ascorbate.

From comparison of Columns 4 and 5 of Table I, it appears as though the fat-free body weight contains a uniform proportion of ascorbic acid in the quantity of 33.3 mg/kg. Further, as seen in Figure 3, the rate of utilization of ascorbic acid also is directly related to the fat-free body weight. The utilization rate was obtained from 2 assumptions: 1) sufficient time had elapsed for the C^{14} labeled ascorbic acid to become uniformly distributed in the body and 2) the proportion of excreted oxalate-C^{14} arising from labeled ascorbic acid was constant. The rate of utilization was obtained by dividing the specific activity of the ascorbic acid into the total activity excreted as oxalate during the 24-hour collection period. As an example, in the case of Subject I, 0.112 μc/day was present as oxalate and the specific activity of the ascorbate isolated that day was 0.0078 μc/mg ascorbic acid, giving a rate of 14.4 mg ascorbic acid utilized on that day. The same subject who, on the 3rd day had received 500 mg of cold ascorbic acid, continued to show an ascorbate utilization of only 15.0 mg/day.

Discussion. Studies of body composition and the use of C^{14} isotopes has here resulted in a method for stating the actual utilization of ascorbic acid by healthy men. After ingesting a single 20 μc quantity of D-glucuronolactone-6-C^{14}, the urine of healthy men contains C^{14} activity which can be isolated and found in the highly purified phenylhydrazine derivative of ascorbic acid. Such activity is not found following the ingestion of D-glucuronic-6-C^{14} acid, indicating that the above lactone, but not its acid, can function as a source of ascorbic acid.

In subjects who ingest 20 μc of L-ascorbic-1-C^{14} acid, the daily urinary oxalate arising from metabolism of the labeled ascorbate is subsequently excreted as a constant proportion of the total C^{14} activity remaining in the body. Thus, it can be inferred that the portion of the daily oxalate which arises from metabolism of ascorbate is formed and excreted at a constant rate. Note, since Crawhall et al. (7) showed that about 40% of the daily oxalate came from breakdown of glycine and here it is shown that another 50% comes from ascorbate, it is possible to speculate that nearly all of the urinary oxalate comes from the metabolism of glycine and ascorbate.

The ingestion of a single, comparatively large 0.5 g quantity of "cold" ascorbic acid or its precursors by a subject whose

body ascorbic acid pool has been previously labeled, as described above, results in increased excretion of C^{14} ascorbate of lowered specific activity. These effects are transitory in that within 2 days the total ascorbate excretion returns to previous levels and the ascorbate specific activity is lower than it was prior to dilution of the body ascorbate pool.

Simultaneously, the total activity and the specific activity of the oxalate decreases, but the proportionality of the total oxalate activity to the specific activity of the ascorbate remains the same. From these effects, it can be inferred that the utilization breakdown of ascorbic acid in the body occurs at a constant rate irrespective of an increased rate of supply of ascorbate to the body.

Further, in 6 men of diverse body weight and degree of fatness, it was found that the ascorbate utilization, as expressed in terms of C^{14} oxalate excretion, occurred at a rate of 0.207 mg per day per kilogram of fat-free body weight. Rarely, if ever, do adult males exceed 90 kg in lean body mass. Therefore, 18 mg per day intake would match the greatest quantity of ascorbate metabolized by the largest healthy man.

Summary

1. Results of studies with healthy men revealed that close to one-fourth of D-glucuronolactone-6-C^{14} was converted to L-ascorbic acid whereas, on the other hand, no activity could be detected in the ascorbate derivative isolated from the urine of subjects receiving D-glucuronic-6-C^{14} acid.

2. One-half of the urinary oxalate arises from the breakdown of ascorbic acid and is excreted at a constant rate.

3. The size of the ascorbic acid pool and its rate of utilization was directly related to the size of the fat-free body weight.

TABLE I. Body Ascorbic Acid in Healthy Men

Subject	Day of collection*	(1) Specific activity of urinary ascorbic acid μc/mg	(2) Activity remaining at end of collection μc	(3) Total body ascorbic acid mg	(4) Total body ascorbate / Body wt mg/kg	(5) Total body ascorbate / Fat-free body wt mg/kg
1	1	0.0077	18.61	2416	29.4	34.0
2	3	0.0079	12.95	1640	20.5	32.9
3	1	0.0110	20.04	1823	29.7	34.4
4	1	0.0021	18.50	$\left(\frac{8800}{4}\right)$ 2200†	22.5	34.0
5	1	0.0036	18.70	$\left(\frac{5200}{4}\right)$ 1300†	27.6	31.8
2	1	0.0028	18.00	$\left(\frac{6600}{4}\right)$ 1650†	20.5	32.9

* All collections were for 24 hours except Subject No. 3 which was for 8 hours.

† Assuming 25% conversion of D-glucuronolactone-6-C^{14} to body ascorbic acid.
Subjects 1, 3 and second test on Subject 2: L-ascorbic-1-C^{14} acid.
Subjects 4, 5 and first test on Subject 2: D-glucuronolactone-6-C^{14}.

REFERENCES

1. Baker, E. M., Bierman, E. L., Plough, I. C., Metabolism, 1960, v9, 478.

2. Archer, H. E., Dormer, A. E., Scowen, E. F., Watts, R. W. E., Clin. Sci., 1957, v16, 405.

3. Schaffert, R. R., Kingsley, G. R., J. Biol. Chem., 1955, v212, 59.

4. Jackel, S. S., Mosbach, E. H., King, C. G., Arch. Biochem. and Biophys., 1951, v31, 442.

5. Allen, T. H., Krzywicki, H. J., Worth, W. S., Nims, R. M., U. S. Army Med. Rsch. Nutr. Lab. Rpt. No. 250, 24 Sept. 1960.

6. Hellman, L., Burns, J. J., J. Biol. Chem., 1958, v230, 923.

7. Crawhall, J. C., Mowbray, R. R., Scowen, E. F., Watts, R. W. E., Lancet, 1959, November 14, 810.

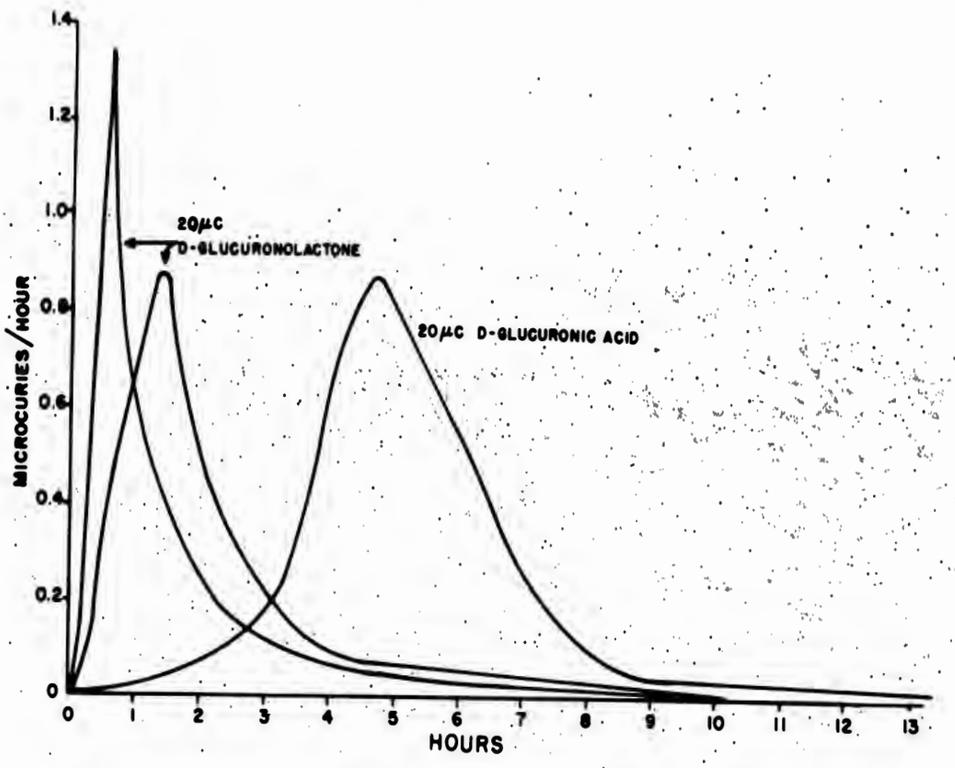


Fig. 1. Rate of $C^{14}O_2$ expiration following ingestion of 20 μc quantities of D-glucuronolactone-6-C^{14} and D-glucuronic-6-C^{14} acid.

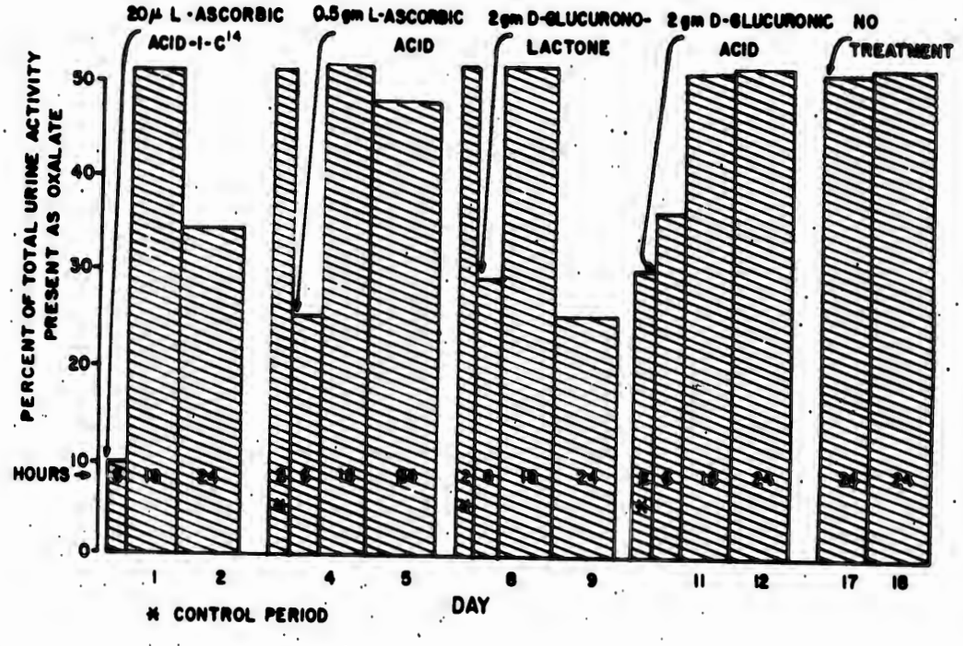


Fig. 2. Percent of total activity excreted by the kidneys as oxalate in Subject No. 1 on the days following ingestion of 20 μc of L-ascorbate-1-C^{14}. The hours of duration of each collection period are indicated in each block of the figure. See text for indicated treatments.

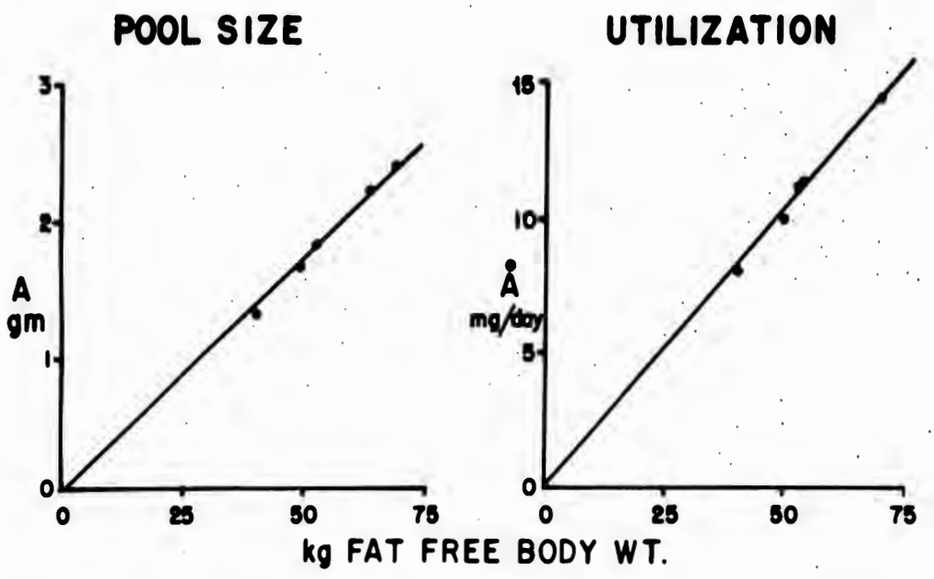


Fig. 3. Total body ascorbic acid and its rate of utilization are directly related to fat-free body weight of healthy men.

FAST RESPONSE SOLID STATE PME DETECTOR FOR LASER SIGNALS

A. BOATRIGHT and H. METTE
U. S. Army Signal Research & Development Laboratory
Fort Monmouth, New Jersey

I. Need for New Concepts in Fast Light Detectors

The invention of the optical maser (laser) by Maiman[1] in 1960 has provided us with a strong coherent light source that may be used as the carrier in wideband optical frequency transmission systems. The advantage of transmitting high volume information by means of modulated light is of particular significance for use over large distances in outer space.

An optical communication system consists, in principle, of three separate parts, i.e. (a) the light source, (b) the light modulator, and (c) the light demodulator. Considerable progress has been made recently in the development of CW optical lasers[2] and of light modulators[3] that are capable of operating up to the microwave region at reasonable power levels. The detection and demodulation of UHF and microwave modulated light, on the other hand, is still less satisfactorily solved. Photomultiplier tubes, e.g. that are frequently used as laser detectors, are seriously limited in their frequency response by their anode-dynode (and ground) output capacity and high anode resistances. It appears that even the best PM tubes cannot operate at frequencies above 10 mc without shunting out their output impedance and sacrificing some responsitivity. Furthermore, photomultipliers possess sharp response peaks in the visible and become very insensitive at the frequency of infrared lasers, (1.2 μ for the He-Ne laser). Therefore, new approaches of fast response light detection have to be explored.

The following solid state light detector, based on the photomagnetoelectric effect in germanium, was developed as part of the internal program at the Solid State and Frequency Control Division, USASRDL, Fort Monmouth, New Jersey, and is capable of overcoming some of the difficulties encountered with photomultipliers.

II. Operating Principle of the PME Detector

a. Description of the PME Effect

The photomagnetoelectric (PME) effect is defined as the voltage produced by incident light on a sample in the presence of a magnetic field, (Fig. 1).

$$U_{PME} = M\,H\,\Theta\,L \qquad \text{(Eq. 1)}$$

where H is the magnetic field strengths, Θ the incident light intensity and L the distance between the electrodes. The PME coefficient, M, is a materials property and is dependent on the front and back surface conditions and may also vary with the wavelength of the incident light. Furthermore, at high magnetic fields and light intensities, M is expected to be dependent on H and Θ. Since the PME voltage is a function of the product of light intensity times magnetic field, PME devices may be designed to measure either magnetic fields[4] or light intensities [5,6] by leaving the other parameter constant. For both types of applications, high sensitivities are achieved only with devices having high values of the PME coefficient, M. For the following we shall be concerned only with the application of the PME effect for the fabrication of a sensitive, fast response light detector.

b. Methods for Optimizing the PME Coefficient

Procedures of achieving high PME coefficients of fast response times may be derived from the physical interpretation of the PME effect. Fig. 2 shows the cross section of a PME sample. Carrier pairs are generated by incident light at the front surface of the semiconductor and diffuse in the direction of the back surface until they recombine. For sample thicknesses smaller than the diffusion lengths (order of 1 mm) most carriers recombine at the back surface. If, on the other hand, a magnetic field is applied perpendicular to the diffusion vector, the charge carriers are separated and establish an electric field perpendicular to the diffusion and magnetic field vectors.

A number of theories[7,8] for specialized cases of the PME effect have been advanced recently,* particularly for thick samples. None of these theories, however, permits a quantitative prediction of the size of the effect for the following reasons:

(1) M depends very strongly on the front surface recombination velocity, S_f, and for thin samples also on the back

*For a comprehensive list of literature on the PME effect see reference 5.

surface recombination velocity, S_b, and no independent method exists at present to obtain numerical values for these quantities. Variations in S_f influence the PME so strongly that, e.g. slight mechanical wiping of the surface with a tissue decreases the effect by several orders of magnitude. In other materials, e.g. Silicon, S_f is believed so large that no measurable PME effect exists.

(2) All previous theories are made with simplifying assumptions implying, e.g. that the concentration of optically generated carriers be small compared to the dark carrier concentration. Furthermore, the influence of the surface states is neglected. These conditions are - especially for thin samples and high light intensitites - not likely fulfilled.(9)

(3) It appears uncertain how far the present diffusion theory is still applicable in cases where the sample thickness is much smaller than the diffusion lengths. It was, therefore, decided to attempt instead an empirical procedure for achieving high PME voltages. This approach is based on the following considerations:

The PME field (Fig. 2) produced across the sample may be regarded as the sum of the Hall rotations of a longitudinal field (Dember field), established between the front and back surfaces for both types of carriers. Seeking high PME voltages at a given magnetic field strengths is therefore equivalent to seeking high Dember fields in the crystal. Now, Dember fields are the result of a carrier concentration gradient between the low recombination front surface and the high recombination back surface and high PME effects are therefore ultimately obtained by achieving high carrier concentration gradients between the front and the back of the sample. This can be done simultaneously in two ways:

(1) Increasing the concentration difference by careful chemical cleaning of the front surface and mechanical sanding of the back surface.

(2) Increasing the gradient by decreasing the distance between the front and back, i.e. decreasing the sample thickness.

While the first method is widely known and frequently applied, the effect of a decrease in sample thickness on the PME output voltage is usually underestimated. However, thick samples (1-2 mm) do not produce higher PME fields than 12 mv/cm at 9 kgauss and strong light intensitites, even with very clean surfaces, while thin samples of 20 microns thickness under the same light intensities and field strengths were found to produce PME fields of almost

1 v/cm. It is obvious that procedure (2) will only be valid to some optimum thickness w_{opt} below which the effect will decrease with the thickness. The size of w_{opt} depends both on the surface recombination velocities S_f and S_b and the absorption depths of the light and therefore can hardly be theoretically predicted. Since the surface recombination velocities S_f and S_b cannot be reproduced accurately in subsequent etchings of the same sample, it is not possible at present to measure the PME effect as a function of sample thickness alone and the experimental determination of w_{opt} is therefore difficult. However, in our experiments there appeared to be a steady trend of increasing PME voltage with decreasing thickness down to 10 microns thickness, indicating that w_{opt} must be below that value.

In addition to increasing the PME effect, this decrease in sample thickness has also the effect of shortening its response time, τ . There are two reasons for this:

(1) By increasing concentration gradient between front and back surface the diffusion velocity increases.

(2) By decreasing distance between front and back surface the effective diffusion path decreases.

Hence, one expects that the response time of the PME effect will decrease approximately with the square of the thickness w. A sharp decrease in τ with the sample thickness is observed experimentally. In thick samples (1 mm) we have found PME response times of between 20 and 100 μsec. Thin samples of 20 microns thickness of the same material, on the other hand, showed response times of well below one microsecond.

III. Fabrication of the PME Detector

a. Choice of Materials

PME voltages higher than one millivolt so far have only been observed in germanium and indium antimonide. However, since the controlled photoetching technique as a convenient means of obtaining single crystalline films is at present only known for n-type germanium, only this material was considered for the present application. Furthermore, the PME effect in germanium is higher than in InSb in the visible and near infrared, and this spectral range was of greater interest to us in view of laser application. The PME effect in InSb on the other hand, is most sensitive in the infrared up to 7.6 microns light wavelengths, and InSb has therefore been proposed for infrared PME detectors.[5,6]

The selection of materials was also determined by the desire to operate the device at room temperature and make it

as temperature insensitive as possible. The PME output voltage in germanium as a function of temperature has been previously investigated by Goldstein, Mette and Gärtner,[10] who found that the effect for both types of carriers possesses a maximum that shifts toward higher temperatures with increasing resistivity. This maximum will have its position at room temperature for 20-30 ohm cm material. N-type germanium in this range was therefore chosen as the best starting material for the fabrication of the detector.

b. Photoetching of Thin Ge Films

The preparation of thin n-type germanium films by controlled photoetching has been previously described.[11,12,13,14] The automatic thickness control achieved by this method is based on the findings that electrolytical etching of n-type germanium in alkaline solution occurs preferred at the thicker portions of the crystal if penetrating light (of a wavelength near the absorption edge) is irradiated into the crystal and produces holes that migrate to the surface.

Before etching, the samples were cut to slabs of 1 x .5 cm^2 area and lapped to five to seven mil thickness. Two copper leads were attached to the short ends of the sample, the contacts checked for ohmic resistance and the sample then mounted with alkali resistant plastic base wax on a 1" x 1/2" square glass or lucite slide, (Fig. 3). The use of two separate current leads to the anode not only provides a more even distribution of the etching current over the sample, but these leads serve later as the PME leads to the very thin and fragile crystal to which application of leads afterwards would be very difficult.

With this arrangement we were able to obtain quite satisfactorily mounted PME samples down to 6 micron thickness, but in most of the later cells 10-20 micron samples were preferred for their greater mechanical strengths.

c. Chemical After Etching of the Front Surface

After completion of the photoetching process the samples were taken from the electrolyte bath and thoroughly rinsed with distilled water. If the sample then is placed between the poles of a magnet it will already produce a rather high PME voltage upon illumination. However, it was found that this PME output can be further enhanced by 50% or more by cleaning the surface with a chemical after etching in slightly diluted CP4. The back surface during this process as in the photoetching remains coated with wax and therefore maintains the desired high back surface recombination velocity.

The chemical etching of the .5 cm^2 surface of a

sample which is only 20 microns or less thick, is difficult and requires special precautions since the etching tends to "bite" into the sample from the side or break through in spots, causing etch pits or holes. Good results were finally obtained in an arrangement whereby a small amount of etching solution was periodically released upon the surface of the sample and rinsed away after 1/2 second each time by dipping the cell in distilled water.

d. Mounting of the Cell

The PME cell is now ready to be placed in a magnetic field. In electromagnets of 10 kgauss field PME voltages of up to 1 volt are obtained with strong light intensities, and 1 cm electrode distance. Smaller and more portable units may be obtained by placing the cell between the poles of a small permanent magnet, (Fig. 4). Since high field strengths are of greater significance for obtaining high PME voltages than good homogeneity, a small Arnold C-magnet was chosen with tapering pole pieces that concentrate the field near the sample. The total unit weighs four ounces and operates with an effective field strength of 2500 gauss across an air cap of .7 mm. Since this same field strength can be obtained with smaller magnets in narrower gaps, and the PME voltage depends only on the length, but not on the width of the sample, it is possible to obtain even smaller units having the same voltage sensitivity, the only difference being a higher cell resistance of the smaller unit.

If the cell is to be used at low light intensitites, the leads should be twisted so as to minimize external pickup.

IV. Photoelectric Characteristics of the PME Cell

a. Spectral Response

According to the theory by Van Roesbroeck,[7] the ratio photoconductivity/PME effect should be independent of the light wavelength and can be used to calculate the carrier lifetimes. One notable exception, however, was predicted theoretically by Gärtner[8] and observed experimentallyby Brand, Baker and Mette[15] near the absorption edge in germanium, where at low back surface and high front surface recombination velocities this ratio becomes dependent on the wavelength and may even reach negative values.

In the present detector, however, we are using only highly etched front surfaces and, therefore, obtain only positive values of the PME effect. The response of a typical PME detector is plotted in Fig. 5 as a function of the incident light wavelengths. Except for a steep increase near the absorption edge near 1.7 microns, the response is found to vary only very little from there to the blue part of the visible region. This behavior is in good agreement with the spectral response of the photoconductivity reported by Briggs.[16]

b. Temperature and Time Dependence

Previous measurements of the temperature dependence of the PME effect in this Laboratory[10] have shown that the effect is very low at high and low temperatures but that a broad sensitivity maximum exists at intermediate temperatures, the exact position of which depends on the sample resistivity. For 25 ohm cm material this maximum occurs at room temperature and the change in the PME effect here is smaller than one percent per degree within a $\pm$ 10°C change.

In contract to this temperature change in the output voltage, the sensitivity of the cell decays initially irreversibly with time, probably by oxidation of the front surface. This decrease may amount to as much as 20 to 30 percent within 12 hours after the chemical etching has been applied, but the cell sensitivity then remains unchanged provided the surface is not mechanically or chemically affected. Still better time stability of the surface is achieved by enclosing the cell in a vacuum tube.

c. Sensitivity vs Light Intensity

In Fig. 6 is a plot of the output voltage three cells having different surface recombination velocities, as a function of the light intensity. The light source used was a standard lamp calibrated for 2870°K color temperature. The light intensity in milliwatts/cm^2 refers to the total spectrum of the lamp and no attempt was made to correlate these numbers with the spectral distribution of the cell. However, since the spectral response is almost constant up to 1.7 microns wavelengths, the correction introduced by the spectral conversion is small. As may be seen from Fig. 8, the voltage output of the cells increase almost linearily up to light intensities of five to 50 milliwatts/cm, however, the linear range is larger for cells with less sensitive surfaces than for sensitive (low recombination) surfaces. Two types of cells may, therefore be designed: (1) for the measurement of small light intensitites, the front surface recombination velocity should be as small as possible, but the linear range then is not very large; (2) if the front surface velocity is less sensitive, the cell is better suited for high light intensity applications due to their extended linear range. However, in either case, the voltage will saturate at very high light intensities and the cells, therefore, are not destroyed by overexposure.

All the measurements were made in effective magnetic field strengths of 2500 gauss. By using higher field strengths, the sensitivity of the cells will increase at a rate indicated in Section g.

d. Response Time

Fig. 7 shows the oscillograph tracing of the light pulse of an ILA light flasher, Model SC 4 with a time resolution of .1 μsec/div. The curve is believed to approximate closely the true shape of the light pulse. One notices, in particular, a very fast rise time of the order 10^{-8} seconds, indicating that the response time of the sample must be faster than 100 millimicroseconds with its exact lower limit not yet determined. The pulse hight of the curve is 200 *mV*.

e. Internal Resistance and Capacitance

An unique feature of the PME cell is that its impedance is almost independent of the incident light intensity. Only at very high light intensities does the cell resistance decrease by as much as 20 to 30 percent. The cell therefore, is particularly useful in combination with an amplifier circuit that requires constant impedance. The resistances of most cells used in the previous experiments were between one and two kilohms depending on the resistivity of the material used and the cell dimensions.

Since the cell does not possess p-n junctions, its capacitance is very small and caused mainly by the lead wires. By choosing proper lead dimensions, it may be kept below one μμf.

f. Noise

Two major sources of noise must be considered for the PME detector: (1) thermal (or Johnson) noise and (2) shot noise due to the statistical generation and recombination of carrier pairs. The shot noise in a PME sample is expected to depend on both front and back surface recombination velocities for which, as pointed out in Section II, no numerical values may be predicted. However, in contrast to junction type photocells, the PME cell has no dark current and its internal resistance does not change with decreasing light intensity. The shot noise, therefore, will diminish with decreasing light intensity, and therefore Johnson noise rather than shot noise will consititute the lower limit for the detection of weak signals. To test this assumption, a cell of 1 kilohm output resistance was connected to the input of a very low noise Hamner N-357 preamplifier and illuminated with weak dc light. The noise of the cell could not be distinguished from the amplifier noise which amounted to about five microvolts. This compares favorably to the Johnson noise which at room temperature, and for 1 mc bandwidth, is calculated to be 4 microvolts. Low noise therefore is one of the outstanding features that will enable the cell together with a low noise amplifier to detect very weak light signals.

g. Magnetic Field Dependence

The magnetic field dependence of the cell was measured and it was found that below 5000 gauss the PME voltage increases with the field linearily: above this point a slight saturation begins to appear until its deviation from linearity reaches 20 percent at 12k gauss. This behavior is in agreement with other magnetoelectric effects such as the Nernst effect[17] in germanium where deviation from linearity begins in the same range.

V. Applications of the Cell, and Conclusions

Table I, (Fig. 11) gives a survey on the most important features of the cell. The size of the PME effect for germanium

TABLE I — Properties of PME Cell

Sensitivity:	2.4 mv/mw/cm^2 at 2500 gauss
Spectral Response:	0.4.....1.8 microns
Temperature Dependence:	$<$ 1% per °C at 25° ± 10°C
Magnetic Field Dependence:	Linear to 5k gauss; 20% Deviation at 12k gauss
Response for Delay Time:	$<$ .1 μsec
Internal Resistance:	1kΩ
Capicitance:	$<$1 μμf
Overload:	Completely overexposure proof
Noise:	$<$ 5 μV at 25°C and f = 1 mc
Weight:	4 oz. including magnet

FIG. 11. TABLE OF PROPERTIES FOR PME DETECTOR

reported here is substantially higher than ever reported before in literature. This increased PME sensitivity enables the cell to serve as a sensitive fast response light detector for such spectral ranges where the responsitivity of photomultipliers becomes poor. As an example, Fig. 8 shows the oscilloscope tracing of a ruby laser signal taken at 1 μsec/cm time resolution with a portable cell. The spikes of the laser signal appear similar to graphs taken previously by other authors[1] with a photomultiplier tube. It shall be further noted that although the response time limits of the PME cell have not been investigated as yet in detail; there is nevertheless reason to believe that, using thinner samples, the response time of PME cells may be decreased below that achieved presently in photomultiplier tubes.

Sill unexplored, but possible in principle, are applications based on the simultaneous change of both the light intensity and the magnetic field. Examples of such devices could include choppers, mixers of electrical and optical signals and modulators.

In conclusion it has been shown that PME devices of considerable sensitivity, and with fast response times can be fabricated and used as fast response demodulators of modulated light pulses and as detector for the investigation of the fine structure of laser signals.

VI. References

(1) T.H. Maiman, R.H. Hoskins, I.J. D'Haenens, C.K. Asawa, and V. Evtuhov, Phys. Rev. 123, 1151 (1961)

(2) Nelson, D.F. and W.F. Boyle, CW Ruby Laser. Am. Phys. Soc. Meeting, 24-27 Jan 62, New York, N.Y.

(3) A microwave light modulator using the Pockels electro optical effect was recently announced by Sperry Gyroscope, Great Neck, N.Y. The device is able to achieve 30 percent modulation at 1000 mc frequency with 6 watts input power.

(4) H. Mette, Rev. Sci. Instr. 28, 1096 (1957).

(5) P.W. Kruse, J. Appl. Phys. 30, 770 (1961).

(6) J. Tyndall, Brit. Comm. and Electronics, p. 56 (1959).

(7) W. van Roosbroeck, Phys. Rev. 101, 1713 (1956).

(8) W. Gärtner, Phys. Rev. 105, 823 (1957).

(9) Preliminary measurements by the authors on the PME effect in the presence of strong surface fields demonstrate the influence of slow surface states up to several seconds. (Unpublished).

(10) S. Goldstein, H. Mette and W. Gärtner, J. Phys. Chem. Solids 8, 78 (1959).

(11) W.H. Brattain and G.C.B. Garrett, Bell Syst. Tech. J. 34, 129 (1955).

(12) A. Uhlir, Jr. Bell Syst. Tech. J. 35, 333 (1956).

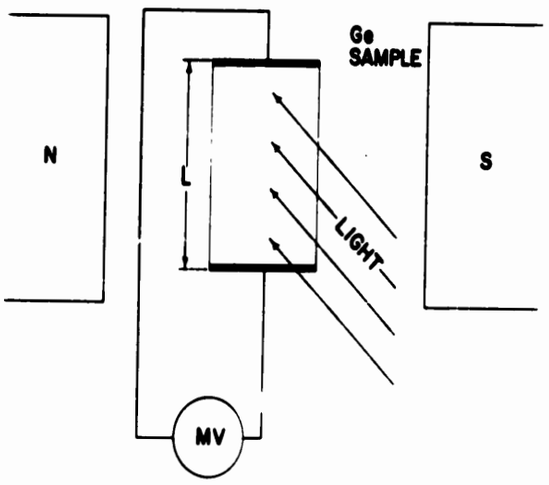
(13) W.A. Albers, Jr. and J.E. Thomas, Jr. Bull. APS Ser II, Vol 3, p. 219 (1958).

(14) A. Boatright and R. Binder, "Photoetching of Germanium Films," USASRDL Technical Rpt N5 2212, June 1961.

(15) F.A. Brand, A.N. Baker, and H. Mette, Phys. Rev. 119, 922 (1960).

(16) H.B. Briggs, Phys. Rev. 87, 1130, (1952).

(17) H. Mette, W. Gärtner, and C. Loscoe, Phys. Rev. 115, 537 (1959).



$U_{PME} = M \cdot H \cdot \Theta \cdot L$

M = PME COEFFICIENT H = MAGNETIC FIELD STRENGTH
Θ = LIGHT INTENSITY L = DISTANCE BETWEEN ELECTRODES

FIG. 1. DEFINITION OF THE PME EFFECT

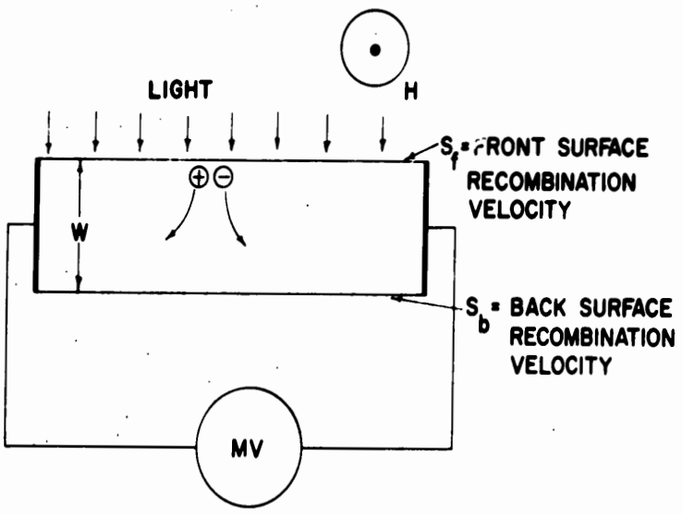


FIG. 2 : PHYSICAL INTERPRETATION OF THE PME EFFECT

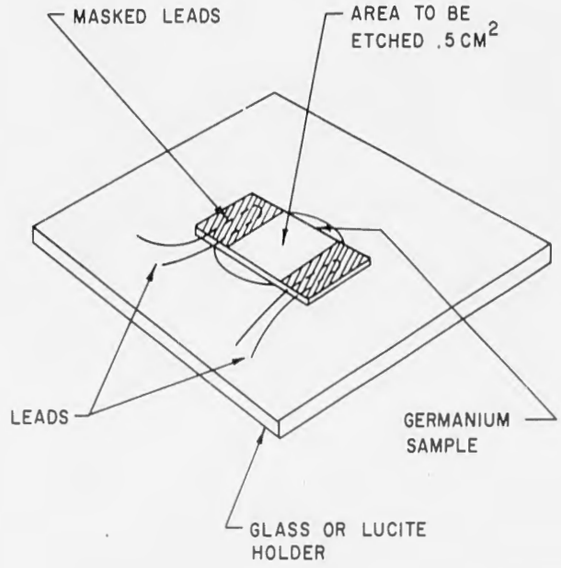


FIG.3. MOUNTING OF PME CELL

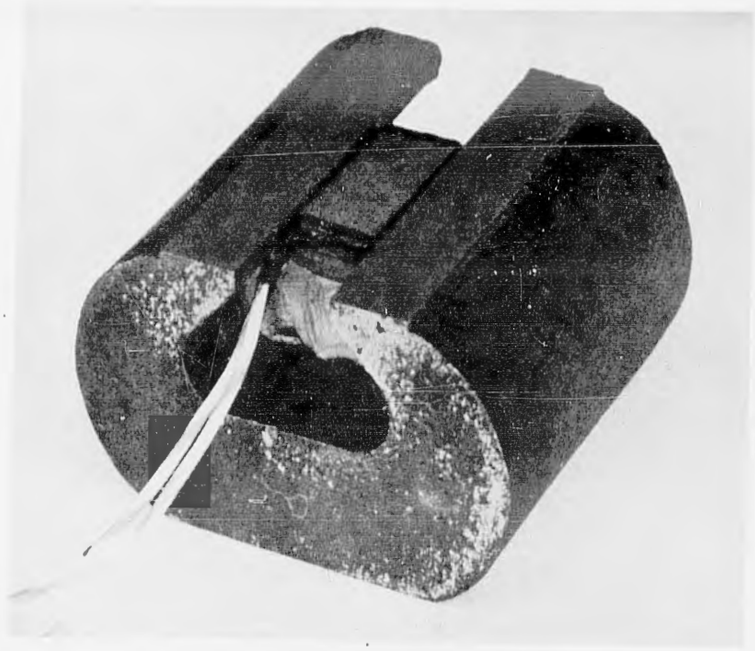
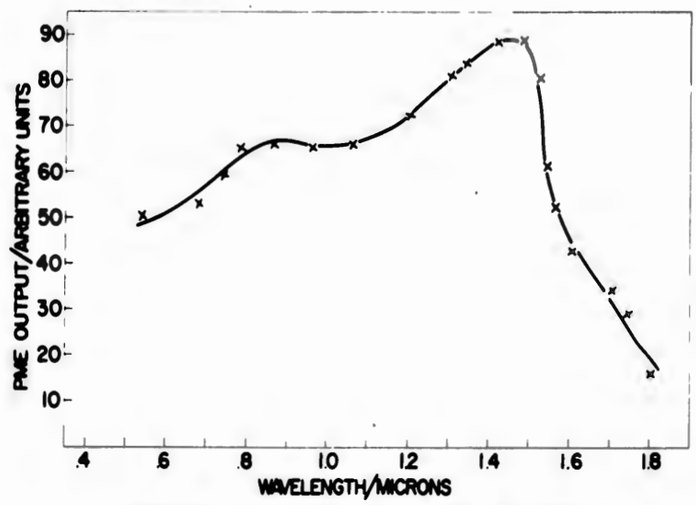


FIG. 4 PHOTOGRAPH OF PME CELL



SPECTRAL RESPONSE OF THE PME DETECTOR

FIG. 5

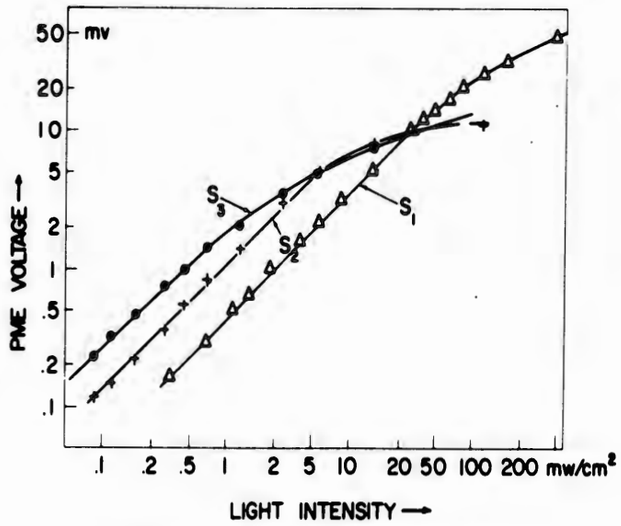


FIG.6. OUTPUT VOLTAGE vs. LIGHT INTENSITY FOR PME DETECTOR WITH S_1=HIGH, S_2=MEDIUM AND S_3=LOW SURFACE RECOMBINATION VELOCITY

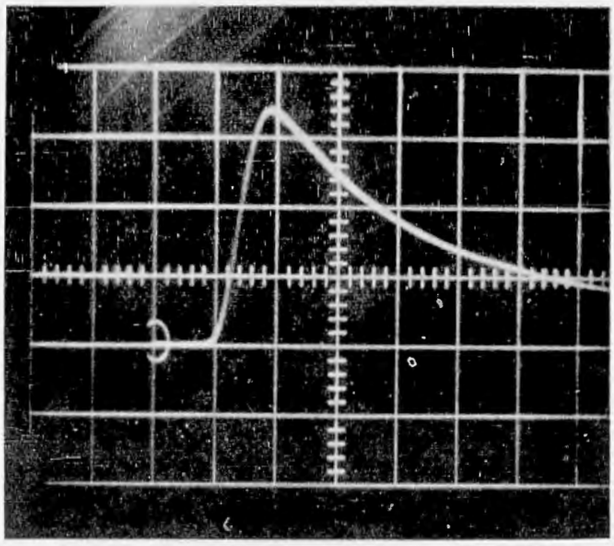


FIG. 7. .1 MICROSECOND LIGHT PULSE TAKEN WITH PME CELL - TIME SCALE .1 MICROSECOND/CM

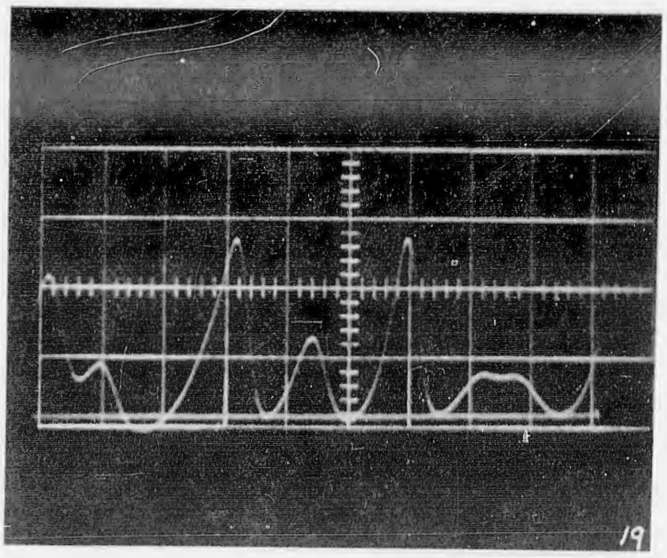


FIG. 8. LASER SIGNAL TAKEN WITH PME CELL. TIME SCALE 1 MICROSECOND/CM

GENERATION AND CONTROL OF AN OUTDOOR ELECTROMAGNETIC ENVIRONMENT THROUGH AUTOMATIC MEANS

CAPTAIN GEORGE D. BROSIOUS (RA)
UNITED STATES ARMY ELECTRONIC PROVING GROUND
FORT HUACHUCA, ARIZONA

With the arrival of the television set in the American home in 1947, the electronics layman was introduced to a new problem in life, that of interference. This introduction was performed by the "ham next door," "milady's" vacuum cleaner, the electric drill of the "man of the house," etc. Now the inconvenience of television interference has been reduced; however, the layman is advised in newspapers and magazines of the existence of a new and greater problem, that of RFI. This thing called RFI (Radio Frequency Interference) is blamed for such disasters as mid-air collisions of aircraft and premature destruction of multi-million dollar guided missiles. Thus, although the layman may not be aware of the intricate nature of RFI, he is aware of its existence.

The Army's approach to this problem, in general, followed much the same pattern as that involving our "electronics layman." During World War II, the inconveniences caused by RFI could, in most cases, be circumvented by either changing frequencies, or physically relocating certain communications facilities. However, it was realized that a combined allied operation, such as "Overlord," would impose special RFI problems; and Operation "Feeler" was put into effect to examine the problem and recommend solutions. Although this operation resulted in an involved and lengthy effort, there was still confusion on the beaches of Normandy, caused by communications breakdown. Since World War II, the Army has undergone many reorganizations as a result of changing concepts. However, one trend has been constant - that of the increasing numbers and complexity of electronics equipments. Faced with the unknown effect of such increases on field operations, Project Monmouth was initiated to determine whether or not a problem existed, and if so, the best methods to approach its solution.

The results of Project Monmouth indicated that there was a problem, and further recommended that a field facility be established to determine the extent to this problem and recommend methods for its

solution. The requirement for the establishment of such a facility was given to the United States Army Electronic Proving Ground, at Fort Huachuca, Arizona. In March 1960, a contract was let for the implementation of this requirement. This contract has resulted in the establishment of the Electromagnetic Environmental Test Facility. The overall facility consists of an outdoor electromagnetic laboratory [1] located in the desert of Southern Arizona; and a Radio Frequency Interference Prediction Model [2] which will be capable of extending the findings of the field test facility to any tactical deployment in any area of the world.

The outdoor electromagnetic laboratory, hereinafter referred to as the "field facility," exists in an area of approximately 2400 square miles to the east of Gila Bend, Arizona. As of this writing, the field facility possesses the capability of testing the communications links of a division in the environment created by the communications equipment and selected forward radars of a corps. This capability is achieved by the accomplishment of the following functions: (1) Generation of an environment equivalent to that which the equipment under test would experience if that equipment were deployed under battlefield conditions. (2) Control of the environment to facilitate rapid testing as well as the provision of a frequency monitoring capability to ensure accurate results without interfering with electronic activities outside the test area. (3) Testing of communications links in such a manner that the existence, extent, and cause of interference may be determined objectively. (4) Evaluation of test results to enable accurate analysis of the situation, as well as providing the basis for corrective recommendations.

This paper concerns itself primarily with the generation and control functions. The first step in the generation of the environment was to determine the size and nature of the environment to be created. In view of the recommendations of Project Monmouth, and the fact that the effectiveness of an army is proven on the battlefield, it was determined that the environment should be that of tactically deployed Army units. The optimum size selected was that of an Army Corps. A tactical scenario was provided by the Army in which a hypothetical corps, consisting of three infantry divisions on line and an armored division in reserve, was deployed in the area of the field facility. An examination of appropriate troop lists, T/O & E's, and normal electronic equipment deployments resulted in the location of approximately 15,000 electronic emitters in the corps area. It was immediately obvious that, from the standpoint of economic and engineering feasibility, this number had to be reduced to enable field implementation.

The reduction rules were based on a normal equipment deployment, depicted in Figure 1. This figure is a schematic representation of the corps area, in which emitter population is indicated by the various bands - Band "A" representing handy-talkie and walkie-talkie radio sets used at platoon level, Band "B" representing mobile

FM radio sets used at company level, Band "C" representing mobile FM and low power AM radio sets used at company and battlegroup level, etc. Since test points within a single band would experience the same general environment, nothing would be gained by placing several test points parallel to the front lines. Conversely, a line perpendicular to the front line shows a gradient, insofar as types of equipment are concerned, and test points along this line would experience differing environments. In accordance with the tactical situation, two test sites were selected: the Forward Test Site (FTS) to test the communications links of company and battlegroup, and the Rear Test Site (RTS) to test the communications links of division forward and main. Based upon this philosophy, the requirement placed on the environment is that it be realistic at these two points only. To satisfy this requirement, and enable physical realizability, the following reduction rules were evolved: (1) Observation of radio silence by reserve elements; (2) Signal strength at the test sites; (3) Application of the "netting" principle.

The application of the first rule eliminated all the emitters in the armored division held in reserve, since it was assumed that units in reserve would be observing radio silence.

Emitters were eliminated under the second rule if the effect of their transmissions was not noted at either test site. Originally, the range which was used as a basis for this elimination was based on the maximum given in appropriate Technical Manuals; however, since that time experience has shown that in many cases the actual range is in excess of the stated maximum. As a result, additional emitters will be added in the future to ensure a truly realistic environment at the test sites. In addition to elimination based only on range, the effects of directional antennas and terrain masking were considered.

The "netting" principle, as applied to military communications, allows many stations to communicate with one another, but on a one-at-a-time basis. Thus, at any instant of time, only one of these many stations may be transmitting, and the remainder are utilized in the receive mode. The consideration of this principle in the field facility allows the elimination of all stations being utilized in the receive mode, thus requiring that only one station in each net be represented in the facility. In practice, this method of reduction was not applied until the first two reduction rules had been considered. The reason for this approach was to ensure that the one emitter from each net that was retained had the most significant effect at the test site. In cases of doubt, or when one emitter could be utilized in more than one net, an additional emitter was always retained.

After the application of the reduction rules, the 12,000 communications emitters normally assigned to the corps depicted in the tactical scenario were reduced to 412. Similar rules were applied to non-communications electronic emitters to reduce the corps total of

approximately 3000 to 83 Since, as of this writing, only 13 of the required non-communications emitters have been added to the environment, there is no experience factor from which the realism of the reduction of non-communications equipment may be judged.

Although the 495 resultant emitters is a feasible quantity, from a logistic standpoint, this quantity of emitters is still spread over the total 2400 square mile area of the field facility. Control of these emitters over such a large area was not feasible by normal methods, and measures had to be taken to group equipments in such a manner that they could be controlled with relative ease. The basis of the grouping theory is that an emitter may be moved insofar as its effect on either test site is not changed by more than a matter of several decibels. This means that emitters which are placed a considerable distance from a test site may be moved approximately one quarter of the distance between that emitter and the test site, without changing the overall effect more than the previously described limit. When this 25% factor is applied to lateral movement, it results in an allowable movement of approximately 15 degrees. In practice, a hypothetical group control point is selected and then this theory is applied to develop the sector in which all included transmitters may be controlled from the control point. This process is indicated in Figure 2, where the sector of control for this point is determined by an azimuth of 30° from the test site and the radial length of the sector is "d/.75 - d/1.25" or .583d. It should be noted, however, that close to the test sites, it is necessary to locate equipments in accordance with their normal tactical deployment, inasmuch as a small change in range or azimuth would change the effect that the particular emitter would have on the test site more than the allowed limit. By the application of this theory of grouping, the 495 emitters were assembled into 26 different groups.

The physical system used to achieve this grouping is indicated in Figure 3. This figure is a photograph of an environmental generator van (EG van). The EG vans act as the control points for each of the various groups. Each group is represented by the environment generators (EG's) which are located around the EG van. These EG's are seen here as white boxes, referred to as "bird houses," which are located at varying distances and azimuths from the EG van. Inside these "bird houses" are located the actual transmitters which create the environment in the field facility.

The next problem, and perhaps the largest, was the selection of the type of equipment which would be used for the emitters in the field facility. The choice was between the use of simulators and the use of actual tactical equipment. Although simulators presented an advantage in the relative ease of control which could be built into them, it was determined that the actual tactical equipment would be used due to the inherent realism of its emissions and the relative immediacy of its procurement. Although difficulty in the provision of power and control was anticipated, this disadvantage was outweighed by

the advantage of realistically reproducing all radiations, spurious and otherwise, so that the complete interference capability of the equipment was actually tested. For further realism, these sets were maintained in accordance with specifications contained in appropriate technical manuals. The problem still existed of how to ensure consistent operation of the sets during lengthy periods of field testing. Since one of the basic premises of the facility is rapid testing by means of automatic control, it was necessary that certain output functions of the equipment be continuously monitored, to ensure that their operation was within the same specifications mentioned above. To enable the accomplishment of these functions, certain modifications were made to the RF sections, modulation sections, and where applicable, the frequency shift keying (FSK) sections of the equipment. These modifications resulted in relatively few changes to existent circuitry, although in some cases they required the addition of passive components. It should be noted that the modifications made were of such a nature that if it were desired to remove a particular piece of equipment from service as an emitter in the field facility and restore it to its original condition, it could be done with the expenditure of little effort.

The control of these emitters was slightly more complex. It required automatic "off-on" control in addition to the continuous monitoring of the set to maintain operation within the specifications. In addition, it was necessary to provide the capability for automatic turn-off of the equipment, should its operation fall below preset specifications. Figure 4 is a photograph of an AN/VRC-10 vehicular radio set prepared as an environment generator. Note that the military equipment appears unchanged but that certain control equipment has been added to make up the complete environment generator. On the left of the military equipment is the control box which was prefabricated for this particular transmitter. Note that two meters are located in the upper half of the control box. The meter on the left represents a capability for continuous control of RF power output. The meter on the right represents a facility for continuous monitoring of the modulation of the transmitter. For transmitters with FSK capability, a third meter is utilized to provide a facility for continuously monitoring the FSK input to the transmitter. Each one of these meters is actually a meter relay, providing the capability for setting a level according to specifications. If the actual reading falls below the preset level, the relay closes, disabling the transmitter. The reasons for the use of a meter relay are twofold: (1) It enables rapid and simple setting of specified limits, and (2) if a generator does malfunction, it allows the repairman to observe which portion of the transmitter is not operating properly. Located below the meters are the control circuits and cable connections which allow the application of power from DC generators to the actual transmitter.

The control box shown here was prefabricated for use with the AN/VRC-10 radio set. A different but similar control box is used for all other types of communications equipment employed in the field

facility. At the present time, automatic controls are not provided for the radars installed in the field facility; however, it is expected that in the near future such automatic controls will be fabricated and installed.

To enhance the realism of the facility, it was necessary to establish frequency assignment plans in accordance with those which would normally be utilized by a similar tactical unit in the field. This implies that frequencies which were assigned to specific equipments in the field facility were those which would be available for that particular piece of equipment had the frequency assignment started from Army level and worked on down to the particular unit in question. The basis of these frequency assignments was JANAP-194A and standard Army frequency assignment practice. It should be noted here that certain changes had to be made due to the non-availability of some frequencies which were utilized by civilian and other military electronics equipments in the vicinity.

Although the fabrication and deployment of the environment was a major task in itself, to achieve the goals of the field facility, it was necessary to provide an overall control function. This control would enable rapid testing to determine whether interference was present and once discovered, to determine its cause. To accomplish this end, the control function has been sub-divided as follows: (1) An automatic control system which provides control of the environment generators as well as recording test results, and (2) a frequency monitoring system which insures proper frequency usage as well as providing data to indicate the actual environment during each test.

A general view of the Automatic Control System is presented pictorially in Figure 5. The Forward Test Site (FTS) is the hub of this control system. It represents a unified control over the entire field facility. A card system is utilized to facilitate the automatic facet of the control system and enables the completion of a single test in 30 seconds. The commands from the Forward Test Site are passed, via this card system, to each of the 26 Environment Generator (EG) sites. These EG Sites accept the commands from FTS, and via their own internal card system, pass commands to the Environment Generators. Each of the EG sites is capable of controlling up to 40 environment generators, giving the field facility a maximum capability of 1040 controlled EG's. The EG itself has an internal control system, which acts on the commands from the EG sites to turn various transmitters off and on in the specified order. The control devices required to implement this control system are more complex and are described in detail in the following paragraphs.

The control devices for the Automatic Control System are located at the Forward and Rear Test Sites, at 24 locations throughout the field facility in environment generator vans, and at various positions surrounding the vans in environment generator shelters. The control devices have been subdivided into three levels, to accompany

the similar subdivision of the Automatic Control System. This subdivision is indicated in Figure 6. The control devices located at the forward test site are designated first level control devices, hereinafter referred to as the Master Control. These devices provide master control for the system by transmitting control signals to second level control devices located at the EG van sites and the test sites. The control devices mounted in the EG vans and the control devices mounted in the EG control rooms at the test sites are second level control devices, hereinafter referred to as "Monitors for Feedback Control." These devices are used to provide on-off signals to the EG control boxes. The control devices mounted in rear test site, proper, are also second level control devices, used to control the Interference Detection Equipment (IDE). These devices are hereinafter referred to as the "Control Monitor." The control box associated with each Army transmitter is a third level control device used in conjunction with the Monitors for Feedback Control to provide on-off signals to that transmitter.

The Master Control system is based on the use of the punched card system, in conjunction with appropriate card readers, communications systems and malfunction indicators. The punched card system used is predicated upon a master deck located at Master Control (FTS), and auxiliary card decks which are located in each of the EG sites. The purpose of the master deck is to initiate the card advance command to the EG sites, instructing each site to advance one card in its auxiliary deck. The auxiliary deck contains the predetermined on-off schedule for the various EG's subordinate to the EG site at which the particular deck is located. A separate deck is delivered to each EG site for every series of tests, and the required on-off sequence punched on each deck is based on the overall test plan.

Figure 7 is a photograph of the control console located at the Forward Test Site, with its associated timing, control, and other ancillary equipment. A "Card Check Display Panel" is provided in the control console on which the number of each EG site appears. When a malfunction (wrong card in the card reader at one of the EG sites) occurs, a visual indication gives the number of the EG site which has the malfunction, allowing the operator to make the necessary correction so that the test may continue. The operations console also includes a digital clock which enables instantaneous recording of the time a particular test was actually performed. The right side of the control console contains a "Display Panel" and "Master Sequencer." The Display Panel provides a visual indication of test number, date, and other test data, which are automatically punched on each test card. The Master Sequencer controls the test sequence. When given a command to proceed in the automatic mode, a signal is generated by the Master Sequencer to advance all card readers located at the EG sites. To the right of the control console, a rack is located with the "Timer Panel" and the "Tone Coder-Decoder" mechanisms. The Timer Panel controls the time sequence of each test cycle. A test cycle normally consists of 30 seconds, subdivided as follows: a "set-up" period of 9 seconds,

an interference test period of 20 seconds, and an interference data punch period of 1 second. The length of these periods is adjustable from the Timer Panel. The Tone Coder accepts the commands originating in the master deck, and codes them into tone signals for transmission to the 26 EG sites and to the Control Monitor at the Rear Test Site. The Tone Decoder, on the other hand, accepts tone signals from the EG sites and Rear Test Site for decoding into DC signals which activate proper controls at the Master Control (FTS). The use of tone signals between Forward Test Site and the various EG sites allows the use of "Spiral-4" cable, since the tones utilized are between 400 and 3000 cycles per second. The use of standard Army field cable is desirable because of the length and roughness of many of the cable runs.

Referring to Figure 6, the second level control devices are the Control Monitor located at the Rear Test Site, the Interference Detection Equipment located at both Forward and Rear Test Sites, and the Monitors for Feedback Control located at the EG vans and also at Forward and Rear Test Sites. The Control Monitor consists of a digital clock, display panel, and a coder-decoder group. Basically, each of these units is designed to function in the same manner as their counterparts installed at the Forward Test Site. The Interference Detection Equipment provides the capability for measuring the intelligibility of both voice and radio teletype test links (20 seconds) and for recording this information on punched cards (1 second). The Monitors for Feedback Control accept the tone commands from the master control devices and convert these commands into 26 volt DC signals which actuate the third level control devices, in accordance with the instructions contained in the auxiliary decks.

The Monitors for Feedback Control also control the rather complex recording of the malfunction data transmitted by the third level control devices. The recording of this malfunction data is accomplished by (1) accepting the data from the third level control devices, (2) placing this data into a memory unit, and (3) upon call, "dumping" the memory so that the information may be punched on one of the test cards from the auxiliary deck. The memory unit is required here to insure that malfunctions through the entire test period are recorded, rather than a sampling at the beginning and/or end of the period. Also included as a Monitor for Feedback Control is a display panel similar to that described for the Forward Test Site. Rather than indicating EG sites, it indicates the various environment generators deployed around a particular EG site. This panel provides both a visual and audible alarm in the event any of the environment generators malfunction.

The third level control devices consist of the control boxes and transmitters which make up the Environment Generators, and have been previously described.

For the purpose of explaining the dynamic operation of the Automatic Control System, the following sequence of events which leads

to the completion of one specific test is presented: Let us assume that Test Run #3 has been in operation for approximately 15 minutes, and we are now at Test #30. As card #30 drops into the card reader at Forward Test Site, the set-up timer goes into operation, allowing approximately nine seconds for EG warm-up and malfunction checks. During this same period, a card advance signal is sent to each of the EG sites and each site is queried as to whether the card just advanced had the digit "0" as the last digit of its serial number. Assuming that only sites 1, 2, and 3 are involved in this particular test run, if sites 1 and 3 have card #30 in their card readers, no indication will be received at the Forward Test Site. However, if site 2 has card #29 in its card reader, a visual and audible alarm will be received at the Forward Test Site and the test automatically halted. At this time the console operator, noting the malfunction, will manually advance one card at site 2, receiving a completed card check, and therefore allowing the test to continue. As the nine second time period set on the set-up timer ends, the IDE test timer begins its operation. During the IDE period of approximately 20 seconds, the interference detection equipment performs its function of analyzing the test link to determine the degradation. In addition, at the EG sites, the cards of the auxiliary deck have commanded certain environment generators to turn themselves "on." If any of the environment generators commanded "on" have a malfunction either in RF power, modulation, or frequency shift keying, a malfunction indication is given at the EG site, and the malfunctioning EG is automatically disabled. All malfunction information is placed in the memory unit at each EG site throughout the test period of twenty seconds. The operator at each EG site knows those EG's which are not operating properly; and by checking the meters at the specific EG location, he can determine which portion of the circuit requires maintenance. As the IDE test timer completes its twenty second interval, several commands are given. The interference information which has been collected over the twenty second interval is "dumped" by the IDE and punched into test card #30 of the master deck, and all EG sites are commanded to put their EG's on stand-by. The EG malfunction data for this test will not be punched until test card #32 is in the card reader at each of the EG sites. At this time, the malfunction information, which has been held in the memory units since test #30, is "dumped" and punched on the card for test #32. Upon receipt of "interference data punch completed" signals from the Forward and Rear Test Sites, test #30 is considered complete and test #31 will automatically commence. It should be emphasized that many variable factors which could affect the time and sequence of these events have not been considered here, in order that the general operation of the Automatic Control System could be demonstrated.

Successful operation of a controlled electromagnetic environmental test area requires knowledge of the ambient environment, including external and internal contributions. With such knowledge the data analysists can make suitable allowance in the evaluation of test data and equipment performance. Without such knowledge, an element of control is lost; because the interference contribution of ambient

sources is unknown. Frequency monitoring vans have been designed to exercise effective control over authorized frequency assignments and to permit adequate performance evaluation of the devices under test. Five frequency monitoring vans (four semi-permanent vans located on the periphery of the environment and one mobile van capable of relocating at any desired point within the facility) are presently utilized to analyze the range of frequencies in use during any test period. This existing capability provides monitor, field intensity measurement, analysis, direction finding, counting, and calibration facilities in the frequency range .5 to 1,000 Mc. It is expected that in the next year this frequency range will be extended to 10.0 kMc.

The generation and control functions described above represent the basic approach taken to establish the field facility as it exists. It should be indicated at this time that future plans call for a more sophisticated approach to the achievement of both the generation and control functions. This sophistication will ultimately result in a refined and expanded facility, capable of adding heretofore unattainable information to that required for the successful control of radio frequency interference.

LITERATURE CITATIONS

(1) L. F. Babcock, "A facility for the investigation of radio-interference problems," Unclassified Proceedings of the Seventh Conference on Radio Interference Reduction and Electronic Compatibility, November 1961.

(2) A. J. Hochn, "Engineering aspects of the EMETF interference-prediction model," Unclassified Proceedings of the Seventh Conference on Radio Interference Reduction and Electronic Compatibility, November 1961.

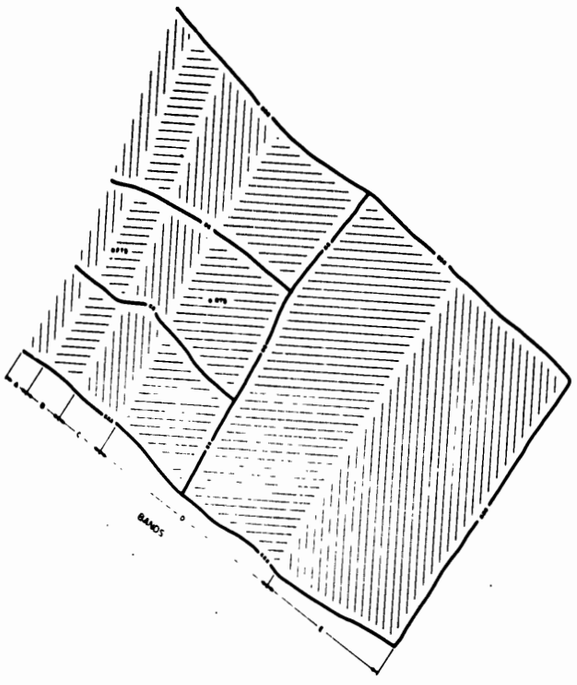


Figure 1 - Emitter Population

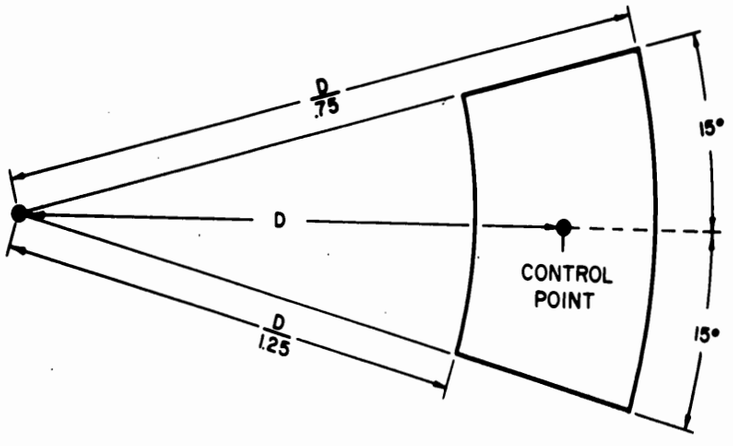


Figure 2 - Sector of Control

Figure 3 - EG Van

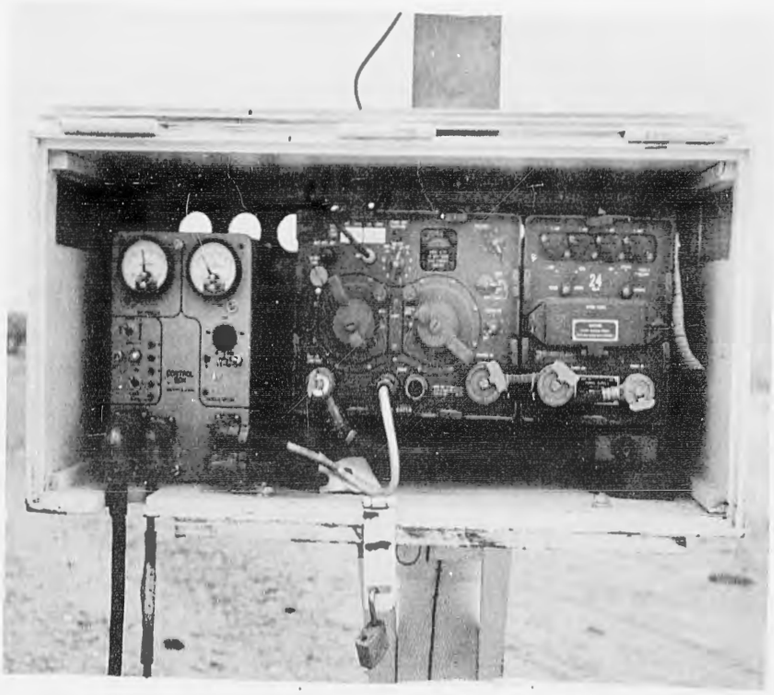


Figure 4 - EG (AN/VRC-10)

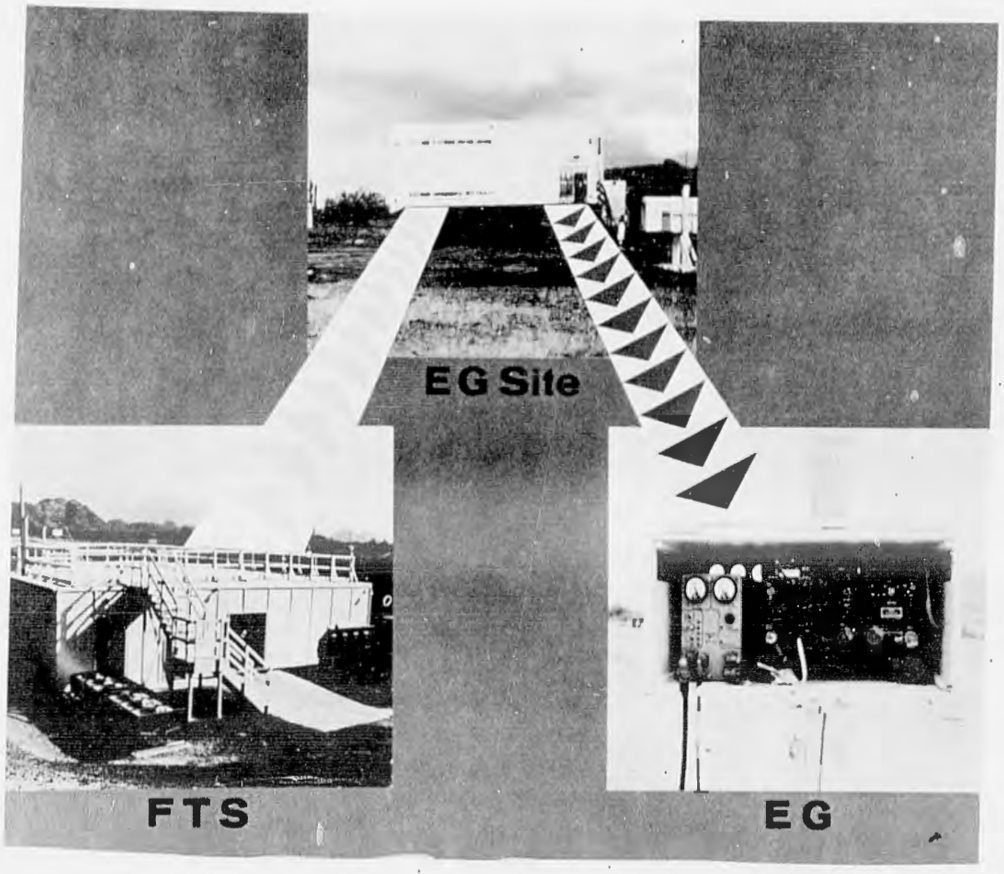


Figure 5 - Automatic Control System

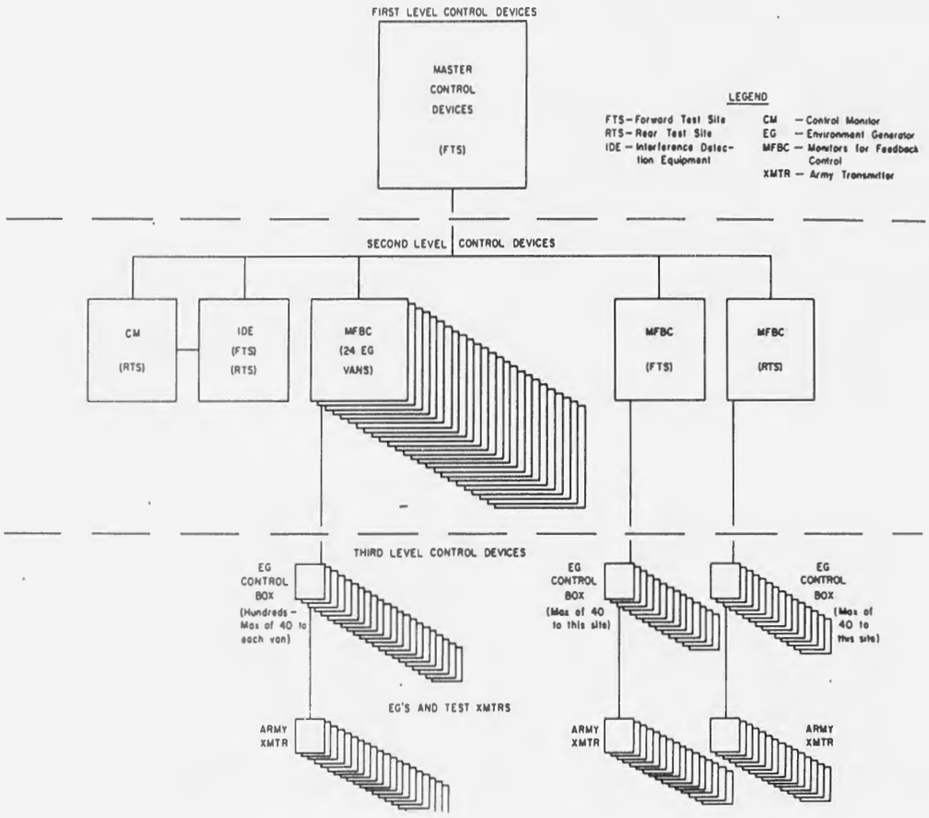


Figure 6 - Control Devices

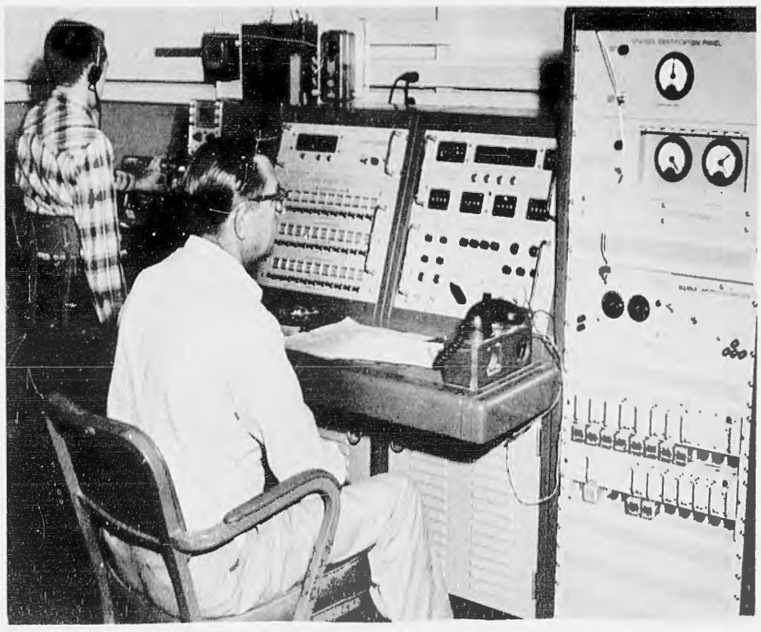


Figure 7 - Control Console (FTS)

OPTIMUM LINEAR FILTERING FOR LINE SCAN IMAGING SYSTEMS

LESLIE G. CALLAHAN, JR.
US ARMY RESEARCH OFFICE (DURHAM)
DURHAM, NORTH CAROLINA

INTRODUCTION: Many optical processes of image formation, image transfer, and image analysis may be represented as one or a succession of linear transformations of a two-dimensional signal from an object plane into a two-dimensional response in an image plane. This representation is the spatial equivalent of one-dimensional time transformations performed by electrical networks. Therefore, many of the concepts and techniques normally associated only with linear electrical systems can be extended to problems involved in the analysis and synthesis of optical imaging systems. For example, a radiation pattern in the object plane can be represented by two-dimensional Fourier components just as a voltage input to an electric circuit is commonly represented by one-dimensional Fourier frequency components. Further, imperfect lens, grain on photographic film, and other imperfections in optical components which tend to degrade the performance of optical systems can be considered in the same perspective as electrical noise.

In recent years, these and other equivalences have been noted and used by a number of investigators in the solution of various types of optical problems (1-22). However, aside from the work of Elias, Jones, O'Neill, and Wessley, relatively little attention has been directed at the optimization of line scan imaging systems (10, 11, 13, 15, 22). Although both Mertz and Shade used two-dimensional electrical techniques for the analysis of line scan systems, they did not concern themselves with the optimization problem in the context of modern circuit theory (14, 21).

Line scan imaging systems denote the class of imaging systems such as television, facsimile, and airborne infrared mapping devices which record input spatial signals as sequences of parallel lines in accordance with a specified scanning pattern. Interest in this class of systems has been largely motivated by the desire to reduce two-dimensional spatial data to one-dimentional electrical data to

facilitate recording or transmission over electrical communication channels. Consequently, the notion of an aperture moving with a constant velocity along the scan lines, and producing a continuous one-dimensional, temporally modulated signal at the output of the system, has generally been accepted as a model of the scanning system by previous investigators. Influenced by this model, most processing of data produced by existing physical scanning systems is effected in only one dimension even though it is intuitively apparent that processing in two dimensions should result in improved performance. The purpose of this paper is to present a new model for the analysis of line scan imaging systems which permits the quantitative comparison of optimum one and two-dimensional linear filtering operations. In this model the sampling process is viewed as a one-dimensional sampling of a static two-dimensional signal. Consequently, its utility is primarily restricted to the design of line scan imaging systems which are not required to reproduce motion in the object plane, i.e., an airborne IR scanning system. Optimum filters are derived using both a statistical and a determinative criterion, i.e., least square error and spatial resolution.

IMAGING SYSTEMS AS FILTERS: The purpose of an imaging system is to transform, or map, input electromagnetic signals from element sources in a two-dimensional field of view, or object plane, into a corresponding output representation, or image plane. In the ideal system, elements reduce to points, or two-dimensional spatial impulse functions, and a one-to-one transformation is effected between the input point sources and output point images. In practical systems, a signal impulse from an input point source is transformed into an output signal with physical dimensions analogous to the impulse response of physical electrical filters, Fig. 1. Consequently, overlap and blurring occur between point images in the image plane.

A system is said to effect a linear transformation if a point in the image plane represents a linear weighted average of the spatial signal emanating from the neighborhood of a corresponding point source in the object plane. Thus, if

(a) $a(x,y)$ = system impulse response at the point (x,y)

(b) $f_1(x,y)$ = input at point (x,y) (1)

(c) $f_2(x,y)$ = output of system at point (x,y)

then the output and input signals are related as

$$f_2(x,y) = \int_{-\infty}^{\infty}\int a(x-u, y-v) f_1(u,v)\,du\,dv \qquad (2)$$

It is noted that if

$$a(x,y) = \delta(x)\,\delta(y) \qquad (3)$$

i.e., a two-dimensional unit impulse function,

then

$$f_2(x,y) = f_1(x,y) \tag{4}$$

or the output is identical to the input.

Fourier analysis in two dimensions permits the representation of a spatial signal by a superposition of plane waves of the form $e^{2\pi j(k_1 x + k_2 y)}$. Here k_1 and k_2 represent spatial frequencies or number of waves per unit length, along the x and y coordinate axis respectively. As in the case of a temporal frequency spectrum for electrical signals, the wave number spectra for spatial signals contains large high wave number components when its corresponding Fourier spatial signal undergoes rapid change. For example, a small object will have a wave number spectrum which decreases slowly with increasing magnitudes of the wave numbers, until wave numbers of the order of the reciprocal of its dimensions are reached, and will then decrease more rapidly. If the object is narrow and tall, the width and height of its spectrum will be roughly the reciprocal of the width and height respectively of its spatial form, as indicated in Fig. 2.

A linear system responds to an individual plane wave component by producing an output signal of the same wave numbers or spatial frequency but with variable amplitudes and phases dependent on the ability of the system to pass that particular component. In this sense an imaging system can be considered or a spatial filter which selectively passes or modifies component frequencies in the input plane. Thus, if $A(k_1,k_2)$ is the spatial frequency response, or transfer function, of an optical system, and $F_1(k_1,k_2)$ the frequency content of the input signal, the output at any point in the image plane is

$$f_2(x,y) = \int_{-\infty}^{\infty} F_1(k_1,k_2) A(k_1,k_2) e^{2\pi j(k_1 x + k_2 y)} dk_1 dk_2 \tag{5}$$

or equivalently the relationship between the spectra is

$$F_2(k_1,k_2) = F_1(k_1,k_2) A(k_1,k_2) \tag{6}$$

From Eqs. (2) and (6), the relationship between the spatial impulse and frequency responses can immediately be written as Fourier transforms or

$$\text{(a)}\quad A(k_1,k_2) = \int_{-\infty}^{\infty} a(x,y) e^{-2\pi j(k_1 x + k_2 y)} dx\, dy$$

$$\text{(b)}\quad a(x,y) = \int_{-\infty}^{\infty} A(k_1,k_2) e^{2\pi j(k_1 x + k_2 y)} dk_1 dk_2 \tag{7}$$

Since the Fourier transform of the spatial impulse function contains all spatial frequencies, the transfer function can be interpreted as characterizing the action of the system on each individual frequency component in a manner analogous to filtering in one-dimensional electrical systems.

OPTICAL SIGNALS: Assuming monochromatic illumination, a coherent system is defined as one in which the relative phases of the

propagating electromagnetic waves in various parts of the system are invariant with time. Until recent experiments with laser devices, this definition required that an object plane be simultaneously illuminated by a single point source for it was impossible to synchronize the phases from two or more point radiators. By definition, interference and diffraction effects due to phase differences are preserved in a coherent system, and the amplitudes of individual wave components are additive. Thus, with a coherent source and a system which preserves coherence, the input, output and system functions must all be expressed in terms of complex amplitudes for linear relations to hold between them.

When the object plane is self-luminous, or illuminated by a number of non-synchronous point radiators, the system is said to be incoherent. In this case, intensities are additive. Consequently, the input, output and system functions must be expressed in terms of intensities for linear relations to hold between them.

O'Neil, Hopkins, and Parrent discuss the intermediate case of "partial coherence," (15, 17, and 20). In this situation, an optical system is linear in "mutual intensity." An exhaustive discussion of this work is beyond the purview of this paper, and the reader is referred to the work of these investigators. For purposes of this discussion only, the conditions of linearity are assumed. The general term "spatial signal" will be used to denote either amplitudes, intensities, or mutual intensities and no consideration will be made as to the type of illumination except to assume that the system performs in a linear manner.

OPTICAL NOISE: In any practical electrical or optical system, two types of signals are present, the desirable signals or measurable attributes from physical objects of interest, and the undesirable signals or noise. The theory of one-dimensional electrical noise and its physical interpretation is firmly established in the study of electric circuits. Its extension to the analysis of electro-optical systems is beginning to receive attention in the literature. Jones has shown that eight types of electrical noises may be introduced by an infrared radiation detector when it is used as a transducer between two-dimensional electromagnetic energy and one-dimensional electrical signals (12).

On the other hand, there has been no satisfactory description of optical two-dimensional noise to date (15). However, in analyses of infrared imaging systems, many investigators have introduced the concept of two-dimensional background noise defined as undesirable signals in the input intensity distribution generated by radiation sources other than the objects of interest (8 and 13). Other investigators have attempted to relate noise theory to the study of granularities in photographic processes (11). A composite of noise sources used by these investigators is depicted in Fig. 3. The input and output desirable signals are represented by f_1 and f_2 ;

background, detector, amplifier, and granularity noise by n_1, n_2, n_3, n_4 respectively; the spatial impulse response of the optical input and output filters by a_1 and a_4; the detector impulse response by a_2; and the impulse response of the electrical circuitry or transmission path by a_3.

LINE SCAN FILTER MODEL: The basic model adopted in this paper is indicated in Fig. 4. A single scanning aperture, or possibly an aggregate of electro-optical collection devices which record a signal for processing, such as a_1, a_2, a_3 and a_4 in Fig. 3, is represented as a pre-filter. Thus, if $f_1(x,y)$ represents a spatial signal in an object plane, and $a_1(x,y)$ represents the overall spatial impulse response of the pre-filter, then from Eq. (2) the output for a linear system is

$$f_2(x,y) = \int_{-\infty}^{\infty}\int a_1(x-u, y-v) f_1(u,v)\,du\,dv \equiv a_1 * f_1 \qquad (8)$$

where the $*$ represents continuous convolution.

The sampler in Fig. 4 represents the scanning pattern which limits the observed data to a set of parallel lines. Thus, if the scan pattern permits continuous observation in the x coordinate, but at discrete positions in the y coordinate with spacing Δy, the output of the sampler on any line may be expressed as

$$\tilde{f}_2(x, m\Delta y) = \int_{-\infty}^{\infty}\int a_1(x-u, m\Delta y - v) f_1(u,v)\,du\,dv \qquad (9)$$

where m indicates the integer index of the sample line along the spatial coordinate, and the tilde denotes sampled functions which will be termed semi-discrete functions.

The post-filter is considered a two-dimensional semi-discrete, linear operator which has available sets of parallel data lines for simultaneous processing, and is depicted in Fig. 5. The input-output relationships of the post-filter are defined in terms of a two-dimensional semi-discrete convolution operation as

$$\tilde{f}_3(x, m\Delta y) = \int_{-\infty}^{\infty} \sum_{i=-\infty}^{\infty} \tilde{a}_2(x-u, m\Delta y - i\Delta y) f_2(u, i\Delta y)\,du\,\Delta y \qquad (10)$$

or $$\tilde{f}_3 = \tilde{a}_2 * \tilde{f}_2$$

The impulse response of a post-filter which processes continuously in only one dimension is immediately defined when

$$\tilde{a}_2(x, m\Delta y) = \begin{cases} \hat{a}_2(x) & \text{if } m = 0 \\ 0 & \text{if } m \neq 0 \end{cases} \qquad (11)$$

The circumflex in Eq. (11) indicates a normal continuous one-dimensional impulse response function.

If semi-discrete frequency functions are defined as

$$\tilde{F}(k_1, k_2) \equiv \int_{-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \tilde{f}(x, m\Delta y) e^{-2\pi j(k_1 x + k_2 m\Delta y)}\,dx\,\Delta y \qquad (12)$$

inversion formulae similar to those cited in Eq. (7) for continuous functions are derived in Ref. 4. as

$$\tilde{f}(x, m\Delta y) = \int_{-\infty}^{\infty} \int_{-\frac{1}{2\Delta y}}^{\frac{1}{2\Delta y}} \tilde{F}(k_1, k_2) e^{2\pi j (k_1 x + k_2 m \Delta y)} dk_2 \, dk_1 \quad (13)$$

Further, Eq. (10) may be expressed simply as

$$\tilde{F}_3(k_1, k_2) = \tilde{F}_2(k_1, k_2) \tilde{A}_2(k_1, k_2) \quad (14)$$

where $\tilde{F}_3$, $\tilde{F}_2$ and $\tilde{A}_2$ are the semi-discrete transforms of $\tilde{f}_3$, $\tilde{f}_2$ and $\tilde{a}_2$ respectively. The relationship between the continuous and semi-discrete spectra is

$$\tilde{F}(k_1, k_2) = \sum_{m=-\infty}^{\infty} F\left(k_1, k_2 + \frac{m}{\Delta y}\right) \quad (15)$$

<u>SYSTEM OPTIMIZATION</u>: Many different criteria, both statistical and deterministic are utilized in the optimization of linear electrical filters. Their application to the design of optical systems has just begun to receive attention in the literature. Many of the limitations of deterministic theories posed by Wiener and others apply equally to optical filters and have been discussed by Elias, Jones, and O'Neill (8, 9, 13, 15). On the other hand, statistical description of spatial signals and noise as true representation of physical processes have not attained the same level of acceptance as those used for electrical signals and noise. Therefore, the author decided to investigate the performance of one and two-dimensional post-filters under two different criterion, the mean square error as a statistical criterion, and spatial resolution as a deterministic criterion more closely associated with conventional optics.

It is shown in Ref. 4. that if Fig. 4 is modified as depicted in Fig. 6, and a system error is defined as

$$\tilde{\epsilon} = \tilde{f}_1 - \tilde{f}_3 = \tilde{f}_1 - \left(\tilde{f}_2 + \tilde{n}\right) \tilde{*} \, \tilde{a}_2 \quad (16)$$

then the optimum one and two-dimensional post-filters and their corresponding least mean square errors are

$$\tilde{A}_2(k_1, k_2) = \frac{\sum_m \bar{A}_1\left(k_1, k_2 + \frac{m}{\Delta y}\right) G_{f_1}\left(k_1, k_2 + \frac{m}{\Delta y}\right)}{\sum_m \left|A_1\left(k_1, k_2 + \frac{m}{\Delta y}\right)\right|^2 G_{f_1}\left(k_1, k_2 + \frac{m}{\Delta y}\right) + \tilde{G}_m(k_1, k_2)} \quad (17)$$

$$E\left[\tilde{\epsilon}^2\right] = \int_{-\infty}^{\infty} \int_{-\frac{1}{2\Delta y}}^{\frac{1}{2\Delta y}} \sum_m G_{f_1}\left(k_1, k_2 + \frac{m}{\Delta y}\right) dk_2 \, dk_1 - \int_{-\infty}^{\infty} \int_{-\frac{1}{2\Delta y}}^{\frac{1}{2\Delta y}} \frac{\left|\sum_m \bar{A}_1\left(k_1, k_2 + \frac{m}{\Delta y}\right) G_{f_1}\left(k_1, k_2 + \frac{m}{\Delta y}\right)\right|}{\sum_m \left|A_1\left(k_1, k_2 + \frac{m}{\Delta y}\right)\right|^2 G_{f_1}\left(k_1, k_2 + \frac{m}{\Delta y}\right) + \tilde{G}_m} dk_2 \, dk_1 \quad (18)$$

$$\hat{A}_2(k_1) = \frac{\int_{-\infty}^{+\infty} \sum_m \bar{A}_1\left(k_1, k_2 + \frac{m}{\Delta y}\right) G_f\left(k_1, k_2 + \frac{m}{\Delta y}\right) dk_2}{\Delta y \int_{-\frac{1}{2}\Delta y}^{\frac{1}{2}\Delta y} \sum_m \left[G_{f_1}\left(k_1, k_2 + \frac{m}{\Delta y}\right) \left| A_1\left(k_1, k_2 + \frac{m}{\Delta y}\right) \right|^2 + \tilde{G}_m \right] dk_2} \tag{19}$$

$$E\left[\hat{\epsilon}^2\right] = \int_{-\infty}^{\infty} \int_{-\frac{1}{2}\Delta y}^{\frac{1}{2}\Delta y} \sum_m G_{f_1}\left(k_1, k_2 + \frac{m}{\Delta y}\right) dk_2\, dk_1 \tag{20}$$

$$- \int_{-\infty}^{\infty} \frac{\left| \int_{-\frac{1}{2}\Delta y}^{\frac{1}{2}\Delta y} \sum_m \bar{A}_1\left(k_1, k_2 + \frac{m}{\Delta y}\right) G_{f_1}\left(k_1, k_2 + \frac{m}{\Delta y}\right) dk_2 \right|^2}{\int_{-\frac{1}{2}\Delta y}^{\frac{1}{2}\Delta y} \left[\sum_m G_{f_1}\left(k_1, k_2 + \frac{m}{\Delta y}\right) \left| A_1\left(k_1, k_2 + \frac{m}{\Delta y}\right) \right|^2 + \tilde{G}_m(k_1, k_2) \right] dk_2} dk_1$$

This derivation is based on the definition of ensemble correlation functions for spatially stationary signals of the type

$$\tilde{R}_{f_1}(x, m\Delta y) \equiv E\left[\tilde{f}_1(x+\alpha, m\Delta y + p\Delta y)\, \tilde{f}_1(\alpha, p\Delta y)\right] \tag{21}$$

where E denotes the ensemble expectations. From Fourier relationships of the type indicated in Eqs. (12)and (13), it is easy to see that if $\tilde{G}_\epsilon(k_1, k_2)$ is power density function, or transform of R_ϵ, then the relationship between the mean square error and $\tilde{G}_\epsilon(k_1, k_2)$ is from Eq. (21)

$$E\left[\tilde{\epsilon}^2\right] = \tilde{R}_\epsilon(0, 0) = \int_{-\infty}^{\infty} \int_{-\frac{1}{2}\Delta y}^{\frac{1}{2}\Delta y} \tilde{G}_\epsilon(k_1, k_2)\, dk_2\, dk_1 \tag{22}$$

Thus, as shown in Ref. 4., the optimization process reduces to finding the algebraic relationship between the power density functions of the signal, noise and filters which minimize ϵ.

As an application of the theory presented above, let the pre-filter consist of a single, two-dimensional aperture limited by system self-noise generated by a continuous two-dimensional white noise process. Let the pre-filter have a top hat spatial frequency response, i.e.,

$$A_1(k_1, k_2) = \begin{cases} 1 & \sqrt{k_1^2 + k_2^2} < \frac{\beta}{2} \\ 0 & \text{otherwise} \end{cases} \tag{23}$$

This response is characteristic of some types of physical coherent scanning systems. Let the input signal be composed of detail with a constant spectra and no background noise. For example, let

$$G_{f_1}(k_1, k_2) = C \tag{24}$$

It is shown that in Ref. 4. that a two-dimensional white noise process can be made statistically equivalent to a semi-discrete noise process on corresponding sample lines if

$$G_n(k_1, k_2) = \begin{cases} \Delta y\, \tilde{G}_m(k_1) & |k_1| < \frac{1}{2}\Delta y \\ 0 & |k_1| > \frac{1}{2}\Delta y \end{cases} \tag{25}$$

Therefore, if N represents a constant one-dimensional white noise spectrum, solution of Eqs. (17) to (20) require the substitution

$$\tilde{G}_n(k_1, k_2) = \Delta y\, N \tag{26}$$

Finally, to simplify the solution, it is assumed that

$$A_1(k_1, k_2) = 0 \qquad |k_1| > \frac{1}{2\Delta y} \tag{27}$$

and

$$G_{f_1}(k_1, k_2) = 0 \qquad \text{whenever} \quad A_1(k_1, k_2) = 0 \tag{28}$$

Substituting Eq. (23) to (28) in Eq. (17), it is seen that the optimum two-dimensional post-filter is

$$\tilde{A}_2(k_1, k_2)_{opt} = \frac{BC}{BC + N} \qquad \sqrt{k_1^2 + k_2^2} < \frac{B}{2} \tag{29}$$

and from Eq. (19), the optimum one-dimensional filter is

$$\hat{A}_2(k_1)_{opt} = \frac{B^2 C}{BC + N} \tag{30}$$

From Eqs. (18) and (20), the corresponding least square errors are

$$E\left[\tilde{\epsilon}^2\right]_{min} = \frac{\pi C B^2}{4}\left(1 - \frac{C}{C - \frac{N}{B}}\right) \tag{31}$$

and

$$E\left[\hat{\epsilon}^2\right]_{min} = C B^2 \sqrt{\frac{N}{N+BC}}\, \tan^{-1}\sqrt{\frac{N}{N+BC}} \tag{32}$$

When the noise is very small, the approximate least square errors are

$$E\left[\tilde{\epsilon}^2\right] \simeq \frac{\pi B N}{4} \qquad [N << BC] \tag{33}$$

and

$$E\left[\hat{\epsilon}^2\right] \simeq B N \qquad [N << BC] \tag{34}$$

Thus, the ratio of the least square errors is

$$\frac{E\left[\tilde{\epsilon}^2\right]}{E\left[\hat{\epsilon}^2\right]} \simeq \frac{\pi}{4} \tag{35}$$

Next a cone frequency response is considered for the pre-filter. This response has application to non-coherent scanning systems when an optical aperture limits the resolution, and is

expressed as

$$A_1(k_1,k_2) = B - \sqrt{k_1^2+k_2^2} \qquad \left[\sqrt{k_1^2+k_2^2} < B\right] \quad (36)$$

When the noise is very small, the approximately least square errors are

$$E\left[\tilde{\epsilon}^2\right]_{min} \simeq \frac{\pi^2}{\sqrt{2}} B\sqrt{NBC} \qquad \left[N << \frac{BC}{6}\right] \quad (37)$$

$$E\left[\hat{\epsilon}^2\right]_{min} \simeq \frac{\pi B^2 C}{3} + 16BN \qquad \left[N << \frac{BC}{6}\right] \quad (38)$$

Thus, the ratio of the least square errors is

$$\frac{E\left[\tilde{\epsilon}^2\right]_{min}}{E\left[\hat{\epsilon}^2\right]_{min}} \simeq \frac{3\pi}{\sqrt{2}}\sqrt{\frac{N}{BC}} \quad (39)$$

It is noted that in the case of no noise, both the one and two-dimensional least square errors with the top hat pre-filter and the two-dimensional least square error for the cone pre-filter reduce to zero. However, in the case of the one-dimensional least square error for the cone pre-filter, there is a constant term which depends only on the signal.

The notion of "resolving power" or resolution has traditionally been associated with the evaluation of the performance of optical systems. Resolution most generally denotes the ability of a system to separate closely related forms or entities or the degree to which they can be discriminated. The classical definition of optical resolution in incoherent imaging systems, such as a telescope, when objects are at a great unknown distance from the collecting aperture, is based on the Rayleigh criterion. For the ideal circular aperture whose spatial frequency response is limited only by diffraction effects, the limiting resolution distance occurs when the maximum amplitude of the impulse response from one point source coincides with the first zero of the impulse response from an adjacent point source. When the two point sources are illuminated equally, the Rayleigh criterion can be simply expressed as the "effective width" or distance from the maximum impulse response to the first zero. Simple definitions of resolution are also used in the evaluation of a pulsed one-dimensional electrical system such as radar. In radar systems with a fixed carrier frequency, one-half the pulse width is commonly used as a measure of the ability to discriminate between responses from closely spaced targets. Thus, the notion of resolution is intrinsically related to the "effective width" of the impulse response.

Beyond these simple resolution measures, more complicated definitions have been developed both in optics and in electrical

systems which take into account background noise, contrast ratios, pulse frequency modulation, and other physically important considerations. To avoid the difficulties attendant to these formulations, and to still permit a quantitative comparison of one and two-dimensional post-filters, a spatial resolution measure will be developed which does not take into account system self-noise, background noise, or background contrasts. The general approach adopted here is that resolution is directly related to the "effective volume" of the overall system impulse response. Thus, if $\tilde{a}_3(x, m\Delta y)$ represents the overall response, or the convolution of the two-dimensional pre-filter and post-filter impulse responses, and with a maximum value at $(0,0)$, a resolution measure $\tilde{\rho}$ is defined as the rectangular solid leaving the same volume as the overall impulse response function squared, or

$$\tilde{\rho} \equiv \frac{\int_{-\infty}^{\infty} \sum_{m} |\tilde{a}_3(x, m\Delta y)|^2 dx \Delta y}{|\tilde{a}_3(0,0)|^2} \tag{40}$$

By use of Parseval's theorm and a two-dimensional version of the Schwarz inequality, it is proven in Ref. 4. that a pre-filter with the cone frequency response indicated in Eq. (36) has a resolution of

$$\tilde{\rho} \geq \frac{1}{\pi B^2} \tag{41}$$

Equality occurs in Eq. (41) when

$$\tilde{A}_2(k_1, k_2) = \frac{B}{B - \sqrt{k_1^2 + k_2^2}} \tag{42}$$

Similarly, with a cone pre-filter and a one-dimensional post-filter, the resolution is

$$\hat{\rho} \geq \frac{3}{2\pi B^2} \tag{43}$$

Equality occurs in Eq. (43) when

$$\hat{A}_2(k_1) = \left[2B - \sqrt{k_1^2 + B^2} - \frac{k_1^2}{B} \log \frac{B + \sqrt{k_1^2 + B^2}}{k_1^2} \right] \tag{44}$$

If the post-filters indicated in Eqs. (42) and (44) are considered optimum in the sense they produce equalities, then the relative advantage of two-dimensional post-filtering with a cone pre-filter is

$$\frac{\tilde{\rho}}{\hat{\rho}} = \frac{2}{3} \tag{45}$$

CONCLUSION: When a line scan imaging system is not required to reproduce motion, it is more advantageous to treat the system as a semi-discrete sampler, and to permit two-dimensional spatial filtering of the recorded data than to treat the system as a one-dimensional temporal modulation which records only one-dimensional signals.

BIBLIOGRAPHY

1. Born, M., and Wolfe, E., Principles of Optics, 1st Ed., New York, New York: Pergamon Press, 1959.
2. Brown, William M., Analysis and Optimization of Coherent High Resolution Radars by Circuit Theory Concepts, Report Number 2900-144-T, Willow Run Laboratories, The University of Michigan, Ann Arbor, Michigan, May 1960.
3. Brown, William M., Analysis of Linear Time Invariant Systems, Book in preparation by McGraw Hill.
4. Callahan, Leslie G., Jr., Optimum Linear Filtering for Line Scan Imaging Systems, Ph. D. Dissertation, The University of Pennsylvania, Philadelphia, Pennsylvania, 1961.
5. Cheatham, T. P., Jr., and Kohlenberg, A., "Analysis and Synthesis of Optical Systems" Boston University, Physics Research Laboratory, Technical Note 84, Part I, March 1952.
6. Cheatham, T. P., Jr., and Kohlenberg, A., "Optical Filters - Their Equivalence to and Differences from Electrical Networks," IRE National Convention Record, pp 6-12, 1954.
7. Cutrona, L. J., Leith, E. N., Palermo, C. J., and Percello, L. J., "Optics Data Processing and Filtering Systems," IRE Trans. on Information Theory, Vol. IT-6, Number 3, pp 386-400, June 1960.
8. Elias, P., "Optics and Communication Theory," J. Opt. Soc. Amer., Vol. 43, pp 229-232, April 1953.
9. Elias, P., "Optical Systems as Communication Channels," Symposium on Info Networks, Brooklyn Poly Inst., April 12-14, 1954.
10. Elias, P., Grey, D., and Robinson, D., "Fourier Treatment of Optical Processes, J. Opt. Soc. Amer., Vol. 42, pp 127-134, February 1952.
11. Jones, R. Clark, "New Method of Describing and Measuring the Granularity of Photographic Materials," J. Opt. Soc. Amer., Vol. 45, pp 799-808, 1955.
12. Jones, R. Clark, "Noise in Radiation Detectors," Proceedings of the IRE, Vol. 47, Number 9, Infrared Issue, September 1959.
13. Jones, R. Clark, The Study of IR Backgrounds by the Wiener Spectrum Method, Polaroid Corporation, Cambridge, Massachusetts, unpublished report prepared and edited by Infrared Laboratory, Willow Run Laboratories, The University of Michigan, Ann Arbor, Michigan, December 1959.
14. Mertz, P., and Gray, F., "A Theory of Scanning and its Relationship to the Characteristics of the Transmitted Signal in Telephotography and Television," Bell Syst. Tech. Journal 13, pp 464-515, 1934.

15. O'Neill, E., Selected Topic in Optics and Communication Theory, Boston University, Technical Note 133, October 1957.

16. O'Neill, E., "Spatial Filtering in Optics," IRE Trans. on Info Theory, Vol. IT-2, pp 56-65, June 1956.

17. O'Neill, E., The Analysis and Synthesis of Linear Coherent and Incoherent Imaging System, Boston University, Physics Research Laboratory, Technical Note 122, September 1955.

18. O'Neill, E., The Modulation Function in Optics, Boston University, Opt. Res. Lab., Technical Note 110, January 1954.

19. Otterman, J., Aperture Correction for Instrumentation Systems, Ph. D. Dissertation, University of Michigan, 1956.

20. Parrent, George B., Jr., "Imaging of Extended Polychromatic Sources and Generalized Transfer Functions," J. Opt. Soc. Amer., Vol. 51, pp 143-151, February 1961.

21. Shade, O., "Image Gradation, Graininess and Sharpness in Television and Motion Picture Systems," J. Soc. Mot. Pict. and Telev. Engrs., Vol. 56, pp 131-171 (1951); Vol. 58, pp 181-222 (1952); Vol. 61, pp 97-164 (1953); Vol. 64, pp 593-617 (1955).

22. Wessley, H., "Optimum Linear Spatio-Temporal Filter," Proceedings of IRIS, Institute of Science and Technology, University of Michigan, Ann Arbor, Michigan, Vol. 3, Number 2, pp 135-141 (1958).

FIGURES

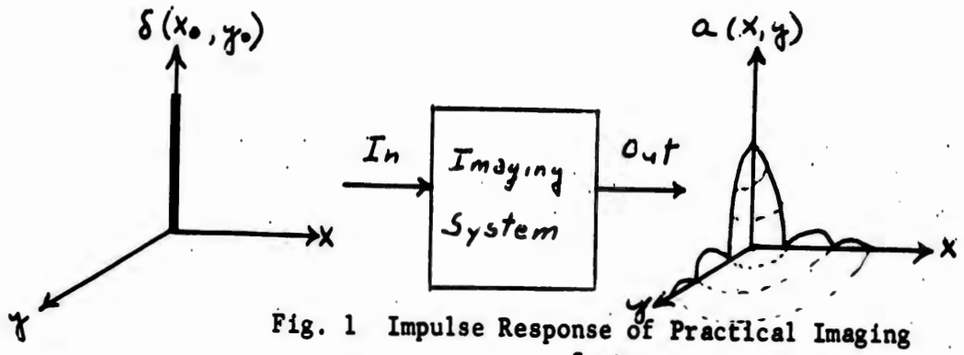


Fig. 1 Impulse Response of Practical Imaging System

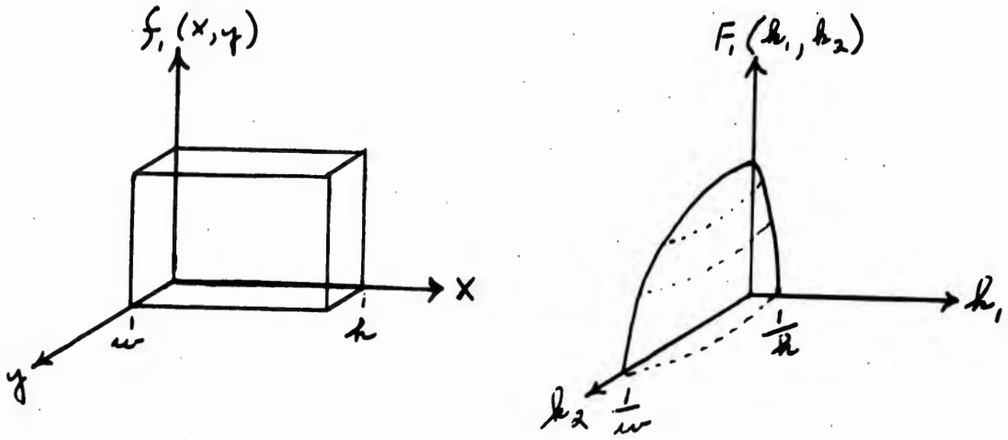


Fig. 2 Spatial Frequency Spectra

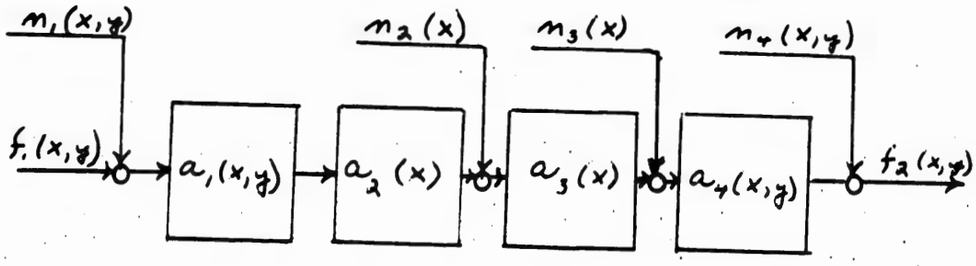


Fig. 3 Noise Sources in Electro-Optical Imaging Systems

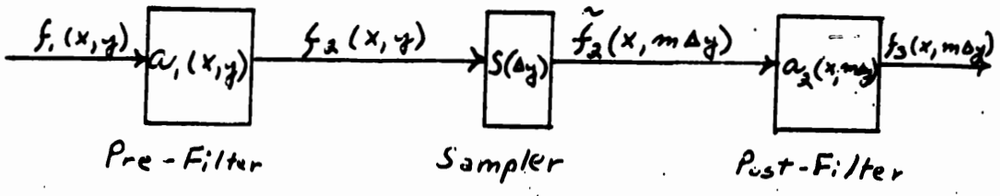


Fig. 4 Line Scan Filter Model

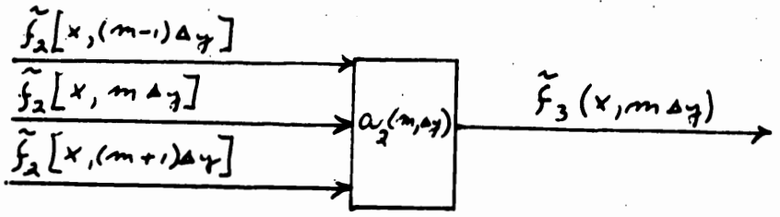


Fig. 5 Two-Dimensional Semi-Discrete Linear Filter

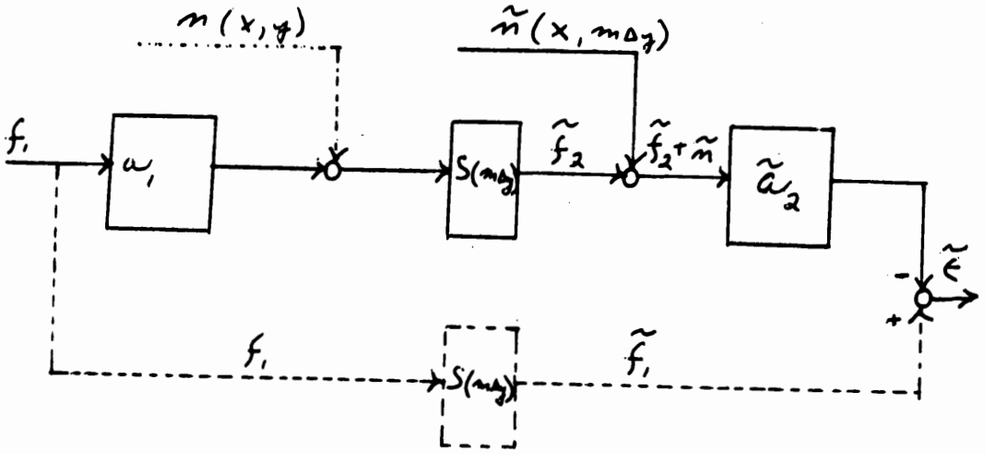


Fig. 6 Least Square Optimization Model

APPLICATION OF THE PRINCIPLES OF FRACTURE MECHANICS TO RECOILLESS RIFLE PROBLEMS

CARL M. CARMAN
FRANKFORD ARSENAL
PHILADELPHIA, PENNSYLVANIA

The consistent demand for reduced weight in recoilless rifles necessitated the use of materials having a high strength-to-weight ratio. It is well known that such materials are subject to brittle failure as exemplified by early attempts to use these materials in certain rocket motor cases. The standard "V" notch Charpy impact test, so successfully used for studying lower strength materials, fails to give meaningful data when applied to these higher strength steels. The Charpy impact values are uniformly low and the transition temperature becomes ill-defined.

MATHEMATICAL MODEL[1]

Since a primary consideration in selecting materials for the recoilless rifle application is their ability to resist crack propagation, the application of fracture mechanics to this problem appeared feasible. Fracture mechanics is a branch of classical mechanics describing the behavior of solids in the presence of a flaw and is based on a mathematical model which assumes infinite strain elasticity theory. This suggests that the basic mathematical model should, at first, neglect the influence of plastic strains and should regard the crack simply as an interior free surface in an elastic solid. The choice of a model is also influenced by the fact that the events at the leading edge of the crack are complex. A careful allowance for the influence upon the stress field of plastic strains and advanced initiations is not practical.

[1]This discussion of the mathematical model has been abstracted from the publications of Dr. G. R. Irwin(1,2,3).

In terms of this model, the force conjugate to speed of crack extension is the rate of dissipation of strain energy with crack extension termed $\mathcal{G}$. It is also possible and convenient to define a stress field parameter K, proportional to the elevated stresses near the crack edge, in such a way that

$$E\mathcal{G} = K^2 \text{ for plane stress} \qquad \text{(Eq. 1)}$$

$$E\mathcal{G} = (1 - \nu^2)\, K^2 \text{ for plane strain} \qquad \text{(Eq. 2)}$$

where E is Young's modulus and ν is Poisson's ratio.

From the mathematical viewpoint there are three crack-surface-displacement modes of crack extension; the opening mode with crack-extension force $\mathcal{G}_I$ and stress intensity K_I, the forward sliding mode with parameters $\mathcal{G}_{II}$ and K_{II}, and the parallel sliding mode with parameters $\mathcal{G}_{III}$ and K_{III}.

Fracture failures of structural metals are associated primarily with opening mode or tensile type stress system. When the fracture is entirely a flat transverse tensile break with negligible plastic contraction parallel to the leading edge of the crack, use of the parameters $\mathcal{G}_I$ and K_I is appropriate.

For a crack traversing a plate under conditions such that plastic contraction of the plate thickness accompanies crack extension, oblique shear fractures form adjacent to the free surfaces and, in the absence of a strain rate sensitivity in the flow stress, these tend to grow in proportion to the $\mathcal{G}$ value as the crack lengthens. For a relatively high degree of crack toughness several plate thicknesses of crack extension from a sharp notch or fatigue crack produce enough growth of the shear lip borders so that a single oblique shear fracture thereafter covers the entire fracture face.

The mathematical representation must, for simplicity, consist of a linear elasticity crack stress field analysis. Thus, the plastic yielding which causes various degrees of oblique shear fracture is not representable in the stress field of the basic model. The best that can be done is to assume an opening mode crack whose length is such that the stress field of the model most nearly fits the real crack stress field in the elastic region surrounding the crack.

The accuracy with which the idealized stress field can be fitted to the real stress field increases with decrease of plastic zone relative to crack and specimen dimensions. Thus, conceptually, the misfit at any fixed fraction of crack length from the leading edge can be made as small as desired by increasing crack and specimen dimensions.

Irwin has applied the Westergaard system of stress analysis to a Griffith type crack in a finite width specimen and has derived the following equation for the stress field parameter, K:

$$K = \sqrt{\sigma^2 W \tan \frac{\pi a}{W}} \qquad \text{(Eq. 3)}$$

where σ = gross section stress
W = width of the specimen
a = 1/2 the crack length

This equation is subjected to the same limitations as the mathematical model, namelv that of elastic stress field analysis. It is possible to alter this expression to account for inelastic strains at the tip of the advancing crack front. Such corrections are subject to criticism from both the theoretical and practical viewpoints. In view of the practical importance of the situations where the plastic deformation is not small compared to the crack length, it is advantageous to have available an instrument which could measure the effective crack length (actual crack length plus the effect of plastic deformation) versus applied load. Such an instrument is available as a result of recent work at the Naval Research Laboratory and Frankford Arsenal.

It is the basic premise of fracture mechanics that the crack edge elastic stress field is known as a function of the applied stress and crack length. In a similar manner the displacements in the elastic stress field are also known as a function of these two variables.

EXPERIMENTAL PROGRAM

The experimental equipment is shown in Figure 1. The two clips mounted on the specimen measure the specimen compliance and the load cell measures the load applied to the specimen. Prior to the testing of each specimen, the clips and their associated strain gage circuits are calibrated for each 0.001 inch displacements. During test this information is fed into the x-axis of the x-y plotter. The output of the load cell is fed into y-axis of the same instrument.

The material used for the experimental phase of this investigation was the new high strength gun steel, 4330V (Mod+Si), developed by Frankford Arsenal and has the following composition: C 0.31, Mn 0.87, S 0.004, P 0.007, Cr 0.87, Ni 1.74, Mo 0.45, V 0.10 and Si 1.45. The steel was received as 0.120 inch thick hot rolled sheet. The standard tensile values of this sheet after oil quenching from 1650°F and tempering at 600°F are given in Table I.

Table I

Tensile Values of 4330V (Mod+Si) Steel

Direction	Yield Strength (0.10% Offset) PSI	Yield Strength (0.20% Offset) PSI	Tensile Strength PSI	Elong. %
Longitudinal	195,000	206,000	248,700	7.9
	195,000	206,000	247,800	7.3
Transverse	195,000	206,000	247,100	7.4
	195,000	206,000	247,500	6.5

Duplicate longitudinal and transverse centrally notched fracture toughness specimens were machined from this sheet. Specimen widths of two, three, four, five, and five and three-quarters inches were used. The two, three, and four inch wide specimens were tested in specially designed pin loaded grips with the electronic compliance gage. Due to high values of load required, the five and five and three-quarters inch wide specimens were tested using Templin wedge grips and so the compliance gage could not be used.

A tracing of a typical load-deflection curve is shown in Figure 2. It will be observed that this curve rises linearly to a certain value of the load and then shows a gradual departure from linearity. The point of departure from linearity is identified with a discernible burst of crack growth from the notch tip. The fracture toughness, calculated at this point, can be closely identified with that developed under known conditions of plane strain[4]. These data obtained for the two, three, and four inch wide specimens are given in Table II.

This is an attempt to measure plane strain values on a plane stress specimen so that the value of the correction for the plastic strain zone at the tip of the crack is unsettled. For the present, the values of K_{Nc} have been computed using the plane stress correction. If these values had been computed using the plane strain calculation the plastic zone at the crack tip would be smaller in the order of $\frac{1}{2\sqrt{2}}$. The values of K_{Nc} are fairly consistent over the range of specimen sizes tested.

In Figure 3 the driving force and crack resistance values are plotted against the absolute crack extension $(a - a_o)$. It will be observed that the crack resistance values increase rather rapidly with the initial crack growth. The higher values of fracture resistance on these curves are shown dotted since they represent the lower bound of fracture toughness values as determined from the fast fracture of the five and five and three-quarters inch wide specimens.

It will be observed that the driving force curves for the three and four inch wide specimens become tangent to the resistance curve at very low or quite conservative values of fracture toughness. The driving force curves for the five and five and three-quarters inch specimens become tangent to the resistance curve at its more steeply rising portion and at higher values of fracture toughness. These values of fracture toughness are considered to be more characteristic of the material tested.

The fast fracture data obtained for these specimens is given in Table III.

These data show that this steel is extremely tough. The high values of the net section stress divided by the yield strength show that most of the specimens were too narrow for reasonable measurements of fracture toughness and lay beyond the scope of the mathematical model. Also the fractures of these specimens were characterized by necking. The transverse specimens five and five and three-quarters inch wide showed lower net section stress, and it was possible to compute a fracture toughness value from these data.

These calculations gave a K_c value ranging from 310,000 psi $\sqrt{in}$ to 350,000 psi $\sqrt{in}$

When such high values of fracture toughness are obtained it is desirable to see if such values are predictable. Irwin[5] has proposed the following equation for estimating the plane stress fracture toughness based on plane strain fracture values:

$$K_c^2 = K^2_{Ic} \left[1 + 1.4\ \beta^2_{Ic}\right] \qquad \text{(Eq. 4)}$$

where β_{Ic} is the relative plastic zone size under plane strain conditions and is equal to $\frac{K^2_{Ic}}{B\sigma_{ys}^2}$. Using an average transverse value of K_{Nc} equal to 111,200 psi $\sqrt{in}$ from Table II, these calculations predict a fracture toughness value of 337,000 psi $\sqrt{in}$ in the transverse directions. This predicted value agrees closely with the experimental data.

The recoilless rifle is a carefully made ordnance item which is fully inspected and can tolerate only a short crack. Under these conditions, the plastic zone size is not small compared with the crack dimensions and the simple Griffith relationship $K_c^2 = \pi\sigma^2 a$ does not apply. However, the analysis may be modified to more closely conform to the actual situation as shown in the following equations. The radius of plastic zone size is given by:

$$r_y = \frac{K_c^2}{2\pi\sigma_{ys}^2} \qquad \text{(Eq. 5)}$$

so that the elastic stress field sees a crack whose half length is equal to

$$a + \frac{K_c^2}{2\pi\sigma_{ys}^2} \qquad \text{(Eq. 6)}$$

so that

$$K_c^2 = \frac{\pi\sigma^2 a}{1-1/2\left(\frac{\sigma}{\sigma_{ys}}\right)^2} \qquad \text{(Eq. 7)}$$

Examination of this equation shows that it requires considerably more fracture toughness to affect fracture arrest when the stress level is near the yield strength.

From these laboratory data it would appear that this steel should give satisfactory service in recoilless rifles. However, it is not possible to directly specify fracture toughness values for recoilless rifle service since sufficient data are not available to correlate fracture toughness and service performance at this time. Therefore, models of the tube-chamber section of the 90mm T234 weapon were made and tested hydrostatically to failure. This experiment was justified since these high strength materials are not particularly strain rate sensitive. A picture of one of these failed models is shown in Figure 4. It will be seen that failure occurred by general yielding and that fracture initiation and propagation were by shear. This shows that if failure were to occur in one of these weapons, fracture would be of the ductile or high energy absorbing type as predicted from the various fracture toughness experiments.

It is recommended that the compilation of data be continued on high strength steels to eventually permit correlation of laboratory fracture toughness measurements with service life and performance of recoilless rifles.

Table II

Plane Strain Fracture Toughness of 4330V (Mod+Si) Steel

Direction	Specimen Width (inch)	Crack Initiation Load (lbs)	Stress (psi)	a_o (in)	$\frac{\pi a}{W}$	K_{Nc}
Longitudinal	2	22,860	98,500	.336	.528	115,200
	2	24,560	109,600	.335	.526	132,400
	3	28,700	81,800	.516	.540	116,200
	3	29,000	82,600	.519	.544	120,100
	4	35,000	71,700	.674	.529	113,800
	4	49,290	101,000	.680	.534	170,800
Transverse	2	23,160	99,000	.335	.526	116,300
	2	21,740	92,100	.337	.529	107,400
	3	20,000	56,500	.536	.660	79,500
	3	25,500	72,600	.534	.624	103,700
	4	40,790	91,900	.672	.528	150,400
	4	33,900	69,500	.672	.528	109,900

a_o = 1/2 the initial crack length

K_{Nc} = fracture toughness values corresponding to plane strain conditions

Table III

Fracture Toughness Data for 4330V (Mod+Si) Steel

Direction	Specimen Width (in)	Breaking Load (lbs)	Gross Section Stress (psi)	a (in)	$\frac{\pi a}{W}$	K_c psi $\sqrt{in}$	Net Section Stress / Yield Strength
Long.	2	30,600	131,900	0.507	0.797	Does not fit model	1.31
	2	30,000	133,900	0.548	0.860	"	1.44
	3	43,600	124,200	0.792	0.829	"	1.28
	3	43,700	124,500	0.782	0.819	"	1.27
	4	61,800	126,600	0.973	0.764	"	1.20
	4	56,800	135,200	0.920	0.723	"	1.22
	5	75,000	126,100	1.210	0.760	"	1.19
	5	76,000	126,700	1.228	0.772	"	1.21
	5.75	86,000	126,800	1.381	0.755	"	1.19
	5.75	84,000	125,900	1.359	0.742	"	1.17
Trans.	2	28,200	120,500	0.515	0.809	"	1.21
	2	28,200	119,500	0.501	0.787	"	1.17
	3	38,500	108,800	0.798	0.836	"	1.13
	3	37,100	105,700	0.804	0.879	"	1.17
	4	52,100	117,300	1.019	0.800	"	1.17
	4	57,500	117,800	-	-	-	-
	5	67,000	111,700	1.224	0.769	312,000	1.07
	5	69,000	116,900	1.221	0.767	344,800	1.17
	5.75	83,000	121,300	1.297	0.709	345,000	1.08
	5.75	82,000	118,800	1.341	0.733	350,000	1.09

a = 1/2 total extent of starting crack + slow crack extension
K_c = fracture toughness

REFERENCES

1. Irwin, G. R., "Relation of Crack Toughness Measurements to Practical Applications", AWS - ASME Meeting, Cleveland, 1962.

2. Irwin, G. R., "Fracture", Springer 'Encyclopedia of Physics', VI, 1958, PP 551-590.

3. Irwin, G. R., J. A. Kies and H. L. Smith, "Fracture Strengths Relative to Onset and Arrest of Crack Propagation", Proc. ASTM, 58, P 640, 1958.

4. Boyle, R. W., A. M. Sullivan and J. M. Krafft, "Determination of Plane Strain Fracture Toughness with Sharply Notched Sheets", AWS - ASME Meeting, Cleveland, 1962.

5. Irwin, G. R. and J. A. Kies, "Fracture Theory as Applied to High Strength Steels for Pressure Vessels", Golden Gate Metals Conference, San Francisco, Calif., Feb. 4-6, 1960, and ASM, Metals Progress, 78, No. 2, Aug. 1960.

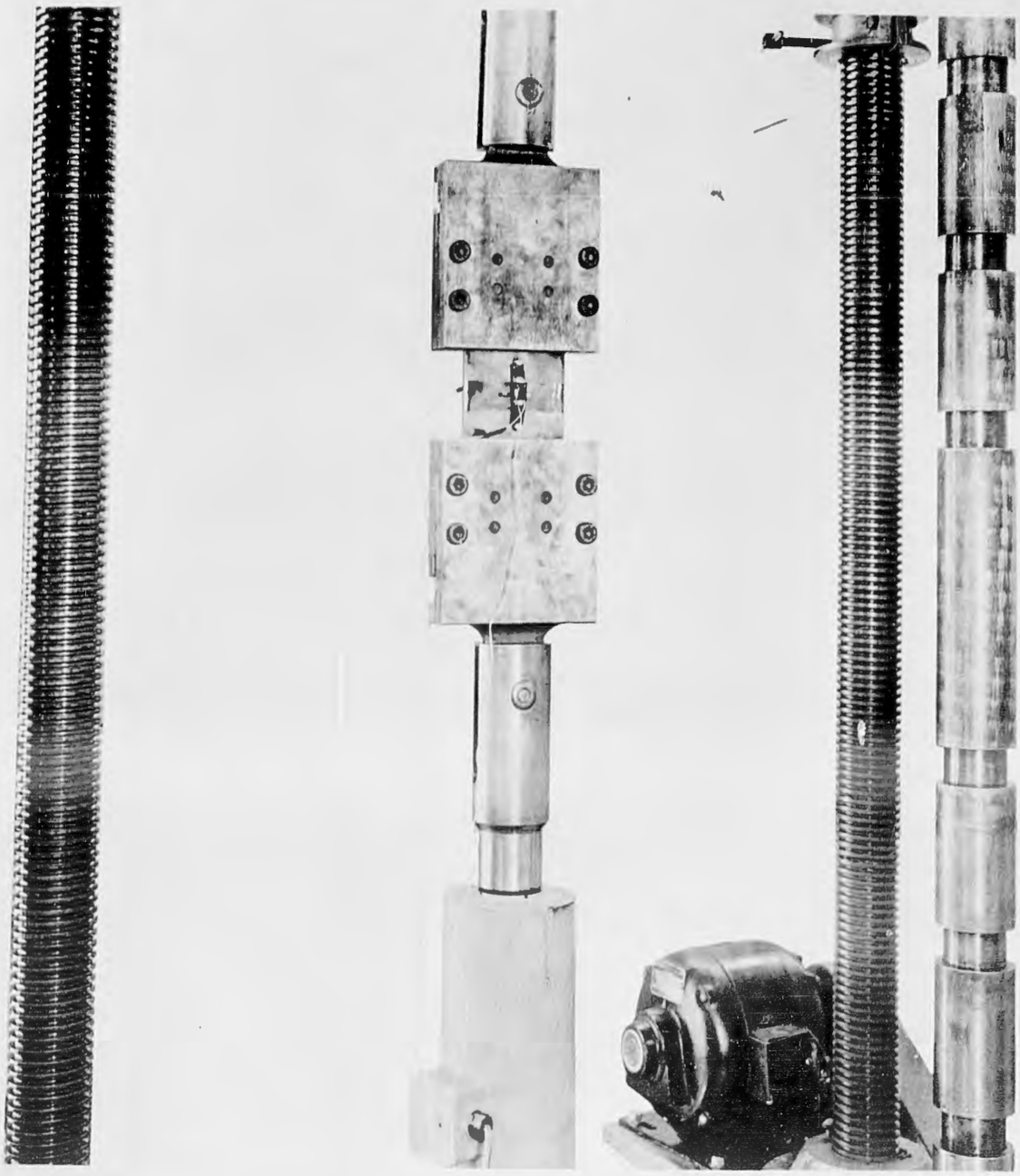


Figure 1. Photograph of the specimen arrangement for electronic compliance gage.

36·231·50752·0RD·62·

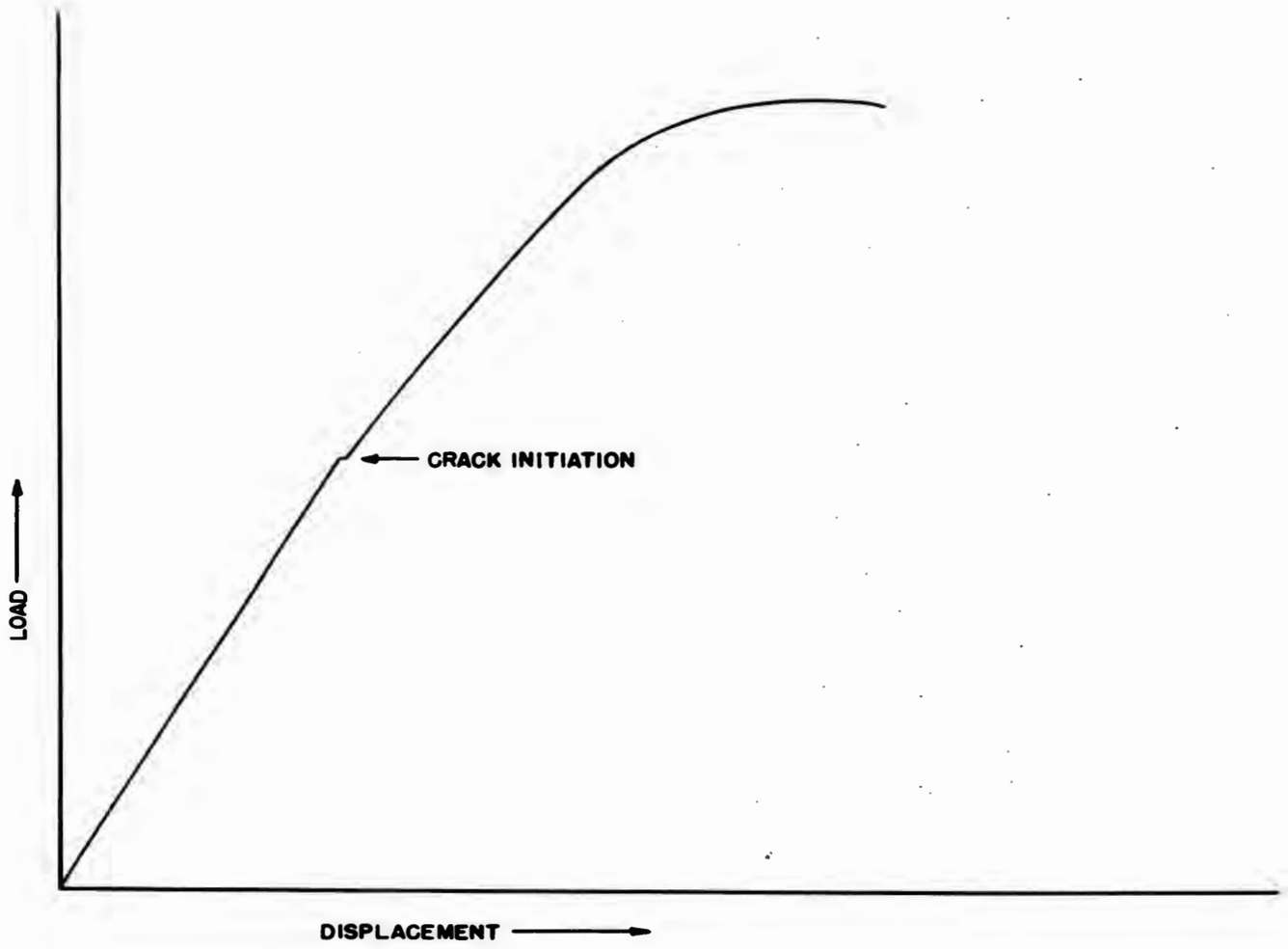


Figure 2. Tracing of load-displacement curve showing crack initiation.

36.231.5094.0.ORD.62.

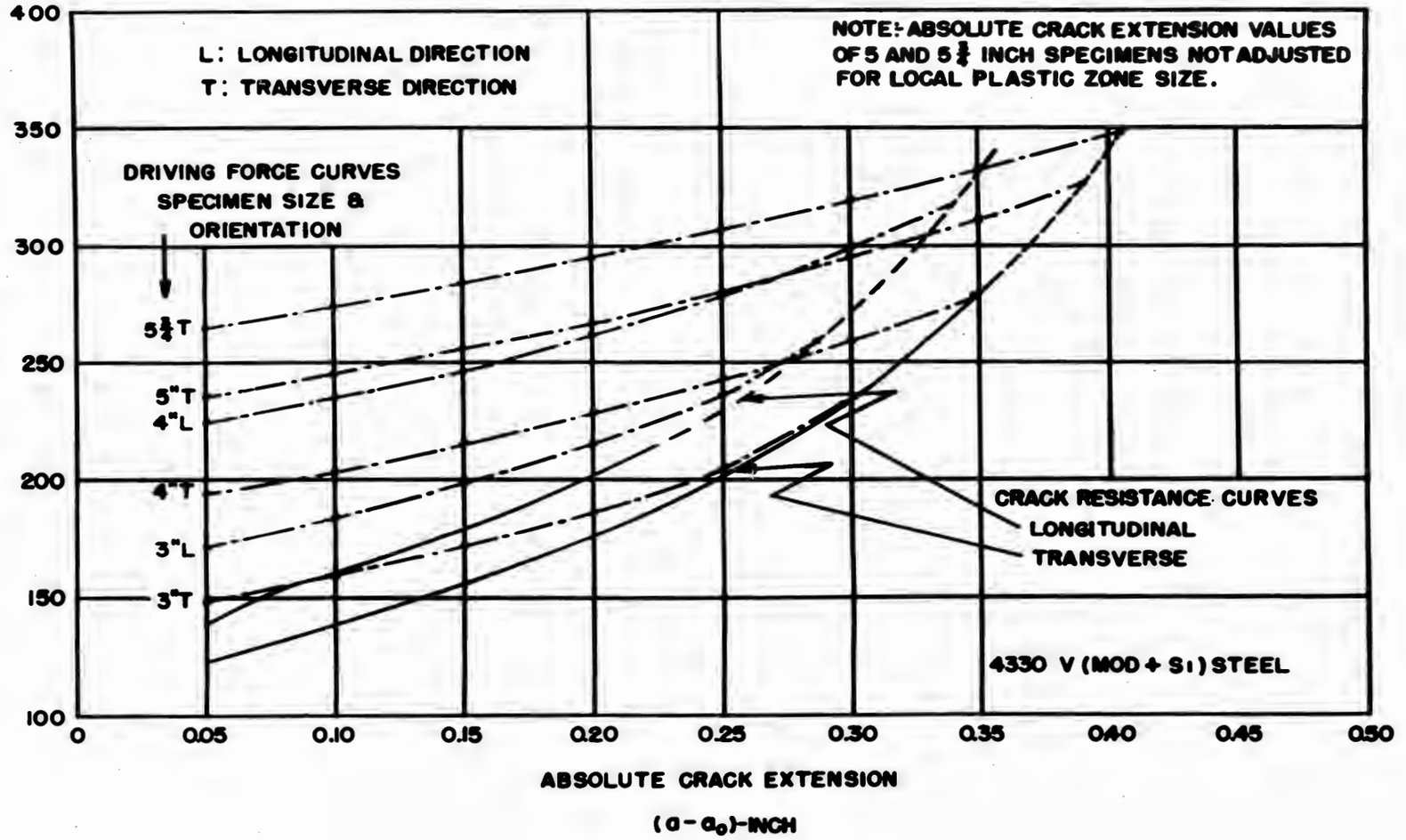


Figure 3. Crack resistance and driving force curves for 4330V (Mod+Si) steel.

Figure 4. Hydrostatic burst test of recoilless rifle tube chamber model.

FEASIBILITY OF NON-METALLIC MINE DETECTION BY EARTH'S MAGNETIC FIELD DISTORTION

STANLEY L. CARTS, JR.
U. S. ARMY ENGINEER RESEARCH AND DEVELOPMENT LABORATORIES
FORT BELVOIR, VIRGINIA

INTRODUCTION:

Magnetic Detection systems have long been of considerable interest in Mine Detection Research. Unlike most other detection methods, magnetic methods are uneffected by soil moisture. "Active" magnetic systems such as the SCR-625 mine detector were used in World War II for detection of metallic mines. The term "active" refers to devices which generate their own magnetic field. Such systems, however, are very susceptible to enemy counter measures. In recent years, mine detection interest has been directed toward not only active methods but also passive methods, while the type of mine of most serious concern has changed from metallic to non-metallic. Passive (sometimes called "earth's field) methods do not generate their own magnetic field but instead measure local distortions in the existing earth's magnetic field.

Theoretical studies (1, 2, 3) have shown that non-metallic as well as metallic mines should be detectable by passive means, since the earth's field will be distorted by probable differences between the magnetic susceptibility of the mine and that of the soil. The magnitude and dimensional distribution of this distortion will be dependent upon the size, shape, and burial depth of the mine; the relative susceptibilities of the soil and of the mine material; and the strength of the earth's field.

Modern advances in magnetometer design have improved instrument sensitivity to the point where the main question concerning passive magnetic detection is not one of instrument sensitivity (magnetometers with sensitivities of better than 10^{-8} oersted are available), but rather of restrictions imposed by the soil, for example, the prevalence of many areas with unfavorable magnetic soil conditions. It was recognized as early as 1945 (4) that soil magnetic properties and natural magnetic anomalies in soils could drastically inhibit the operation of magnetic detectors

(1, 3, 4, 5, 6, 7, 8), but no quantitative studies relating to their character were attempted. Information on these conditions, i.e., magnetic susceptibility, magnetic homogenuity, and the characteristics of naturally occurring magnetic field anomalies, was needed in order to evaluate the feasibility of a passive magnetic detection system. These restricting conditions depend on many factors, some of which are not ordinarily classified by geologists and agronomists. The simplest definitive test is the measurement of the performance of the passive magnetic detection system under various field conditions, and determination of the relation of these conditions to broader existing classifications. To provide this information investigations were designed and initiated by USAERDL. Because of the scope of the investigations both internal and contracted programs were used.

INSTRUMENTATION:

In making such studies, it is important that the instrumentation used provide data reflecting as near as possible the actual magnetic conditions existing above the soil surface. In this way, the information can be used to evaluate the principle involved rather than a particular measuring device. To perform this function, a special instrumentation system was developed at USAERDL.

A Magnatest FM-201 Precision Magnetic Field Meter was modified for use to measure distortions in the vertical component of the earth's magnetic field. (9, 10). This instrument is supplied with dual detection probes having 200 foot leads. At USAERDL these probes were built into an aluminum and lucite holder which accurately positioned and aligned the probes on a common vertical axis. This assembly was arranged as shown in Figure 1, for use as a vertical gradiometer to measure the difference in the vertical component of the earth's field at the soil surface and at a point 4 feet above the surface. It was considered that the field 4 feet above the surface was relatively free of distortions originating in the immediate surface soil (6). The difference in field strength at the two probes, therefore, indicated directly magnetic micro-distortions as small as one gamma (10-5 oersted). The gradiometer was mounted on an aluminum boom and was manually rotated slowly at a constant rate in a circular scan over the surface of the ground. Surface microrelief along the scan path was measured electrically by a small wheel at the other end of the boom. A potentiometer was arranged so that changes in the position of the wheel as it passed over the ground produced changes in potentiometer resistance. This equipment shown in Figure 1 was constructed by Dr. John Cook at Southwest Research Institute based on designs originated at USAERDL. (11)

TEST PROGRAM:

The magnetometer assembly was used to measure in situ magnetic field distortions from buried mines and natural magnetic micro-anomalies. A 50-cubic-inch anti-personnel mine buried flush with the soil surface was used to provide a standard mine signal.

Test sites representative of world soils and climate effects were chosen in the United States and the Panama Canal Zone. This data was augmented by soil susceptibility data collected at the test sites and taken from world soil samples available at the Waterways Experimental Station, Vicksburg, Mississippi and Beltsville Agriculture Station, Beltsville, Maryland.

Soil susceptibility data from over 700 different world soils was collected and magnetic micro-anomaly measurements were made of 245 representative test sites. Analysis of this data and the computations of magnetic field strengths and configurations for pertinent conditions are covered in detail by several USAERDL and contractor reports. (11, 12, 13, 14). It shall suffice here to give only a summary discussion of the resultant determinations showing restricting factors and their relationship to broader known conditions.

Table 1 lists the factors which were found to produce the major restrictions on the passive magnetic method of detecting non-metallic mines.

Table I. Soil Factors and Their Influences Upon a Passive Magnetic Method of Non-Metallic Mine Detection

Factor	Influence
Low Soil Susceptibility	Inadequate magnetic contrast between mine and soil.
Soil Inclusions (Rocks and Roots)	Anomalous mine signals.
a. Igneous rock	Large polarized signal even from small pebbles.
b. Sedimentary rock	Usually negative, unpolarized signal, size of which depends upon size of rock and depth.
c. Metamorphic rock	If basic, influence is like igneous rock; if acid, influence is like sedimentary rock.
d. Non-metallic material	Negative signal, magnitude of which depends upon size, burial depth, and soil susceptibility.
Mineral Concentrations	Anomalous signals.
Soil Matrix Variations	Frequent changes in background level which limit usable instrument sensitivity.
Surface Microrelief	Anomalous signals.
a. Mound or bump	Positive signal, size of which depends upon size of bump and soil susceptibility.

b.	Depression	Negative signal, size of which depends upon size of hole and soil susceptibility.

A general discussion of these factors follows.

SOIL MAGNETIC SUSCEPTIBILITY:

Several significant facts develop from a study of world soil susceptibilities. The magnetic susceptibility of a particular soil type is usually not constant from point to point in a local area but instead fluctuates about a mean value. As a result, the magnetic field above the soil surface is distorted in a corresponding manner. This is listed as "soil matrix noise" in Table I. Field tests have shown a definite trend toward greater variability of susceptibility with lower susceptibility soil. These variations become such a problem in the lower susceptibility soils that instrument sensitivities greater than 1 gamma are unusable (1). To be detected and recognized in the presence of this background noise, a mine signal must be 2 gamma or greater. With this required signal amplitude, we can calculate a threshold susceptibility in each latitude zone below which detection of non-metallic mines is not feasible because of insufficient susceptibility contrast to provide an adequate mine signal.

The importance of soil susceptibility is illustrated in Figure 2 which shows the relationship between soil susceptibility and theoretical mine signal at a height of 6-inches above a 50-cubic-inch non-metallic mine for four latitude zones. One observes that in the $\lambda - 60^{\circ}$ to $\lambda - 90^{\circ}$ zone (curve D) a soil susceptibility of at least 70 ucgs units is required to produce a 2-gamma signal, the minimum acceptable for reliable detection; in the $\lambda - 60^{\circ}$ to $\lambda - 40^{\circ}$ zone (curve C) a susceptibility of 80 ucgs units is needed; in the $\lambda - 25^{\circ}$ to $\lambda - 40^{\circ}$ zone (curve B) a susceptibility of 125 ucgs units is needed; and in the 0° to $\lambda - 25^{\circ}$ zone (curve A) a susceptibility of 300 ucgs is needed.

The susceptibility of a soil is a direct function of its magnetic mineral content. This in turn is dependent on the amount of magnetic material originally present in the parent material from which the soil was developed, and the effects which weathering and erosion have had in the chemical reduction or physical concentration of the magnetic grains. Because of the increased chemical weathering induced by warmer climate, one would usually expect the susceptibility of soils developed from similar parent material to steadily decrease from Frigid regions into Tropic regions. A plot of susceptibility of soils developed from four main parent material groups v.s. latitude (Figure 3) does indeed show such a general decrease from the Frigid regions into the temperate zones, but then shows a pronounced increase of susceptibility as one proceeds further into warmer areas.

It is an interesting side light to our investigation that this data supports recent suggestions of formation of a highly magnetic mineral, maghemite, under tropical conditions. (15, 16, 17).

By consideration of Figures 2 and 3, and a knowledge of the geographical distribution of the four parent material groups we can show that 34% of the world land surface is expected to present insufficient soil-mine susceptibility contrast for reliable detection.

ANOMALOUS SIGNALS:

Simple measurement of the magnetic field distortion due to a buried mine does not in itself constitute detection. If the mine signal is mixed with or surrounded by similar signals caused by natural anomalies mine detection is still not complete. The presence of excessive numbers of anomalous magnetic signals is the most widespread and most limiting soil magnetic property encountered. Three sources of anomalous responses have been identified:

a. Soil inclusive materials (rocks and roots) with magnetic susceptibilities different from the soil matrix.
b. Changes of magnetic mineral concentration within the soil.
c. Irregular soil surface relief.

The most common anomalies found in field studies were those consisting of pieces of soil parent rock distributed throughout the soil. Probably as much as 90% of the anomalies affecting a passive magnetic detector can be attributed to stones in the soil matrix. These rock pieces may have either a higher or lower susceptibility than the soils developed from them. Whether they are higher or lower depends on the original mineralogy of the parent material and the soil forming processes involved.

Field studies indicated, with a few exceptions, absence of anomalies resulting from inclusions in soils developed from transported unconsolidated sediments. This held true for both marine and alluvial unconsolidated sediments, such as comprise most of the Atlantic Coastal Plain. Tests made in this province on both the Atlantic and Gulf coasts produced similar anomaly-free results (9, 10). Alluvial soils in stream valleys are usually free of rock-caused anomalies except in areas where cobbles and boulders are mixed with the soil. Usually, the sorting action of streams is apparent, and except for the present stream bed, the material is separated according to size.

Beach sands are extremely variable in rock content. Where the beach is closely associated with mountainous terrain, rock-caused anomalies are likely to be present on the beach. In areas such as the

Atlantic Coastal Plain, where the beach is preceded by gradual slopes or rock-free soils, the beaches are usually free from this type of false response.

In contrast to the transported soils, residual soils usually contain rock-caused anomalies. Soils developed from consolidated sediments usually have a higher susceptibility than their parent material. When pieces of such parent rock are present in the soil, they produce magnetic field distortions similar to non-metallic mine signals. This situation is illustrated in Figure 4. Such conditions are quite prevalent in limestone, sandstone, and shale soils. These conditions are of particular concern in mine detection because they can produce a negative anomaly of the same order of magnitude as a non-metallic mine.

In residual soils developed from basic igneous and metamorphic rock, pieces of parent rock present even more complex effects. Such rocks are usually of high susceptibility and pieces of them exhibit magnetic polarization. This material produces spectacular anomalies often in the order of thousands of gamma. These anomalous effects can be in the form of negative peaks, positive peaks or both, depending upon the orientation of the source rock. Anomalies of this type are illustrated in Figure 5.

The effects of pieces of parent rock in residual soils derived from acid rocks are extremely variable and difficult to predict. In general, the effects are intermediate between the effects observed in soils from consolidated sediments and those from basic igneous and metamorphic rocks.

It was noted that, as the soil depth decreased, the number of fragments of parent rock distributed within the soil increased. As a result, anomalies in the earth's magnetic field just above the soil surface were more prevalent in mountainous terrain where soils are usually shallow, than in level areas where soils are usually deeper. It was also obvious that the frequency of such anomalies appeared to be much higher in the soils developed from the high-susceptibility basic rocks, than in the soils from the lower susceptibility rock. This is true because although even a very small pebble of high-susceptibility rock can cause a large anomalous signal, a rock of low susceptibility must be at least the same general size as a mine to produce a troublesome false signal.

Local variation of magnetic mineral content within the soil can also produce magnetic field distortions which interfere with mine detection. The effect of variation in mineral concentration becomes more serious as the magnetic mineral content increases. In areas where beach sands are derived principally from volcanic rocks, the magnetic mineral content is usually high and areas of nearly pure magnetite are not uncommon. Wave and wind action sort the heavier magnetic minerals from the lighter quartz and shell grains, and

alternate layers of magnetite and quartz or shell grains result. The effect is strikingly illustrated in Figure 6. Similar conditions are reported for the Phillipine Islands, areas of Japan, and volcanic islands in the Pacific. In such areas, the effect on detection of buried mines is detrimental, as extreme susceptibility variations occur within a few inches of lateral displacement. The Panama beach shown in Figure 6 has variations of susceptibility from 50 to 300,000 ucgs units within a lateral distance of a few inches.

Although magnetite is the magnetic mineral most frequently found in such areas, maghemite, ilmenite, and olivine also occur as concentrations on beaches. Concentrations of magnetic materials have not been found in soils other than beach sands, possibly because most other soils are not so susceptible to sorting by wind and wave action.

Measurements made over high susceptibility soils showed that surface relief irregularities could cause anomalous signals greater than non-metallic mine signals. Theoretically, when the upper horizon of the soil is considered an infinite horizontal homogeneous layer, the earth's field is highly uniform. This field will be distorted, however, in the neighborhood of discontinuities in the layer of soil. A positive correlation of position and amplitude between geomagnetic anomalies and microtopography has been generally noted in field studies by all investigators. Surface relief variations often produce anomalous signals similar to the signals from buried non-metallic mines when the volume and form of the irregularity are similar to a mine. This condition is expected because a non-metallic mine has extremely low susceptibility, essentially that of air. Therefore, signals of similar nature would be produced by a mine buried flush with the surface and by a mine-sized depression in the soil surface. If the mine is buried below the surface, the mine signal is smaller than the signal from a mine-sized depression. Although this effect is most easily observed in high-susceptibility soils, it holds true for soils of any susceptibility.

In some areas, visual inspection of the soil surface will disclose anomalous responses obviously caused by surface irregularities. In general, however, visual inspection cannot effectively alleviate this problem, because (a) the mine may be buried on a rise or in a depression (under such conditions the mine signal will be mixed with the signal produced by surface effects), and (b) small bumps and dips of anti-personnel mine size are often obscured by vegetation coverage.

There appears to be no way to relate surface relief to any specific soil or geologic classifications; although surface relief is reflected to some extent in land use. For this reason, the detrimental effects of microrelief were not considered in the feasibility evaluation of a passive magnetic detection system even though they are acknowledged as a source of false signals.

When world relief, geology, and soil maps are used in conjunction with the anomaly data concerning the occurrence of the various anomaly types we find that approximately 62% of the world can be expected to exhibit anomalous magnetic signal which are so excessive that use of a passive magnetic system would not be feasible.

CONCLUSIONS:

In conclusion, we find that 34% of the world lacks sufficient mine-soil susceptibility contrast to provide an adequate mine signal, and 62% of the world can be expected to present such a high frequency of false signals that the use of the technique is not feasible. In the 34% and 62% of the world there is a common 22% of the world which exhibits both restrictions. The chart below summarizes these figures and shows that in 74% of the world detection of buried non-metallic mines by passive magnetic methods is not feasible.

Table II

Restriction	Inclusive of Other Effects	Exclusive of Other Effects
Insufficient Mine-soil contrast	34	12
Excessive anomalies	62	40
Both restrictions	-	22
TOTAL:		74

More sensitive instrumentation will not improve the world-wide feasibility of passive magnetic mine detection systems because severe restrictions are imposed on the use of passive magnetic phenomenon by natural magnetic soil properties and not by inadequate instrument sensitivity.

While these studies were performed primarily for mine detection purposes, the basic data accumulated applies equally well to many other usages. The detailed data collected is available at USAERDL and in the reports mentioned earlier. We believe that this information can be of aid to other investigators working in geomagnetic, geochemical, and other related fields.

REFERENCES

1. WIEGAND, D. E., Detection of Buried Explosives by Magnetic Means, Armour Research Foundation, Contract DA-44-009-ENG-487, 18 November 1949.

2. SPIES, OTTO R., Anomaly of the Earth's Magnetic Field Due to a Spherical Inclusion in the Ground, Franklin Institute, Contract DA-44-009-ENG-2125, 6 January 1955.

3. Final Report, Contract DA-44-009-ENG-407 (as amended), Electro-Mechanical Research, Inc., 20 May 1948.

4. Tests on Performance of SCR-625 Mine Detector as Related to Soils in the Vicinity of Washington, D. C., U. S. Geological Survey for Chief of Engineers, U. S. Army, 16 June 1945.

5. Report on Use of SCR-625-C Mine Detector and Data to Serve As Basis for Prediction of Performance, Military Geology Unit, U. S. Geological Survey.

6. THARP, JAMES, Field Study of Magnetic Susceptibility of Some Soils of Luzon, As Measured by Mine Detector SCR-625-C, 26 August 1945.

7. ROBERTS, RALPH J.; SAMPSON, EDWARD; STRIKER, M. M.; and STEWART, THOMAS E., Performance of the SCR-625 Mine Detector over Different Rocks and Soils, USAERDL, Fort Belvoir, Va., 1949.

8. BLOOM, ARNOLD L., Mine Detection Studies, Varian Associates, Contract DA-44-009-ENG-2568, 1 June 1955-15 January 1957, Final Technical Report.

9. "Magnetic Measuring Equipment Using Foerster Probes", Unpublished Report No. 20, Magnatest Corporation, Chicago, Ill.

10. Operating Instructions for the Magnatest FM-201 Precision Magnetic Field Meter, Magnatest Corporation, Chicago, Ill.

11. COOK, JOHN, Soil Magnetism Studies, Southwest Research Institute, Contract DA-44-009-ENG-3646, January 1960.

12. CARTS, STANLEY L. and DUEY, HERBERT H.; Magnetic Properties and Anomalies in Soils of Virginia and West Virginia and Significance to Land Mine Detection, USAERDL Report 1592-TR, 9 September 1959.

13. CARTS, STANLEY L.; ORR, DONALD G.; and MACCORMAC, JACK R.; The Effects of Soil Magnetic Properties and Natural Magnetic Micro-Anomalies of Typical Tropical Soils on Passive Magnetic Land Mine Detection Methods, USAERDL Report 1665, 19 May 1960.

14. CARTS, STANLEY L. and WEBB, PHILIP K.; World-Wide Feasibility of a Passive Magnetic Method of Detecting Buried Non-metallic Land Mines, USAERDL Report, 1683-RR, 31 July 1961.

15. LEBORGNE, E., "Abnormal Magnetic Susceptibility of Surface Soil," Annales de Geophysique, Vol. 11, No. 4 1955, pp. 1-21.

16. STAVROU, A., "Investigations on the Magnetic Properties of Greek Red Earth", Beitrage Zur Geophysik, Vol. 66, No. 3 (1957) pp. 214-255.

17. MATSUSAKA, YOSHITO and SHERMAN, G. DONALD; "Magnetism of Iron Oxide in Hawaiian Soils", Soil Science, Vol. 91, No. 4, April 1961, pp. 239-245.

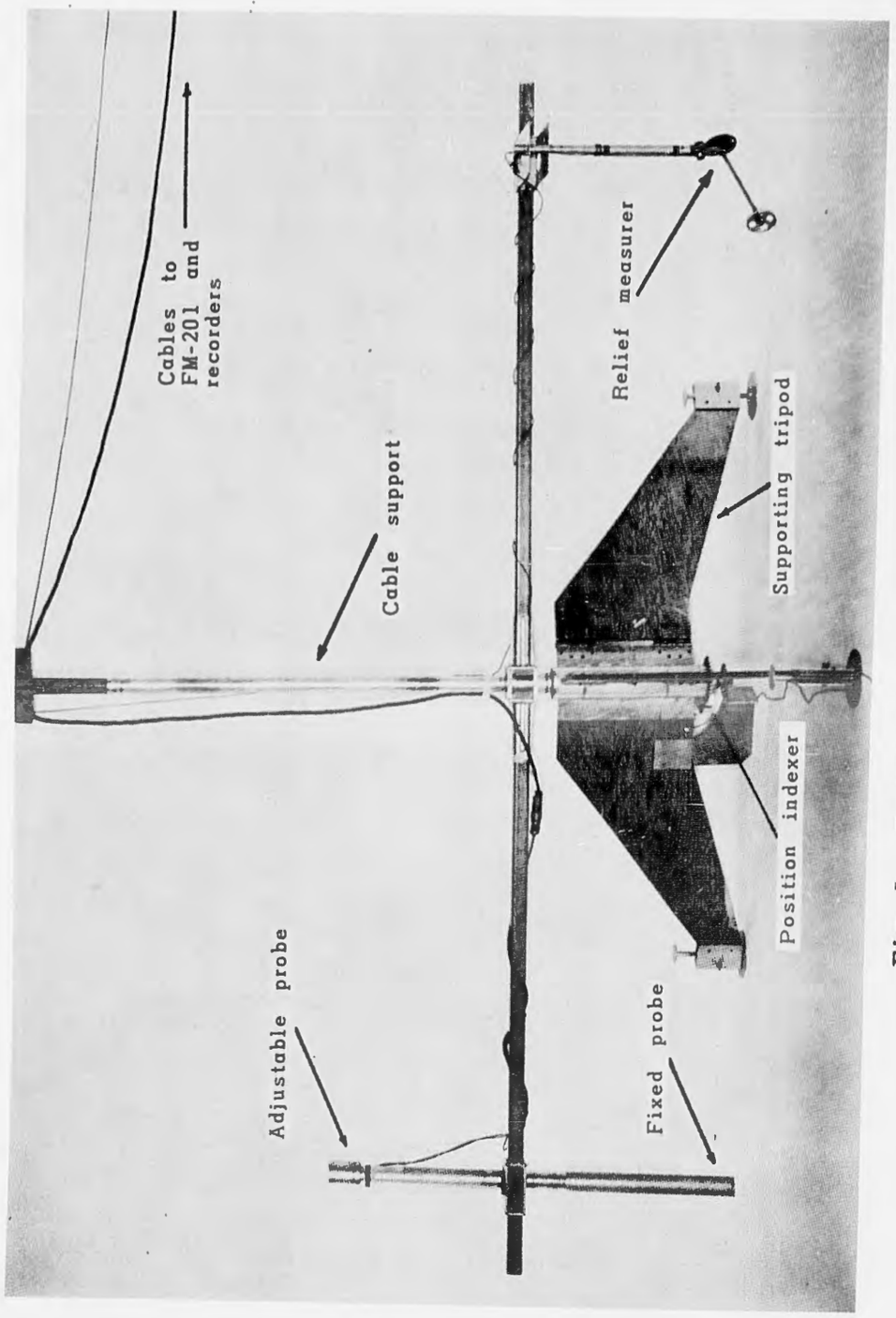


Fig. 1. Instrumentation arrangement.

G5356

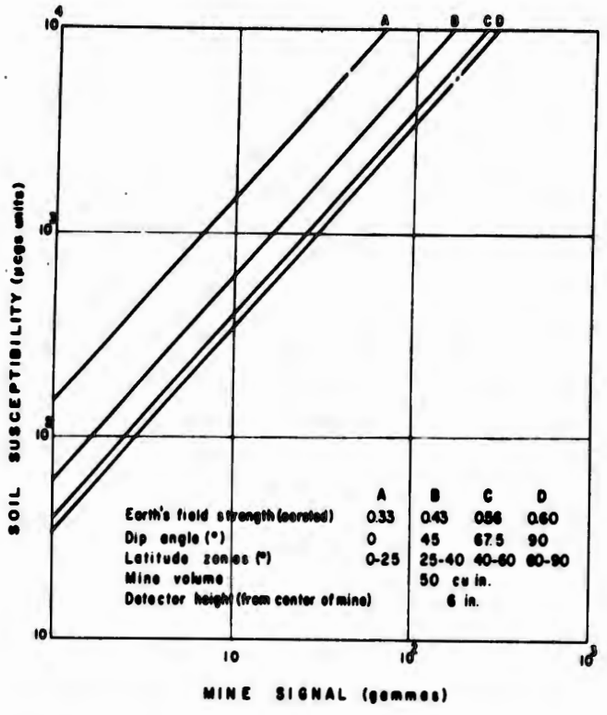


Fig. 2. Effect of soil susceptibility on non-metallic mine signal

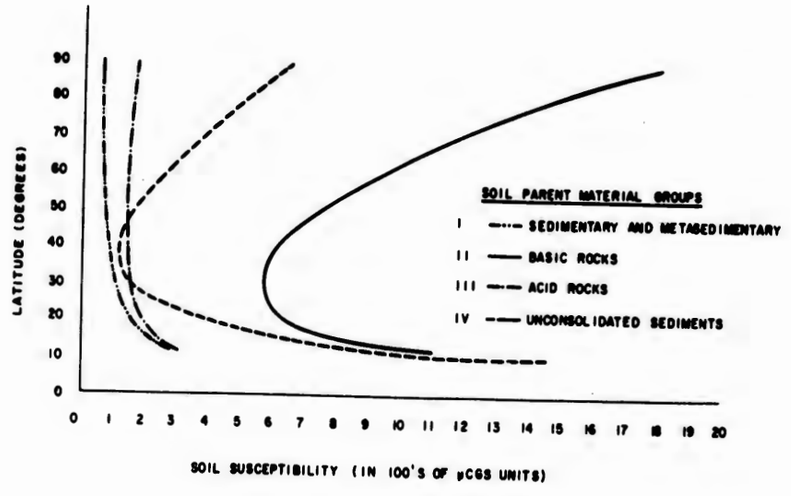


Fig. 3. Latitude vs. susceptibility of soils developed from four parent material groups.

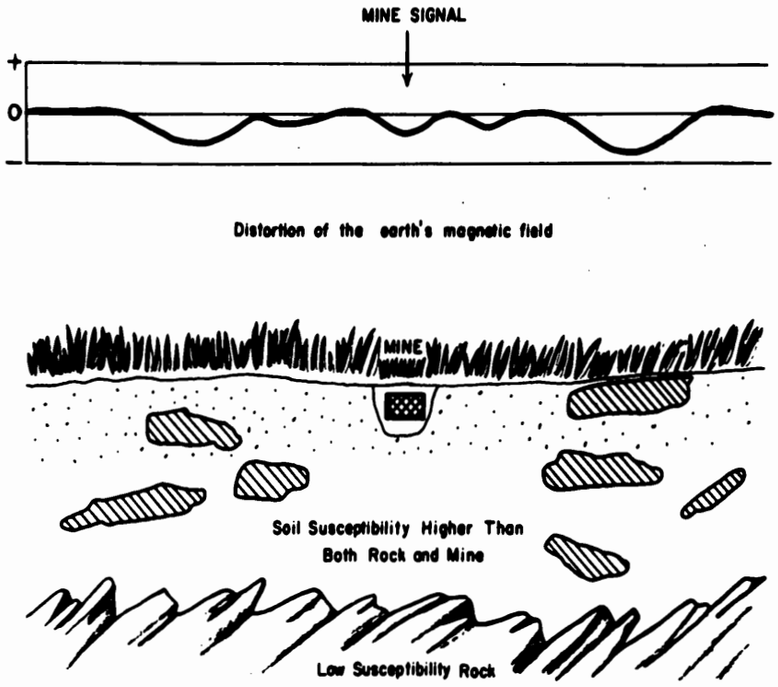


Fig. 4. Magnetic field distortions above residual soil developed from sedimentary rocks.

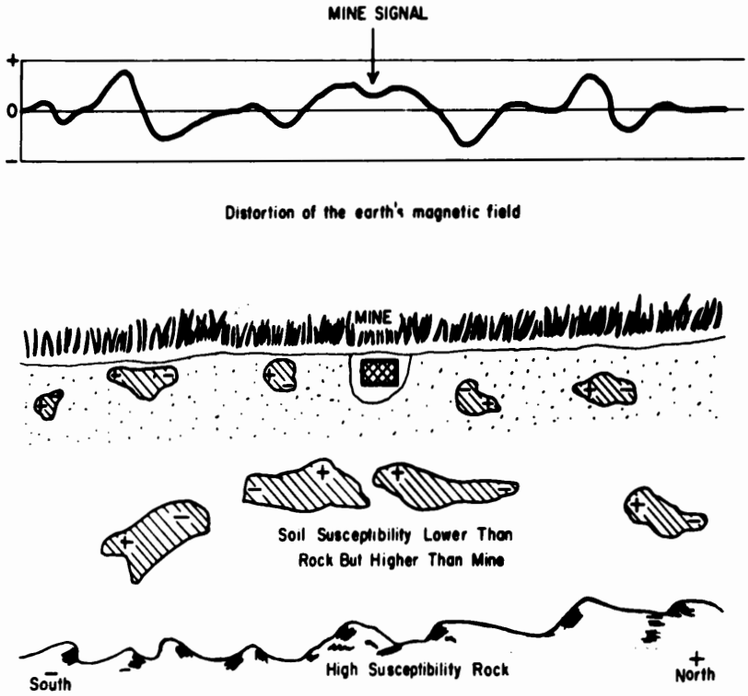


Fig. 5. Magnetic field distortions above residual soil developed from basic igneous or basic metamorphic rock.

Fig. 6. Panama beach sand containing magnetite concentrations.

THE RADIATION CHEMISTRY OF LIQUID DINITROGEN TETROXIDE

T. C. CASTORINA, A. O. ALLEN and E. R. JOHNSON
NUCLEONICS RESEARCH SECTION, PICATINNY ARSENAL
DOVER, NEW JERSEY

Introduction

The nitro group is considered to be one of the most important plosophoric groups in Ordnance chemistry. When substituted on an organic molecule in sufficient numbers it supplies the necessary oxygen balance for the complete conversion of the carbons and hydrogens to carbon dioxide and water. The problem has been to effect this optimum number of substitutions without degrading the parent molecule in the process. Conventional methods of synthesis are rapidly exhausting all avenues of approach for a solution to this problem. New methods of synthesis must be investigated and one possible method is preparative radiation chemistry.

High energy radiation may produce the already familiar entities, carbonium and nitronium ions, free radicals and excited molecules in environments devoid of high pressures, high temperatures and stringent oxidizers. This might make possible modifications and improvements in existing methods of synthesis that depend on their reactions. Even more important, however, is the possibility that high-energy radiation, e.g., cobalt-60 gammas, may afford a unique method of forming different, active intermediates capable of undergoing new types of reactions with nitrogen dioxide. This would open a whole new field of nitrolysis.

Before a study can be made of a two-component system, as would be the case for preparative radiation chemistry, each of the components should be studied separately. For a two-component system consisting of benzene and nitrogen dioxide only one of the components needs investigation. The radiation chemistry of benzene is well established and benzene has been found to be relatively radiation resistant. However, the radiation chemistry of nitrogen dioxide in its dimeric form, as a liquid, has not been reported in the literature. The radiolysis of nitrogen dioxide gas in a nuclear reactor has been studied by Harteck and Dondes[1] and more recently the Russians[2] have reported on the gamma radiolysis of gaseous nitrogen dioxide at varying intensities and pressures. For the present study, an

investigation of the cobalt-60 gamma radiolysis of nitrogen dioxide in the liquid phase is considered more appropriate to subsequent investigations involving its reaction with aromatics and possibly aliphatics.

Results and Discussion

Nitrogen dioxide (NO_2) is an odd-electron molecule and hence paramagnetic. NO_2 absorbs light over a wide range and the brown-red color is due to its presence. The diamagnetic dimer, dinitrogen tetroxide (N_2O_4) is transparent over a wide range of wavelengths. N_2O_4, has a freezing point of -11.2°C and a boiling point of 21.2°C. The vapor pressure in equilibrium with the liquid at 25°C is one atmosphere. In the saturated vapor phase N_2O_4 dissociates 8.6% at -11.2°C and 15.9% at 21.2°C. In the liquid phase[3] the dissociation is homolytic at -11.2 to the extent of 0.018% and 0.12% at 21.2°C. The N_2O_4 used in this study was obtained from a gas cylinder and purified by repeated passage through phosphorous pentoxide-saturated glass wool. The purification also included the passage of purified oxygen through the liquid at 0°C for the conversion of any of the lower oxides of nitrogen to N_2O_4. Samples were prepared in 5 cc pyrex break-seal containers and irradiated at 20°C in a cobalt-60 gamma field of approximately 4×10^5r/hr.

The irradiated samples were degassed at liquid nitrogen temperature (-192°C) and the permanent gases measured at STP. The permanent gases were analyzed mass spectrometrically and checked for oxygen content by the platinum filament combustion method. The mass spectrometric analysis indicated the gas composition to consist of only N_2 and O_2 and the relative amounts of each constituent were calculated from their respective mass peaks. The amount of O_2 was also determined by its reaction with H_2 at the catalyzing surface of the heated (barely perceptible glow point) platinum wire. The results listed in Table 1 show both methods to be complementary and in good agreement.

Table 1

Percent Composition of N_2 and O_2

Expt No.	Pt. Filament Analysis		Mass Spectrometric Analysis	
	$\%O_2$	$\%N_2$[(a)]	$\%O_2$	$\%N_2$
1	3.6	96.4	3.5	96.5
2	3.4	96.6	3.6	96.4
3	5.7	94.3	6.5	93.5
4	5.4	94.6	6.4	93.6
5	3.6	96.4	5.7	94.3
6	25.8	74.2	26.0	74.0
7	43.5	56.5	44.8	55.2
8	50.7	49.3	50.9	49.1

(a) By difference

The calculation of the G value (i.e., the number of molecules produced or destroyed per 100 electron volts absorbed) was based on the ferrous sulfate dosimeter and corrected for the electron density difference of N_2O_4 as compared to the 0.8N H_2SO_4 aqueous solution. Justification for this was furnished by Harteck and Dondes[4], who showed that the absorption of gammas via the Compton effect in the gas is the same as in the liquid phase and, therefore, that the energy absorbed was proportional to their respective weights and electron densities. Further verification of this was obtained by varying the volumes of N_2O_4 gas to N_2O_4 liquid ratios by a factor of 600. The independence of the G value to the varying amounts of gas in equilibrium with the liquid is demonstrated in Table 2. The ratio of 0.004 represents a liquid sample with approximately 0.02 cc of gas confined in the 1 mm capillary end of the break-seal. Since the G value remains constant it is implied that the mechanism of liquid N_2O_4 radiolysis is not affected by the presence of the gaseous phase.

Table 2

Independence of G $(N_2)_{STP}$ to Variation of Vol. $(N_2O_4)_g$: Vol. $(N_2O_4)_l$ Ratio

Sample No.	Dose ($X10^{-7}$ r)	$G(N_2)_{STP}$	V N_2O_{4g}: V N_2O_{4l}
8-30-1	4.92	0.048	0.031
8-28-6	4.92	0.047	0.050
8-30-2	4.92	0.053	0.100
8-28-3	5.27	0.051	0.307
8-28-4	5.27	0.051	0.348
9-20-2	4.21	0.048	0.877
9-20-1	4.21	0.047	2.318
9-20-4	4.21	0.049	0.952
9-20-5	4.21	0.047	1.637
9-20-6	4.21	0.046	0.114
9-20-3	4.21	0.048	0.035
7-5-5	2.18	0.045	0.004
		0.048 ± 0.002	

To determine the criticality of temperature control during long periods of irradiation, a series of samples of $(N_2O_4)_l$ were irradiated at varying temperature. The G value was observed to remain constant over the temperature range from 0° to 30°C. The results of these experiments are summarized in Table 3. However, in spite of this apparent insensitivity of G value to temperature fluctuations in this range, all subsequent samples were irradiated at 20°C $\pm$ 1°.

Table 3

Independence of $G(N_2)$ to Radiolysis of $N_2O_4(l)$ at Varying Temperatures

Sample No.	Dose ($\times 10^{-6}$ r)	Temp °C	$G(N_2)$	Average $G(N_2)$	Average Deviation
9-14-3	39.36	0	0.048		
9-14-4	39.36	0	0.048		
9-14-1	39.36	0	0.047	0.047	±0.001
9-14-2	39.36	0	0.047		
9-14-5	39.36	0	0.046		
8-30-3	39.36	14	0.045		
8-30-4	39.36	14	0.045	0.048	±0.002
8-30-1	49.20	14	0.048		
8-30-2	42.90	14	0.053		
9-20-5	42.14	21	0.047		
9-20-4	42.14	21	0.049		
9-20-1	42.14	21	0.047		
9-20-2	42.14	21	0.048	0.048	±0.001
8-28-6	49.20	21	0.047		
8-28-4	52.68	21	0.051		
7-15-1	26.4	30	0.049		
7-15-2	26.4	30	0.047	0.048	±0.001

In the process of making the above determinations five to six sample batches of $(N_2O_4)_l$ were irradiated together for the same period of time. Upon removal from the source they were placed in refrigeration at 2°C and taken as needed for analysis over a period of two days. In doing this, it was observed that the percentage of oxygen gradually increased with each successive sample analyzed. This was indicative of the existence of a thermally unstable constituent which was decomposing with the evolution of oxygen. Any of the oxides of nitrogen lower than NO_2 were ruled out as possible contributors of molecular oxygen at temperatures in the vicinity of 2° to 25°C. N_2O, for example, is very stable and any NO or N_2O_3 present would, instead, be instantaneously converted to NO_2 in the presence of oxygen. The only possible oxide that is known to be thermally unstable at room temperature is the higher oxide, N_2O_5. The original studies on the rate of dissociation of N_2O_5 to NO_2 plus O_2 are classics in the field of kinetics.

Since the oxygen evolution after irradiation was not anticipated, the oxygen content had inadvertently been permitted to vary randomly from one set of experiments to another. In order to give meaning, therefore, to any of the previous determinations, it became necessary to subtract this random distribution of oxygen and report all G values in terms of nitrogen gas produced as a direct product of radiolysis. To demonstrate the validity of treating the data in this manner an extensive series of samples of $(N_2O_4)_1$, prepared at different times, were irradiated to different total doses. A plot of the c.c. of N_2 at STP versus the dose in rads is given in Fig. 1. The curve obtained shows a definite induction period which may represent the time interval when intermediate species are increasing in sufficient concentrations for final conversion to stable radiolytic products. It is also possible that this induction period is due to the presence of small quantities of impurities which are scavenging the intermediate species, momentarily preventing a formation of the necessary concentration for their conversion to radiolytic products. Following this period of induction a steady state exists as represented by the straight line portion of the curve. The average $G(N_2)$ for this steady state is 0.048. Apparently $(N_2O_4)_1$ is relatively radiation resistant with respect to the N_2 gas produced. The points on the graph correspond to a minimum of duplicate determinations and are not drawn in proportion to their average deviation of $\pm$ 3%.

To establish the identity of the thermally unstable higher oxide of nitrogen, a kinetic study was made of its rate of decomposition in solution of $(N_2O_4)_1$. The determination of the specific rates of oxygen evolution were conducted at three different temperatures on three sets of five samples, each set irradiated for a total of 96 hours equivalent to 4×10^7 rads. For a particular run, the five samples were placed in a constant temperature bath directly after radiolysis and removed with time for oxygen analysis. These results are summarized in Fig. 2. The disappearance of the higher oxide of nitrogen, designated as $[N_2O_4]O$ is plotted on the logarythmic scale versus the time in hours on the abscissa. The straight line obtained on this semi-log plot is indicative of a first order reaction. Eyring and Daniels[5] have reported the thermal decomposition of N_2O_5 to be first order. The y intercept is equal to that quantity of oxygen present upon immediate removal from the gamma source (termed "zero" time). To check this value (33%), samples of $(N_2O_4)_1$ were irradiated for the same period of time and placed in a -78°C bath directly upon removal from the ionizing field. The average value of oxygen content was found to be 3% less than the extrapolated value. Each of the points plotted is corrected for 6•3% of oxygen contributed by the N_2O formation from N_2O_4 (this product of radiolysis will be discussed subsequently). The specific rate constants corresponding to 20°, 22°, and 25°C are 0.200×10^{-5} sec^{-1}, 0.289×10^{-5} sec^{-1} and 0.540×10^{-5} sec^{-1}, respectively, and the average activation energy is calculated to be 34 kcal. The values given by Eyring and Daniels are 0.344×10^{-4} sec^{-1} at 20°C and 25 kcal for the activation energy, based on

1 m M concentrations of N_2O_5 in 5 cc of $(N_2O_4)_1$. The concentrations dealt with in this study are in the order of 20μ M $[N_2O_4]$ 0 per 5 cc of $(N_2O_4)_1$ or 50 times more dilute. This discrepancy of results obtained may conceivably be due to a concentration effect and/or perhaps a magnification of experimental errors in this extremely dilute concentration range, peculiar to the method of analysis employed. To pursue these possibilities, synthetic samples of N_2O_5 in N_2O_4 having concentrations comparable to 20μ M/5cc are currently being prepared. The kinetics of such solutions will be studied and the results compared to those obtained thus far. In addition, attempts are concurrently in progress for the direct identification of this higher oxide by ultra-violet absorption analysis. The absorption spectrum of standard solutions of N_2O_5 in $(N_2O_4)_1$ will be compared with that of the unknown.

It had been noted earlier in this work that the percent of oxygen at near "zero" time after the first 48 hours of irradiation was 3•5%. This not only suggested the possibility that N_2O_4 does not go directly to N_2+O_2 but that the intermediate atomic oxygen specie combines with N_2O_4 to form a thermally unstable adduct which is stabilized by the induced back reaction in the gamma field. This was forcefully demonstrated by the comparison of the "zero" time oxygen content as a function of irradiation time with the rate of post irradiation oxygen evolution as shown in Fig 3. It can be readily seen that the rate of oxygen formation during irradiation is relatively slow in comparison to the thermal decomposition outside the gamma field. As a check on this comparison, equivalent samples that had shown a low oxygen content at "zero" time were thermally equilibrated at 30°C for 7 days. The percent oxygen in every case increased to the stoichiometric value based on the proportion of nitrogen according to: $N_2O_4 \longrightarrow N_2 + 4\ [N_2O_4]\ 0$ and $4\ [N_2O_4]\ 0 \longrightarrow 2O_2$ or 67% oxygen.

Once having established the post irradiation thermal decomposition of one of the products of radiolysis it became possible to investigate the existance of additional radiolytic products, a series of samples were irradiated for 96 hours and thermally equilibrated for 7 days. Control samples (i.e., unirradiated $(N_2O_4)_1$) were included in this determination as in all determinations made in this study. The permanent gases were first removed as usual at -192°C. The sample was then degassed at -130°C. At this temperature N_2O_4 has no vapor pressure so that any gas collected would be either N_2O or CO_2. The gas collected was analyzed mass spectrometrically and the spectrum of the mass peaks was compared to those of the CO_2 and N_2O standards. The peak height intensity ratios of the mass to charge ratios of the various combinations of the fragmentation fractions were used to identify the unknown gas. The values given in Table 4 compare well with those of the N_2O standard. In addition, it should be emphasized that the spectrum of the unknown had a strong 30 peak but contained no trace of a 22 peak. According to the spectrum of

the CO_2 standard the following fragmentation takes place:

$$CO_2 \longrightarrow CO_2^{++} + CO^+ + O^+ + CO_2^+$$

Where the spectrum of the CO_2 standard contains a 22 peak but no 30 peak, that of the N_2O standard has a strong 30 peak but no 22 peak. It was, therefore, concluded that the unknown gas, fractionated at -130°C, was N_2O. This conclusion was further substantiated by the results obtained with the Bechman vapor phase chromatography unit using silica gel as the absorbing medium. The retention times of the unknown and N_2O standard were identical.

Table 4

Mass Spectrometric Analysis
of Product of Radiolysis of $(N_2O_4)1$
Factionated at -130°C

M/e Fractions	Intensity (Peak Height) Ratios of M/e Fractions				
	N_2O[a] Std	12-14-2A	12-14-1A	11-30-4A	11-9-5A
30/28 = NO/N_2	2.65	2.60	2.25	2.39	2.62
30/14 = NO/N	5.53	5.48	5.15	5.45	5.51
14/16 = N/O	2.34	2.37	2.34	2.22	2.32
44/30 = N_2O/NO	2.90	3.28	3.02	3.02	3.10

[a]Based on the N_2O Std, the mass spectrometric fragmentation is:

$$N_2O \longrightarrow NO^+ + N_2^+ + O^+ + N^+ + N_2O^+$$

With the identification of the stable products of radiolysis, the dependency of their formation on dose was determined. Fig. 4 shows this dependency to be linear after an induction period. The results of the individual experiments are given in Table 5. At the steady state concentrations of the respective gaseous products of radiolysis, the material balance based on the valency changes involved in the process, viz;

$$4\ G(N_2) + 3\ G(N_2O) = 2\ G(O_2)$$

is found to be within an experimental error of 5%.

The following mechanism is suggested as a working hypothesis:

Reactions taking place in the spur -

$$N_2O_4 \longrightarrow N_2O_3 + O$$

$$N_2O_3 + N_2O_3 \longrightarrow N_2O_4 + 2NO$$

$$O + O \longrightarrow O_2$$

Table 5

N_2O/N_2 : N_2O/O_2 : O_2/N_2 with Dose

Expt No.	Grams $(N_2O_4)_1$	Dose Rads $(x10^{-6})$	c.c.$(O_2)_{STP}$ per gram $(N_2O_4)_1$ $(x10^2)$	c.c.$(N_2)_{STP}$ per gram $(N_2O_4)_1$ $(x10^2)$	c.c.$(N_2O)_{STP}$ per gram $(N_2O_4)_1$ $(x10^2)$	Percent[1] Oxygen	$G(O_2)$	$G(N_2)$	$G(N_2O)$[2]
1	5.096	9.84	0.92	0.35	0.67	72.0	0.044	0.017	0.033
2	7.800	9.84	1.68	0.38	0.63	81.6	0.081	0.018	0.031
3	4.943	19.68	5.00	1.24	1.41	78.0	0.120	0.034	0.034
4	7.070	19.68	4.99	1.31	1.04	79.3	0.121	0.032	0.025
5	5.464	29.52	7.70	2.93	1.67	72.5	0.125	0.047	0.027
6	7.154	29.52	7.50	2.83	1.47	72.6	0.121	0.046	0.024
7	5.580	40.70	9.87	4.09	1.99	70.8	0.116	0.048	0.023
8	4.671	39.36	9.87	3.95	2.12	71.8	0.120	0.047	0.026
9	4.820	39.36	9.14	3.81	2.12	70.6	0.111	0.046	0.026
10	4.886	49.20	12.96	5.04	2.50	72.0	0.126	0.049	0.024
11	5.740	50.45	12.42	4.95	2.41	71.5	0.118	0.047	0.023
12	7.500	59.04	15.34	6.12	3.20	71.5	0.124	0.050	0.026

(1) $\%O_2$ based on $[O_2]/[O_2]+[N_2] \times 100$ as determined by filament combustion and mass spectrometric analyses of radiolysed samples thermally equilibrated at 27°C for 7 days.

(2) N_2O identified by both mass spectrometric and vaporphase chromatographic analyses.

$$N_2O_3 + O \longrightarrow N_2O_4$$
$$N_2O_3 + O \longrightarrow N_2O_2 + O_2$$

Ractions taking place in the bulk of $(N_2O_4)_1$:

$$N_2O_2 + N_2O_4 \longrightarrow N_2O + N_2O_5$$
$$O + N_2O_4 \longrightarrow N_2O_5$$
$$N_2O_3 \longrightarrow NO + NO_2$$
$$NO + O \longrightarrow NO_2$$
$$N_2O_3 + NO \longrightarrow N_2O_2 + NO_2$$
$$NO_3 + NO_2 \longrightarrow N_2O_5$$
$$N_2O_5 \longrightarrow NO_3 + NO_2$$
$$2NO + O_2 \longrightarrow N_2O_4$$
$$NO_3 \longrightarrow NO + O_2$$
$$NO + NO_3 \longrightarrow N_2O_4$$
$$N_2O + O \longrightarrow N_2 + O_2$$
$$N_2O + O \longrightarrow 2NO$$

Conclusions

1. The material balance obtained in this study suggests the existence of the following stoichiometric relationship:

$$6(N_2O_4)_1 \longrightarrow N_2O + 4\left[N_2O_4\right]O + N_2 + 3/2\ O_2$$

At thermal equilibrium after $4\left[N_2O_4\right]O \longrightarrow 4N_2O_4 + 2O_2$,
$2N_2O_4 \longrightarrow N_2O + N_2 + 3\frac{1}{2}O_2$
and based on the data obtained:
$3N_2O_4 \longrightarrow N_2O + 2N_2 + 5\frac{1}{2}O_2$

2. The total yield of the stable products of radiolysis per 100 ev absorbed is:

$$G(N_2) + G(O_2) + G(N_2O) = 0.048 + 0.120 + 0.025 = 0.293$$

3. The total yield based on the inclusion of the metastable constituent which can be assumed to be N_2O_5, though this will require further substantiation is:

$$G(N_2) + G(N_2O_5) + G(N_2O) = 0.048 + 0.240 + 0.025 = 0.313$$

References

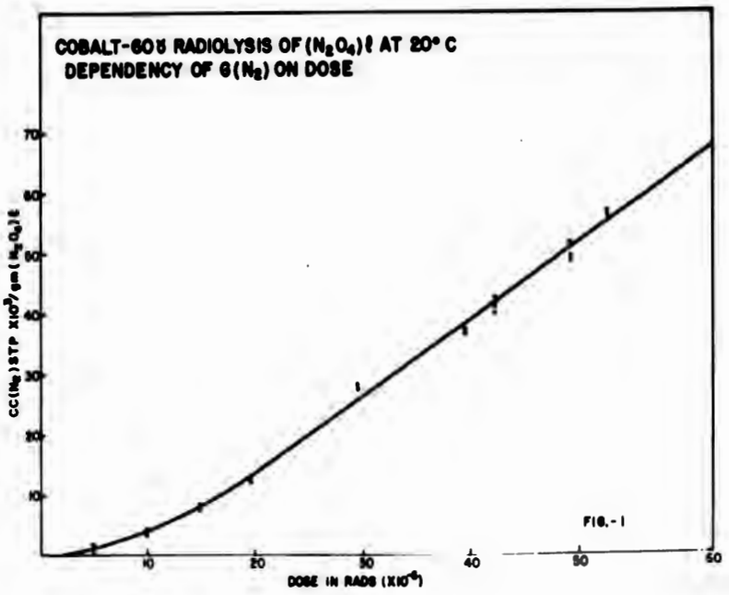
1. Harteck, P. and Dondes, S., NYO-9967, Rensselaer Poly. Inst. (1961) Contract AT (30-3)-321.

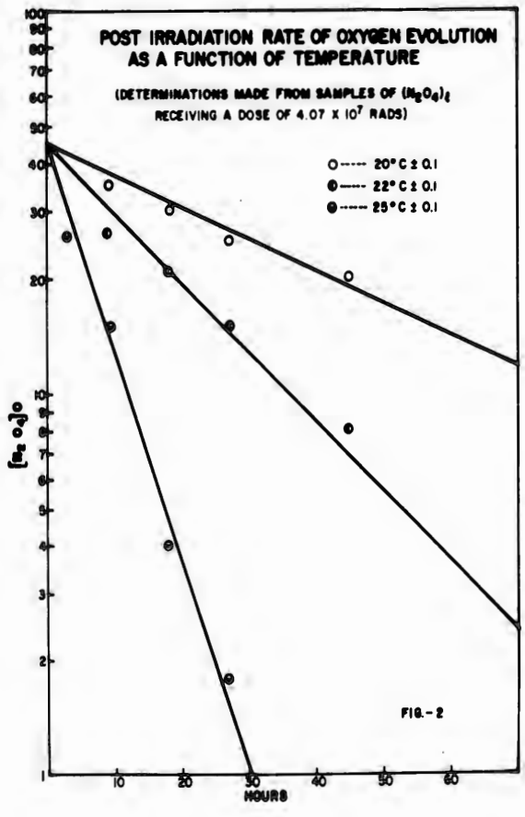
2. Dimitriev, M.T. and Saradzhev, L. V., Russian J. Phys. Chem. 35, No. 4, 354(1961).

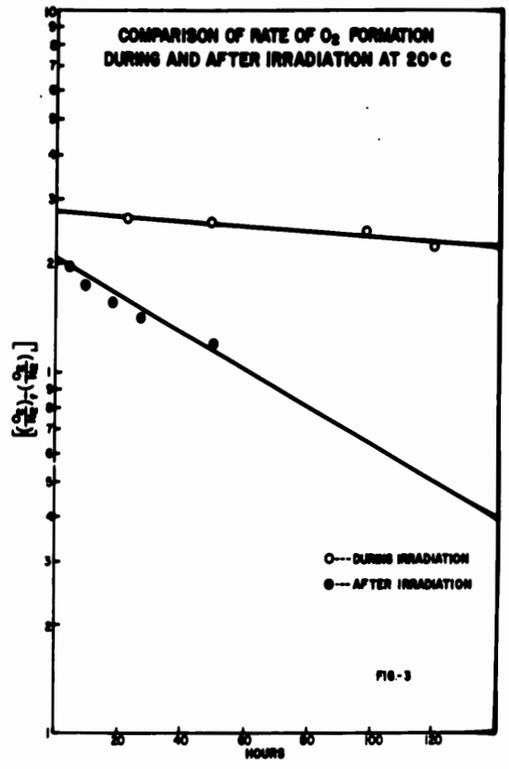
3. Gray, P. and Rathbone, P., J. Chem Soc. of London (1959), 3550-57.

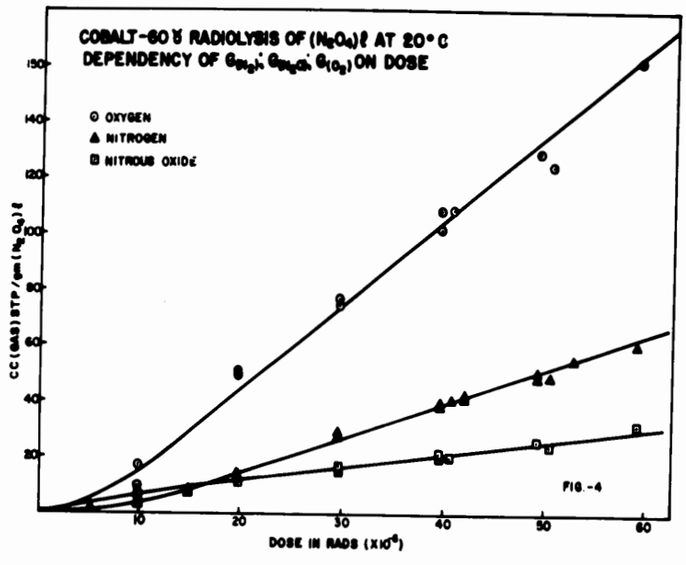
4. Harteck, P. and Dondes, S., Nucleonics 14, No. 3, 66(1956).

5. Eyring, H. and Daniels, F., J. Am. Chem. Soc. 52, 1472-86 (1930).


COBALT-60 RADIOLYSIS OF $(N_2O_4)_\ell$ AT 20° C
DEPENDENCY OF $G(N_2)$ ON DOSE
$CC(N_2)$ STP $X10^3$/gm $(N_2O_4)_\ell$
DOSE IN RADS ($X10^{-6}$)
10
20
30
40
50
60
70
FIG. - 1


POST IRRADIATION RATE OF OXYGEN EVOLUTION
AS A FUNCTION OF TEMPERATURE
(DETERMINATIONS MADE FROM SAMPLES OF $(N_2O_4)_\ell$
RECEIVING A DOSE OF 4.07 X 10^7 RADS)
○----- 20° C ± 0.1
●----- 22° C ± 0.1
●----- 25° C ± 0.1
$[N_2O_4]_0$
HOURS
FIG. - 2


COMPARISON OF RATE OF O_2 FORMATION
DURING AND AFTER IRRADIATION AT 20° C
O--- DURING IRRADIATION
● --- AFTER IRRADIATION
FIG.-3
HOURS


COBALT-60 γ RADIOLYSIS OF $(N_2O_4)\ell$ AT 20° C
DEPENDENCY OF $G_{(N_2O_4)}$, $G_{(N_2O)}$, $G_{(O_2)}$ ON DOSE
O OXYGEN
▲ NITROGEN
⊡ NITROUS OXIDE
FIG.-4
DOSE IN RADS (X10⁻⁶)

AN OPERATIONS RESEARCH MODEL OF MOTOR TRUCK TRANSPORT DERIVED FROM NUCLEAR TRANSPORT THEORY

H. CHARLTON, and J. H. NEBLETT *
U. S. ARMY TRANSPORTATION RESEARCH COMMAND
FORT EUSTIS, VIRGINIA

Today, military mobility is becoming identified with a major problem area; logistics of motor transport. A further simplifying assumption frequently made is that this problem can be measured by the ability of a vehicle to traverse natural terrain. This paper contends that the validity of this implicit assumption has not been demonstrated; rather, the complete system should be analyzed, with simplifying assumptions made only if the model complexity is beyond the capacity of modern mathematics and computer technology.

This paper will present a formulation of a theoretical model. The importance of the parameters will be the subject of a computer analysis after the model has been historically verified.

A general evaluation of an operational situation should be based first on a knowledge of the time-space density of units of supply under large-scale operations involving varying battle conditions and terrain features. Assuming then that general coefficient and probability functions, which will take these factors into account, can be realistically defined; we will investigate the density of carrier units operating at two speeds in a symmetric disc geometry. This problem requires only a four-dimensional phase space. Formulation of the systems research problem will start through a statement of the significance and dimensions of the coefficients and probabilities involved.

Consider a large-scale theater of operation involving regions of varying terrain conditions and various intensities of enemy action. Suppose that it is desired to be able to predict or maintain a supply system (function) in this area. To do this

it is assumed, barring local anomalies and historical accident, that by and large the gross determination of a supply function is governed by general terrain properties, enemy action, personnel behavior, and command decision. This paper will consider only supplies delivered by land carriers such as trucks and will define our unit of supply as one loaded truck. We assume that this unit travels at a speed v and is destination oriented. In the course of a "trip", it may be put out of operation (annihilated), deflected (scattered), or forced to change its speed (transferred) by terrain, enemy action, or decision. Taking such effects into account, the problem is to construct an applicable model that determines a probable supply function.

Although a natural dependent variable or concept is density, the original derivation was by defining the dependent variable as the intensity of units at a point in space traveling in a given direction. Formulation of absorption and scattering in angular dependent intensity requires an integral formulation as, for instance, the intensity in a given direction at a point is influenced by units being scattered into this direction from all other directions at this point; thus, leading originally to a transfer equation from which was obtained the density equation. But for clarity and brevity this paper will proceed directly to the density equations by logical statements rather than by mathematical deduction.

In the specific model considered, it is further assumed that cross-country mobility, v_1, substantially differs from on-road mobility, v_2; for this reason, two speed groups are considered. This leads to a system of two dependent transfer equations, one governing the density for on-road speed and the other governing the density for off-road speed where the average speeds are taken to be representative of the groups.

The specific objective now is to relate the coefficients and probabilities occurring in the transfer equations to operationally meaningful quantities and present the relations expressing these quantities. The coefficients and probabilities will be given for a two-spacial dimensional model.

The general significance of the density (β) and mass absorption coefficient (k) is:

$$\text{Strictly, the density } \beta = \lim_{A>0} \frac{M}{A}$$

where A is an area and M is the mass associated with this area. As we are interested in a stochastic model, average values and

probabilities apply, so define

$$\beta = \frac{\bar{M}}{\bar{A}} = \lim_{A>0} \frac{M}{A}$$

where $\bar{M}$ is the total mass associated with a fixed area $\bar{A}$.

Thus, take $k\beta ds$ to be dimensionless.

Now β = mass/unit area, where mass is defined in terms of intensity of enemy action, terrain properties, etc.

k is thus seen to be a probability that in a unit area an absorption will take place

per unit mass
per unit length

k can be called the mass absorption coefficient.

We now consider separately the different absorption terms listed in table 1.

Annihilation by Enemy Action

β_{av1} = occurrences unit area, where an occurrence could be the detonation of a bomb. As discussed,

$$\beta_{av1} = \frac{\text{total number of expected events (occurrences) in a region}}{\text{area of the region}}$$

Then k_{av1} is the probability per unit occurrence that a unit will be annihilated in terms of length perpendicular to the direction of travel.

β_{av2} and k_{av2} are similarly defined.

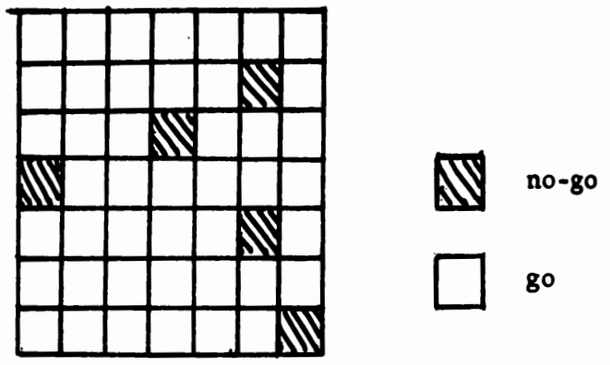
Speed Transfer by Enemy Action

k_{bv1} and k_{bv2} are just the probabilities per unit occurrence that a speed transfer will take place.

Speed Transfer by Terrain

The nature of terrain is fundamentally different in the cases of on-road and off-road.

Off-Road: From terrain features such as slope and soil conditions, a go/no-go function can be evaluated for an area which will determine whether a specific vehicle type will or will not stall in this area. If we consider a region R (see Figure 1) and impose a grid on it giving go/no-go values for different cells,



R

Figure 1.

a density β_{bvl} can be defined as

$$\beta_{bvl} = \frac{\text{number of no-go cells in the region}}{\text{area of the region}} .$$

Then k_{cvl} becomes the probability per unit no-go that a unit will be transferred given in terms of length perpendicular to the direction of travel.

On-Road: Consider a road net in a given region. Impose upon it a rectangular grid (see Figure 2) such that the map in each cell is connected and such that each crossroad falls within a cell. Now, in each cell of this grid, construct a linear approximation to the road map such that the maximum distance from the road to the linear segment is a preassigned value, say ϵ, and such that the segment is tangent to the road at its contact points. Now consider the following two quantities:

Φ = sum of the angles that adjacent segments made with each other then considered as vectors located in first and second quadrant with respect to a preferred direction y.

θ = the absolute value of the signed sum of the angles that the segments make with a preferred direction when located as above.

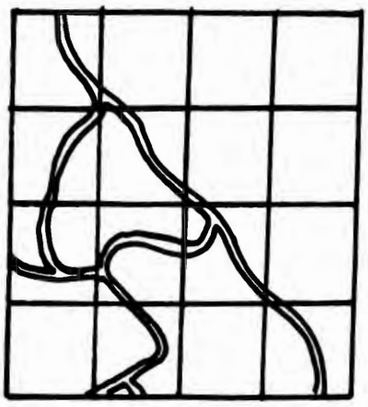


Figure 2.

Further consider the conditions of the road net in a cell. Because of such factors as a bridge out, road demolished, or mud, traffic may be stalled. From these three quantities can be determined estimates of a transfer/no-transfer function for each cell. Thus, as in the case of off-road traffic, a density β_{bv2} and a mass absorption coefficient k_{cv2} can be evaluated.

Scattering by Enemy Action

k_{dv1} and k_{dv2} are just the probabilities per unit occurrence that a scattering event will take place.

Scattering by Terrain

k_{ev1} is the probability per unit no-go that a scattering event will take place. The on-road density β_{ev2} is, however, in general different from β_{bv2}. Consider Figure 2; a scattering event occurs if there are two or more branches issuing from a cell with different tangent vectors. We will assume that the case in which two road branches issue from a cell is fundamentally the same as that in which multibranches issue from a cell. A further refinement would be to make this distinction. Thus we could, in principle, determine a scatter/no-scatter function for each cell and then derive a density β_{ev2} and corresponding coefficient k_{ev2}.

SUMMARY OF
EVENT AND INTERACTION CROSS-SECTIONS
(TABLE 1)

Absorption (losses)	Off-Road	On-Road
Annihilation by enemy action	$k_{av1}\ \beta_{av1}$	$k_{av2}\ \beta_{av2}$
Speed transfer (enemy action)	$k_{bv1}\ \beta_{av1}$	$k_{bv2}\ \beta_{av2}$
Speed transfer (terrain)	$k_{cv1}\ \beta_{bv1}$	$k_{cv2}\ \beta_{bv2}$
Scattering (enemy action)	$k_{dv1}\ \beta_{av1}$	$k_{dv2}\ \beta_{av2}$
Scattering (terrain)	$k_{ev1}\ \beta_{bv1}$	$k_{ev2}\ \beta_{ev2}$
Incidental losses	L_1	L_2
Gains		
Sources	S_1	S_2
Speed transfer (enemy action)	$k_{bv2}\ \beta_{av2}$	$k_{bv1}\ \beta_{av1}$
Speed Transfer (terrain)	$k_{cv2}\ \beta_{bv2}$	$k_{cv1}\ \beta_{bv1}$
Scattering (enemy action)	$k_{dv1}\ \beta_{av1}$	$k_{dv2}\ \beta_{av2}$
Scattering (terrain)	$k_{ev1}\ \beta_{bv1}$	$k_{ev2}\ \beta_{ev2}$
Incidental gains	G_1	G_2

where the coefficients and probabilities are defined in the obvious manner with appropriate dimensions.

Following the spirit of the simple transport model discussed, the specific transport model in which we are interested would have the following terms:

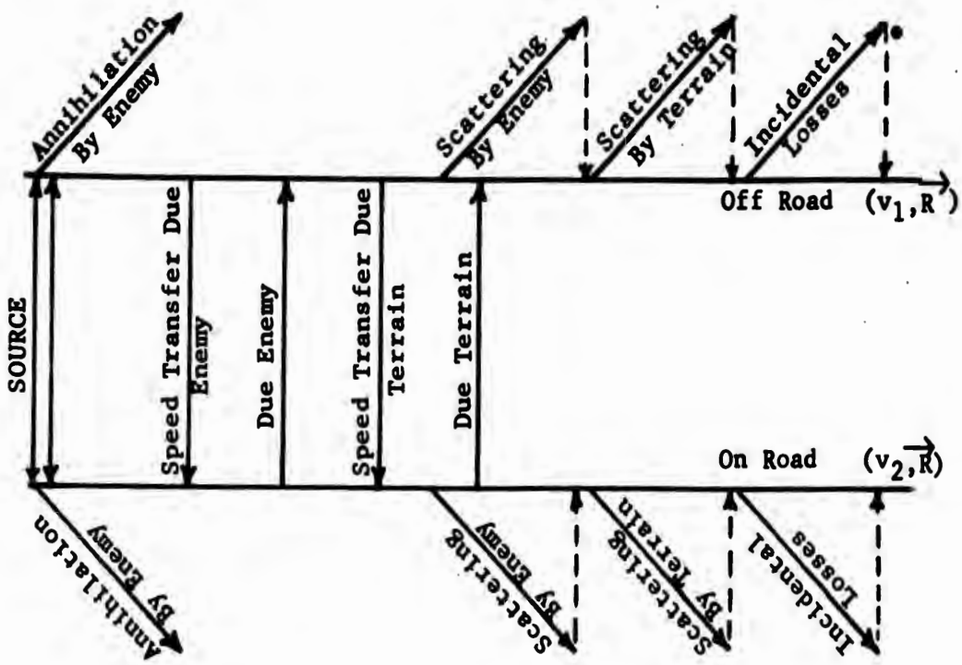


Figure 3. Model.

From inspection of Figure 3 we may write two rate balance equations; one for off-road and one for on-road density. The descriptive form being:

The change in density with respect to time multiplied by the recriprocal of the velocity for that speed is equal to: the negative change in density with respect to distance; minus the sum of the event probabilities of losses times the density; plus the probability of gain from adjacent elements times the density; plus the probability of gain from another speed times the density of the other speed.

Having accounted for all the events of the system we may select from table 1 the corresponding interaction cross-sections for the balance equations: algebraically collecting terms we arrive at the required mathematical model with the density (u_1,u_2) as the dependent variables; and time (t) and position(s) as the independent variables:

$$\frac{1}{v_1}\frac{du_1}{dt} = -\frac{du_1}{ds} - K_1 u_1 + \bar{C}_{t1} u_2 + \bar{C}_{s1} u_1$$

$$\frac{1}{v_2}\frac{du_2}{dt} = -\frac{du_2}{ds} - K_2 u_2 + \bar{C}_{t2} u_1 + \bar{C}_{s2} u_2$$

where

$$\bar{C}_{t1} = \frac{k_{bv2}\,\beta_{av2}}{2\pi} + \frac{k_{cv2}\,\beta_{bv2}}{2\pi}$$

$$\bar{C}_{s1} = \frac{k_{dv1}\,\beta_{av1}}{2\pi} + \frac{k_{ev1}\,\beta_{bv1}}{2\pi}$$

$$\bar{C}_{t2} = \frac{k_{bv1}\,\beta_{av1}}{2\pi} + \frac{k_{cv1}\,\beta_{bv1}}{2\pi}$$

$$\bar{C}_{s2} = \frac{k_{ev2}\,\beta_{ev2}}{2\pi}$$

$$K_{v1} = \left[k_{av1} + k_{bv1} + k_{dv1}\right]\beta_{av1} + \left[k_{cv1} + k_{ev1}\right]\beta_{bv1}$$

$$K_{v2} = \left[k_{av2} + k_{bv2} + k_{dv2}\right]\beta_{av2} + k_{cv2}\,\beta_{bv2} + k_{ev2}\,\beta_{ev2}$$

Arbitrary source and loss terms have been dropped.

BIBLIOGRAPHY

1. Ambarzumian, V. A., Theoretical Astrophysics, Pergamon Press, New York, 1958.

2. Anselone, Philip, A., "The Uniform Convergence of Chandrasekhar's Approximations in Classical Transfer Problems", Mathematical Research Center Summary Report 97, University of Wisconsin, Madison, Wisconsin, October 1959.

3. Chandrasekhar, S., Radiative Transfer, Dover Publications, New York, N.Y., 1960.

4. Glasstone, S., and Edlund, M. C., The Elements of Nuclear Reactor Theory, D. Van Nostrand Company, Inc., New York, N. Y., 1952.

5. Morse, P. M., and Feshbach, Herman, Methods of Theoretical Physics, McGraw-Hill Book Company, New York, N.Y., 1953.

6. "Notes Paralleling S. Chandrasekhar", attributed to P. Anselone.

7. Weinberg, A. M., and Wigner, E. P., The Physical Theory of Neutron Chain Reactors, University of Chicago Press, Chicago, Illinois, 1958.

EXOTHERMIC REACTIONS BEHIND REFLECTED SHOCKS

ARTHUR COHEN, ROBERT A. WIRES, and ROGER A. STREHLOW
BALLISTIC RESEARCH LABORATORIES
ABERDEEN PROVING GROUND, MARYLAND

I. Theory

In the past decade the shock tube has become an important facility for studying fast gas phase reactions at high temperatures. The advantages of shock tube methods over other techniques are:

a) rapid heating (shock front is of the order of a micron thick)

b) uniform conditions which can be varied by changing shock strength

c) reactions are homogeneous

d) thermodynamic properties of heated gas are easily calculated from a shock velocity measurement.

Most of the early shock tube work[1] was conducted in the gas behind the incident shock. Recently[2,3] there has been increased interest in using the gas behind the reflected shock wave. The advantages of using this region over the incident region for chemical studies are:

a) higher temperatures for the same initial conditions

b) stationary test gas which leads to better time resolution

However, there are limitations for using this regime. The calculation of thermodynamic properties from a shock velocity measurement is straightforward only when the assumption of one dimensional, steady state gas flow is made. Under certain conditions, the interaction of the reflected shock with the boundary layer[2] (and pressure waves) in the incident gas make these assumptions questionable. Previous work in this laboratory indicates that the steady state one dimensional conditions are most closely attained near the back wall of the shock tube and in mixtures with small heat capacities. We have used the reflected shock wave to initiate reaction in exothermic systems,

diluted with Ar, at the back wall of a shock tube.

It is well known that the initiation of highly exothermic reactions by shock waves in a gas confined in a tube will lead to a C.J. detonation wave. However, for reactions which require a high enough activation energy, it is possible to find a range for the incident shock velocity, subject to some composition limitations, where reaction will not occur in the incident gas but will be initiated by the reflected shock. This has been done for mixtures of stoichiometric H_2-O_2 in Ar and B_2H_6 in Ar.

II. Experimental

A conventional shock tube consists of a high pressure section separated from the test section by a diaphragm. After puncturing the diaphragm, the shock wave is produced by the piston-like action of the expanding high pressure air. It has been found[4] that ignition of explosive gases can be caused by local reflections of the incident wave from cracks in the tube wall or by bursting the diaphragm in the explosive mixture. Therefore, precaution was taken in the construction of our shock tube. A four inch i.d. stainless steel shock tube with a 50 inch compression section, an 88 inch buffer section and a 164 inch test section was used in these experiments. The buffer section was separated from the test section by a four inch stainless steel ball valve (a six and one half inch sphere with a four inch hole bored through it) which allowed undisturbed passage of the shock when in the open position. When closed the valve was vacuum tight and allowed us to place an inert gas, of the same density as the explosive gas, in the section near the diaphragm and in the bore of the valve. Pressure differential across the valve was adjusted to zero externally and the valve was opened about 10-20 seconds before firing the tube. This effectively prevented any premature detonations initiated by non-ideal diaphragm burst. Also, the test section was constructed to eliminate all internal crevices at the joints and window mountings in order to eliminate premature detonations caused by local reflected shocks. Strip photographs (distance-time) of reflected shock behavior were taken with an eight inch schlieren system through 8" X 0.025" slit windows at the back wall. A rotating drum camera was used to record density gradients (and radiation) on 35mm Tri-X film. The light source was a No. 50 flash bulb. Timing dots are placed on the film through a separate optical system by flashing a Sylvania 1130B cold crater lamp at a frequency of 10,000 cycles/sec from pulses supplied by a secondary frequency standard. Standard distance markers are placed in the schlieren field of view near the shock tube to provide a means of calculating the magnification factor of the optical system. The optical system is synchronized to the operations of the shock tube by using conventional thyratron variable delay circuits.

Strip film negatives are read directly using a Kodak Contour Projector. A typical schlieren photograph of an incident and reflected shock followed by a delayed reaction in the H_2-O_2-Ar system is reproduced in Figure 1a. The effect of shock strength on delay times

is shown in Figure 1b. Shock velocities can be measured with a precision of ±2%. The measurement of a delay time involves, in some cases, an extrapolation of the reaction wave to the back wall. The precision of this measurement is about ±5%. Mixtures were stored for at least three hours in stainless steel tanks before using. H_2, O_2, and Ar were obtained from the Matheson Co. and used without further purification. B_2H_6 was obtained from the Callery Co. and was purified by fractional distillation with liquid N_2. Mass spectrographic analysis indicated the impurities in all gases were less than 0.4%. The leak rate in the tube was always less than 20 μ Hg/min. A machine calculation of reflected shock conditions as a function of incident shock velocity and the initial condition was performed assuming one dimensional flow, vibrational equilibrium, and no reaction behind either the incident or reflected shock.

III. Results in H_2-O_2

It is known that autoignition of exothermic systems in a closed tube can lead to formation of combustion waves. The structure of these waves can be described as a shock wave preceding a flame which is then followed by a rarefaction wave. The "reaction" wave we observe in the H_2-O_2-Ar system conforms to this structure and is due to ignition at the back wall of the shock tube. Analysis of the formation of combustion waves is complicated by the interaction of kinetic and hydrodynamic phenomena. It seems reasonable to assume that if a temperature profile through a reaction zone were known, then ignition and hence wave formation would occur somewhere near the maximum temperature gradient. For reactions which proceed by a chain branching mechanism this corresponds to a time near the end of the induction zone. Duff[5] used a numerical procedure to solve a system of 9 kinetic equations and obtained the reaction profile behind a shock wave moving through a H_2-O_2-Xe mixture. Nicholls[6] has attempted to solve this system of equations by an analytical procedure in order to relate experimentally determined ignition delays in H_2-O_2 mixtures to the rate constants of the system. He obtained an equation for [H] vs time near the end of the induction and assumed that a characteristic time for ignition was determined by the inflection point of the equation. The following expression was obtained:

$$\tau = \frac{1}{2[O_2]k_1} \ln \frac{[O_2]k_1}{[H_2]k_4} \qquad (A)$$

where

τ = characteristic ignition time delay

$[\ ]$ = concentration in mol/liter

$[O_2]$, $[H_2]$ = initial concentration of O_2 and H_2 respectively

k_1, k_4 = rate constants of the reaction shown below

$$H_2 + O_2 \xrightarrow{k_4} 2OH \qquad (4)$$

$$O_2 + H \xrightarrow{k_1} OH + O \qquad (1)$$

Values of τ were obtained from measurements of distance between shock and flame front in a stabilized detonation. Using $k_1 = A_1 \exp(-E_1/RT)$ his data, though scattered, indicate a value near 18 Kcal/mol for E_1. (E_1 is the activation energy of reaction (1).)

Schott and Kinsey[3], using an ultraviolet absorption technique, have recorded the rate of increase in $[OH]$ behind incident and reflected shocks. They considered the following reactions

$$H_2, O_2, M \xrightarrow{R_o} H, O, OH \qquad (0)$$

$$H + O_2 \xrightarrow{k_1} OH + O \qquad (1)$$

$$O + H_2 \xrightarrow{k_2} OH + H \qquad (2)$$

$$OH + H_2 \xrightarrow{k_3} H_2O + H \qquad (3)$$

(R_o, k_1, k_2, k_3 are the rate constants for reactions (0) - (3) respectively)

to describe the reaction profile during the induction period. By assuming a steady state for $[O]$ and $[OH]$ they derive the expression

$$n = \log \frac{[H]}{[H]_o} = \log \frac{[OH]}{[OH]_o} = .8686\, k_1 [O_2]\, t \qquad (B)$$

where t = time after shock passage and $[H]_o$ = value of $[H]$ which satisfies $2k_1 [O_2][H] = R_o$. They measured a time t_i at which the $[OH]$ became detectable (1×10^{-6} mol/liter). By assuming the value of n in (B) at t_i is independent of temperature their results can be expressed as $\log [O_2] t_i = -10.647 + \frac{3966}{T}$ with E_1 = 18.1 Kcal/mol.

We have measured the times τ for the formation of a combustion wave at the back wall after shock reflection. Figure 1 is typical of the schlieren photographs from which data has been obtained. For some film, delay times were measured by extrapolating the reaction wave to the back wall.

All our data was taken in stoichiometric H_2-O_2 mixtures diluted with 70%, 85%, and 94% Ar. Reaction was observed in the

ranges 920°K $<$ T_{rs} $<$ 2000°K and $2.5\times10^{-4} < [O_2] < 1.6\times10^{-3}$ where T_{rs} and $[O_2]$ are the temperature and oxygen concentration behind the reflected shock. Delays ranged from 6 to 880 microseconds. Measurements of the incident and reflected shock velocities indicated that at shock reflection the gas was unreacted. A least squares fit of our data to equation (B) leads to the expression $\log [O_2]\tau = -10.42 + \frac{3760}{T}$. Our delay data are shown in Figure 2 and are compared with the results of Schott and Kinsey (dotted line in Figure 2). They determined a value of 17.5 Kcal/mol for E_1 by comparing their experimental value for k_1 at 1650°K with an accepted value at 793°K. This approach was necessary since the value of n in equation (B) at t_i has some temperature dependence. Near T_{rs} = 1000°K our delays are, on the average, a factor of three greater than theirs. If we assume a relationship independent of temperature, exists between τ and some time t_n of the induction zone (i.e., time for a n-fold increase in $[H]$) then the activation energy of reaction (1) can be obtained from the variation of $\log [O_2]\tau$ with $\frac{1}{T}$. This assumption leads to a value of 17.2 Kcal/mol for E_1 which agrees with Schott and Kinsey's result.

Results in B_2H_6-Ar

The formation of solids behind the reflected shock in B_2H_6-Ar mixtures has been recorded using schlieren photography. Figure 3 contains typical photographs taken in a B_2H_6-Ar mixture and a B_2H_6-N_2-Ar mixture. The reaction is evident by the disappearance of the schlieren light source some time after shock reflection. (The multiple gradients surrounding the reflected shock are due to bifurcation[2].) Mixtures of 10%, 20%, and 30% B_2H_6 were used. Delays from 10 to 500 microseconds were observed in the temperature range 770°K $<$ T_{rs} $<$ 1050°K. Velocity measurements indicate the gas is unreacted at shock reflection.

The data for the decomposition of B_2H_6 was analyzed by assuming the reaction wave is initiated by an adiabatic thermal explosion at the back wall. The following approximate expression for the induction time τ in adiabatic thermal explosions was derived by Todes[7]:

$$\frac{1}{\tau} = \frac{E}{RT}\frac{Q}{\bar{C}T}W \qquad (C)$$

τ = induction time, Q = heat of reaction/mol of reactants, T = initial temperature, E = activation energy/mol, W = reaction velocity, $\bar{C}$ = average heat capacity/volume. Assuming Q = constant, $W = k[B_2H_6]^n[M]^m$ (where $[M]$ is total concentration) and $k = Ae^{-E/RT}$ equation (C) can be written as

$$\frac{[B_2H_6]^n[M]^m\tau}{\bar{C}\,T_{rs}^{\,2}} = \text{constant} + \frac{E}{4.58}\frac{1}{T} \qquad (D)$$

The delay data was compared with equation (D) for various values of n and m. The best fit was obtained for n = 1/2, m = 1. The results are

shown in Figure 4 and lead to an overall activation energy of 24 Kcal/mol. Our efforts to identify the solid products have not been successful. Therefore, at present, it is difficult to compare our results with mechanisms suggested by other investigations of B_2H_6 pyrolysis. It is hoped that observations made during the induction period, using a spectrographic absorption technique, will lead to detailed information concerning the overall pyrolysis reaction.

IV. Conclusions

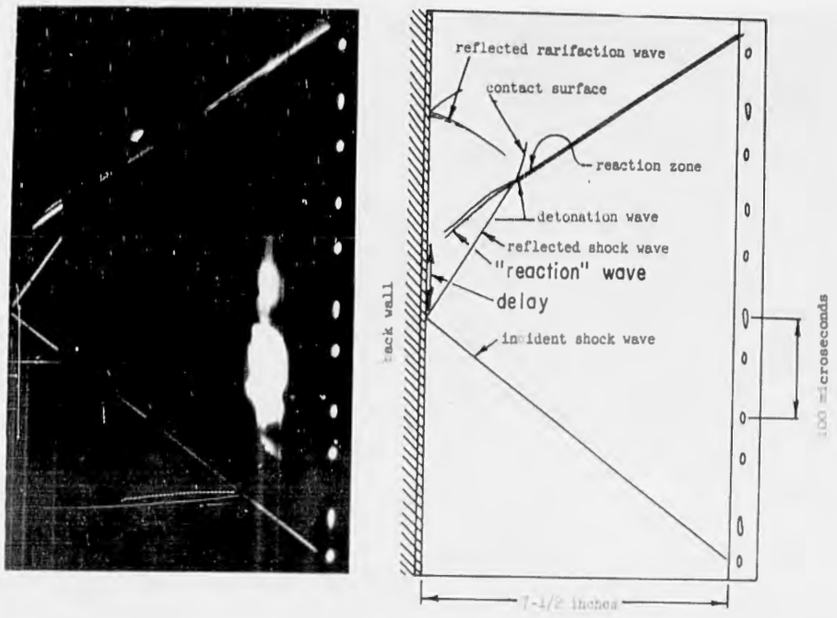
The results presented here indicate that shock tube studies of exothermic reactions can be made at the back wall of a shock tube using the reflected shock regime. Under these conditions the initial reaction is not coupled to the reflected shock and experimental conditions can be controlled as well as calculated from a measurement of the incident shock velocity. An activation energy of 17.2 Kcal/mol has been obtained for the reaction $H + O_2 \longrightarrow OH + O$. This is in agreement with results of Schott and Kinsey[3] and Nicholls[6]. The reaction velocity for the formation of solids in the B_2H_6-Ar system can be represented as $W = k[B_2H_6]^{1/2}\left\{[B_2H_6] + [Ar]\right\}$. The corresponding activation energy is 24 Kcal/mol.

Acknowledgement:

The authors wish to express appreciation to Dr. Eli Freedman for helpful discussions during the preparation of this manuscript.

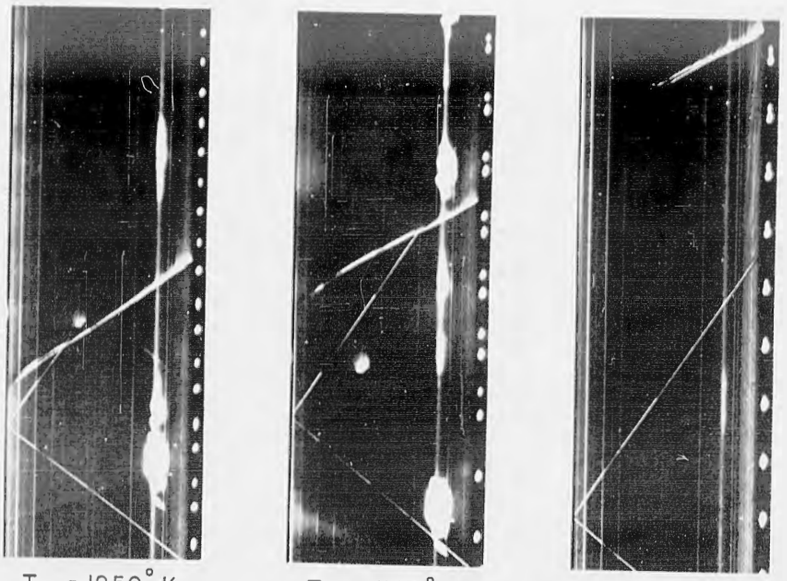
References:

1. H. Pritchard, Quarterly Reviews, 14 46 (1960)
2. R. Strehlow and A. Cohen, J. Chem. Phys. 30 257 (1959)
3. G. Schott and J. Kinsey, J. Chem. Phys. 29 1177 (1958)
4. M. Stienberg and W. Kaskan, "5th Symposium (International) on Combustion", p. 664-71, Reinhold, New York (1955); J. Fay, "4th Symposium (International) on Combustion", 501-07, Williams and Wilkins, Baltimore (1953)
5. R. E. Duff, J. Chem. Phys. 28, 1193 (1958)
6. J. A. Nicholls, "Stabilization of Gaseous Detonation Waves with Emphasis on Ignition Time Delay Zone" AFOSR TN 60-442.
7. W. Jost, "Explosion and Combustion Processes in Gases", p. 18, McGraw-Hill, New York (1946)



The initiation of detonation behind a reflected shock in an $0.2\ H_2$, $0.1\ O_2$, 0.7 argon mixture. $M_s = 2.40$, $T_{rs} = 1260°K$, $t_1 = 35$ microseconds.

a ↑ Fig. 1 b ↓



$T_{rs} = 1250° K$
$\tau = 55\mu$ SEC

$T_{rs} = 1140° K$
$\tau = 158\ \mu$SEC

$T_{rs} = 1010° K$
$\tau = 573\mu$ SEC

THE INITIATION OF DETONATION IN $0.100_2 - 0.20\ H_2 - 0.70$ ARGON MIXTURES AT VARIOUS TEMPERATURES, $P_0 = 2$ CM

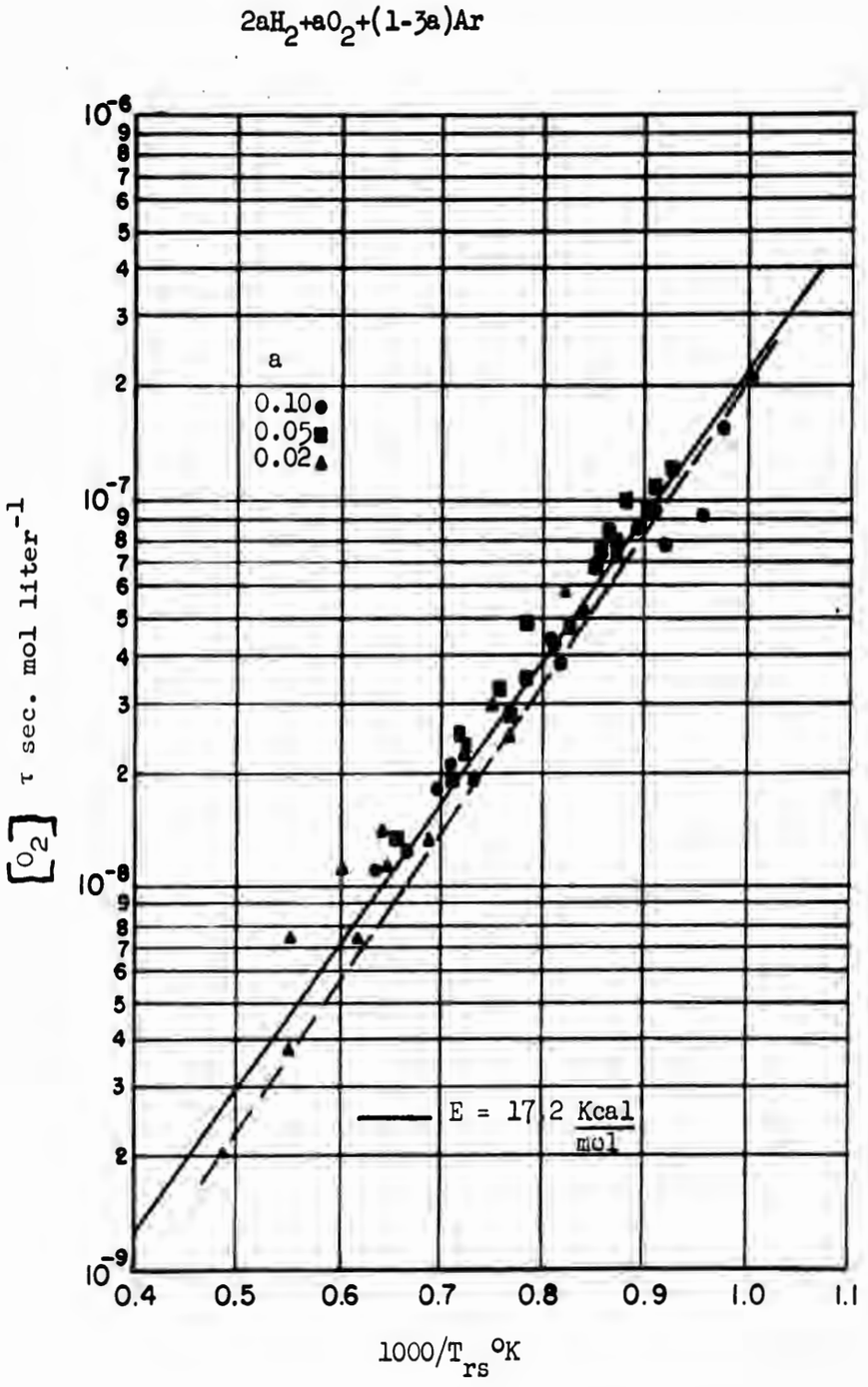
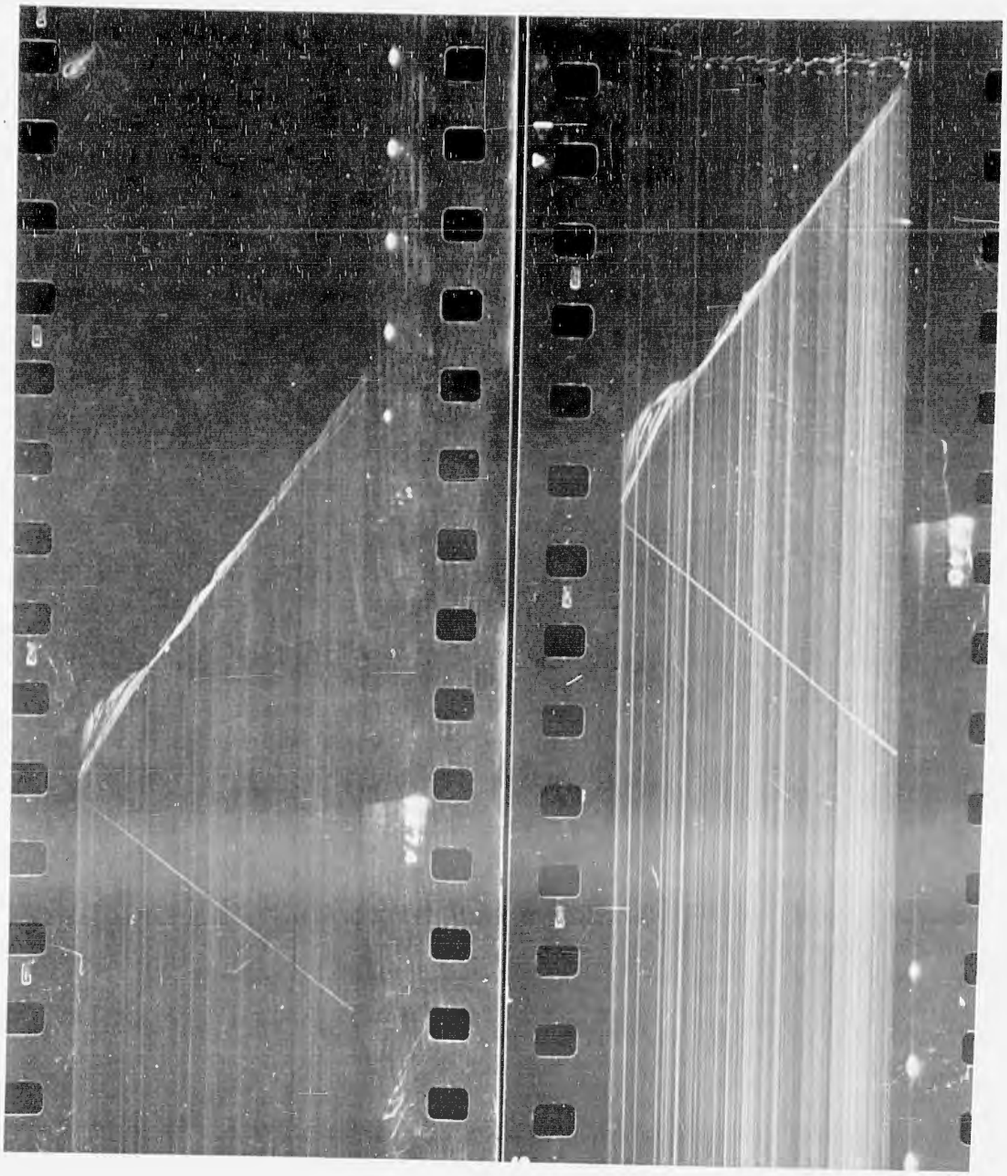


Figure 2. Delay data for stoichiometric hydrogen-oxygen-argon mixtures.



0.2 B_2H_6,0.1 N_2,0.7 Ar		0.2 B_2H_6,0.8 Ar
950	T_{rs} (°K)	949
4	P_o (cm.Hg)	2
73	τ (μsec.)	77

Fig. 3

Initiation of the condensation reaction in diborane-argon and in diborane-nitrogen-argon mixtures.

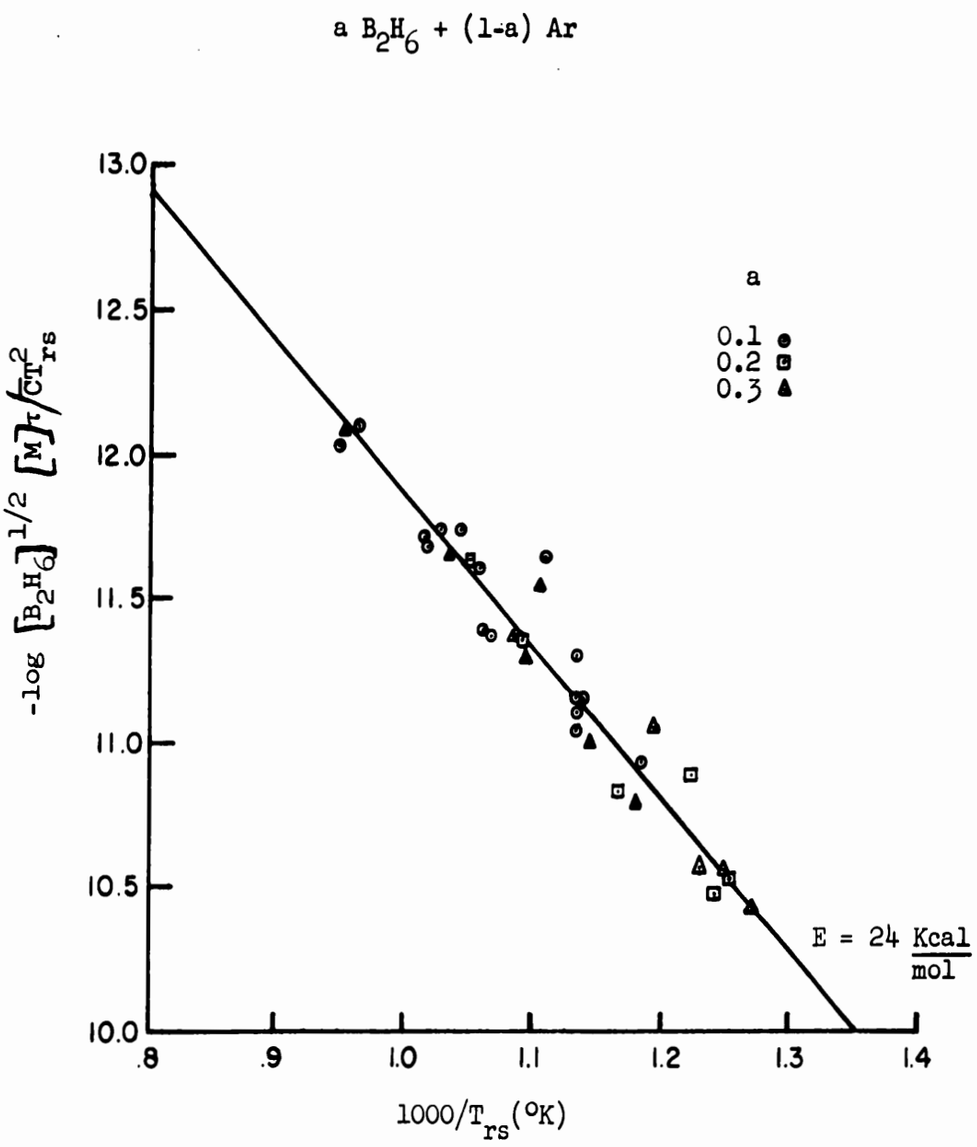


Figure 4. Delay data for diborane-argon mixtures.

NUCLEAR TIMING

RAYMOND H. COMYN, MAURICE APSTEIN
DIAMOND ORDNANCE FUZE LABORATORIES
WASHINGTON, D. C.

Timers are used to control the functioning of many Ordnance devices at the end of predetermined time periods. The duration of these time periods may vary from the few seconds needed by missile safety and arming programmers to the days or months required by demolition or satellite applications. The size, accuracy and ruggedness of timers have become more and more important in recent years, and in many cases the strict requirements can be met only by using an electronic digital approach such as that used in the Copperhead programmer.(1) In this type of timer, an oscillator feeds electrical pulses at a constant rate into a series of binary counters and associated gates(2) which indicate when a predetermined number of pulses have been reached at the end of the desired time period. Each "flip-flop" binary can assume only one of two possible states and shifts from one state to another each time a pulse enters the circuit. Hence, it acts as a frequency divider with the output of one stage forming the input of the next. With this arrangement, the electrical characteristics of circuit components are not so critical as those used in analog circuits and considerable component variation can be permitted in manufacture as well as in operation over a wide range of ambient conditions.

However, the operation of the oscillator is critical since the accuracy of the timer depends on maintaining a stable frequency under all operating conditions. Oscillators using tuned L-C circuits and R-C sawtooth generators have an inherent weakness because of variations in circuit components. Natural resonators such as crystals or tuning fork oscillators have more inherent stability but require careful handling and temperature control.(1) Moreover, the timing error is cumulative and proportional to elapsed time. Therefore, a proposal has been made to use the random emission of particles from nuclear disintegration as a time base which would be independent of ambient conditions. Such a time base

has the interesting characteristic that it becomes proportionately more accurate as the timing interval increases, which makes it particularly applicable for long time delays.

1. Proposed Nuclear Timer

A timer operating on this principle consists of four parts:(3) A radioactive source which emits particles at a predictable average rate; a detector and amplifier which converts the nuclear emission into electrical pulses; and a binary system which counts the pulses and indicates the end of the predetermined time period when a definite cumulative output of pulses has been attained.

Radioactive disintegration is a random phenomenon and the number of particles resulting from radioactive decay will follow a binominal distribution in time. If a fixed proportion of the particles ejected from a definite quantity of radioactive material reaches a detector, the standard deviation of the resulting number of electrical pulses will be equal to $\sqrt{N}$ where N equals the total average number of pulses in a given time.

For large values of N, the number of pulses can be expected to be within $N \pm \sqrt{N}$. It follows that the fractional standard error is $\frac{\sqrt{N}}{N} = \frac{1}{\sqrt{N}}$ and the larger the number of pulses the smaller the fractional standard error. The one and three sigma distribution limits are shown for several values of N in Table 1 below:

TABLE 1

EFFECT OF TOTAL NUMBER OF PULSES ON PERCENTAGE ERROR

N	$\sqrt{N}$	Percentage Error for ± 1 Sigma Limits	Percentage Error for ± 3 Sigma Limits
100	10	± 10%	± 30%
1,000	31.6	± 3.2	± 9.5%
10,000	100	± 1%	± 3%
100,000	316.2	± .32	± .96
1,000,000	1000	± 0.1%	± 0.3%
10,000,000	3162	± 0.03%	± 0.1%

It must be emphasized that the percentage accuracy of the nuclear timer depends on the total number of pulses counted during the desired time period and that this percentage accuracy will be the same whether the total number of pulses are counted within 10 seconds or 10 months.

The counting rate, of course, will depend on the required accuracy and the total time period. These relationships are illustrated in Table 2 below:

TABLE 2

RELATIONSHIP BETWEEN TIME AND COUNTING RATE FOR 0.1% ERROR

Total Time	(3 Sigma Limits) Total Counts Required	Counting Rate Counts per second
10 seconds	9,000,000	900,000
100 seconds	9,000,000	90,000
1 hour	9,000,000	2,499.3
1 day	9,000,000	104.13
1 month	9,000,000	3.472

2. Nuclear Sources

The radioactive source used in a nuclear timer must have a number of well-defined characteristics. First, it must be safe to handle and must neither expose nearby personnel to harmful radiation, nor supply personnel to the collective radiation of a large number of timers in storage. The source must have a long half life to insure a relatively constant emission rate over a period of years. The radioactive material must be physically stable to prevent its escape from the timer during handling. Also, the radioactive emission from the source must be nearly monoenergetic. If the emission has a range of energies as in a beta source, for example, the height of the electrical pulses produced by the detector will vary continuously from the level of electrical noise to their maximum value. This would be satisfactory if the pulse heights and the voltage discrimination level were constant but when either one of these parameters shifts, the rate of pulse generation changes. Unfortunately, the detector sensitivity and gain of the amplifier system both vary to some extent with variations in supply voltage and ambient temperature. However, if the nuclear emission is monoenergetic there will be a considerable voltage difference between the level of electrical noise and pulses. When this occurs, the discrimination level can be set so that minor variations in the level or pulse height will not affect the counting rate.

Two nuclear sources, Americium^{241} and Plutonium^{239} appear to be well-suited to this application. They both emit alpha particles near 5 Mev and low level gamma radiation which can be shielded readily. The respective half lives are 475 and 24,100 years. Although the disintegration rate of Americium^{241} decreases about 0.7% in 5 years, it is readily available and satisfactory for laboratory tests and other uses where long term storage stability is not a problem. Plutonium^{239} is recommended for uses where high accuracy is required over long storage periods. Alpha source offer

another advantage in comparison to gamma sources; if used with detectors which are sensitive only to alpha particles, the timer can be shielded easily from extraneous radiation.

Americium^{241} has been used as an alpha source for all of the experimental work at DOFL. These sources were prepared at the National Bureau of Standards by diffusing Americium^{241} into a Monel metal or stainless steel surface. However, these sources also may be prepared by electroplating or fused oxide techniques.

3. Radiation Detectors

Although there are many detectors capable of detecting the alpha radiation, most have shortcomings which make them unsuitable for Ordnance use. The following types were examined:

SCINTILLATION DETECTORS: In a scintillation detector, the kinetic energy of the emitted particle is converted into light, and the latter is converted into electrical energy by a photosensitive device. Photomultiplier tubes which are commonly used in scintillation detectors are extremely sensitive to shock and require a voltage supply of about 1500 volts. For these reasons, photomultiplier tubes do not appear suited to this application.

The possibility of using photodiode semiconductors was investigated. However, no photodiode with the required sensitivity could be found.

PROPORTIONAL DETECTORS: The proportional detector is a gas filled tube which requires an operating voltage between 700 and 1500 volts. The output pulse height is determined by the length of the ion track produced by a nuclear particle and factors governing gas multiplication. The latter is sensitive to pressure and, therefore, the pulse height in a sealed tube would be sensitive to ambient temperature. A few preliminary tests using Americium^{241} in a proportional tube showed that the output pulse was very pressure sensitive and that this detector did not appear suited to this application.

GEIGER MULLER DETECTORS: These also are gas filled tubes operating at relatively high voltages but differ from the proportional detector in that the output is not dependent on the energy of the nuclear particle since the latter serves only to trigger a discharge. As a result the output pulse is always about 1 volt and this device also possesses the advantage that no further amplification is required. However, any radiation that produces ionization will produce a pulse and the detector cannot distinguish between source and background radiation.

Commercial Geiger tubes specified to operate over the military ambient temperature range of -65°F to +165°F were selected

for these tests. The detectors were made with a thin mica end window for alpha detection and contained bromine as a quenching agent. The initial tests were run with an Americium^{241} source clamped firmly to the end window. In all tests, the Geiger tube and the source were allowed to come to equilibrium in a controlled temperature cabinet while the required instrumentation remained at room temperature. Although the counting rate remained practically constant[4] at 7600 pulses per second at both room temperature and +165°F, it dropped to 5300 pulses per second at -65°F. It was thought that the quenching agent might have affected the counting rate at low temperatures and additional tubes were procured with chlorine in place of bromine for the quenching agent.

The counting rates for the chlorine quenched tubes[4] were 6400 counts per second at +165°F, 6100 at room temperature and 4500 at -65°F. These data showed that this system was not suitable for use in the proposed timer.

4. Solid State Detectors

Solid state detectors[5, 6, 7] have a number of advantages for timing applications when used with alpha sources. In addition to their small size they operate at relatively low voltages and are not normally sensitive to gamma and beta radiation. In fact, they are not affected appreciably by cosmic radiation and, therefore, the pulses generated with this system will depend almost entirely on the alpha source.

Semiconductor alpha detectors operate by converting the kinetic energy of the alpha particle into a proportional electrical signal. In one type of semiconductor detector the particles enter a p-n diode junction in silicon and liberate electron hole-pairs by collisions. A bias of a few volts (5-50V) sweeps these pairs out of the junction into the electronic circuit. Silicon-gold surface-barrier detectors operate in a similar fashion. The p-n diodes usually are prepared by diffusing phosphorus into high resistivity p-type silicon while the surface barrier units are made by evaporating a thin gold film onto the silicon surface.

SILICON P-N JUNCTION DETECTORS: Initial tests were made with Americium^{241} deposited directly on the surfaces of silicon p-n detectors purchased from two commercial sources. Data collected after a few days showed that the detectors began to deteriorate after receiving about 10^{10} or 10^{11} alpha particles. This means that nuclear timers with high pulse rates would probably require a mechanical shutter to protect the detector during storage.

After the damage to the detector became apparent the pulse height distribution from one of the detectors was determined using a multi-channel analyzer. The 100-channel analyzer was used with the amplifier gain adjusted to give the pulses accumulated in

channel 1 a height between 0 and 1 volt. Pulses in channel 2 had voltages between 1 and 2 volts, etc.

After one of the detectors (Supplier B) shown in Figure 1 had been exposed to 7 x 10" alpha particles, its pulse height distribution was measured (Figure 2). Pulse height distributions were then measured using a similar unused detector with Americium^{241} deposited on the surface. These data, Figure 3 when compared with Figure 2, show the effect of alpha radiation on the height of pulses from the detector. Apparently, prolonged exposure of these detectors to high intensity alpha radiation reduces the height of the pulses and as the exposure continues, fewer of these pulses are large enough to pass the discriminator and enter the counting circuit.

The test fixture shown in Figure 4 was constructed of Invar, an iron nickel alloy with a low coefficient of expansion, and was designed to perform the following functions: (1) Hold the source and detector at a fixed relative distance over the temperature range -65°F to +165°F; (2) shield the detector completely when not in use; (3) collimate the beam of alpha particles; (4) partially shield the detector for a low count rate suitable for pulse height measurements. Since there was some uncertainty as to the physical stability of the high intensity alpha source the apparatus was completely sealed after the source was placed inside with no mechanical connections of any kind between the source chamber and the outside. The device (Figure 4) had three stable positions and by rotating it 360°, the small shields could be moved in and out of position in between the source and detector. One of the shields was blank and the other has a small hole to reduce the counting rate.

The number of pulses emitted by the detector in 100 seconds was measured at -65°F, +75°F and +165°F. The detector and source holder were soaked at these temperatures for over 20 minutes before each test was made with the instrumentation remaining at room temperature. During these tests the bias voltage on the detector was held at 24 volts and the discriminator was set to pass all pulses above 0.03V. The signal voltage at -65°F and room temperature was 0.20V and dropped to 0.19V at +165°F. The noise level was 0.02V at the two lower temperatures and rose to almost 0.03V at +165°F. The data are shown in Table 3 below:

TABLE 3

TOTAL PULSES MEASURED IN 100 SECONDS USING AMERICIUM 241 SOURCE AND A P-N DIODE DETECTOR

+75-F Before Temperature Tests	+75°F After Temperature Tests
2,653,281	2,655,176
2,656,394	2,653,375
2,653,837	2,658,984
2,654,592	2,658,831
2,655,679	2,655,851
Average: 2,654,757	Average: 2,656,443
Std. Deviation = 1283	Std. Deviation = 2425

-65°F	+165°F
2,612,457	2,663,661 *
2,613,873	2,651,841
2,614,585	2,654,908
2,614,447	2,651,482
2,614,061	2,651,108
2,611,200	2,656,294
2,610,722	2,654,745
2,612,866	2,650,179
Average: 2,613,026	Average: 2,652,936
Std. Deviation = 1475	Std. Deviation = 2334

* Omitted from calculations

TABLE 4

SUMMARY OF DATA FROM TABLE 3

		Std.Deviation Calculated from Test Data	Std.Deviation Calculated from Theory*
-65°F	2,613,026	1475	1616
+75 Before Temperature Tests	2,654,757	1283	1630
+75 After Temperature Tests	2,656,443	2425	1630
+165°F	2,652,936	2334	1629

*Std. Deviation = $\sqrt{N}$

Based on a 99% confidence level the data from Table 3 and 4 measured before and after the temperature tests show that the detector did not change during these tests. Also, the variances calculated from these test data agree, considering the small sample size with the variances calculated from theory. However, the cold temperature data show an unexpected lower pulse rate. Since there is no theoretical reason to expect a difference in temperature coefficient of the source, this error is believed due to a variation in the proportion of the total pulses reaching the detector, or gain variations in the amplifier and discriminator, and is being investigated further. Either of these causes could easily have accounted for the difference, and should be susceptible to compensation by conventional electronic techniques.

GOLD BARRIER SILICON DETECTORS: Preliminary tests have been run using gold barrier silicon detectors and Americium^{241} sources. In order to eliminate detector deterioration and to reduce testing difficulties, the disintegration rate of the source was kept below 100 counts per second. Sources containing such a small quantity of Americium^{241} are physically stable when made by diffusion techniques and require no special handling precautions. The source holder shown in Figure 5 was used to collimate the stream of alpha particles. The source holder and collimator were made of one piece of metal and this arrangement together with the relatively large detector area insured that small changes in the relative position of source and detector with ambient temperature would not change the pulse rate. The size of this detector and source holder were selected for convenience. Eventually, it is expected that the size of detector and source can be reduced to a package 1/4" diam. and 1/4" long. Tests were run with the source and detector at -65°F, +75°F +165°F. The instrumentation remained at room temperature. The data in Table 4 were obtained by counting the pulses falling within a 5 volt range during a one minute period.

TABLE 4

PULSE HEIGHT DISTRIBUTION USING AN AMERICIUM^{241} SOURCE AND GOLD BARRIER SILICON

Voltage Range	Pulses Per Minute in Indicated Range		
	-65°F	+73°F	+165°F
15 - 20	3	4	10
20 - 25	0	4	6
25 - 35	2	1	19
30 - 35	1	0	88
35 - 40	3	2	3
40 - 45	1	4	0
45 - 50	2	3	0
50 - 55	2	4	0
55 - 60	6	1	0
60 - 65	15	3	0
65 - 70	12	4	0
70 - 75	35	23	0
75 - 80	115	42	0
80 - 85	253	220	0
85 - 90	6	165	0
90 - 95	0	0	0

These data show that the height of pulses from this type of detector are affected markedly by ambient temperature. Apparently, the pulse heights dropped markedly at +165°F and many of the electrical pulses caused by the alpha radiation fell into the noise region below 15 volts. This detector does not appear to be as satisfactory as the p-n diode detectors for timing applications.

5. Further Detector Investigations

Two additional types of semiconductor detectors are being investigated. One is a transistor type of nuclear particle detector.[8] In this device a transistor pulse current flows after a particle produces ionization in entering the collector depletion region. Typical gains are between 200 and 300, and the signal to noise ratios are comparable to the values obtained with good amplifiers and diodes of similar area.

A second promising semiconductor detector has been fabricated by diffusing boron (p+) and phosphorus (n+) into opposite faces of a slab of high resistivity silicon. The capacitance of this n^+ip^+ detector approaches a constant value upon the application of a few volts bias. It appears to have good characteristics for low noise systems.

A study is being made to determine methods of eliminating the low amplitude pulses (apparently caused by some alpha particles) which fall between noise pulses and the maximum values generated by most alpha particles. The cause of these low amplitude pulses is not readily apparent since they are found in p-n detectors with Americium^{241} on the surface, in p-n detectors used with collimated alpha particles, and with gold barrier detectors and collimated alpha particles emitted from a source 0.2" distant. Evidently they are not caused by repeated collisions with air molecules, or by low angle alpha particles. In each of the cases mentioned, the ratio of pulses in a one volt channel to the total alpha pulses is between .0008 and .0011. The cause of this phenomenon is probably either in the silicon itself or in lower energy alpha particles that have been attenuated in some manner in the source, perhaps by depth of diffusion. In any case, this difficulty must be corrected before a highly accurate nuclear timer of this type can be made.

6. Electronic Circuits

Since the alpha particles generate a very limited quantity of electrical energy in the detector a charge sensitive preamplifier is required in addition to amplifier, discriminator, and binary circuits. At present, conventional laboratory instrumentation is being used to amplify and count the pulses from the detector. As soon as the present investigation has been completed and a detector has been selected semiconductor circuits will be constructed. Later, these will be miniaturized using the DOFL 2-D approach.[10]

Integrated semiconductor circuits being investigated for another project may be used for the nuclear timer. These include low-power circuits[11] that are reported to operate in the microwatt range. It is hoped that these circuits will prove practical and that a complete nuclear timer can be built to operate on less than 100 microwatts. A nuclear timer operating on this power level would be particularly attractive for long timing applications.

In any case it appears that it will be possible to package a nuclear timer in a very small space, probably less than two cubic inches.

7. Conclusions and Recommendations

1. The proposed timer appears feasible but additional work is required before its accuracy can be determined. Additional investigations also are required to eliminate temperature effects.

2. The estimated size of the radiation source and detector is 1/4" dia. and 1/4" long.

3. The eventual size of the complete timer should be below 2 cu. in.

4. Americium^{241} and Plutonium,^{239} both alpha emitters, are satisfactory radioactive sources for this application. Plutonium^{239} is preferable when long term storage stability is required.

5. A silicon p-n diode is the most promising of the detectors tested.

6. Transistor and n^+ip^+ type detectors should be studied.

7. The cause of the low voltage pulses produced in silicon p-n diodes by alpha particles from an Americium^{241} source should be investigated.

Acknowledgments

The authors are indebted to many DOFL Employees who have assisted in these investigations. The contributions of Douglas Holland, Cita Marie Wright and George Blagman have been particularly helpful.

This work is being funded by the Novel Research Program of the Army Research Office.

References

(1) DOFL Report R320-59-11 Safety Mechanism, Arming Programmer, and Time Fuze for the Copperhead System, Gerald W. Kinzelman, 20 November 1959

(2) DOFL Report R920-61-5 Digital Circuits Using "NOR" and Split Input Binary Stage Building Blocks, Amiel Goodman, 1 December 1961

(3) DOFL Report TM-60-5 The Use of Chemical Reactions in S & A Devices, Raymond H. Comyn, 11 February 1960

(4) DOFL Report R320-61-19 Nuclear Timing, Raymond H. Comyn and Douglas F. Holland, 9 August 1961

(5) IRE Transactions on Nuclear Science, Vol. NS-8, No. 1, 1 January 1961, Introduction to Semiconductor Particle Detectors, W. L. Brown

(6) IRE Transactions on Nuclear Science, Vol. NS-8, No. 1, 1 January 1961, Improvements in Encapsulated Silicon Junction Alpha Detectors, R. S. Jackson, P. P. Webb, and R.L. Williams

(7) Eighth Scintillation and Semiconductor Counter Symposium, March 1962, The State of the Art in Nuclear Particle Detectors, J. W. Mayer

(8) IRE Transactions on Nuclear Science Vol. NS-8, No. 1, 1 January 1961. Transistor Form of Nuclear Particle Detector, R. L. Williams and P. P. Webb

(9) Eighth Scintillation and Semiconductor Counter Symposium, March, 1962, Properties of an n^+ip^+ Semiconductor Detector, F. P. Ziemba, et al

(10) DOFL Report TR 951 Status of Microminiaturization, Eleanor F. Horsey and Philip J. Franklin, 1 June 1961

(11) CBS Laboratories Report, Stamford, Connecticut, Microwatt-Microelectronics, W. W. Gaertner, M. Schuller, C. Keizman, and C. Levy.

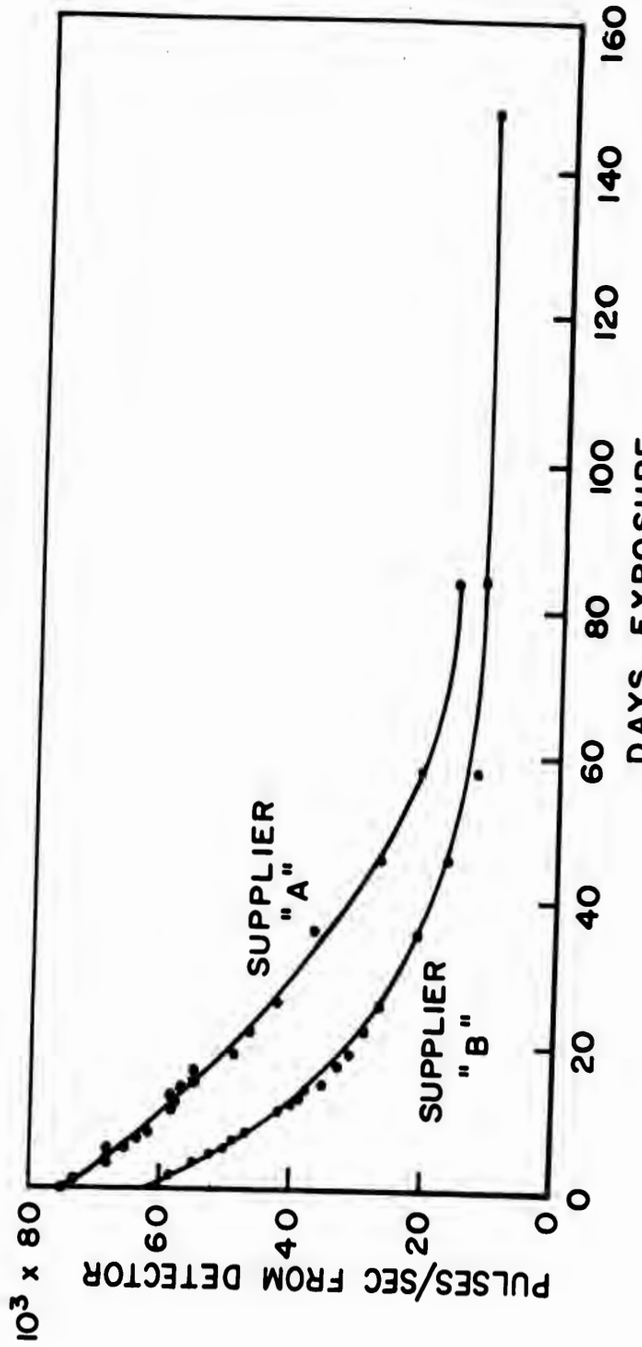


Figure 1

Effect of High Intensity Alpha Radiation on Performance of Silicon p-n Detectors

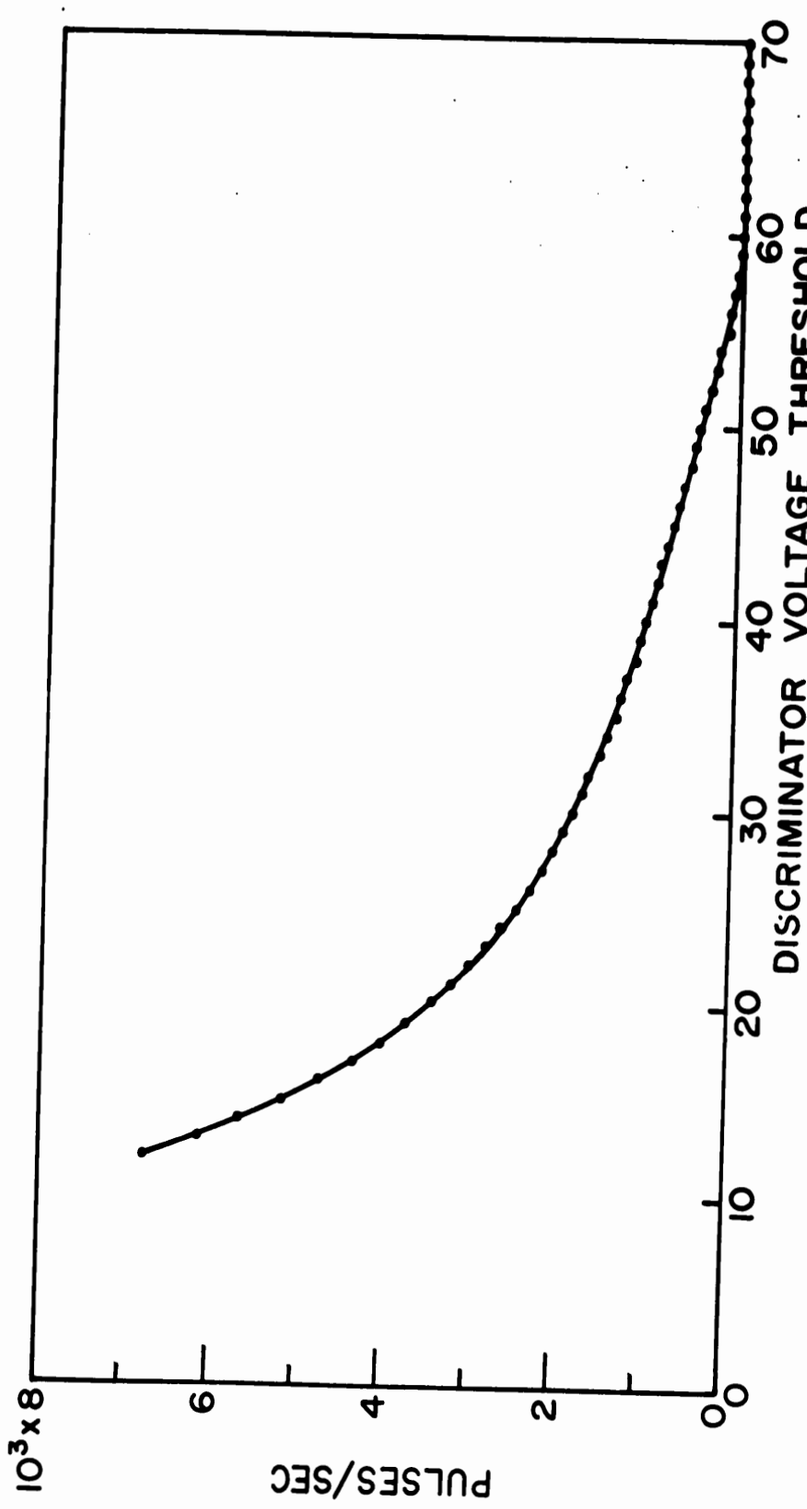


Figure 2
Distribution of Pulse Heights After Exposure of Silicon Detector to 7 x 10^{11} Alpha Particles

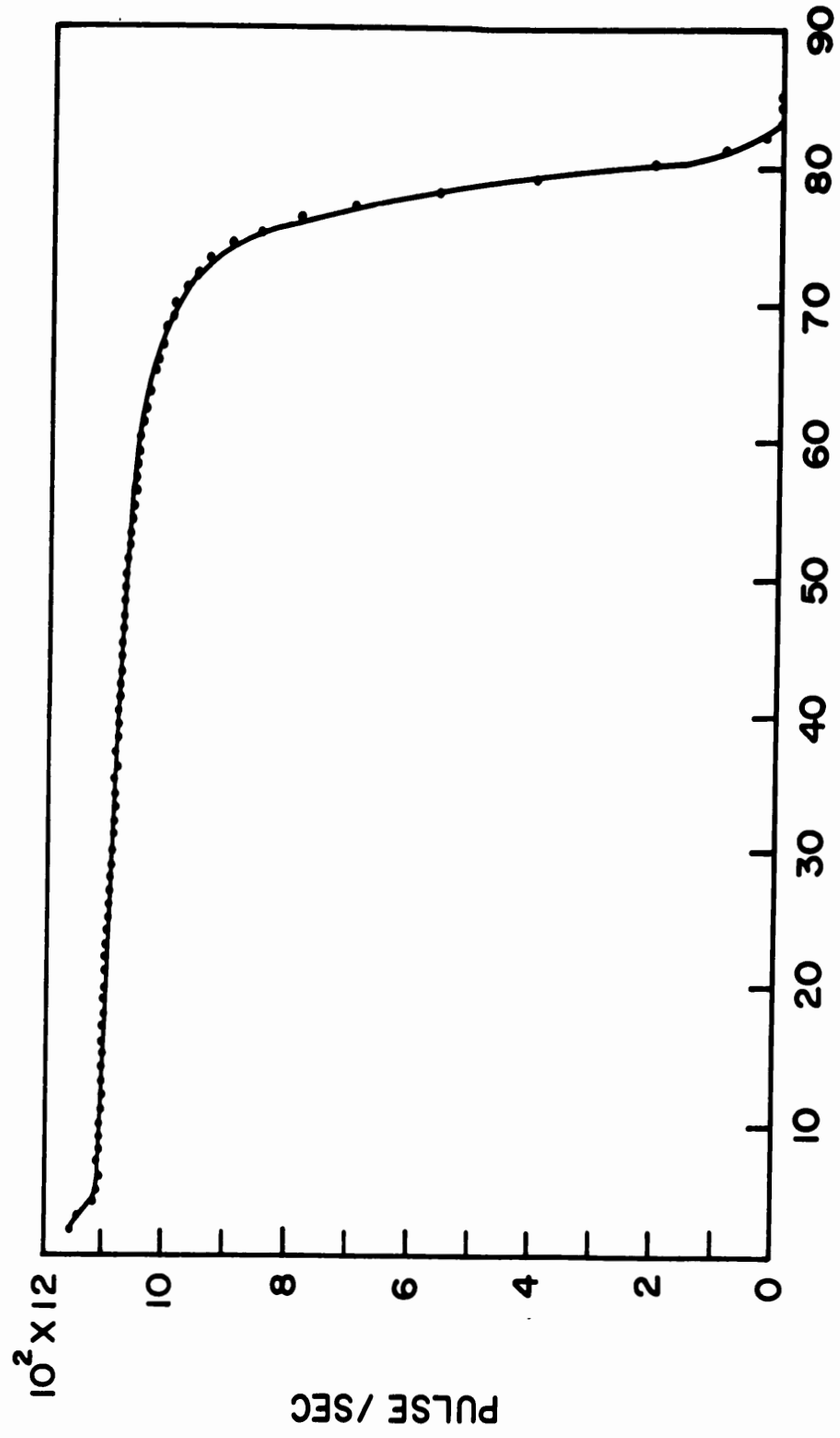


Figure 3
Distribution of Pulse Heights Before Prolonged Exposure of Silicon Detector to Alpha Radiation

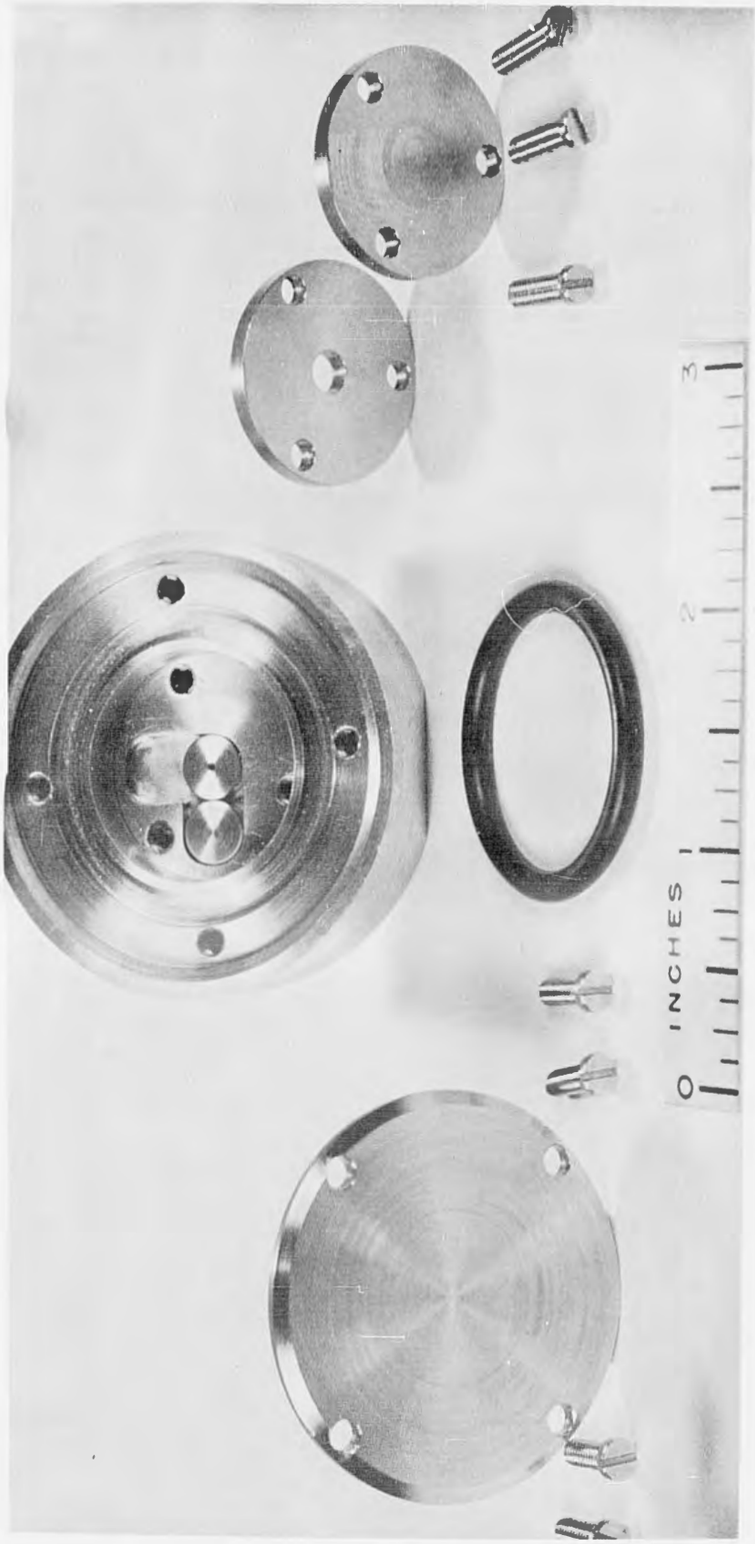


Figure 4. Test Fixture

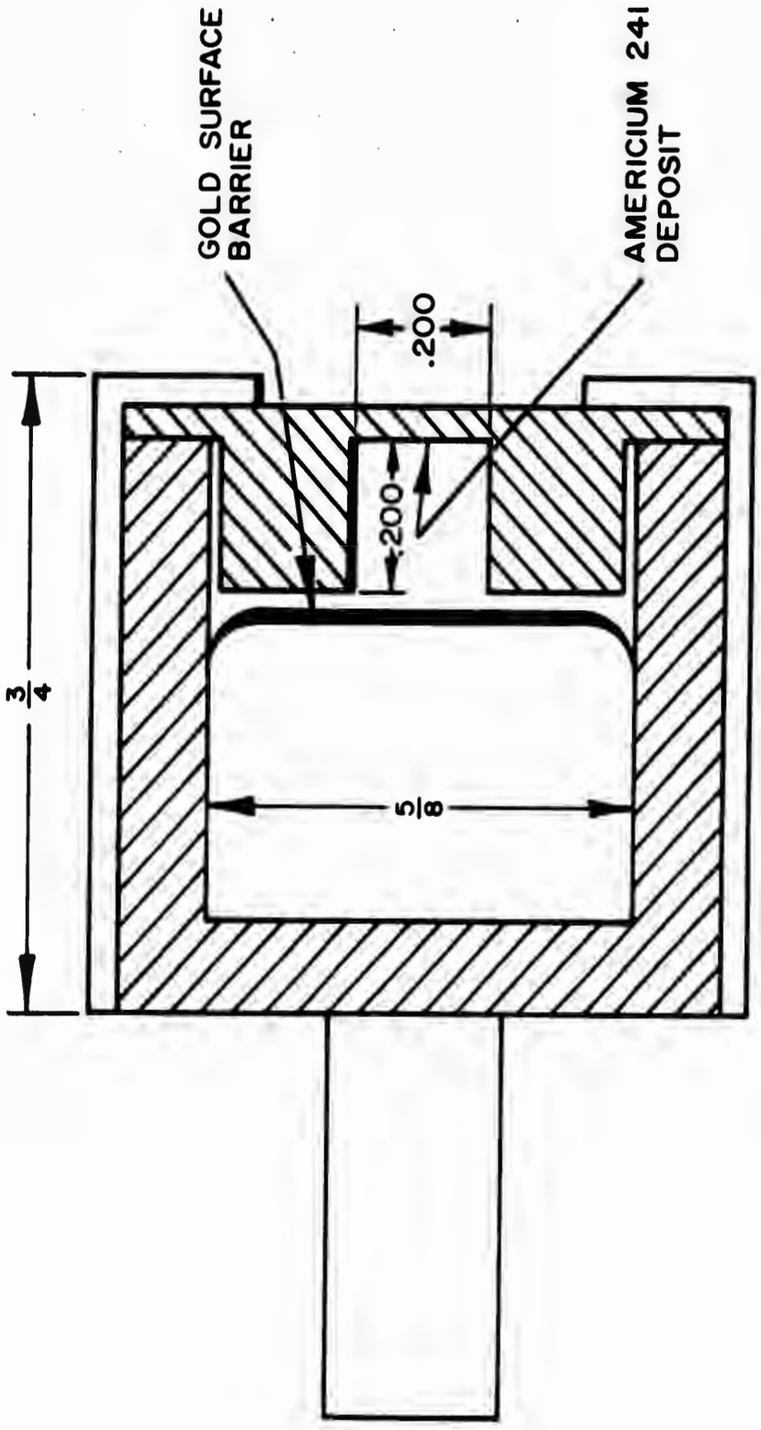


Figure 5

Source Holder for Collimating Alpha Particles

PERITONEAL DIALYSIS IN THE TREATMENT OF RENAL FAILURE: I. THE POSTOPERATIVE AND POST-TRAUMATIC GROUP OF PATIENTS

THOMAS E. DAVIS, CAPTAIN, MC
DEPARTMENT OF METABOLISM
WALTER REED ARMY INSTITUTE OF RESEARCH
WASHINGTON, D. C.

Acute renal failure is an important military problem. In World War II forty percent of one group of severely wounded patients developed acute renal failure with a fatality rate of ninety percent in the severely oliguric (1). Early in the Korean War, the mortality rate of severely wounded oliguric patients dropped from ninety percent to fifty-three percent after the establishment of the Renal Insufficiency Center and improved general management (2). This drop was due to hemodialysis (Artificial Kidney).

Hemodialysis has two limitations in the treatment of multiple casualties. The casualties must be evacuated to a Renal Failure Center and a trained team cannot perform more than one dialysis in a eight-hour period. A method of treatment which can be applied by the staff of any Field Hospital and permits simultaneous treatment of several patients is required. Peritoneal dialysis may meet this requirement.

At Walter Reed General Hospital since February 1960, peritoneal dialysis has been performed in thirty-eight patients. It is our purpose in this paper to present our results with particular reference to patients with post-traumatic and postoperative renal failure.

MATERIAL AND METHODS

I. Patients Dialyzed

Thirty-eight patients were dialyzed. Of this total group, fifteen had acute renal failure while twenty-three were of the

chronic variety. In the acute group, there were two sub-groups of special interest. There were five patients in one sub-group with post-traumatic renal insufficiency following heat stroke, third degree burns, shock due to bleeding and severe tissue trauma due to accidents. A second sub-group consisted of four patients who developed renal failure in the postoperative period. Two patients in the post-traumatic group underwent abdominal surgery also.

II. Methods of Dialysis

(a) The patient was in bed or on another appropriate flat surface with the bladder emptied.

(b) A surgical prep of the abdomen was performed.

(c) One thousand milliliters of dialysis fluid was run into the abdomen through a 15 gauge needle at the site where the peritoneal catheter was to be placed.

(d) Following local anesthesia, an incision was made in the abdominal wall in the midline one-third of the way down from the umbilicus to the symphysis pubis. An alternate area was either lower quadrant.

(e) A Duke #17 trocar was used to enter the peritoneal cavity and the catheter was passed through the trocar toward either gutter. When a free flow of fluid in and out of the catheter was assured, the skin was tightened around the catheter by a mattress suture.

(f) Fluid was run in as rapidly as possible and allowed to remain in the abdomen for one hour, and then drained. In Figure 1, a patient is shown being dialyzed.

(g) Commercially available dialysis solutions were used and were warmed to body temperature prior to their use. Dialysis fluid with an electrolyte composition approximately that of normal extracellular fluid was used to prevent electrolyte depletion and bring abnormal serum electrolytes toward normal. The composition of the fluid is shown in Table I. The osmolality of the solutions was dependent on the concentration of glucose. In patients with overhydration, hyperosmolar solutions were used. Tetracycline and heparin was added to one liter of each set of two liters. Potassium in concentrations equivalent to normal serum was added in normokalemic patients.

(h) The length of dialysis varied from patient to patient depending on the indication. Eight to seventeen hours were required in hyperkalemic patients while the length of dialysis varied from twenty-four to one hundred and forty-four hours in patients with azotemia and overhydration.

III. Studies Obtained

(a) Blood

Blood urea nitrogen, serum sodium, potassium, CO_2, chloride, white blood count, hemoglobin, and hematocrit were obtained at the beginning and end of dialysis and daily during the procedure. On a number of occasions these determinations were performed at six hourly intervals. Blood sugar, uric acid, creatinine, calcium, phosphorus, albumin and globulin were obtained as indicated.

(b) Peritoneal Fluid

(1) Chemistries

All peritoneal fluid was analyzed for urea nitrogen, potassium and glucose and, as indicated, for uric acid.

(2) Cultures

Appropriate cultures were obtained daily.

(c) Other Studies

Initial electrocardiograms were obtained and were performed serially in patients with hyperkalemic ECG changes. In other patients, they were performed as indicated. The patients were weighed daily.

INDICATIONS FOR DIALYSIS

Our indications for peritoneal dialysis and the number of dialyses done for each indication for the entire group are shown in Table II. The number of dialyses done for the post-traumatic and postoperative groups alone are shown in Table III. In the entire group, azotemia alone or in combination with hyperkalemia or overhydration accounted for eighty-three percent of the dialyses.

In most patients, dialysis was instituted before evidence of clinical uremia appeared. This was done anticipating that uremia would manifest itself before the onset of diuresis. Some patients had distinct clinical evidence of uremia on admission and were dialyzed promptly. At the present time, patients are dialyzed when the blood urea nitrogen reaches 150 mg/100 ml.

Hyperkalemia can cause sudden cardiac arrest and must be promptly relieved. A serum potassium above 7 mEq/L with or without ECG changes constituted an indication for dialysis. This was felt necessary because any complication such as a seizure might produce a significant increase in the serum potassium and cause cardiac arrest.

Overhydration, especially when pulmonary edema was present, was another indication for prompt dialysis.

RESULTS

A. Treatment of Hyperkalemia

Hyperkalemia is the only chemical abnormality likely to cause death in the first week of oliguric renal failure (3). This is especially true in patients who have suffered extensive tissue damage as the result of trauma. Excess serum potassium has a toxic effect on the heart which is recorded on the ECG before any signs or symptoms appear. Infusion of calcium, glucose and insulin are only temporary measures and ion exchange resins fail in patients with severe potassium intoxication. Maxwell performed thirty-seven dialyses in which hyperkalemia was present and after the procedure the serum potassium was normal in thirty-one. In the others the serum potassium was lowered 2 mEq/L but was still above normal. The gratifying feature was the rapidity of the ECG improvement. Usually the ECG showed definite changes toward normal during the first exchange (4). Our studies confirm Maxwell's findings. In the postoperative and post-traumatic group, five patients with hyperkalemia were dialyzed and are shown in Figure 2. The serum potassium was within normal limits at the end of dialysis in four of the five patients. In the fifth patient, the serum potassium at the end of dialysis was 5.9 mEq/L. In three of the five patients, the serum potassium was normal by twelve hours. In eight of ten other patients with serum potassium above 7 mEq/L, the serum potassium was normal or had fallen 2.1 mEq/L or more by the end of dialysis. In seven of the ten patients,

the serum potassium was normal or near normal by twelve hours.

In those patients followed with serial ECG's, there was definite improvement to near normal usually within three to five hours and a return to the pre-hyperkalemic ECG by five to ten hours. The response to dialysis in one patient who rapidly developed oliguria, hyperkalemia with ECG abnormalities, hyponatremia and azotemia with mental confusion after surgery is shown in Figure 3. As depicted, after institution of peritoneal dialysis there was a prompt fall in the serum potassium with return of the ECG to the pre-hyperkalemic tracing and improvement in the hyponatremia. The expected blood urea nitrogen rise was prevented and his sensorium cleared.

B. Treatment of Azotemia

The syndrome of clinical uremia develops after five to eight days of oliguria in post-traumatic renal insufficiency (3) Whether to dialyze prophylactically or after the symptoms of clinical uremia appear has been frequently discussed (4, 5). Most of our patients were dialyzed before the development of clinical uremia except for those who were uremic on admission. In evaluating our results we took into consideration the patient's clinical response, the fall in blood urea nitrogen, the amount of urea removed, and the peritoneal clearance of urea.

Our findings in five patients with post-traumatic renal failure are listed in Table IV. Four of the five patients were critically ill on admission and all but Patient 14 showed considerable improvement with dialysis. This improvement consisted of a significant fall or prevention of the expected rise in blood urea nitrogen in four of five patients. It also consisted of the relief of the uremic syndrome in two and the prevention of uremia in two other patients. Three of the five patients died of their underlying disease despite dialysis. Approximately one gram of urea was removed per hour in most of the patients and the peritoneal clearance of urea ranged from 13.9 to 20.7 ml/min.

A patient in this group who had sustained severe multiple injuries due to an auto accident is depicted in Figure 4. Because of leakage during the initial peritoneal dialysis, three hemodialyses were performed. Despite the leakage, which complicated nursing care and accurate determination of fluid balance, the blood urea nitrogen was successfully decreased. A final dialysis was

performed by the peritoneal route and resulted in a significant fall in the blood urea nitrogen. The patient diuresed and recovered.

Even though recent abdominal surgery is an often given contraindication for peritoneal dialysis, it was successfully carried out in six patients. The results are shown in Table V. The time in the postoperative period when dialysis was instituted varied from the second to the fifth day. In all patients there was significant improvement with dialysis, although this was temporary in Patient 18. Improvement consisted of the prevention of the expected blood urea nitrogen rise in four patients and the significant fall of blood urea nitrogen in three of four dialyses in two other patients. It also consisted of the prevention of uremic syndrome in four patients, the temporary relief in one and the permanent relief in another. The elevated serum potassium was significantly decreased in three of these patients who were discussed in an earlier section. Four of the six patients died in spite of successful dialyses. The causes of death varied from hepatic failure in one to overwhelming infection in another.

The peritoneal clearances of urea ranged from 6.2 to 20.7 ml/min. which is comparable to the clearances in other patients who have not undergone surgery (6). This suggests that peritoneal dialysis is as efficient in postoperative patients as others.

C. Treatment of Overhydration

Overhydration is frequently found in oliguric patients and usually results from over-exhuberant fluid therapy by physicians. Some fluid overload is present in most patients with renal failure and if it progresses to pulmonary edema, dialysis done promptly will be life saving. During peritoneal dialysis, negative fluid balance can be achieved rapidly by changing the osmolality of the dialysis solutions. This is achieved by using two liters of seven percent dextrose dialysis solution (677 mOsm/L) for rapid removal, and a mixture of one liter of one and one-half percent dextrose solution (380 mOsm/L) and one liter of seven percent dextrose solution for slower removal. The patient's exact fluid balance is known at all times because it is measured with each exchange. By this means, it is possible to remove as much fluid as desired at a safe rate.

None of the post-traumatic or postoperative group had

significant fluid overload but this can be a frequent problem. Our results in six patients with marked fluid, four of whom had pulmonary edema, are noted in Table VI. Three of the four patients with pulmonary edema were promptly relieved. In the fourth patient, pulmonary edema was not improved despite the net removal of five liters of excess fluid. At autopsy in this patient, severe intrapulmonary hemorrhage was found secondary to his underlying disease and explains the lack of response. In two other patients, anasarca was promptly relieved with dialysis.

COMPLICATIONS

In the thirty-eight patients dialyzed, infection resulting from dialysis was seen in three patients. Peritonitis occurred in two patients while septicemia and infection of the subcutaneous tissues of the abdominal wall occurred in another. No infection resulted from the dialysis in either the post-traumatic or post-operative group.

Leakage from the peritoneal cavity through surgical drains in the abdomen was a problem in two postoperative patients. This complicated nursing care and accurate measurement of fluid balance. In one patient, the leakage caused us to switch to hemodialysis but later peritoneal dialysis was successfully resumed.

DISCUSSION

Acute renal failure was recognized as an important military problem by the Allies and the Germans during World War I. During World War II and in the Korean War, eighty to ninety percent of severely wounded oliguric patients succumbed (1, 2, 3). It was apparent during the Korean War that some means of decreasing this mortality was necessary and, therefore, the Renal Insufficiency Center was established in Korea. With the use of hemodialysis, the mortality rate was decreased to fifty-three percent.

Unfortunately, a Hemodialysis Center has certain limitations due to the time delay of transportation to the center, to the necessity of an expertly trained team, and overtaxing of facilities when a large number of casualties require dialysis. Since hyperkalemia develops rapidly in this group of patients and may cause sudden cardiac arrest, any delay in dialysis may be fatal. Pulmonary edema is seen early in acute renal failure and also

demands prompt treatment. Clinical uremia takes longer to develop and is not as urgent a problem.

Peritoneal dialysis is particularly suited to the treatment of acute renal failure in the military situation. First of all, it may be carried out by the staff of any Field Hospital and allows for the simultaneous treatment of a number of patients. Peritoneal dialysis may be started within thirty minutes after the decision to dialyze is made while at least two to three hours are necessary to ready the patient and the artificial kidney for dialysis. When treating hyperkalemia, time may be critical.

As shown, there is a prompt decrease in the serum potassium and reversal of the ECG abnormalities with peritoneal dialysis. The improvement of pulmonary edema can be prompt with the use of hyperosmolar dialysis solutions. Also, as shown, clinical uremia can be effectively combated with peritoneal dialysis with a significant fall in blood urea nitrogen and improvement in the clinical symptoms.

A further advantage is that this procedure may be carried out in the immediate postoperative period without loss of efficiency and, if necessary, in the presence of peritonitis (4). Peritoneal dialysis also may be done in the face of bleeding secondary to renal failure which prohibits the use of the artificial kidney unless regional heparinization of the machine is used (4). Severe hypotension which may be a relative contraindication to the use of the artificial kidney is immaterial in consideration of peritoneal dialysis.

The complications of peritoneal dialysis are few and in this group of patients leakage from surgical drains complicated the procedure in two patients and it was necessary to resort to hemodialysis in one patient because of this complication. Infection, bleeding into the peritoneal cavity, wound dehiscence, dehydration and hypotension did not occur in our patients.

TABLE I

Composition of Peritoneal Dialysate Fluid*

	mEq/L	Gm/L	mOsm/L
Na	140.5		140.5
Cl	101.0		101.0
Ca	3.5		1.7
Mg	1.5		1.0
Lactate	44.5		44.5
Dextrose		15.0 or 70.0	83 or 388
Total	291.0		371.7 or 676.7

*To one liter of each set of two liters 25 mgs of Tetracycline, 5 mgs of Heparin and 8 mEq of Potassium Chloride was added. In patients with hyperkalemia, no potassium chloride was added.

TABLE II - INDICATIONS FOR PERITONEAL DIALYSIS

Number of Patients	38
Number of Dialyses	65
Azotemia	45
Hyperkalemia	5
Dialyzable Toxins	2
Overhydration	3
Azotemia and Hyperkalemia	7
Azotemia and Overhydration	2
Hyperkalemia and Overhydration	1

TABLE III. INDICATIONS FOR PERITONEAL DIALYSIS IN THE POST OPERATIVE AND POST TRAUMATIC PATIENTS

Number of Patients	9
Number of Dialyses	13
Azotemia	9
Hyperkalemia	3
Hyperkalemia and Azotemia	1

TABLE IV

Effect of Peritoneal Dialysis in Post-traumatic Renal Failure

Patient	Cause	Duration of Dialysis Hrs.	Blood Urea Nitrogen Before Mg%	Blood Urea Nitrogen After Mg%	Urea Removed Gm.	C Urea Ml/Min.	Remarks
22	Shock due to hemorrhage	44	150	49	40	17.2	Anuric for 4 days before admission. Survived.
30	Traumatic Shock	45 49	146 266	114 97	35 93	--- 20.7	Multiple fractures and ruptured bladder due to accident. Survived.
14	Heat Stroke	35 55	150 180	141 160	37 38	---- 14.2	In shock on admission; on pressor agents during hospitalization. Died
18	Traumatic Shock	20 70	160 226	66 202	--- ---	---- ----	Sucked up into jet engine. Ruptured spleen and duodenum. Peritonitis on admission. Died.
32	70% 3rd degree burns	14 23	98 100	97 94	14 18	16.1 13.9	Severe 3rd degree burns. Died
Average		39	164	113	36	16.4	

D'VIS

TABLE V
Peritoneal Dialysis in the Postoperative Patient

Patient	Diagnosis	Postop. Day	Length of Dialysis Hrs.	Blood Urea Nitrogen Before Mg%	Blood Urea Nitrogen After Mg%	C_{Urea} Ml/Min	Serum Potassium Before mEq/L	Serum Potassium After mEq/L	Remarks
7	Thrombotic Thrombocytic Purpura	2	7	75	83	14.0	8.9	5.6	Dialyzed for hyperkalemia & this was successful. Died.
19	Hepatorenal Syndrome	3	144	51	50	13.5	7.9	4.8	Portocaval shunt. Dialyzed for hyperkalemia. Died in hepatic failure.
33	Acute Tubular Necrosis	2	20	50	56	20.6	8.6	5.0	Dialyzed for hyperkalemia. Successful & pt survived
30	Acute Tubular Necrosis	3 19	45 69	146 266	114 97	---- 20.7	6.6 3.5	4.5 4.9	Auto accident with multiple fractures, ruptured bladder. Survived.
11	Cortical Necrosis	4	64	109	98	6.2	5.4	3.9	Post-aneurysmectomy, renal failure. Died.
18	Acute Tubular Necrosis	5 10	20 70	160 226	66 202	---- ----	5.1 6.5	5.0 ---	Ruptured spleen, duodenum and hematoma due to trauma. Temporary improvement.

TABLE VI

Peritoneal Dialysis in the Treatment of Overhydration

Patient	Diagnosis	Length of Dialysis Hrs	Fluid Removed Liters	Change in Clinical State
4	Nephrotic syndrome	14 24	8.0 8.6	Peripheral edema relieved
27	Nephrotic syndrome	73	8.2	Peripheral and pulmonary edema relieved
10	Chronic glomerulo-nephritis	48	7.2	Peripheral and pulmonary edema relieved
13	Chronic pyelo-nephritis	13 10	5.5 3.1	Pulmonary edema relieved
29	Goodpasture's syndrome	46	5.1	Peripheral edema relieved
6	Goodpasture's syndrome	34 30	2.0 5.0	Peripheral edema relieved Pulmonary edema not relieved

Bibliography

1. The Board for the Study of the Severely Wounded, The Physiologic Effects of Wound, p.122: Office of The Surgeon General, Department of the Army, 1952.

2. Smith, L. H., Post, R. S., Teschan, P. E., Abernathy, R. S., Davis, J. H., Gray, D. M., Howard, J. M., Johnson, K. E., Klopp, E., Mundy, R. L., and O'Meara, M. P.: Post-traumatic renal insufficiency in military casualties. II. Management, Use of an artificial kidney, prognosis. Amer. J. Med. 18: 187, 1955.

3. Meroney, W. H., Herndon, R. F.: The management of acute renal insufficiency. J.A.M.A. 155: 877, 1954.

4. Maxwell, M. H.: The indications for and the limitations of peritoneal dialysis. Bull. Acad. Med. of New Jersey 6: 341, 1960.

5. Teschan, P. E., Baxter, C. R., O'Brien, T. F., Freyhof, J. N., and Hall, W. H.: Prophylactic hemodialysis in the treatment of acute renal failure. Ann. Intern. Med. 53: 992, 1960.

6. Boen, S. T.: Kinetics of peritoneal dialysis. Medicine 40: 243, 1961.

Figure 1. Patient undergoing peritoneal dialysis

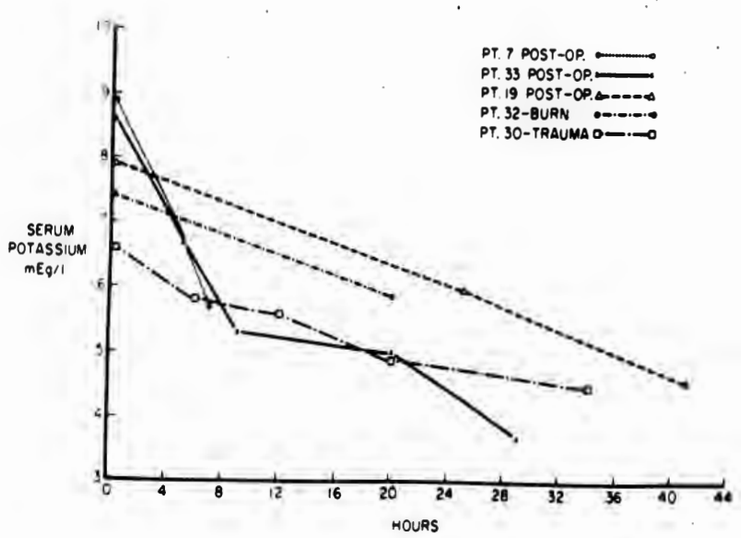


Figure 2. Response of the elevated serum potassium in postoperative and post-traumatic patients to peritoneal dialysis

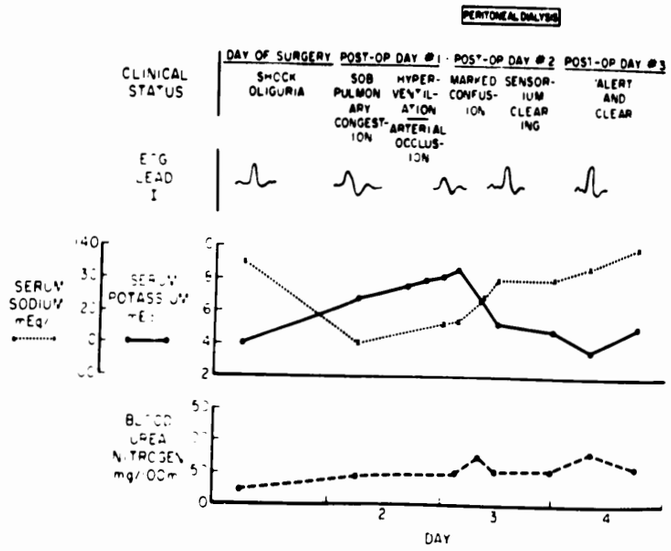


Figure 3. Response of a patient with renal failure, hyponatremia, hyperkalemia, and mental confusion to peritoneal dialysis

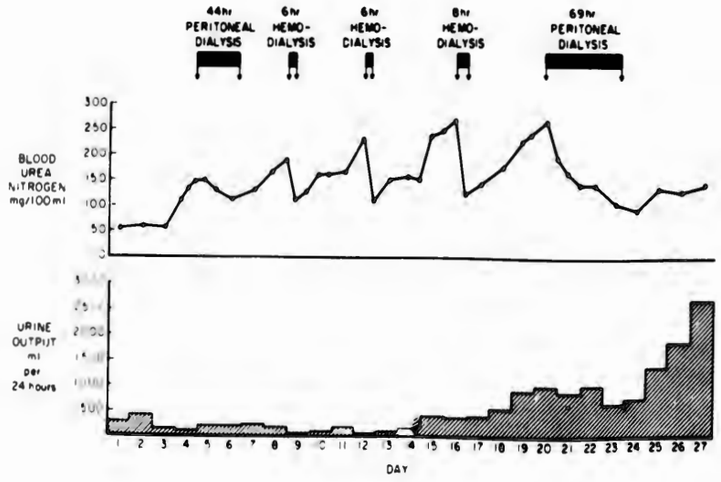


Figure 4. Effect of peritoneal dialysis and hemodialysis on a patient with post-traumatic renal failure

COLD INJURY AND COLD ACCLIMATIZATION IN MAN

T. R. A. DAVIS, M. D. and MAJOR R. J. T. JOY, MC, USA
U. S. ARMY RESEARCH INSTITUTE OF ENVIRONMENTAL MEDICINE
NATICK, MASSACHUSETTS

INTRODUCTION

The chronic exposure of laboratory animals to cold causes profound physiological and biochemical changes which have been accepted as defining a true adaptation or acclimatization to cold (58). After a few weeks of constant cold exposure, the rat has stopped shivering in the cold (19, 57); has replaced the heat initially contributed by shivering with quantitatively altered biochemical mechanisms for increased heat production (27, 49, 56); has increased his tissue responsiveness to calorigenic hormones such as nor-epinephrine (34, 40); will maintain body temperature and survive when exposed to extreme cold (15, 55); and has histological evidence of an adaptive process (30).

Since the demonstration in 1952 by Blair (10) that cold acclimatization of animals markedly reduces or eliminates the occurrence of frostbite when the animal is re-exposed to freezing temperatures, we have systematically explored the possibility of inducing cold acclimatization in man with the belief that this might provide a measure of physiologic protection against cold injury. Cold injury has been recorded as a major military disease since the march of Xenophon and the Ten Thousand (64). However, when cholera, typhus and plague were the major causes of disease in troops (66), little attention was paid to the climatic diseases and they were accepted as "the lot of the soldier" (38). As modern medicine reduces the impact of infectious agents, trenchfoot and frostbite are assuming more and more importance as the causes of large numbers of casualties.

In our last three major wars cold injury was an important source of casualties in American troops. In World War I there were 2,000 cases* of trenchfoot with a loss of 97,000 man-days (47 days per patient) (42) which, when we consider the short time and the small forces we had in combat, is a significant loss of manpower. In World War II the United States Army had 90,500 cases of cold injury (approximately 2/3 trenchfoot and 1/3 frostbite) - the European Theater alone had 71,000 cases of cold injury (62). From 1942 to 1945 the Army lost 7,500,000 man-days (83 days per patient), and at times, cold injury became an epidemic disease which crippled tactical operations (62). In Korea from 1950 to 1952 there were 6,300 cases of cold injury - the majority of which was frostbite, causing 283,500 lost man-days (45 days per patient) (47). The problem becomes even more meaningful when it is realized that 87% of all cases will be riflemen from the infantry rifle companies below the grade of E-5, (50) reflecting the close correlation of cold injury with outdoor exposure, combat action, immobility, and supply difficulties (54). In summary then, we are talking about a disease of major importance to combat commanders, which will remove troops from combat operations for a long period of time, which will return only 15% to full duty again (62), and which selectively strikes the thin fighting edge of the division slice.

Continued improvement in protective clothing has reduced the number of cold injury casualties, (48) but the disease is far from eliminated. There are some natural limits to the weight of protective clothing that can be worn - and the current 30 pound load may approach such limits. It is our contention that there must be an attempt to provide a "better" soldier, one so "conditioned" to the environmental extremes, that he is "immunized" to some extent to cold injury, or has greater endurance in the cold, or - at the very least - can operate in the cold with a lighter clothing burden. If it were possible to adapt or acclimatize soldiers so that they would be better able to live and fight in the cold, then the tactical commander would have his operational capability greatly extended. In effect, this would be another kind of protective clothing, built into the soldier, costing nothing, weighing nothing, not hampering his movement or comfort - in short - a "physiological armor."

It has long been common knowledge that men become "accustomed to" cold weather or to work in refrigeration plants. However, because of the variability of the exposure and the unknown effects of physiological habituation there has been considerable doubt that a true physiological-biochemical change occurs in men systematically exposed to the cold. Investigation of this phenomenon in man has yielded contradictory and equivocal results, (2, 3, 5, 6, 8, 11, 12, 14, 16, 25, 39, 46, 52). Essentially there

* Figures are rounded to the nearest hundred.

have been four different methods of approach: Some investigators have used cold chambers, (2, 16, 25, 32, 33, 59), some have exposed subjects to cold climates, (3, 6, 11, 14, 39, 41, 45, 46, 52), others have studied subjects during different seasons of the year, (8, 24, 26, 43), and still others have studied the indigenous peoples of cold regions (1, 12, 31, 44, 51, 63).

METHODS

One of the major problems in studying human cold acclimatization has been the lack of good indices of the acclimatized state. Skin or rectal temperatures, (2, 3, 14, 25, 32, 33, 59), basal and cold exposed metabolic rates, (16, 33, 41, 52, 59) and "tolerance" or "comfort" (25, 26, 43) have been used to define acclimatization in the great majority of investigations. The peripheral circulation, (12, 44), hormonal changes, (11), fluid balance, (3, 8, 24), hematologic changes, (6, 11, 46), the critical temperature, (51), changes in plasma proteins, (6), cardiovascular responses, (12), and miscellaneous changes in serum chemistry, (6, 46), have been used as indices in other studies. This array of methods of measurement demonstrates the difficulty of defining human cold acclimatization and has in large measure contributed to the conflicting scientific reports about cold acclimatization in man. The decline in or extinction of shivering with cold acclimatization in animals (15, 17, 18, 19, 28, 57) and the occasional non-quantified note of its loss in human cold exposure (5, 25, 31, 35, 52, 53) supported our belief that a loss of the shivering response is an important adaptive mechanism in cold acclimatization, and we have therefore used it as the index of physiologic change in the studies reported here.

All results reported in this paper were obtained during standard cold exposure tests. For any given set of experiments, the subjects were exposed to a fixed room temperature dressed only in shorts and reclining on a mesh cot. The standard exposures ranged from 30 minutes to two hours at ambient temperatures between 1.4°C. and 13°C. Prior to cold room exposure the subjects were fasted for 15 hours and basal control measurements made during a one hour period at an ambient temperature of 28°C. Shivering was measured by integrating muscle action potentials using either a Grass amplifier with integration by a Philbrick Model MK Operational Four-fold Absolute Electronic integrator or an Offner Electromyograph with a Model 1501 integrator. Plate electrodes were placed on the thighs and upper arms and the shivering data is reported as millivolts (mv) of electrical activity per minute. It is not possible to compare absolute shivering data in millivolts among all groups studied because the previous thermal experience of the subjects at the beginning of any given study could not be known. Some individuals had a very high initial shivering response and have been

presumed to have a low degree of cold acclimatization, while other men had low initial values and a smaller absolute magnitude of decrease. However, the percentage change from beginning to end of any given study is similar for either type of subject. Thus, when shivering responses are compared among different groups of subjects, the mv values are changed to percentages.

Oxygen consumption was measured by analyzing expired air with a Beckman E-2 Oxygen Analyzer and measuring respiratory minute volumes with a Phipps and Bird flowmeter. Basal metabolic rates were obtained during the one hour control period prior to cold exposure thus allowing calculation of the cold induced increase in oxygen consumption. All values are reduced to STP and reported as Calories/m^2/hr., as ml/min. oxygen consumed or as percent increase above basal. Rectal and skin temperatures were sensed by copper-constantan thermocouples in combination with a Brown recording potentiometer. Skin temperatures were obtained from the forehead, upper arm, upper thigh, calf, dorsum of the hand, dorsum of the foot, pad of the index finger, and pad of the great toe, except in the Greenland study, when only the latter four areas were measured. Rectal temperatures were secured at 12.5 cms insertion and all ambient air temperatures were measured at a height level with the recumbent subject. All the men used as test subjects were active duty soldiers who volunteered to perform as subjects in these experiments.

Statistical analysis was performed throughout on a paired data basis using standard methods. The Fisher t-test and a test of trend (4) were used with the absolute values to obtain significance. The null hypothesis was rejected at the 5% level of confidence. Plus and minus signs always refer to one standard deviation unless otherwise stated.

RESULTS

This paper will discuss, in order, the investigation of cold acclimatization in men followed throughout the year for seasonal differences, in men deliberately cold exposed in a cold chamber, in men living in Arctic and sub-Arctic climates, the effects of heat acclimatization upon pre-existing cold acclimatization, and the retention of cold acclimatization.

Seasonal Cold Acclimatization

Seasonal acclimatization to cold was studied in two groups of subjects. One group of 6 subjects was studied from October to February and another group of 5 subjects was studied from February to September. Figure 1 illustrates the change in heat production and shivering in the group studied during the Autumn-Winter period when exposed once monthly to a standard cold exposure of one hour to an

ambient chamber temperature 14.1 ± 0.5°C. This was their only deliberate exposure to cold. During the period of study, the subjects performed their ordinary military duties as laboratory assistants, exposed to cold only as is the ordinary individual living in Kentucky. Heat production decreased from an initial value of 59 percent above basal to 28 percent above basal ($p<0.05$). Shivering simultaneously declined 95 percent between October and January ($p<0.01$) with no shivering in five of the six subjects by December. Basal metabolic rates obtained at 28°C. did not change during the period of study.

Figure 2 illustrates the results obtained from the Spring-Summer group of subjects. The same measurements made in the Autumn-Winter group were repeated in response to a mean standard cold exposure of 13.7 ± 0.9°C. for one hour. Instrument failure precluded the study of oxygen consumption. As expected, shivering increased in all subjects during the summer months with a maximum occurring in September ($p<0.01$). Rectal and skin temperature responses during the exposure period were unchanged throughout the seasons in both groups ($p>0.10$).

These studies demonstrate, we believe, that there are seasonal changes in the physiologic responses to cold exposure in a temperate climate which are maximal in late winter but which are not retained, and are lost during the summer (21).

Chamber Cold Acclimatization

In this series of studies two groups were exposed nude except for shorts and slippers 8 hours daily in a Cold Chamber. They performed no exercise but spent the time quietly sitting and reclining, reading and watching television. Measurements of shivering, oxygen consumption and skin and rectal temperature were secured at intervals during the period of exposure on each subject, for one hour after an initial exposure to the cold that day of one hour. One group of 6 subjects was exposed in September (Summer group) and the other during the month of March (Winter group) to chamber temperatures of 13.5 ± 1.5°C. and 11.8 ± 0.5°C. respectively. Although cold-induced heat production in both groups remained significantly ($p<0.01$) elevated above basal, the Winter group showed no change ($p>0.10$) throughout the exposure period. On the other hand, the Summer group started with higher heat production values which decreased ($p<0.05$) after the 14th day and, at the end of the period of acclimatization, approximated the heat production values of the Winter group (Figure 3). Basal metabolic rates did not change in either group. Shivering activity was decreased by the 8th day of exposure in both groups ($p<0.01$), and remained depressed for the remaining time (Figure 3).

Rectal temperatures in both groups fell 0.7°C by the 14th day ($p<0.01$) (Figure 4). Although mean skin temperature in the Summer group showed a small fall ($p<0.05$), no such change was observed in the Winter group ($p>0.05$). The quantitative changes described are evident, but not the least striking manifestation of acclimatization was the observed change from marked subjective discomfort of the subjects early in the study, to rather complete tolerance of and indifference to the cold at the end of the study.

These physiologic adjustments to a prolonged period of nude cold exposure demonstrate in man findings previously elicited in rats chronically exposed to cold (53, 17, 15, 19, 20, 30, 54). It is our contention that this is quantitative evidence demonstrating that a physiologic cold acclimatization occurs in man and that this acclimatization can be artificially induced (22).

The Effects of Arctic Living

In order to test the effect of sub-arctic and arctic climates upon cold acclimatization two studies were performed (20, 36). In the first, seven subjects were tested at Fort Knox immediately before and after a 52 day sojourn in Anchorage, Alaska in January and February. Nineteen days of their stay were spent in field conditions two hundred miles north of Anchorage, living in tents and participating in a winter maneuver. Table I records the microclimate to which they were exposed. The cold exposure test utilized for this group consisted of a 60 minute exposure to 12.8°C. ± 0.5°C. with the previously described variables being measured. The effect of their Alaskan sojourn upon their shivering rate, oxygen consumption and heart rate are plotted in Figure 5. The decrease in shivering of the group was statistically significant ($p<0.01$). Basal and cold-induced heat production, rectal and skin temperatures and heart rate were unaffected.

Because the Alaskan subjects were not exposed to a microclimate significantly different from that found in temperate zone living (Table I), another study was done in true Arctic weather, with the additional variable of a deliberate reduction in protective clothing. Ten other subjects were taken to Greenland during the months of October and November and there exposed for 7 hours per day while wearing summer type underwear, O.G. trousers and shirts, thermal boots, gloves, and the shells to their parkas. Exercise was limited to moderate walking. Their microclimate is recorded in Table I, and it can be seen that the effective temperature was much lower than in Alaska, due to the windchill effect. They were tested before and after their stay in Greenland, for 30 minutes at 1.4°C. ± 0.5°C. The results of this type of exposure for a period of 43 days are demonstrated in Figure 6. An effective cold acclimatization was produced, with a significant difference in shivering

between their pre-exposure and post-exposure test ($p<0.01$). Again, the cold-induced oxygen consumption, basal metabolic rate, skin and rectal temperatures and heart rate were not significantly changed.

The Effect of Heat Acclimatization Upon Cold Acclimatization

Because of the known usefulness of heat acclimatization in producing an improved physiological response to hot climates (7, 9, 23), we next investigated whether or not a deliberate acclimatization to heat would interfere in any way with a pre-existing cold acclimatization.

The results in Figure 7 were obtained from six subjects cold acclimatized in a chamber at 13.5°C. ± 1.5°C. for 28 days during the month of October and six subjects seasonally acclimatized between October and February. Cold acclimatization and weekly testing for one hour at 13.5°C. were performed as before. Seven days after the last cold exposure, the subjects were exposed to heat in a climatic room at 41.0°C. (chamber group) and 49.0°C. (seasonal group) eight hours daily for twenty-one days. During the period of heat acclimatization, the subjects exercised on a treadmill at 3.5 mph 20 minutes in each hour for a total daily hike of 9.3 miles. Measurements of pulse rate, blood pressure and rectal temperature made before and after the exercise periods indicated that the criteria for acclimatization to heat had been satisfied (7).

Following heat acclimatization, the shivering response was again measured during the standard cold exposure previously employed for these groups. There was no difference between pre-heat and post-heat shivering ($p>0.05$) or heat production ($p>0.10$) or rectal temperature ($p>0.05$) or surface temperature ($p>0.05$).

These results would appear to demonstrate for pre-existing cold acclimatization, as for pre-existing heat acclimatization (26, 58), that both are independent processes of physiological adjustment to environment, do not participate in the same processes of adaptation, and may coexist simultaneously.

The Retention of Cold Acclimatization

The differences between seasonal and experimental (chamber) cold acclimatization and the failure of heat acclimatization to reverse the changes effected by cold acclimatization, excited speculation concerning the long-term retention of experimental or cold-chamber induced acclimatization. Accordingly, five subjects who had been chamber acclimatized to 13.5 C. eight hours daily for 28 days were followed for 11 months by giving the standard exposure test once a month. A group of controls not previously chamber acclimatized were also tested once a month to the same standard conditions of

cold exposure. Figure 8 illustrates the results and it can be seen that there is a parallel response in both groups to the standard exposure until June. In July, August and September the chamber group continues to exhibit a low and unchanged level of shivering ($p > 0.10$), while the control group increased their rate of shivering ($p < 0.01$). These results suggest to us that there are differences between experimental and seasonal cold acclimatization.

DISCUSSION

Whether measured by survival (10, 55, 56), changes in endocrine function (13, 29, 34), heat production (15, 56), shivering (15, 18, 19, 57), oxygen uptake of tissue (19, 61), or changes in specific enzyme (27, 37, 49, 58, 65), a true acclimatization to the cold has been well established for laboratory animals.

The presence or absence of cold acclimatization in man has not been as easily demonstrated or universally accepted and many observers have not thought that it can occur. As a consequence of the studies reported here, some of the reasons for the differences of opinion can be seen; the presence of clothing inhibits acclimatization and an inadequate period of exposure will not permit adequate establishment of acclimatization. The reliability and reproducibility of criteria has also provided conflicting results, for example, body temperature appears to decrease significantly only if the acclimatizing stimulus is as sufficiently great as in prolonged nude cold exposure. In the studies using natural and experimental exposure, no firm evidence was obtained which would suggest that a peripheral temperature change would occur consistently as a result of cold exposure, results which both agree (2) and are at variance (14) with the experience of others. In the experiments reported here, a significant fall in oxygen consumption occurred only in those subjects who could be considered by the violence of their shivering response at the beginning of exposure to be minimally cold acclimatized, as would occur in late summer. Subjects who were measured in the summer had a consistently higher initial cold-induced heat production than those measured in the winter (Figures 1 and 3). After acclimatization, heat production in both groups stabilized at 30 to 40% above basal.

Figure 9 compares the percent change in shivering of several of the groups described. The slopes were derived by plotting the initial shivering value (before cold exposure) as 100 percent, and then plotting the post-cold shivering value as a percent of initial value. The Alaskan and seasonal values, obtained under the same test temperature conditions, had the same rate of change indicating that the cold experience for the Alaskan group, adequately dressed for the environment, was not essentially different from that at Fort Knox at the same time of the year. There is a marked difference when the

Alaskan and seasonal values are compared to the winter Chamber Study (which had essentially similar test conditions). This is not an unexpected finding, since nude living in the chamber must have provided a greater degree of cold exposure and a prolonged period of low skin temperatures.

Comparison of the rate of change of the shivering response of these studies with the Greenland data must be approached with caution, since the test times and exposure temperatures were different. However, since shivering is less with higher test temperatures, this suggests that the rate of change would probably be as great as the chamber group had the Greenland subjects been tested at 13.0°C. From this data we have drawn the conclusion that the standard seasonal response to a temperate or sub-arctic climate, when one is adequately clothed, will not produce a persistent or markedly effective cold acclimatization. Since the use of cold chambers is not a practical way to acclimatize large numbers of troops, the Greenland experiment suggests a fruitful compromise; exposure to cold weather while lightly clad will produce a cold acclimatization akin to chamber acclimatization in both magnitude and direction.

The results obtained in the seasonal group studied from October to February and in the Chamber group exposed at the end of summer show a substantial economy of energy expenditure, amounting to about 50% of the cold-induced heat production. It is likely that shivering increases convectional heat loss in nude man, therefore, the reduction in shivering resulting from acclimatization is accompanied by a reduction in such heat loss. Moreover, Swift (60) has pointed out that shivering is accompanied by hyperventilation, which may also increase heat loss from the respiratory tract. Both of these possibilities imply that shivering is an inefficient and wasteful method of heat production in the cold. Although these studies do not demonstrate consistent changes in skin temperature, there is a significant lowering of rectal temperature if the cold exposure is sufficiently great. Thus, some attempt has been made to decrease heat loss by decreasing the core-skin temperature gradient.

In the chamber group acclimatized at the end of winter, and in the winter seasonal, Alaskan, and Greenland subjects, the lack of any significant change in oxygen consumption in the cold at a time when the subject is not producing heat by shivering may represent in large measure a non-shivering heat production which has previously been demonstrated in acclimatized animals (17, 19). This non-shivering thermogenesis may well be due (experiments to demonstrate this are now in progress) to a mechanism similar to that shown for the rat (40) namely, that there is an increased tissue responsiveness to the normal levels of endogenous calorogenic hormones, such as norepinephrine, which produces a greater quantitative and qualitative rate of biochemical heat production.

These experiments suggest that a seasonal cold acclimatization occurs in man, but that it is not retained over the summer. On the other hand, chronic nude exposure in a cold chamber is well retained at least through one summer season. It appears that the ability to retain cold acclimatization is highly dependent upon the rate and degree of cold exposure. This would suggest that retention is a function of the degree of acclimatization achieved. The inability of heat acclimatization to affect either artificially or naturally induced cold acclimatization indicates that both acclimatizations can co-exist in an individual. The loss of the seasonally acquired cold acclimatization over the summer period is considered by us to be due not only to a lesser degree of cold exposure in clothed men, but also to the absence of an adequate acclimatizing stimulus. Therefore, we believe that heat and cold acclimatization are not mutually exclusive; the loss of one during change of seasons is due to the absence of a continuing adequate stimulus and not to the presence of the opposite acclimatization.

CONCLUSIONS

These studies demonstrate that man has a physiological adaptation to the cold; that a decline in quantitative cold-induced shivering is a consistent and valid index of its presence or absence; that experimental cold acclimatization in a chamber can be induced at a faster rate and to a greater degree than that occurring seasonally; that acclimatization in appropriately clothed individuals is similar in rate and degree regardless of the ambient climate, suggesting that the exposure of the skin to the cold is the inducing mechanism; that heat acclimatization does not affect either artificial or natural cold acclimatization; and that artificial cold acclimatization is retained through the summer months while seasonal acclimatization is not. We believe cold acclimatization to be a phenomenon requiring controlled conditions for its definitive and measurable occurrence. More so than heat acclimatization, it is an elusive physiological property of man not capable of easy laboratory measurement or conceptual definition.

Future human experimentation in this field must carefully consider the effects produced by previous thermal experience, by clothing, by exposure time, by seasonal change, and by the variability of skin and rectal temperatures and oxygen consumption as indices of acclimatization.

Future research from this laboratory will explore the mechanism of non-shivering thermogenesis in man, the optimal methods of inducing cold acclimatization in troops, and the ability of cold acclimatization to increase the cold tolerance of soldiers and protect them from cold injury.

REFERENCES

1. Adams, T. and Covino, B. G. Racial variations to a standardized cold stress. J. Appl. Physiol., 12, 9, 1958.

2. Adams, T. and Heberling, E. J. Human physiological responses to a standardized cold stress as modified by physical fitness. J. Appl. Physiol., 13, 226, 1958.

3. Adolph, E. F. and Molnar, G. W. Exchanges of heat and tolerances to cold in men exposed to outdoor weather. Amer. J. Physiol., 146, 507, 1946.

4. Alexander, H. W. A general test for trend. Psych. Bull., 43, 533, 1946.

5. Balke, B., Cremer, H. D., Kramer, K. and Reichel, H. Untersuchung Zur Kalteanpassung. Klin. Wochenschr., 23, 204, 1944.

6. Barlow, J. S. Measurements of changes in hematocrit, plasma specific gravity, and plasma sodium in troops taking part in Arctic winter exercises. Defense Research Board Canada, DRNL Tech. Paper #13, 1954.

7. Bass, D. E., Kleeman, C. R., Quinn, M., Henschel, A. and Hegnauer, A. H. Mechanisms of acclimatization to heat in man. Medicine, 34, 323, 1955.

8. Bazett, H. C., Sunderman, F. W., Doupe, J. and Scott, J. C. Climatic effects on the volume and composition of the blood in man. Amer. J. Physiol., 129, 69, 1940

9. Bean, W. B. and Eichna, L. W. Performance in relation to environmental temperature. Reaction of normal young men to simulated desert environment. Federation Proc., 2, 144, 1943.

10. Blair, J. R. Trans. First Conf. on Cold Injury, Josiah Macy Jr. Foundation, New York, 1952, pp 207-233.

11. Bly, C. G., Johnson, R. E., Kark, R. M., Consolazio, C. F., Swain, H. L., Lawdani, A., Maloney, M. A., Figueroa, W. G. and Imperiale, L. E. Survival in the cold. U. S. Armed Forces Med. J., 1, 615, 1950.

12. Brown, G. M., Bird, G. S., Boag, T. J., Boag, L. M., Delahaye, J. D., Green, J. E., Hatcher, J. D. and Page, J. The circulation in cold acclimatization. Circulation, 9, 813, 1954.

13. Brown-Grant, K. Changes in the thyroid activity of rats exposed to cold. J. Physiol., 131, 52, 1956.

14. Carlson, L. D., Burns, H. L., Holmes, T. H. and Webb, P. P. Adaptive changes during exposure to cold. J. Appl. Physiol., 5, 672, 1953.

15. Cottle, W. H. and Carlson, L. D. Regulation of heat production in cold-adapted rats. Proc. Soc. Exper. Biol. and Med., 92, 845, 1956.

16. Daniels, F. Jr., Fainer, D. C., Bonmarito, C. L. and Bass, D. E. Acclimatization to cold in man. Federation Proc. 10, 32, 1951.

17. Davis, T. R. A. and Mayer, J. Demonstration and quantitative determination of the contribution of physical and chemical thermogenesis on acute exposure to cold. Amer. J. Physiol., 181, 675, 1955.

18. Davis, T. R. A. and Mayer, J. The nature of the physiological stimulus for shivering. Amer. J. Physiol., 181, 669, 1955.

19. Davis, T. R. A., Johnston, D. R., Bell, F. C. and Cremer, B. J. The regulation of shivering and non-shivering heat production during acclimation of rats. Amer. J. Physiol., 198, 471, 1960.

20. Davis, T. R. A. and Joy, R. J. T. Natural cold acclimatization in man in the sub-arctic. Physiologist, 3, 47, 1960.

21. Davis, T. R. A. and Johnston, D. R. Seasonal acclimatization to cold in man. J. Appl. Physiol., 16, 231, 1961.

22. Davis, T. R. A. Chamber cold acclimatization in man. J. Appl. Physiol., 16, 1011, 1961.

23. Eichna, L. W., Bean, W. B., Ashe, W. F. and Nelson, N. Performance in relation to environmental temperatures. Bull. Johns Hopkins Hosp., 76, 25, 1945.

24. Forbes, W. H., Dill, D. B. and Hall, F. G. The effect of climate upon the volumes of blood and of tissue fluid in man. Amer. J. Physiol., 130, 739, 1940.

25. Glaser, E. M. Acclimatization to heat and cold. J. Physiol., 110, 330, 1949.

26. Glickman, N., Inouye, T., Keeton, R. W. and Fahnstock, M. K. Comparison of physiological adjustments of human beings during summer and winter. Heating, Piping, 20, 113, 1948.

27. Hannon, J. P. and Vaughan, D. A. Effect of exposure duration on selected enzyme indexes of cold acclimatization. Amer. J. Physiol., 200, 94, 1961.

28. Hart, J. S., Heroux, O. and Depocas, F. Cold acclimatization and the electromyogram of unanesthetized rats. J. Appl. Physiol., 9, 404, 1956.

29. Hart, J. S. Metabolic alterations during chronic exposure to cold. Federation Proc., 17, 1045, 1958.

30. Heroux, O. Histological evidence for cellular adaptation to non-freezing cold injury. Canad. J. Biochem. and Physiol., 37, 811, 1959.

31. Hicks, C. S. and Matters, R. F. The standard metabolism of Australian aborigines. Australian J. Exper. Biol. and Med. Sci., 11, 177, 1933.

32. Horvath, S. M., Golden, H. and Wager, J. Some observations on men sitting quietly in extreme cold. J. Clin. Invest., 25, 709, 1946.

33. Horvath, S. M., Freedman, A. and Golden, H. Acclimatization to extreme cold. Amer. J. Physiol., 150, 99, 1947.

34. Hsieh, A. C. L. and Carlson, L. D. Role of adrenaline and noradrenaline in chemical regulation of heat production. Amer. J. Physiol., 190, 243, 1957.

35. Irving, L. Human adaptation to cold. Nature, 185, 572.

36. Joy, R. J. T., Poe, R. H., Berman, F. R. and Davis, T. R. A. Some physiological responses to Arctic living. Arch. Environ. Health, 4, 22, 1962.

37. Kline, D., McPherson, C., Pritchard, E. T. and Rossiter, R. J. Effect of exposure to a cold environment on labeling of phospholipid in rat liver slices. Proc. Soc. Exper. Biol. and Med., 92, 756, 1956.

38. Larrey, D. J. Memoirs of Military Surgery and Campaigns of the French Armies. (Trans. by R. H. Hall). Baltimore, Joseph Cushing, 1814.

39. LeBlanc, J. Evidence and meaning of acclimatization to cold in man. J. Appl. Physiol., 9, 395, 1956.

40. Le duc, J. Catacholamine production and release in exposure and acclimation to cold. Acta Physiol. Scandinav., 53, 1-101, Suppl. 183, 1961.

41. Lewis, H. E., Masterson, J. P., and Rosenbaum, S. Stability of basal metabolic rate on a polar expedition. J. Appl. Physiol., 16, 397, 1961.

42. Love, A. G. Medical Department of the United States Army in the World War. Vol. 15, Statistics, Wash., D. C., U. S., Government Printing Office, 1925.

43. MacLean, D. L. and Partridge, R. C. Seasonal change in the body's response to ventilative conditions. J. Indust. Hygiene, 22, 244, 1940.

44. Meehan, J. P. Racial and individual differences in the peripheral vascular response to a cold stimulus. Amer. J. Physiol., 179, 657, 1954.

45. Milan, F. A., Elsner, R. W. and Rodahl, K. Thermal and metabolic responses of men in the Antarctic to a standard cold stress. J. Appl. Physiol., 16, 401, 1961.

46. Miller, A. J. Study of certain blood properties in healthy men transferred from a subtropic to a subarctic environment. J. Aviat. Med., 20, 201, 1949.

47. Orr, K. D. and Fainer, D. C. Cold injuries in Korea during Winter of 1950-51. Medicine, 31, 177, 1952.

48. Orr, K. D. and Associates. Summary of Activities, Cold Injury Research Team, Korea, 1951-52. USAMRL Report #113, April 1953, Fort Knox, Kentucky, 1058 pp.

49. Patkin, J. K. and Masoro, E. J. Effects of cold acclimation on lipid metabolism in adipose tissue. Amer. J. Physiol., 200, 847, 1961.

50. Trenchfoot Survey, No. C-114, I and E Division, Army Service Forces, 2 June 1945.

51. Scholander, P. F., Anderson, K. L., Krog, J., Lorentzen, F. V. and Steen, J. Critical temperature in Lapps. J. Applied Physiol., 10, 231, 1957.

52. Scholander, P. F., Hammel, H. T., Anderson, K. L. and Lyning, Y. Metabolic acclimation to cold in man. J. Appl. Physiol., 12, 1, 1958.

53. Scholander, P. F., Hammel, H. T., Hart, J. S., Le Messurier, D. H. and Steen, J. Cold adaptation in Australian aborigines. J. Appl. Physiol., 13, 211, 1958.

54. Schuman, L. M. Epidemiology of cold injury in man. Trans. Second. Conf. on Cold Injury, Josiah Macy Jr. Foundation, New York, 1954, pp 11-85.

55. Sellers, E. A., Reichman, S., Thomas, N. Acclimation to cold: natural and artificial. Amer. J. Physiol., 167, 644, 1951.

56. Sellers, E. A., Reichman, S., Thomas, N. and You, S.S. Acclimatization to cold in rats: metabolic rates. Amer. J. Physiol., 167, 651, 1951.

57. Sellers, E. A., Scott, J. W. and Thomas, N. Electrical activity of skeletal muscle of normal and acclimatized rats on exposure to cold. Amer. J. Physiol., 177, 372, 1954.

58. Smith, R. E. and Hoijer, D. J. Metabolism and cellular function in cold acclimatization. Physiol. Rev., 42, 60, 1962.

59. Stein, H. J., Eliot, J. W. and Bader, R. A. Physiological reactions to cold and their effects on the retention of acclimatization to heat. J. Appl. Physiol., 1, 575, 1949.

60. Swift, R. W. The effects of low environmental temperature upon metabolism. J. Nutrition, 5, 213, 1932.

61. Weiss, K. A. Adaptation of rats to cold air and effects on tissue oxygen consumption. Amer. J. Physiol., 177, 201, 1954.

62. Whayne, T. F. and De Bakey, M. E. Cold Injury, Ground Type. U. S. Govt. Printing Office, Wash. D.C., 1958, 570 pp.

63. Wyndham, C. N. and Morrison, J. F. Adjustment to cold of bushmen in the Kalahari desert. J. Appl. Physiol., 13, 219, 1958.

64. Xenophon. Anabasis. (Trans. by W. H. D. Rouse), New American Library, New York, 1959, 191 pp.

65. You, S. S. and Sellers, E. A. Increased oxygen consumption and succin-oxidase activity of liver tissue after exposure of rats to cold. Endocrinology, 49, 374, 1951.

66. Zinsser, H. Rats, Lice and History. Atlantic Monthly, Boston, 1935, 301 pp.

TABLE I

Mean daily environmental conditions and cold exposure in Alaska and Greenland. "Indoors" is barracks or tent temperature. Climatic conditions were recorded in the immediate vicinity of the subjects. "Cold exposure" is time spent outdoors, away from shelter.

	Alaska				Greenland
	Camp	Field	Camp		Camp
Time - days	25	19	8		43
Indoor Temperatures-°F	74	62	72		75
Outdoor Temperatures-°F	16	5	25		12
Wind velocity - mph	1	0.5	2		10
Outdoor Temperature adjusted for wind-°F*	-11	5	-11		-40
Cold Exposure - hrs/man/24 hrs	4	7	3		7

*corrected for cooling isochrones according to the Siple formula (Amer. Philosoph. Soc., 89: 200, 1945).

TABLE 1

Mean Daily Environmental Conditions and Cold Exposure in Alaska and Greenland.

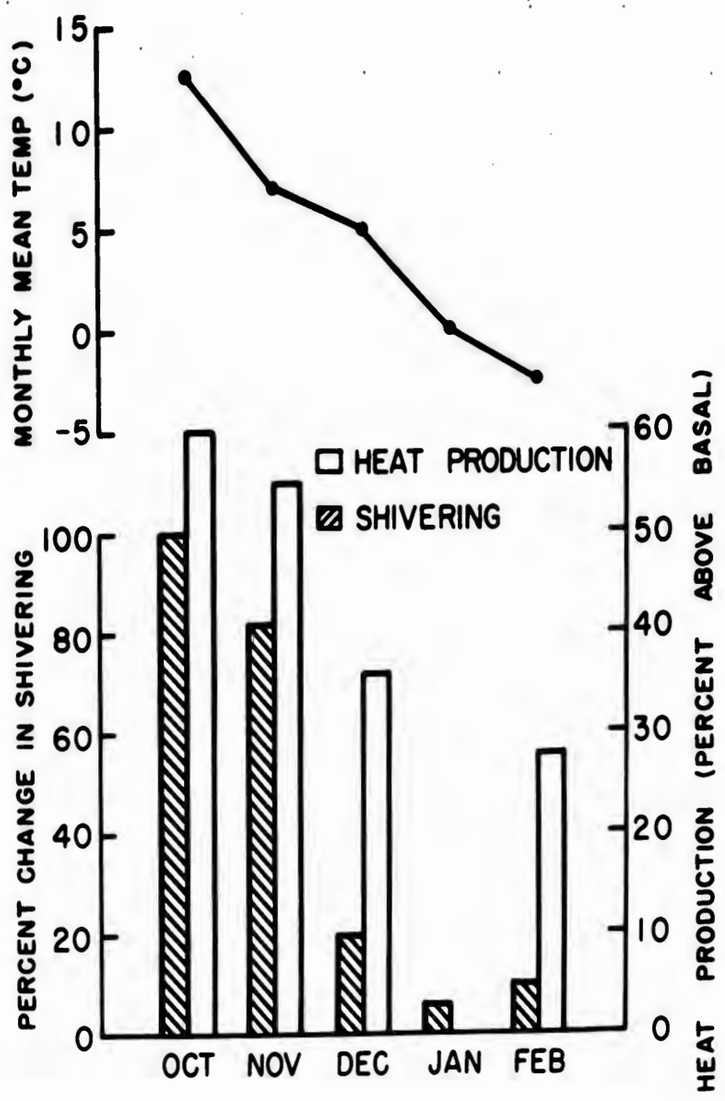


FIGURE 1

Shivering and heat production response to a standard cold exposure (1 hr.) of 6 subjects as a result of seasonal climatic change. Cold exposures were made once monthly. Top curve is the mean monthly outdoor temperature at Fort Knox, Kentucky. Shaded bars represent shivering expressed as a percentage of the October value, and open bars represent heat production expressed as percent elevation above basal.

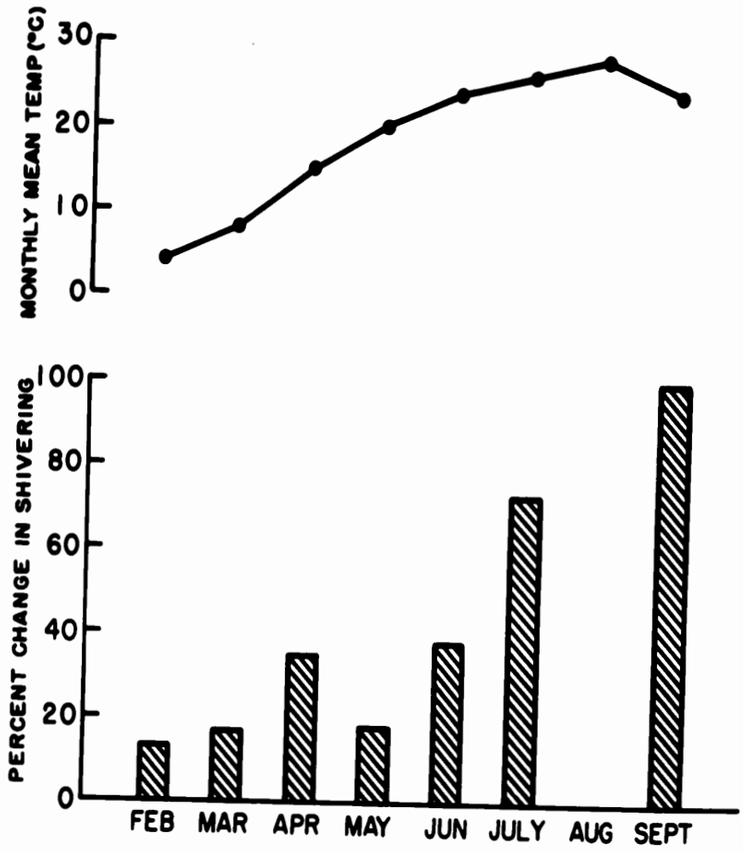


FIGURE 2

Response of 5 subjects to seasonal cold acclimatization beginning in February. Cold exposures were made monthly at 13.7 ± 0.9°C. Top curve is the mean monthly outdoor temperature at Fort Knox. Shaded bars represent shivering, expressed as a percentage of the September value.

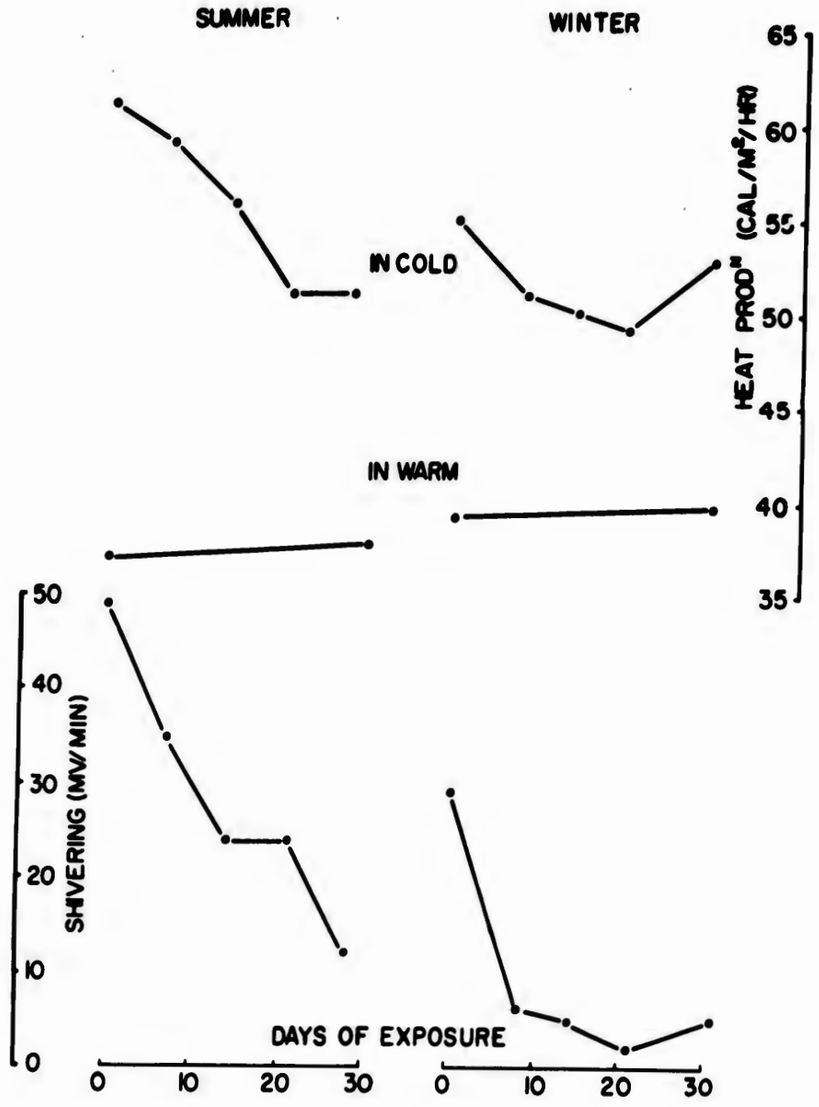


FIGURE 3

Alterations in mean shivering and mean heat production in nude subjects exposed 8 hours daily in a cold room. Measurements were made under standard conditions of exposure.

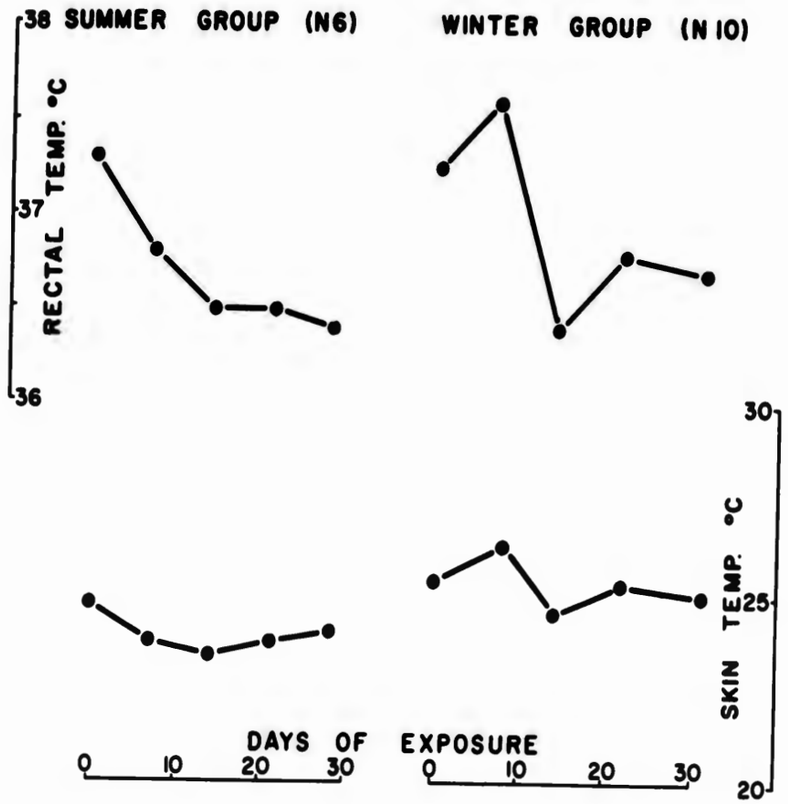


FIGURE 4

Alterations in mean rectal and mean skin temperature in subjects exposed 8 hours daily in a cold room. Measurements were made under standardized conditions of exposure.

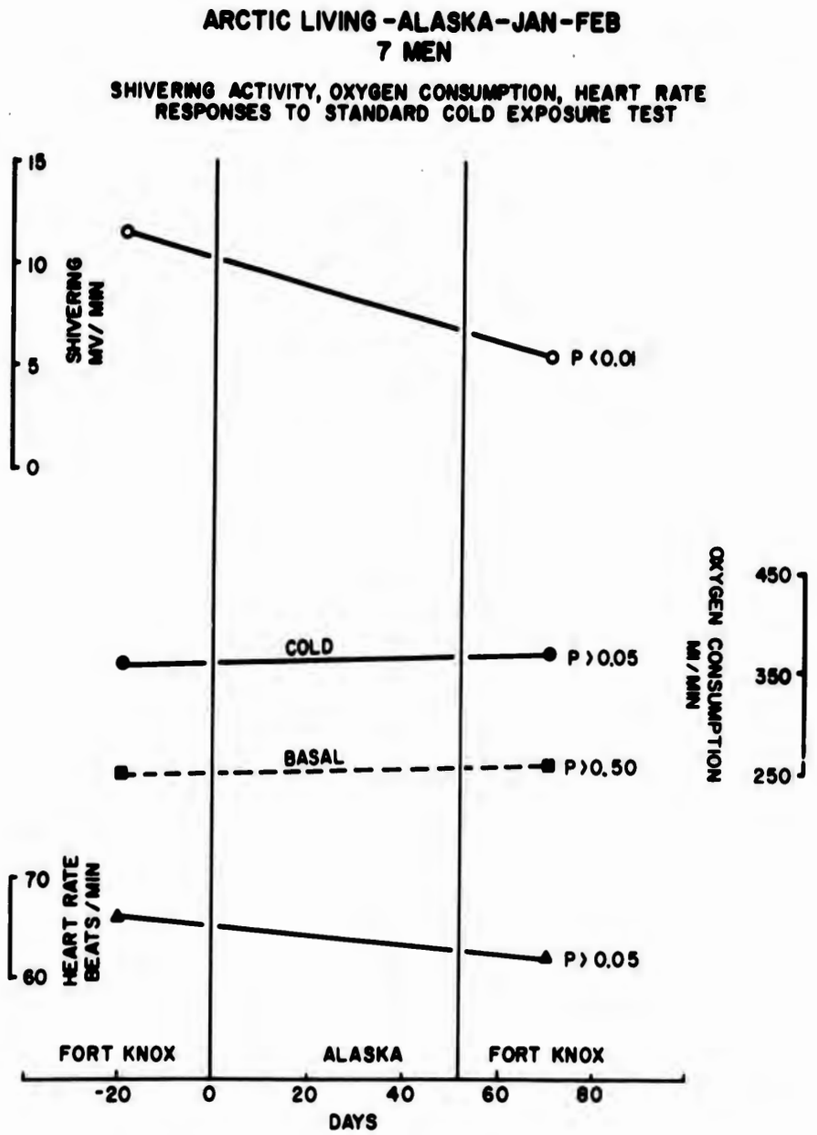


FIGURE 5

Responses to standard cold-exposure test (13.0°C for 90 min.) before and after living in Alaska for 52 days.

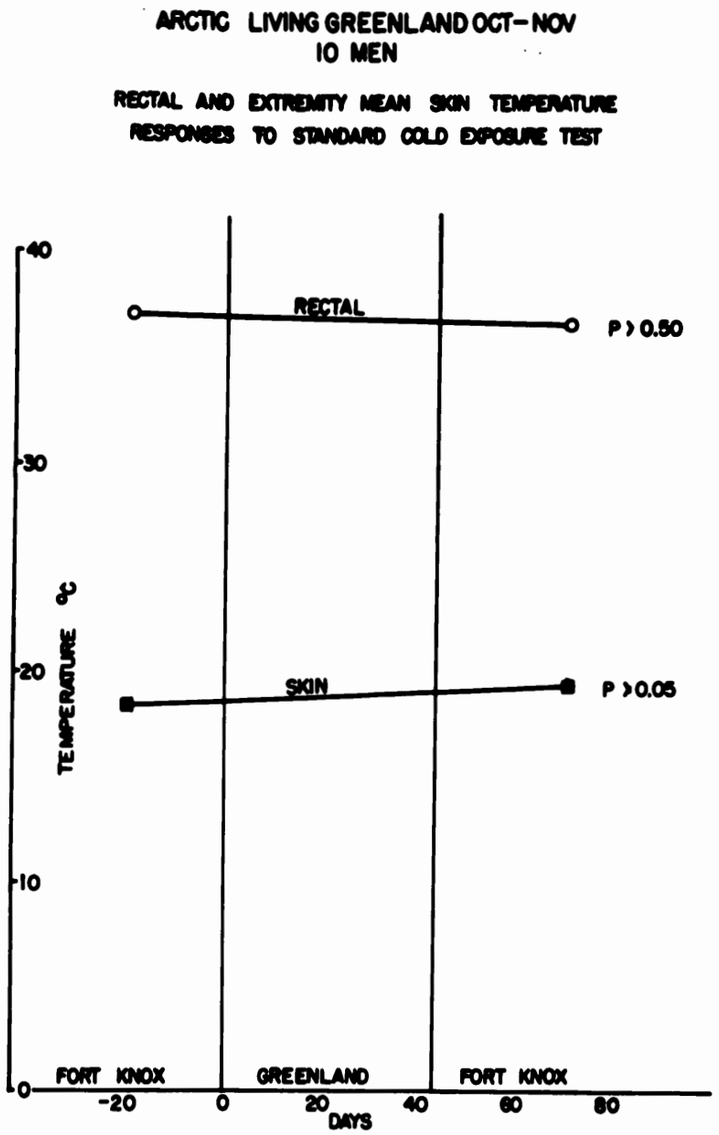


FIGURE 6

Responses to standard cold-exposure test 1.4°C for 30 min.) before and after living in Greenland for 43 days.

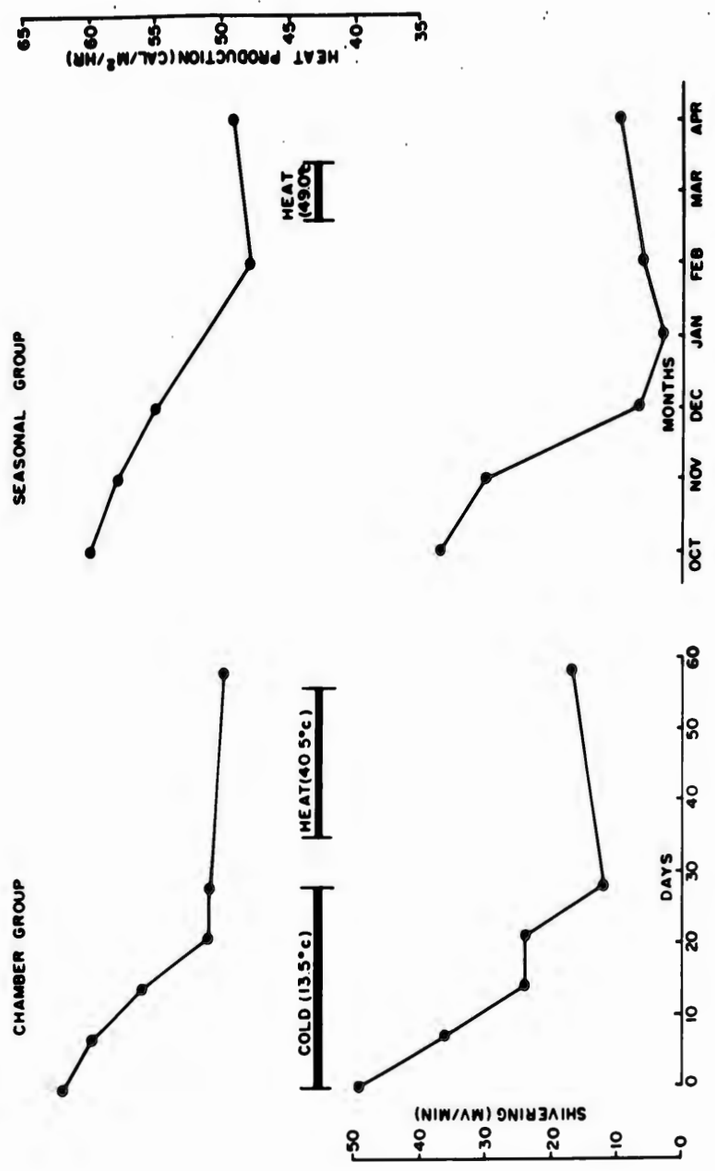


FIGURE 7

Responses to standard cold exposure test (13.5°C for 60 min.) for chamber acclimatized and seasonally acclimatized subjects. The 21 days of heat acclimatization after cold acclimatization does not change the responses to standard cold exposure test.

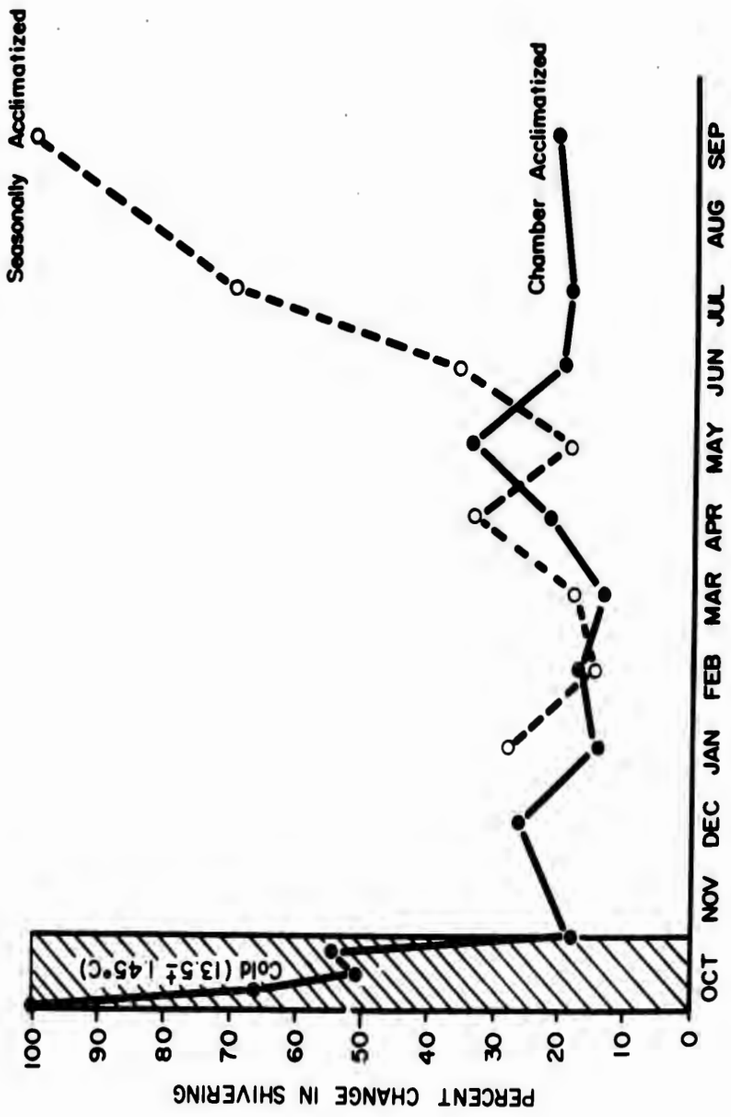


FIGURE 8

Responses to standard cold exposure test (13.5°C for 60 min.) performed once monthly on chamber acclimatized group and seasonally acclimatized group. No difference in the shivering response is seen until June. In July, August and September the chamber acclimatized group maintained their low level of shivering.

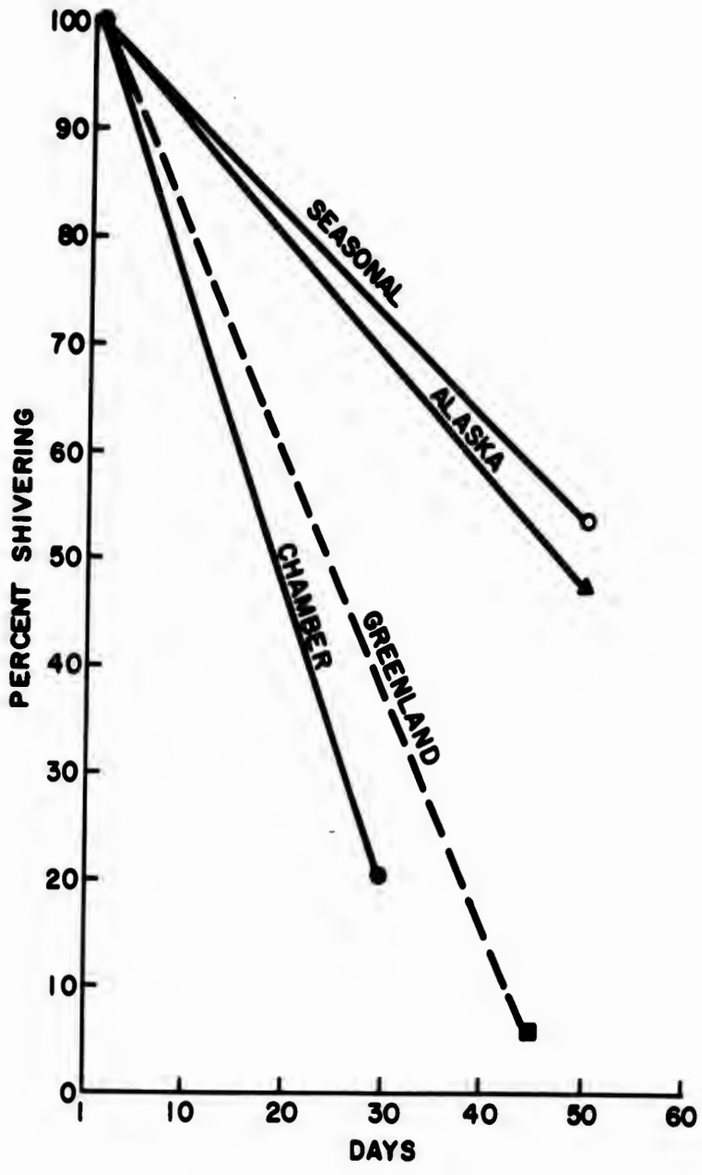


FIGURE 9

Comparison of the rate of decrease of shivering among 4 groups of subjects. Compares the percent change in shivering of several of the groups described. The slopes were derived by plotting the initial shivering value (before cold exposure) as 100 percent, and then plotting the post-cold shivering value as a percent of initial value.

THE DOPLOC DARK SATELLITE TRACKING SYSTEM

A. L. DEBEY and V. W. RICHARD
BALLISTIC RESEARCH LABORATORIES
ABERDEEN PROVING GROUND, MARYLAND

INTRODUCTION

The DOPLOC (DOppler Phase LOCk) dark satellite tracking system is a radio reflection Doppler system deriving its name from the Doppler frequency phase-locked tracking filter technique used. Dark, i.e. non-radiating, satellites are illuminated by a ground-based transmitter and signals reflected from illuminated satellites are received at one or more ground-based receiving sites.

The frequency of the transmitted signal is compared with the frequency of the signal received via reflection from the satellite and the difference, or Doppler frequency, constitutes the basic data from which orbital parameters are determined. A unique method has been developed for the determination of a complete set of orbital parameters from Doppler data recorded in the course of a single pass of a satellite. The ability of the DOPLOC system to provide complete orbital parameters, using only a few minutes of Doppler observations obtained from a dark satellite during the course of a single pass, provides a solution to a number of difficult and important space vehicle tracking problems. For example, it is desirable to know the orbital parameters as quickly as possible after launching a satellite. The orbital parameters of a newly launched satellite can be computed from DOPLOC data within minutes after the beginning of its free flight. Again, after attempting to deflect or steer a satellite into a different orbit, it is desirable to know the new orbital parameters within minutes after giving the steering command. It is also desirable to keep track of satellites after their transmitters have ceased operating or in cases where the transmitting frequencies are not suitable for tracking or are unknown.

Considerable impetus was given to the development of reflection Doppler techniques and instrumentation when, under a directive issued by the Advanced Research Projects Agency (ARPA) in June 1958, a DOPLOC system was installed as part of a national passive satellite detection fence across the southern portion of the United States.

This directive placed the responsibility on the Ballistic Research Laboratories to develop a long range passive Doppler tracking system, to construct, instrument and supervise the operation of tracking stations, to analyze received signals, and to produce orbital parameters of non-radiating satellites passing over a specified area of the continental United States. In response to the ARPA directive, the DOPLOC dark satellite tracking system was developed.

A three-station interim DOPLOC complex was installed in the center portion of the Satellite Detection Fence consisting of an illuminator transmitter at Fort Sill, Oklahoma, and receiving stations at Forrest City, Arkansas and White Sands Missile Range, New Mexico (Fig. 1). Although this complex ceased operation in June 1961, the equipment, Doppler tracking techniques, output data and results described in this paper are illustrative of the basic DOPLOC system.

DOPLOC SYSTEM DESCRIPTION

Transmitting Station

A 50 kw, 108 mc, continuous wave illuminator transmitter was used at the Fort Sill site, which was midway between the two receiving sites it served. Three transmitting antennas were provided, each with a fan beam shaped radiation pattern 8° by 76° and 16 db gain (Fig. 2). The fan beam pattern of the center antenna was vertically directed with its broad dimension oriented east-west. The beams of the other two antennas were directed 20° above the northern and southern horizon, respectively, with their broad pattern dimension horizontal. The transmitter could be quickly switched to any one of the three antennas.

Receiving Stations

Three antennas, each with a fan beam pattern $8^{\circ} \times 76^{\circ}$, were used at each receiving station. These antennas were oriented for maximum beam pattern volume overlap with the transmitting antennas. The central axes of the receiving and transmitting antennas intersected at about 900 miles range. In this manner, the antennas were oriented to detect the approach of a satellite as it came over the horizon, intercept it again overhead, and again as it receded toward the horizon (Fig. 3). The successive passage of a satellite through the three antenna beams provided three segments of the Doppler "S" curve as shown in Fig. 3.

Doppler Data

Doppler frequency is defined to be the frequency obtained by heterodyning a locally generated signal against the signal received from the satellite, followed by a correction for the frequency bias introduced as a result of the difference between the frequency of the local oscillator and that of the transmitter. If the Doppler frequency, as defined, is plotted as a function of time for a satellite in orbit, one obtains an "S" curve of the form shown in Fig. 4(a). The asymmetry of the curve is typical for a tracking system with a ground-based transmitter and receiver separated by an appreciable distance. This asymmetry contributes greatly to the unambiguous

determination of the orbital parameters from single pass data. Only for a satellite whose orbital plane bisects the base line will the Doppler data produce a symmetrical "S" curve. Fig. 4(a) illustrates an analog plot of the Doppler frequency received from tracking a satellite-borne transmitter. With a ground-based transmitter it is not economically feasible to illuminate the entire volume of space under surveillance; therefore, it is necessary to limit the number of observations to minimize the cost and complexity of the transmitter and antennas. The interim system installed in the Fence provided three sections of the "S" curve as shown in Fig. 4(b). Studies made subsequent to the interim Fence operation have shown that an optimum system should use a scanning fan beam antenna, providing a number of discrete observations at regular intervals, as shown in Fig. 4(c). Data in any of these forms may be used as input for the orbital parameter computing procedure, but, in order to handle the Doppler data rapidly and accurately, the Doppler frequency must be automatically counted and digitized at the receiving sites. Automatic, real-time counting of the Doppler frequency requires a signal of high quality, i.e., with very low noise content. High quality Doppler data, which are essentially noise free, are provided in the DOPLOC system by use of a very narrow bandwidth, phase-locked tracking filter following the receiver. The tracking filter must initially acquire the signal and be adjusted precisely in frequency (i.e. phase-locked) to the Doppler signal before it can function as a noise clean-up device. (A description of the automatic search and lock-on system follows.)

Automatic Search and Lock-On

One of the most significant developments of the DOPLOC system was a signal search and lock-on device which automatically placed the tracking filter on any signal occurring within the search frequency range of the equipment. The signals received via reflection from a satellite were very weak, ranging from 10^{-16} to 10^{-19} watts, which put them as much as 30 db (1000 times in power) below the noise at the receiver output. The receiver had a bandwidth of 16 kc to accommodate the range of Doppler frequencies which extended from 2 to 14 kc. The signals were of short duration, averaging about 8 seconds in the center beam. Thus, short duration, weak, noisy signals made it necessary to use a very sensitive, fast search, automatic lock-on (ALO) system. The ALO system developed used a comb filter frequency search unit. To meet the requirements of minimum search time, ten filters of 10 cps bandwidth were placed in the desired audio frequency band and the signal was fed to all of the filters. The ten filters were spaced 100 cps apart to cover a 1 kc band. These ten filters were switched twelve times to cover the full 12 kc band. The rate of switching determined the time available to find a signal as well as the total sweep time. Three switching rates were used, 2.5, 5 and 10 cps. When the 10 cps rate was used, the time for each 1 kc band was 0.1 second and 1.2 seconds for the entire Doppler frequency band. If a signal was present in the noise in the frequency range from 2 to 14 kc, possessing sufficient amplitude and was within the bandwidth of one of the fixed filters, it activated that filter. The filter, upon

receiving a signal, closed its control relay which put an oscillator frequency equal to that of the filter into a circuit termed the "set frequency" control. This circuit compared the fixed filter oscillator frequency with the tracking filter oscillator frequency and generated an error voltage which was applied to the tracking filter to produce a phase-lock between the oscillators. The tracking filter was then switched to the track position where the tracking filter loop was closed and a phase-lock was obtained on the true input signal. This search and lock-on sequence required a maximum of 1.2 seconds, which was adequate for the three antenna system. A newer automatic search system, using a bank of 1200 transistorized filters, has been developed recently to permit scanning the 2 to 14 kc range in 0.1 second. The need for this faster time response arose in connection with the proposed scanning DOPLOC system, where a narrow beamwidth scanning antenna would repeatedly sweep past a satellite, resulting in signal durations of about 0.1 second.

Tracking Filter

The tracking filter provides special noise filtering characteristics, making possible the successful reception of the extremely low energy signals returned from satellites at long ranges. Large improvements in the signal-to- noise ratio of noisy received signals are realized by extreme reduction of the system bandwidth through the use of the tracking filter. Bandwidths adjustable from 1 to 100 cps are available, with 10 cps normally used. The tracking filter is capable of phase-locked operation when the input signal is a maximum of 38 decibels below the noise, (i.e., a noise-to-signal ratio of 6300 when the receiver bandwidth is 16 kc and the filter bandwidth is 1 cps). The tracking filter is an electronic bandpass filter whose center frequency automatically tracks the frequency of the input signal. The filtering action is obtained by use of a frequency-controlled oscillator that is correlated (phase-locked) with the input signal. The basic block diagram of the tracking servo loop is shown in Fig. 5. Tracking is accomplished with an electronic servo system designed to make the frequency-controlled oscillator follow the frequency and phase of the input signal. This electronic servo system has been designed to yield essentially zero tracking error for a constant rate of change of input frequency. An inherent feature of this third-order servo control design is an effective acceleration memory which provides tracking through signal dropouts. Experience with signal reception from satellites has proved the necessity for this memory feature, since the received signal amplitude may vary widely and rapidly. The filter works through null periods very effectively without losing lock. In addition, this memory provides effective tracking of the desired Doppler signal in the presence of interfering signals or when several satellites are within receiving range simultaneously.

The signal-to-noise power improvement furnished by the tracking filter is equal to the ratio of the input source noise bandwidth to the filter bandwidth. The internal noise generated by the filter is negligible at all bandwidths. The relation between

input and output signal-to-noise is shown in Fig. 6, where the threshold sensitivity for a receiver with a 1 cps bandwidth and 3 db noise figure is shown to be 2 x 10^{-20} watts (-197 dbw, or 0.001 microvolts across 50 ohms).

An experimental study has been made of the relation between signal-to-noise ratio and the uncertainty or random error in measuring the frequency of a Doppler signal. The test results showing rms frequency error as a function of signal-to-noise ratio and tracking filter bandwidth are shown in Fig. 7. An integration time or counting interval of one second was used for these measurements. As a typical example, a signal 24 db down in the noise can be read to an accuracy of 0.08 cps when a 10 cps filter bandwidth is used.

Since the key to successful determination of orbits from a minimum number of observations lies in obtaining data with small values of random and systematic error, the high quality data output of the DOPLOC receiving system and tracking filter has been an important feature.

Data Handling

The basic Doppler information, available at the output of the tracking filter, was a constant amplitude, varying frequency sine wave. Recording was accomplished in both analog and digital form, utilizing magnetic tape recorders, strip-chart recorders, digital printers and teletype tape punches. To permit rapid data handling and real time transmission to the computer, it was necessary to digitize and encode the Doppler data at the receiving stations. The digitized Doppler data was transmitted via teletype to the Ballistic Research Laboratories where the received data, recorded on punched paper tape, was in a binary format suitable for feeding directly into the BRL ORDVAC computer used for the calculation of orbital parameters.

Magnetic tape recordings were also made of the raw, unfiltered data at the output of the receiver, which served as back-up data should a failure occur in the tracking filter-digitizing system during a satellite pass. This is a unique advantage of the audio frequency tracking filter used in the DOPLOC system in contrast to systems that use phase-locked filters in the radio frequency portion of the system. The latter systems cannot record the unfiltered signals and, consequently, have no back-up data in the event of a filter malfunction.

TRACKING RESULTS

Doppler Data

Examples of recorded data are shown in Figs. 8 and 9. Fig. 8 is a dual channel, strip-chart record of data obtained on Revolution 140 of satellite 1960 Delta (Discoverer XI). The upper record indicates frequency as a function of time, showing the three segments of the "S" curve. These segments correspond to the times of passage of the satellite through the three antenna beams. The step wave forms show the functioning of the automatic search and lock-on

system. The lower channel indicates the received signal strength. Fig. 9 shows two types of Doppler data output records. The first is the punched paper type in binary code on standard five-level teletype tape. The first punched data block contains Universal Time at the beginning of the run. Subsequent data blocks, recorded once per second, contain the Doppler data to the nearest 0.1 cycle per second. The second type of recording is the printed paper tape in Arabic numerals. Two examples are shown, one of Doppler frequency and the other the inverse of the frequency, or Doppler period. Each line contains Universal Time and the Doppler data at print-out time. The Doppler period print-out gives an order of magnitude higher resolution of Doppler recording than the frequency print-out.

Orbit Computation Procedure

The method of orbit solution consists of a curve-fitting procedure, in which a compatible set of approximations for the orbital parameters are improved by successive differential corrections. The approximations are obtained from a least-squares treatment of an overdetermined system of equations of condition. The imposed limitation of single pass detection permits several assumptions which considerably simplify the computing procedure. Among these is the assumption that the Earth may be treated dynamically as a sphere while geometrically regarding it as an ellipsoid. In addition, it is assumed that no serious loss in accuracy will result if drag is neglected as a dynamic force. With these assumptions, it is apparent that the satellite may be regarded as moving in a Keplerian orbit. Since the system operates at the relatively high frequency of 108 mc, it is feasible to neglect both atmospheric and ionospheric refraction effects. The analog data are used to determine initial position and velocity components from which are computed orbital parameters, position and velocity components versus time and Doppler frequency versus time. The latter data are compared to the observed Doppler data in a mathematical comparison routine. Differential corrections are derived and used to correct the initial point estimates. This iterative process is repeated until the corrections fall below predetermined lower limits after which the process is stopped and the final orbital parameters printed out by the computer.

Convergence of the computation rests primarily upon the adequacy of the initial approximations for position and velocity. It has been established that, for a system consisting of a single receiver and an earth-bound transmitter at opposite ends of a 400 mile base line, convergence is assured when the error in each coordinate of the initial estimate is not in excess of 50 to 75 miles and the velocity components are correct to within 1/2 to 1 mile per second. However, if single pass measurements are available from two or more receivers, the system geometry is greatly strengthened. Convergence can than be expected when the initial approximations are within 50 to 100 miles of the correct value in each coordinate and 1 to 2 miles per second in each velocity component. Several successful methods have been developed for computing sufficiently accurate initial approximations to position and velocity to assure convergence of the primary

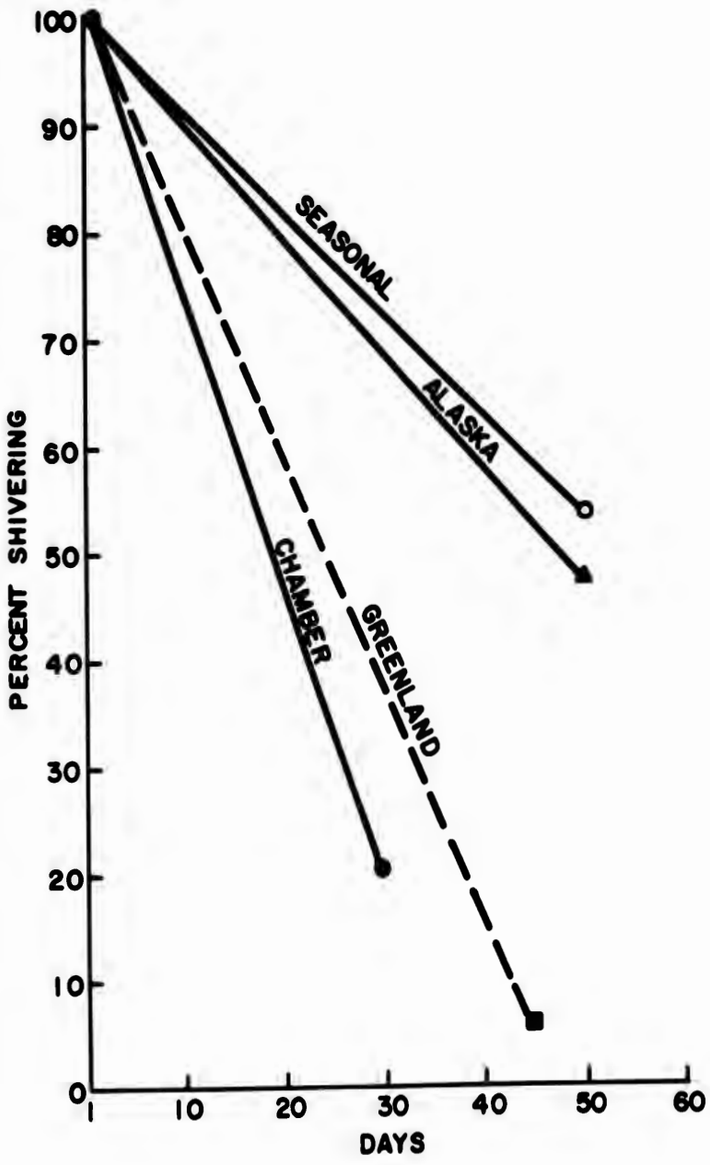


FIGURE 9

Comparison of the rate of decrease of shivering among 4 groups of subjects. Compares the percent change in shivering of several of the groups described. The slopes were derived by plotting the initial shivering value (before cold exposure) as 100 percent, and then plotting the post-cold shivering value as a percent of initial value.

THE DOPLOC DARK SATELLITE TRACKING SYSTEM

A. L. DEBEY and V. W. RICHARD
BALLISTIC RESEARCH LABORATORIES
ABERDEEN PROVING GROUND, MARYLAND

INTRODUCTION

The DOPLOC (DOppler Phase LOCk) dark satellite tracking system is a radio reflection Doppler system deriving its name from the Doppler frequency phase-locked tracking filter technique used. Dark, i.e. non-radiating, satellites are illuminated by a ground-based transmitter and signals reflected from illuminated satellites are received at one or more ground-based receiving sites.

The frequency of the transmitted signal is compared with the frequency of the signal received via reflection from the satellite and the difference, or Doppler frequency, constitutes the basic data from which orbital parameters are determined. A unique method has been developed for the determination of a complete set of orbital parameters from Doppler data recorded in the course of a single pass of a satellite. The ability of the DOPLOC system to provide complete orbital parameters, using only a few minutes of Doppler observations obtained from a dark satellite during the course of a single pass, provides a solution to a number of difficult and important space vehicle tracking problems. For example, it is desirable to know the orbital parameters as quickly as possible after launching a satellite. The orbital parameters of a newly launched satellite can be computed from DOPLOC data within minutes after the beginning of its free flight. Again, after attempting to deflect or steer a satellite into a different orbit, it is desirable to know the new orbital parameters within minutes after giving the steering command. It is also desirable to keep track of satellites after their transmitters have ceased operating or in cases where the transmitting frequencies are not suitable for tracking or are unknown.

Considerable impetus was given to the development of reflection Doppler techniques and instrumentation when, under a directive issued by the Advanced Research Projects Agency (ARPA) in June 1958, a DOPLOC system was installed as part of a national passive satellite detection fence across the southern portion of the United States.

This directive placed the responsibility on the Ballistic Research Laboratories to develop a long range passive Doppler tracking system, to construct, instrument and supervise the operation of tracking stations, to analyze received signals, and to produce orbital parameters of non-radiating satellites passing over a specified area of the continental United States. In response to the ARPA directive, the DOPLOC dark satellite tracking system was developed.

A three-station interim DOPLOC complex was installed in the center portion of the Satellite Detection Fence consisting of an illuminator transmitter at Fort Sill, Oklahoma, and receiving stations at Forrest City, Arkansas and White Sands Missile Range, New Mexico (Fig. 1). Although this complex ceased operation in June 1961, the equipment, Doppler tracking techniques, output data and results described in this paper are illustrative of the basic DOPLOC system.

DOPLOC SYSTEM DESCRIPTION

Transmitting Station

A 50 kw, 108 mc, continuous wave illuminator transmitter was used at the Fort Sill site, which was midway between the two receiving sites it served. Three transmitting antennas were provided, each with a fan beam shaped radiation pattern 8° by 76° and 16 db gain (Fig. 2). The fan beam pattern of the center antenna was vertically directed with its broad dimension oriented east-west. The beams of the other two antennas were directed 20° above the northern and southern horizon, respectively, with their broad pattern dimension horizontal. The transmitter could be quickly switched to any one of the three antennas.

Receiving Stations

Three antennas, each with a fan beam pattern $8^{\circ} \times 76^{\circ}$, were used at each receiving station. These antennas were oriented for maximum beam pattern volume overlap with the transmitting antennas. The central axes of the receiving and transmitting antennas intersected at about 900 miles range. In this manner, the antennas were oriented to detect the approach of a satellite as it came over the horizon, intercept it again overhead, and again as it receded toward the horizon (Fig. 3). The successive passage of a satellite through the three antenna beams provided three segments of the Doppler "S" curve as shown in Fig. 3.

Doppler Data

Doppler frequency is defined to be the frequency obtained by heterodyning a locally generated signal against the signal received from the satellite, followed by a correction for the frequency bias introduced as a result of the difference between the frequency of the local oscillator and that of the transmitter. If the Doppler frequency, as defined, is plotted as a function of time for a satellite in orbit, one obtains an "S" curve of the form shown in Fig. 4(a). The asymmetry of the curve is typical for a tracking system with a ground-based transmitter and receiver separated by an appreciable distance. This asymmetry contributes greatly to the unambiguous

determination of the orbital parameters from single pass data. Only for a satellite whose orbital plane bisects the base line will the Doppler data produce a symmetrical "S" curve. Fig. 4(a) illustrates an analog plot of the Doppler frequency received from tracking a satellite-borne transmitter. With a ground-based transmitter it is not economically feasible to illuminate the entire volume of space under surveillance; therefore, it is necessary to limit the number of observations to minimize the cost and complexity of the transmitter and antennas. The interim system installed in the Fence provided three sections of the "S" curve as shown in Fig. 4(b). Studies made subsequent to the interim Fence operation have shown that an optimum system should use a scanning fan beam antenna, providing a number of discrete observations at regular intervals, as shown in Fig. 4(c). Data in any of these forms may be used as input for the orbital parameter computing procedure, but, in order to handle the Doppler data rapidly and accurately, the Doppler frequency must be automatically counted and digitized at the receiving sites. Automatic, real-time counting of the Doppler frequency requires a signal of high quality, i.e., with very low noise content. High quality Doppler data, which are essentially noise free, are provided in the DOPLOC system by use of a very narrow bandwidth, phase-locked tracking filter following the receiver. The tracking filter must initially acquire the signal and be adjusted precisely in frequency (i.e. phase-locked) to the Doppler signal before it can function as a noise clean-up device. (A description of the automatic search and lock-on system follows.)

Automatic Search and Lock-On

One of the most significant developments of the DOPLOC system was a signal search and lock-on device which automatically placed the tracking filter on any signal occurring within the search frequency range of the equipment. The signals received via reflection from a satellite were very weak, ranging from 10^{-16} to 10^{-19} watts, which put them as much as 30 db (1000 times in power) below the noise at the receiver output. The receiver had a bandwidth of 16 kc to accommodate the range of Doppler frequencies which extended from 2 to 14 kc. The signals were of short duration, averaging about 8 seconds in the center beam. Thus, short duration, weak, noisy signals made it necessary to use a very sensitive, fast search, automatic lock-on (ALO) system. The ALO system developed used a comb filter frequency search unit. To meet the requirements of minimum search time, ten filters of 10 cps bandwidth were placed in the desired audio frequency band and the signal was fed to all of the filters. The ten filters were spaced 100 cps apart to cover a 1 kc band. These ten filters were switched twelve times to cover the full 12 kc band. The rate of switching determined the time available to find a signal as well as the total sweep time. Three switching rates were used, 2.5, 5 and 10 cps. When the 10 cps rate was used, the time for each 1 kc band was 0.1 second and 1.2 seconds for the entire Doppler frequency band. If a signal was present in the noise in the frequency range from 2 to 14 kc, possessing sufficient amplitude and was within the bandwidth of one of the fixed filters, it activated that filter. The filter, upon

receiving a signal, closed its control relay which put an oscillator frequency equal to that of the filter into a circuit termed the "set frequency" control. This circuit compared the fixed filter oscillator frequency with the tracking filter oscillator frequency and generated an error voltage which was applied to the tracking filter to produce a phase-lock between the oscillators. The tracking filter was then switched to the track position where the tracking filter loop was closed and a phase-lock was obtained on the true input signal. This search and lock-on sequence required a maximum of 1.2 seconds, which was adequate for the three antenna system. A newer automatic search system, using a bank of 1200 transistorized filters, has been developed recently to permit scanning the 2 to 14 kc range in 0.1 second. The need for this faster time response arose in connection with the proposed scanning DOPLOC system, where a narrow beamwidth scanning antenna would repeatedly sweep past a satellite, resulting in signal durations of about 0.1 second.

Tracking Filter

The tracking filter provides special noise filtering characteristics, making possible the successful reception of the extremely low energy signals returned from satellites at long ranges. Large improvements in the signal-to- noise ratio of noisy received signals are realized by extreme reduction of the system bandwidth through the use of the tracking filter. Bandwidths adjustable from 1 to 100 cps are available, with 10 cps normally used. The tracking filter is capable of phase-locked operation when the input signal is a maximum of 38 decibels below the noise, (i.e., a noise-to-signal ratio of 6300 when the receiver bandwidth is 16 kc and the filter bandwidth is 1 cps). The tracking filter is an electronic bandpass filter whose center frequency automatically tracks the frequency of the input signal. The filtering action is obtained by use of a frequency-controlled oscillator that is correlated (phase-locked) with the input signal. The basic block diagram of the tracking servo loop is shown in Fig. 5. Tracking is accomplished with an electronic servo system designed to make the frequency-controlled oscillator follow the frequency and phase of the input signal. This electronic servo system has been designed to yield essentially zero tracking error for a constant rate of change of input frequency. An inherent feature of this third-order servo control design is an effective acceleration memory which provides tracking through signal dropouts. Experience with signal reception from satellites has proved the necessity for this memory feature, since the received signal amplitude may vary widely and rapidly. The filter works through null periods very effectively without losing lock. In addition, this memory provides effective tracking of the desired Doppler signal in the presence of interfering signals or when several satellites are within receiving range simultaneously.

The signal-to-noise power improvement furnished by the tracking filter is equal to the ratio of the input source noise bandwidth to the filter bandwidth. The internal noise generated by the filter is negligible at all bandwidths. The relation between

input and output signal-to-noise is shown in Fig. 6, where the threshold sensitivity for a receiver with a 1 cps bandwidth and 3 db noise figure is shown to be 2 x 10^{-20} watts (-197 dbw, or 0.001 microvolts across 50 ohms).

An experimental study has been made of the relation between signal-to-noise ratio and the uncertainty or random error in measuring the frequency of a Doppler signal. The test results showing rms frequency error as a function of signal-to-noise ratio and tracking filter bandwidth are shown in Fig. 7. An integration time or counting interval of one second was used for these measurements. As a typical example, a signal 24 db down in the noise can be read to an accuracy of 0.08 cps when a 10 cps filter bandwidth is used.

Since the key to successful determination of orbits from a minimum number of observations lies in obtaining data with small values of random and systematic error, the high quality data output of the DOPLOC receiving system and tracking filter has been an important feature.

Data Handling

The basic Doppler information, available at the output of the tracking filter, was a constant amplitude, varying frequency sine wave. Recording was accomplished in both analog and digital form, utilizing magnetic tape recorders, strip-chart recorders, digital printers and teletype tape punches. To permit rapid data handling and real time transmission to the computer, it was necessary to digitize and encode the Doppler data at the receiving stations. The digitized Doppler data was transmitted via teletype to the Ballistic Research Laboratories where the received data, recorded on punched paper tape, was in a binary format suitable for feeding directly into the BRL ORDVAC computer used for the calculation of orbital parameters.

Magnetic tape recordings were also made of the raw, unfiltered data at the output of the receiver, which served as back-up data should a failure occur in the tracking filter-digitizing system during a satellite pass. This is a unique advantage of the audio frequency tracking filter used in the DOPLOC system in contrast to systems that use phase-locked filters in the radio frequency portion of the system. The latter systems cannot record the unfiltered signals and, consequently, have no back-up data in the event of a filter malfunction.

TRACKING RESULTS

Doppler Data

Examples of recorded data are shown in Figs. 8 and 9. Fig. 8 is a dual channel, strip-chart record of data obtained on Revolution 140 of satellite 1960 Delta (Discoverer XI). The upper record indicates frequency as a function of time, showing the three segments of the "S" curve. These segments correspond to the times of passage of the satellite through the three antenna beams. The step wave forms show the functioning of the automatic search and lock-on

system. The lower channel indicates the received signal strength. Fig. 9 shows two types of Doppler data output records. The first is the punched paper type in binary code on standard five-level teletype tape. The first punched data block contains Universal Time at the beginning of the run. Subsequent data blocks, recorded once per second, contain the Doppler data to the nearest 0.1 cycle per second. The second type of recording is the printed paper tape in Arabic numerals. Two examples are shown, one of Doppler frequency and the other the inverse of the frequency, or Doppler period. Each line contains Universal Time and the Doppler data at print-out time. The Doppler period print-out gives an order of magnitude higher resolution of Doppler recording than the frequency print-out.

Orbit Computation Procedure

The method of orbit solution consists of a curve-fitting procedure, in which a compatible set of approximations for the orbital parameters are improved by successive differential corrections. The approximations are obtained from a least-squares treatment of an overdetermined system of equations of condition. The imposed limitation of single pass detection permits several assumptions which considerably simplify the computing procedure. Among these is the assumption that the Earth may be treated dynamically as a sphere while geometrically regarding it as an ellipsoid. In addition, it is assumed that no serious loss in accuracy will result if drag is neglected as a dynamic force. With these assumptions, it is apparent that the satellite may be regarded as moving in a Keplerian orbit. Since the system operates at the relatively high frequency of 108 mc, it is feasible to neglect both atmospheric and ionospheric refraction effects. The analog data are used to determine initial position and velocity components from which are computed orbital parameters, position and velocity components versus time and Doppler frequency versus time. The latter data are compared to the observed Doppler data in a mathematical comparison routine. Differential corrections are derived and used to correct the initial point estimates. This iterative process is repeated until the corrections fall below predetermined lower limits after which the process is stopped and the final orbital parameters printed out by the computer.

Convergence of the computation rests primarily upon the adequacy of the initial approximations for position and velocity. It has been established that, for a system consisting of a single receiver and an earth-bound transmitter at opposite ends of a 400 mile base line, convergence is assured when the error in each coordinate of the initial estimate is not in excess of 50 to 75 miles and the velocity components are correct to within 1/2 to 1 mile per second. However, if single pass measurements are available from two or more receivers, the system geometry is greatly strengthened. Convergence can than be expected when the initial approximations are within 50 to 100 miles of the correct value in each coordinate and 1 to 2 miles per second in each velocity component. Several successful methods have been developed for computing sufficiently accurate initial approximations to position and velocity to assure convergence of the primary

computation. Fig. 10 illustrates a graphical method for determining initial position estimates from the analog data. The measured values of Doppler frequency and rate of change of frequency at the time of passage through the midpoint of the vertical beam are entered on the graph. Associated values of satellite altitude and distance east of the transmitter site can be read. The north-south position of the satellite is on the great circle path connecting the transmitter and receiver since the data are taken at the time the satellite is in the center of the vertical antenna beam. Velocity components are determined consistent with the assumption of circular motion, the height, which is determined graphically, and the assumed inclination. In addition to the graphical solution, a digital method has been developed which is suitable for machine computation. This method establishes an approximate orbit to provide initial approximations for the more sophisticated primary computation which in turn yields a refined orbit determination.

Results of Orbital Computation

Numerous convergent solutions have been obtained with actual field data from a single receiver. Orbital parameters computed from DOPLOC data obtained for 1960 Delta (Discoverer XI) are shown in Fig. 11. Measurements were recorded for 14 seconds in the north antenna beam, 7 seconds in the center beam, and 8 seconds in the south beam, with two gaps of about 45 seconds each in the data. Thus, observations were recorded for a total of 29 seconds within a time interval of two minutes. Using the graphical method described above to obtain initial approximations, convergence was achieved in three iterations, on the first pass through the computing machine. It will be noted, in the comparison of DOPLOC and Space Track (National Space Surveillance Control Center, Bedford, Massachusetts) results, that there is good agreement in the semi-major axis, eccentricity, inclination and right ascension of the ascending node. This is characteristic of the single pass solution when the eccentricity is small and the computational input is limited to Doppler frequency. Since the orbit is very close to being circular, the mean anomaly at epoch and argument of perigee are difficult for either the DOPLOC System or Space Track to determine accurately. When limited to single pass, single-receiver observations, the DOPLOC system provides an excellent determination of the orientation of the orbital plane, a good determination of the shape of the orbit, and a fair-to-poor determination of the orientation of the ellipse within the plane.

As an example of the accuracy of DOPLOC observations and the strength of the orbital computation method, a solution was determined using only seven frequency observations during the passage of 1958 Delta (Sputnik III). The observations were selected to serve as an example of the type of reduction required for the proposed DOPLOC scanning beam antenna system. The results, presented in Fig. 12, show very good agreement with the published orbit. This example shows that the method is quite feasible for use with periodic, discrete measurements of Doppler frequency. Of course, the proposed system would normally yield several more observations than were available in

this example.

CONCLUSION

It is noteworthy that numerous solutions have been obtained with field data from a single DOPLOC receiver recorded during a single pass of a satellite. Further, these measurements have been confined to three short periods of observation within a two to three minute interval. Additional receivers spread over greater distances would, of course, considerably enhance the accuracy of the results. For example, a system with two receivers and a ground transmitter would reduce error propagation in the computations tc approximately one-tenth of that to be expected from a system with a single receiver. Removing the restriction of single pass determination would further enhance the accuracy of the results.

Computing times have been found to be reasonable. Convergent solutions have required 20 to 40 minutes on the Aberdeen Proving Ground ORDVAC which requires the coding to include floating decimal sub-routines. More modern machines, such as the BRLESC now in operation at the Ballistic Research Laboratories, could perform the same computation in 2 to 4 minutes. Hence, it is realistic to state that the system has the potential capability of orbit determination within five minutes after observation time. In addition, the DOPLOC tracking method has general application and, therefore, need not be confined to satellites in Keplerian orbits. Other applications would be the tracking of projectiles, rockets, guided missiles and space vehicles for the determination of acceleration and velocity components and trajectory parameters.

BIBLIOGRAPHY

1. Richard, V. W., DOPLOC Tracking Filter, BRL Memorandum Report No. 1173, October 1958.

2. deBey, L. G., Tracking in Space by DOPLOC, IRE Transactions on Military Electronics, Vol MIL-4, Numbers 2 and 3, April-July 1960.

3. deBey, L. G., DOPLOC Tracking Determines Orbits of Satellites, Electronic Industries, October 1960.

4. Patton, R. B., Orbit Determination from Single Pass Doppler Observations, IRE Transactions on Military Electronics, Vol MIL-4, Numbers 2 and 3, April-July 1960.

5. Patton, R. B., Richard, V.W., Determination of Orbital Elements and Refraction Effects from Single Pass Doppler Observations, BRL Memorandum Report No. 1357, June 1961.

6. Hodge, A. H., Final Technical Summary, Report Period 20 June 1958-30 June 1961, BRL-ARPA DOPLOC Satellite Detection Complex, BRL Report No. 1136, July 1961.

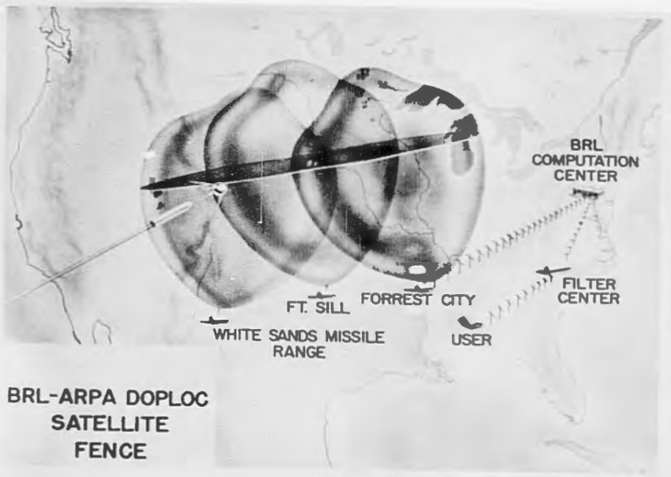
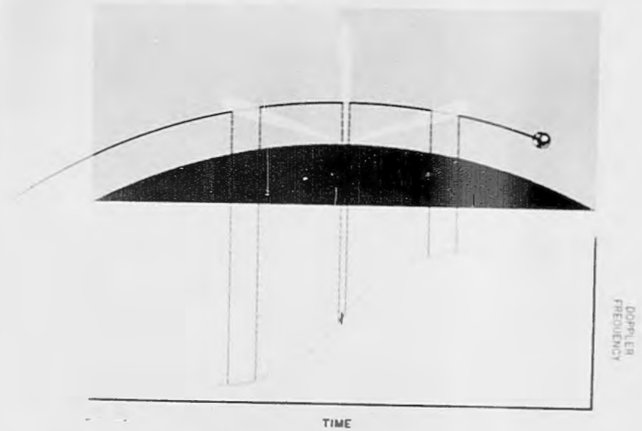


FIG 1

FIG 2. TRANSMITTING ANTENNAS



BASIC INTERIM DOPLOC SYSTEM

FIG 3

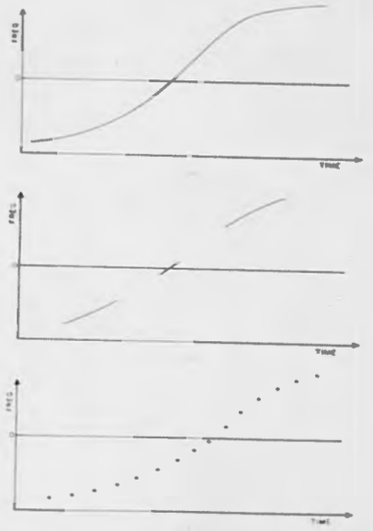
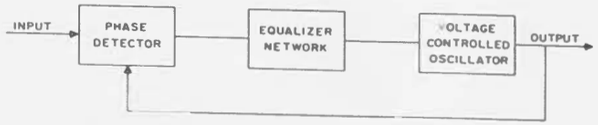
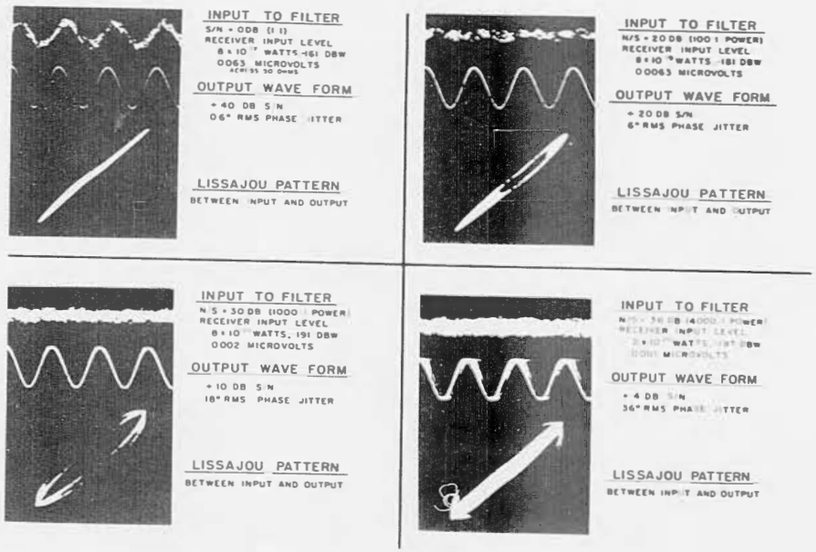


FIG 4 DOPPLER FREQUENCY – TIME CURVES



BASIC TRACKING FILTER SYSTEM

FIG. 5



INPUT & OUTPUT WAVEFORMS VS NOISE TO SIGNAL RATIO
FOR 1 CPS BANDWIDTH

FIG 6

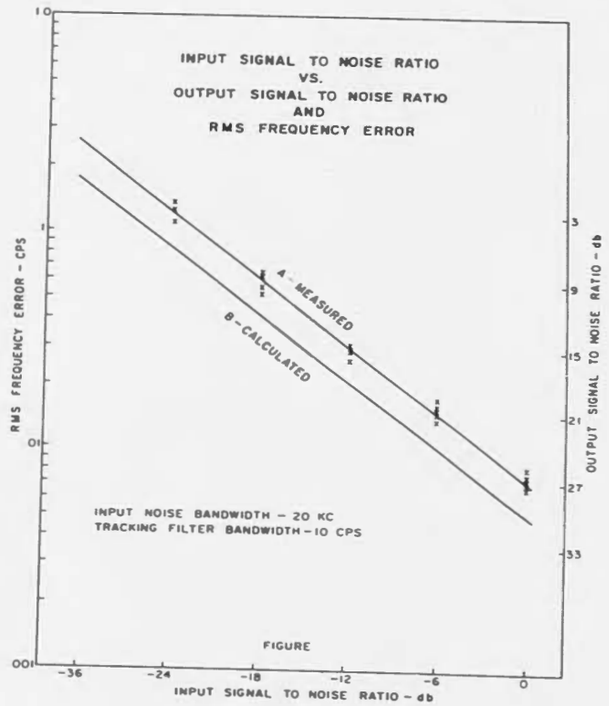
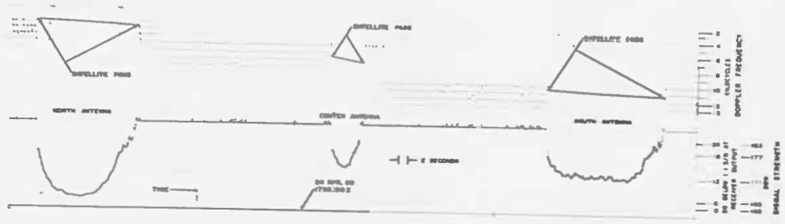
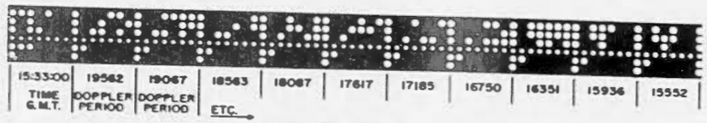


FIG. 7

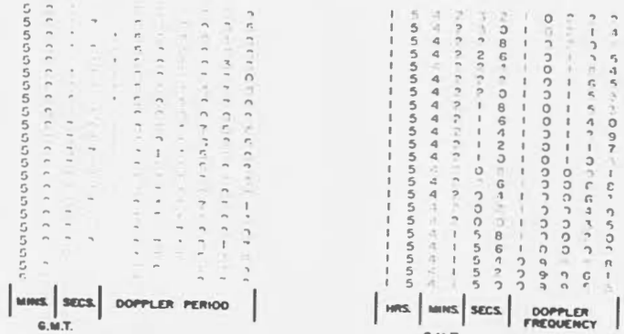


DOPPLER RECORD OF 60 DELTA, REV 140, FORREST CITY, ARKANSAS
17:28:06Z, ALTITUDE 116 MILES, 320 MILES EAST FORT SILL

FIG. 8

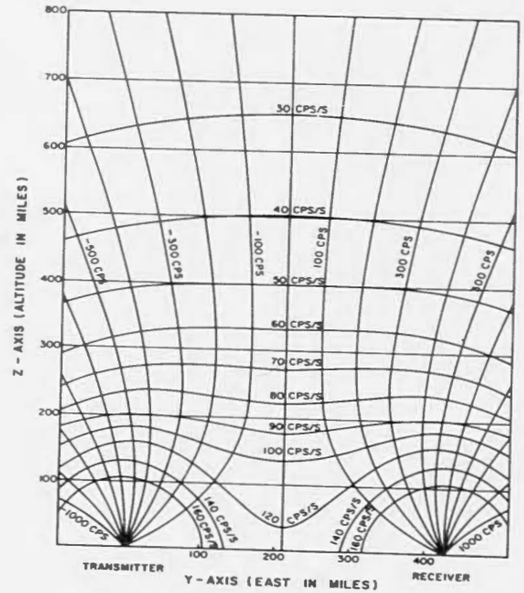


BINARY CODED PUNCHED TAPE



TYPICAL DOPLOC DATA OUTPUT

FIG. 9

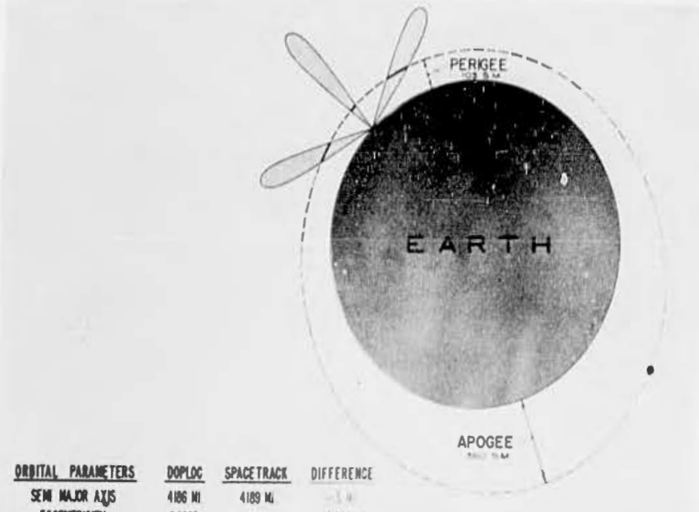


DOPLOC FREQUENCY AND RATE OF CHANGE OF FREQUENCY AS A FUNCTION OF POSITION IN THE YZ - PLANE (FOR 80° INCLINATION)

FIG. 10

TYPICAL SINGLE PASS DOPPLER ORBIT SOLUTION

(BY ARMY - DOPLOC SATELLITE TRACKING SYSTEM)
REVOLUTION NO. 30 OF 1960 DELTA (DISCOVERER XI)



ORBITAL PARAMETERS	DOPLOC	SPACETRACK	DIFFERENCE
SEMI MAJOR AXIS	4186 MI	4189 MI	
ECCENTRICITY	0.0295	0.0291	
INCLINATION	80.15°	80.01°	
RIGHT ASCENSION OF ASCENDING NODE	295.82°	295.84°	
ARGUMENT OF PERIGEE	154.96°	146.28°	

FIG. 11

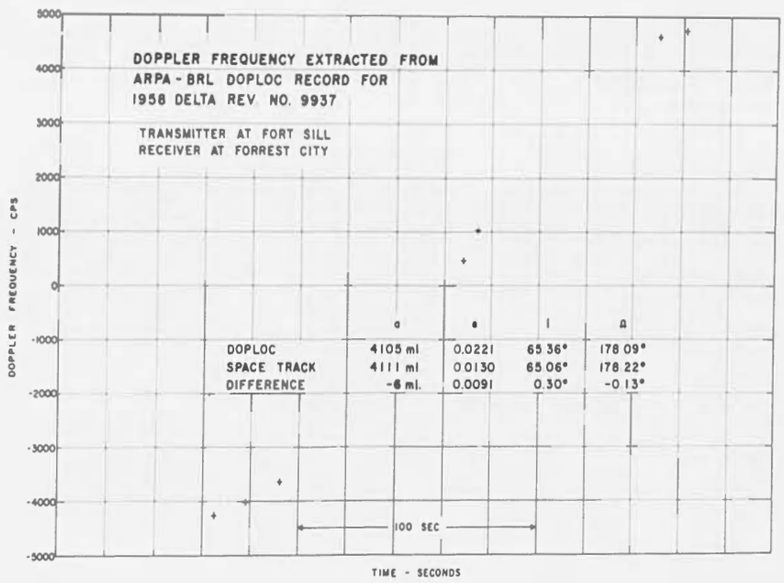


FIG. 12

DEVELOPMENT OF REPETITIVE KILOTON THRUSTS OF MILLISECONDS DURATION FOR SIMULATING THE FIRING AND DETERMINING THE RELIABILITY OF HEAVY WEAPON COMPONENTS

C. W. EGAN and J. P. PURTELL
WATERVLIET ARSENAL
WATERVLIET, NEW YORK

HISTORICAL: In order to provide realistic and economical life estimates for breech components of heavy ordnance weapons and to compare alternate designs, a series of relatively high energy fatigue machines have been built over a period of years which cyclically reproduce stress environments encountered in actual service.

In the earlier models, experiments were conducted with a drop tester in which a falling weight impacted on a piston fitted into a shortened gun tube. This piston transmitted a dynamic thrust to the breech. The effects of the magnitude of the falling weight, the drop height, the foundation mass and the spring constant of the system upon the thrust-time pulse were explored. It was found that peak thrust varied directly with the weight of the falling mass, and with approximately the square root of the drop height. The device was found to require a substantial base and foundation in order to maintain stability. In order to achieve the desired pulse shape and duration, it was necessary to provide a means for varying the spring constant. This was accomplished by use of a liquid spring.

After this preliminary study was completed, a partly automated unit was constructed utilizing a steam-driven pile driver which transmitted the load to the specimen through a liquid-cushioned, force-intensifying, hydro-mechanical linkage. This unit is presently capable of applying dynamic thrusts up to 1,750,000 pounds in a few milliseconds at a cyclic rate of over 100 per minute. It is expected that thrusts up to 2,500,000 pounds will be attainable in the near future.

DETAILS OF TEST MACHINE & SPECIMENS: The fatigue tester has been designed for flexibility in order to accommodate the wide range of breech components to be tested. A weight of 1,000 5,000 or 10,000 pounds may

be repetitively dropped from heights up to 42 inches. Load pulses may be adjusted for magnitude and duration by varying the volume of the fluid chamber, when used, or by substituting fluids of different compressibilities. Specimen supporters used to hold the component in place during testing are shrink assembled to the test machine and may be readily changed when a different type component is to be tested.

The first major type of breech component tested on the fatigue machine is a sliding block type breech ring such as is used in some tank mounted guns. The configuration of the machine for the type of specimen is shown in Fig. 1. Although the breech ring is the primary test specimen, the associated breech block and a short stubby load cell whose lower end is cartridge shaped are used so the load may be transmitted to the ring in approximately the same manner as in firing. The strain gaged load cell is calibrated in place by applying a static pressure to the fluid chamber. Early calibrations on a high capacity universal tester were found to be extremely erratic since it was found that the response of this short load cell depended markedly on the manner in which the load was distributed over the bearing area.

The lack of thermal stress generation is of no major concern for this type of specimen since, in actual firing, the breech ring is effectively insulated from thermal shock by the breech block and tube. Likewise, the thrust only, and not the firing pressure, is reproduced.

A second type of component tested is the coupling of a so-called split-chamber breech in which no breech block is used. Such a breech is shown, ready for testing, in Fig. 2. In this breech, the chamber and tube are subjected to thermal shock during firing while the coupling is considered to be relatively insulated from the thermal shock. In the test machine the fluid chamber for pulse shaping is moved to the specimen chamber itself and a continuous single piston may be used. In this manner, the internal pressure, as well as thrust, is simulated in the breech. Calibration is achieved from static chamber pressure measurements correlated with signals from the monitoring strain gages mounted externally on the chamber. Additional monitoring of thrust is effected with an accelerometer installed on the falling weight.

A third type of specimen is exemplified by mortar baseplates, such as shown in Fig. 3. These specimens are characterized by their low thrust levels and springiness which eliminate the necessity for a pulse shaper. Thrust is monitored by strain gages mounted on a reduced section of the piston. Calibration of the piston gages is carried out on a universal tester.

The fatigue tester is equipped with an insulated cold chamber and may be operated at sub-zero temperatures by circulating air from a dry ice heat exchanger. With this setup, the effect of extreme cold on fatigue response may be checked.

THEORETICAL ANALYSIS: In order to facilitate the design of experiments, an analysis of the high energy impact machine was made. The information obtained from this analysis allows closer control over the desired force-time pulse for a given experiment by means of proper design of pistons and fluid chambers, and eliminates much of the "cut-and-try" procedure of running the test.

The impact machine and the specimen, whether it be a breech ring or a breech chamber coupling, may be represented by a four degree of freedom spring-mass system. The analysis of this system is accomplished by well-known classical methods. [1] The equations of motion are:

$$x_i = -\alpha_{i1} m_1 \ddot{x}_1 - \alpha_{i2} m_2 \ddot{x}_2 - \alpha_{i3} m_3 \ddot{x}_3 - \alpha_{i4} m_4 \ddot{x}_4 \qquad (1)$$

where: $i = 1, 2, 3, 4,$

α's are influence coefficients. The initial conditions (at $t = 0$) are:

$$x_i = 0, \qquad i = 1, 2, 3, 4,$$
$$\dot{x}_1 = \sqrt{2gh}, \text{ where } h = \text{drop height of the falling weight,} \qquad (2)$$
$$\dot{x}_2 = \dot{x}_3 = \dot{x}_4 = 0.$$

The following solution is assumed:

$$x_i^* = a_i \sin \omega t. \qquad (3)$$

Substituting equations (3) into equations (1) yields:

a. The frequency equation, from which the natural frequencies ω_j ($j = 1,2,3,4$) are obtained;

b. The amplitude ratios a_1^j/a_4^j, a_2^j/a_4^j, etc., which correspond to the natural frequencies.

Equations (3), then, are a solution for motion at a natural frequency only. The general solution x_i may now be obtained by a superposition of the solutions for motion at the natural frequencies.

$$x_i = a_i^1 \sin \omega_1 t + a_i^2 \sin \omega_2 t + a_i^3 \sin \omega_3 t + a_i^4 \sin \omega_4 t. \qquad (4)$$

The constants C_i^j are evaluated by substituting equations (4) into equations (1) and using the initial conditions (2).

This analysis has been applied to the mortar baseplate fatigue test, which, because of the absence of an upper piston, reduces to a three degree of freedom system, shown in Fig. 4. The force transmitted to the baseplate is obtained by multiplying x_3 (Calculated from one of equations (4)) by the spring rate of the baseplate (k_3). The resulting force-time curve is plotted along with the experimentally observed curve in Fig. 5. It can be seen that good correlation exists with regard to pulse duration and maximum force.

EXPERIMENTAL PROCEDURES: Typically, a series of breech components of the same or alternate designs are selected for fatigue testing. Since the complicated geometry of the specimens precludes an exact stress prediction, strain measurements are taken at various points to guide the design of future components and to provide a basis for correlation of results on alternate designs.

Before fatigue tests are run, a static test is conducted in order to permit precise measurement of load strains. This static test is accomplished by pressurizing the fluid chamber, when used, to the proper pressure with a pump and intensifier system. For mortar baseplates, where no fluid chamber is used, a universal test machine is used for the static test. In addition to the strain gages, instrumentation for the static tests includes switching and balancing units and a strain indicator.

Instrumentation for the dynamic tests consists of cathode followers and oscilloscopes for the piezoelectric accelerometers and pressure gages, and oscilloscopes or oscillographs with appropriate amplifiers for the strain gages. Oscilloscopes are triggered by the interrupting of a light beam by the falling weight. The same photocell used for triggering actuates a counter which records the number of simulated rounds.

RESULTS: The full scale fatigue tester has produced results in the laboratory which could not practically be obtained in the field because of economic and time limitations. A typical laboratory test requires approximately three to six months and may cost approximately $25,000. A corresponding field test firing live ammunition would typically require more than a year and may cost in excess of a million dollars. Although the monetary savings may be impressive, it is felt that the reduction of lead time which aids the rapid development of new weapons is the more important result.

Despite the relative advantage of laboratory vs. field fatigue testing, in an absolute sense the tests are still too expensive and time consuming to permit the testing of large sample sizes. Accordingly, modern statistical techniques for analysis of small sample sizes must be used. In the Watervliet laboratory, the method of Pope[2] which

involves the conservative selection of a coefficient of variation based upon past experience with similar components is used. With this method, the safe life of the component can be estimated for any desired value of reliability and statistical confidence.

An unusual test series involving very conservatively designed components which had been assigned life ratings on criteria other than test results was recently completed. The object of the tests was to indicate the feasibility of extending the life rating so that possibly the components would not have to be arbitrarily discarded. Five full scale specimens endured simulated firing cycles several orders of magnitude greater in number than needed with no evidence of fatigue damage. A statistical technique for the estimation of safe life was derived as follows:

a. A type of failure distribution was assumed. Both normal and log-normal distributions were assumed for the analysis. The graphical presentation of the analysis, Fig. 6, illustrates the normal distribution approach.

b. A desired rated life X_r and reliability R are chosen and plotted on the mean $\bar{X}$ vs. standard deviation σ chart. For the normal distribution this is a straight line $\frac{\bar{X} - X_r}{\sigma} = f(R)$, where $f(R)$ is a constant taken from statistical tables of the normal distribution.

c. A coefficient of variation $\sigma/\bar{X}$ is selected as in the method of Pope. This also plots as a straight line in the $\bar{X} - \sigma$ plane.

d. Calculations of the probability of observing the laboratory data for various combinations of population mean and standard deviation are made and these are plotted and cross-plotted to obtain the curved lines in Fig. 7.

It can be shown that the reliability-rated life line divides the $\bar{X} - \sigma$ plane into safe and unsafe zones (the safe zone is shown shaded in Fig. 6). The object of the analysis is to estimate whether or not the population as characterized by its particular, but unknown, $\bar{X}$ and σ, is in the safe or unsafe zone. The coefficient of variation selection dictates that the population must lie below the coefficient of variation line. Therefore, the only remaining unsafe region is the area within triangle OPX_r. But a given constant probability curve rejects as improbable (for example, the .001 curve rejects as 99.9% improbable) the fact that the population lies to the left of this curve. The remaining unsafe region, triangle OPX_r, has thereby been eliminated, and it may be concluded that the component will endure firings with a reliability R

with a certain statistical confidence. The calculations may be repeated for different combinations of rated life X_r, reliability R and statistical confidence and summarized graphically as in Fig. 7 enabling the estimation of a safe rated life for a desired reliability and confidence level.

CONCLUSION: The high energy fatigue machine is capable of performing rapid full scale tests on heavy weapon components with a considerable savings of both time and money compared to field tests. It abets the trend to lighter, high performance, low lead time weapons by providing a rapid means of full scale evaluation.

ACKNOWLEDGEMENT:

The authors are indebted to T. MacLaughlin for his theoretical analysis and to A. Doty and G. Ingram who over the years made many contributions to the development of this test device.

REFERENCES:

1. Den Hartog, J. P., "Mechanical Vibrations", McGraw-Hill Book Company, Inc., 1956 (Fourth Edition), pp. 79-83, 122-126

2. Pope, J. A. (Editor), "Metal Fatigue", Chapman and Hall, Ltd., 1959, pp. 123-130

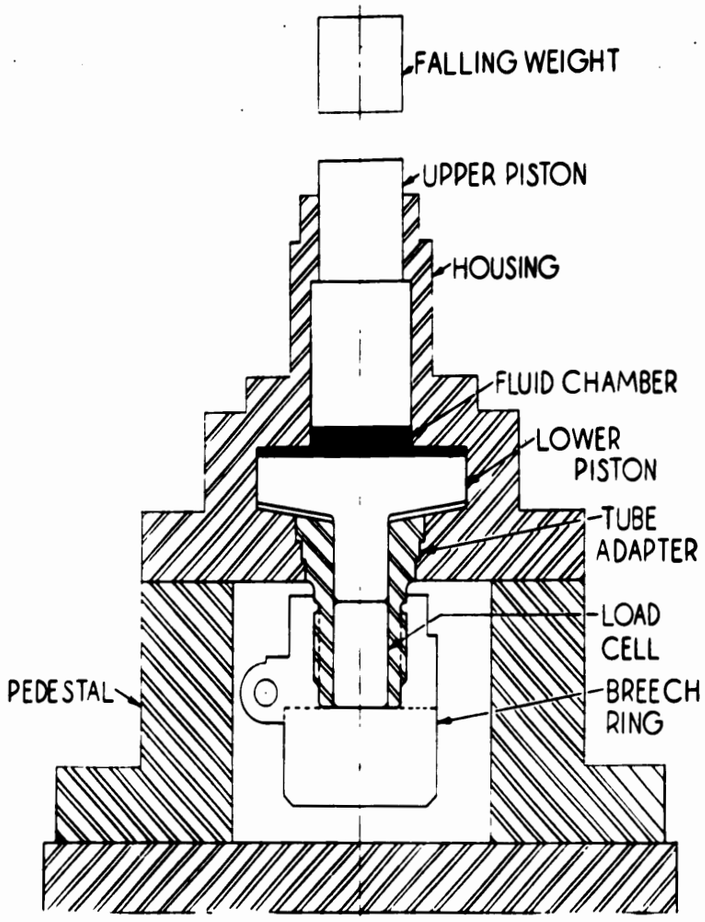


Fig. 1
High Energy Impact Machine -
Breech Ring Fatigue Test

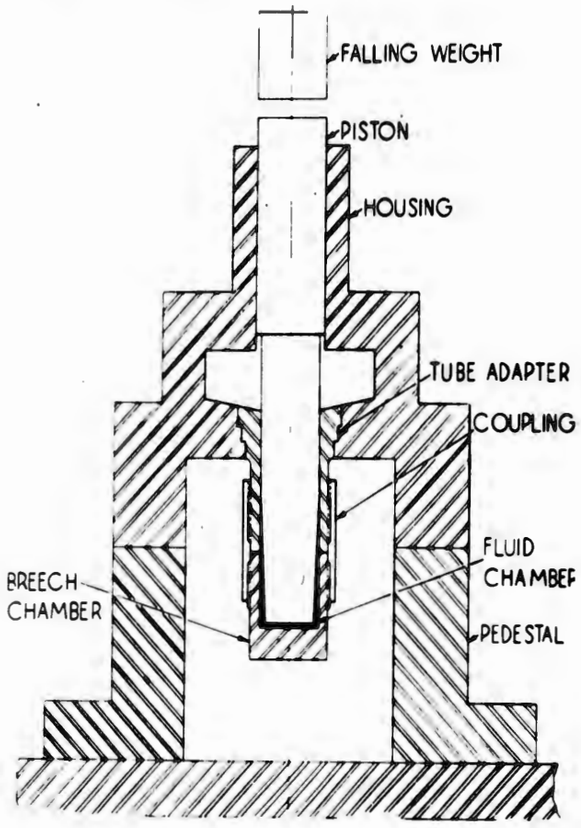


Fig. 2
High Energy Impact Machine - Split Chamber Breech Coupling Fatigue Test

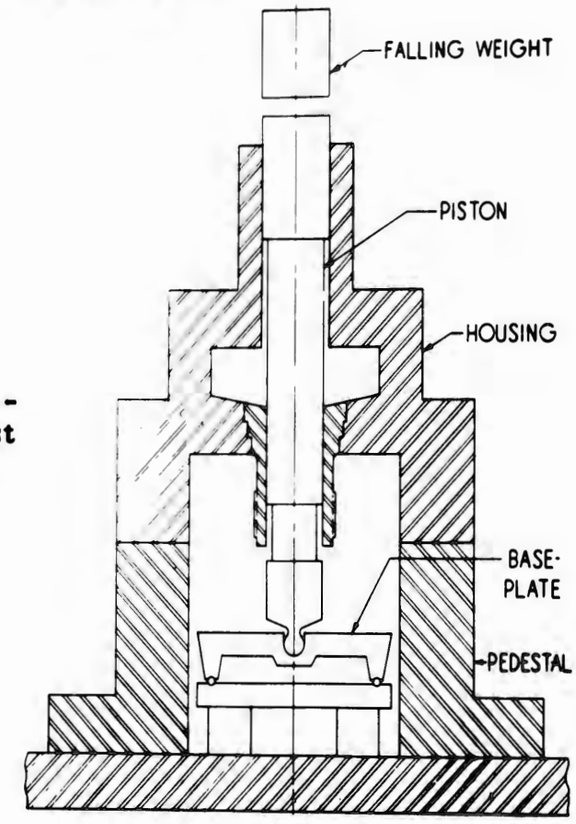


Fig. 3
High Energy Impact Machine - Mortar Baseplate Fatigue Test

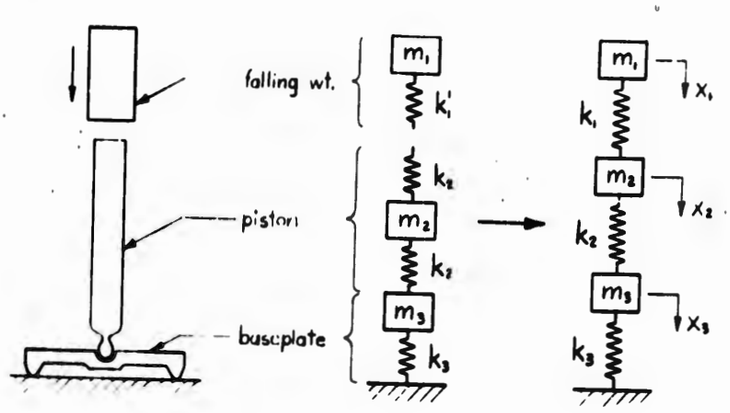


Fig. 4
Spring-Mass Representation -
Mortar Baseplate Fatigue Test

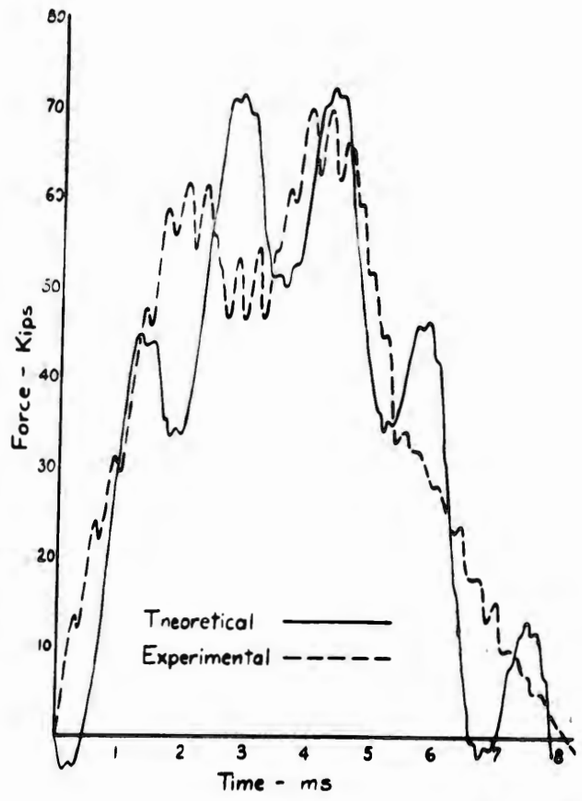


Fig. 5
Load Pulse -
Mortar Baseplate Fatigue Test

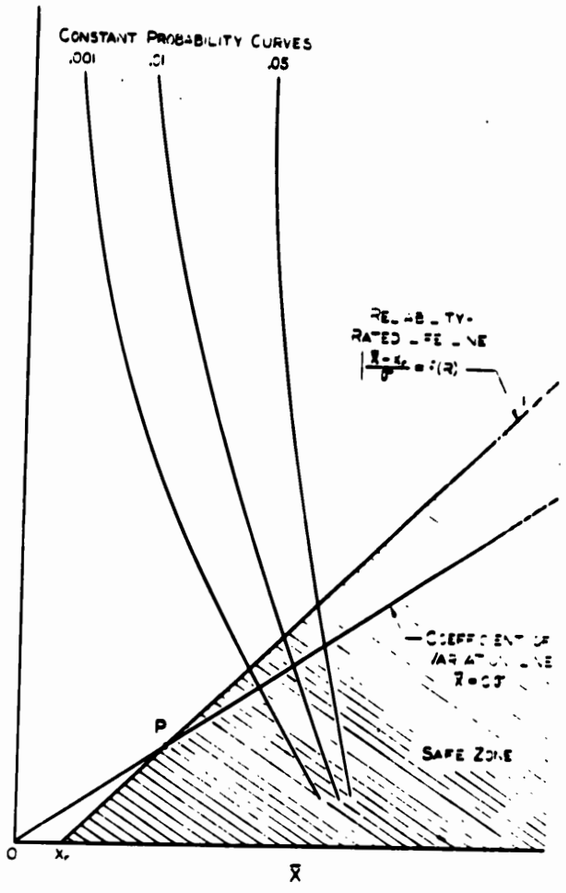


Fig. 6
Graphical Analysis of Fatigue Data Based on Normal Failure Distribution

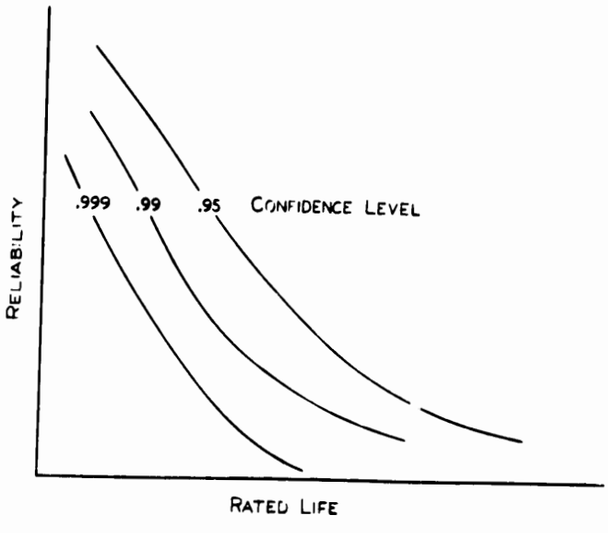


Fig. 7
Hypothetical Reliability vs. Rated Life & Confidence

INTERIM REPORT ON THE DESIGN AND CONSTRUCTION OF THE AOMC 8,000-kw PLASMA FACILITY

J. J. EHRLICH, J. F. PERKINS, C. M. CASON and T. A. BARR, JR.
U.S. ARMY ORDNANCE MISSILE COMMAND
REDSTONE ARSENAL, ALABAMA

The 8,000-kw plasma facility is being constructed for applied high-temperature research. Additionally, the large volume of hot, partially ionized gas which will be produced in the facility will make possible meaningful measurements of electromagnetic energy transmission and generation for a wide band of frequencies, ranging from microwaves through the visible spectrum. In these two areas, this facility will be used in support of the antimissile-missile program.

FACILITY REQUIREMENTS: Of the many methods for producing high altitude, high velocity conditions, only the hyperthermal (arc driven) wind tunnel offers true simulation of all important parameters. (See table I.) The general requirements of such a facility may be described as follows:

1. There must be a mechanism for heating and/or driving a gas stream.
2. The gas stream must have either a very high velocity or a very high temperature or both.
3. The gas stream size, shape, and direction must be controlled.

Ideally, the mechanisms which affect these conditions should be independently adjustable. In practice, the ideal situation is never realized.

The size of a hyperthermal test facility is determined principally by the power available for heating and driving the gas. Therefore, the first considerations were predicated on the availability of six d.c. generators, each being capable of delivering 1,560 kw at 575 volts d.c. Three motor generator sets, using two generators and one drive motor in each set, were assembled and installed along with their transformers, controls and exciters.

To make the facility useful and yet hold it within practical bounds, consideration had to be given to conditions such as vacuum pumping capacity and gas heater pressure, as well as the previously mentioned maximum available power which would partially simulate the reentry environment.

PERFORMANCE REQUIREMENTS: Several trial performance characteristics were determined using Mollier charts for high temperature air (1, 2) and tables of atmospheric parameters (3). For the simulation of reentry, a constant speed of 7,400 m/sec (24,000 ft/sec) for ICBM's, and a constant speed of 4,600 m/sec (15,000 ft/sec) for IRBM's were assumed. These values were determined from minimum energy trajectories and are typical of ICBM's and IRBM's, but do not represent specific characterisitcs of any particular missile. The free stream pressure was taken to be 1.41×10^{-3} mm Hg (296,000-ft altitude) initially, and was assumed to gradually increase to 3.14 mm Hg (122,000-ft altitude). This pressure range, or altitude range is advantageous in two respects:

1. It covers the transition zone between free molecular and viscous fluid flow, and
2. It is a suitable pressure range for the operation of plasma jets.

A flow power of 4,000 kw was assumed as a conservative estimate of the power actually available for driving the gas stream. A linear altitude-time relation was assumed for both ICBM's and IRBM's. Also, it should be noted that because of the difference in reentry angles, both ICBM's and IRBM's take about the same length of time to traverse the same vertical distance. The real difference between the two lies in the velocities which requires that either gas stream diameters and/or flow powers be different for each range missile considered. It appears that changing both variables simultaneously is more acceptable than parametric changes of only one.

The ICBM reentry simulation problem was analyzed first with regard to the limitation imposed by the maximum available power. This limit appears in the following equation:

$$P = \tfrac{1}{2} \rho v^3 A, \qquad \text{(1a}$$

where P is flow power, ρ gas density, v the stream velocity, and A the flow stream cross sectional area. In using equation 1a for the case, P = 4,000 kw, and v = 7,400 m/sec, ρ is the air density at the lowest altitude considered, that is, the largest value of ρ is used. After substituting these values in equation 1a and rearranging terms, the upper limit for the product ρA may be determined:

$$\rho A \leq \frac{2P}{v^3} = 1.97 \times 10^{-5} \text{ kg/m} \qquad \text{(1b}$$

A certain amount of flexibility in the system design now appears, since A is variable within certain limits set by equation la.

The largest practical value of A based on available space considerations is of the order of one square meter. Such a large cross section, however, would seriously limit the running time of the test facility because of the limited pumping capacity.

By using a small amount of scaling, not to be confused with the several orders of magnitude commonly used, an area of 0.1 m^2 or of 0.05 m^2 might be satisfactory. For $A = 0.1\ m^2$, ρ will be equal to or less than 1.97×10^{-4} $\frac{kgm}{m^3}$. From the table of air properties (3) this density gives a lower altitude limit of 210,000 feet. This condition is reached 10 seconds after the initial pressure condition (296,000 ft altitude). From this time on, the flow cross section must decrease as the gas density increases so that the condition $\rho A = 1.97 \times 10^{-5}$ kg/m is satisfied. Figure 1 shows the cross section area as a function of time as determined from equation la, account being taken of the $<$ sign for the limiting value of ρA.

Since the variation of density through the atmosphere is approximately exponential, the power required for simulating reentry will also vary in the same way. Figure 2 shows the flow power as determined from equation 1. The power increases exponentially until the limit of 4,000 watts is reached. It is then held constant for the remainder of the run.

The Mollier diagram of reference 1 may be used to determine the initial state (stagnation conditions in the heater) of the gas at any time during the simulated reentry if it is assumed that the gas goes through an isentropic expansion. The lowest altitude gives the most difficult gas heater condition. The dimensionless entropy at this altitude is 29.0 and the specific enthalpy required to give a velocity of 7,400 m/sec at this altitude is $h = \frac{1}{2}v^2 + h_o = 2.76 \times 10^7$ joules/kg. These conditions require a stagnation pressure of 10^6 atmospheres and a stagnation temperature of 1.4×10^4°K. These requirements preclude the use of simple expansion processes in generating the correct reentry environment for ICBM's.

Similar calculations which have been made for IRBM's show that the cross section area limitation for ρA is reached 15 seconds after the initial altitude. The stagnation condition for this is 10^4 atmospheres at 7,000°K. While these conditions are no more severe than those encountered in the hot-shot type tunnel (4), they are too extreme for continuous operation in the plasma jet facility.

Here again, the use of magnetic drivers or some other scheme to circumvent these high pressures and temperatures is indicated.

CALCULATION OF STAGNATION CONDITIONS: To determine the conditions behind a normal shock on a blunt body placed in such a stream, the conservation equations for mass, momentum, and energy must be solved. For ease of computation, these equations can be put in the following nondimensionalized forms:

$$\frac{\rho_2}{\rho_0} = \frac{\rho_3}{\rho_0}\,\frac{v_3}{v_2} \tag{2}$$

$$\frac{P_2}{P_0} = \frac{1}{RT_0}\,\frac{\rho_2}{\rho_0}\left(\frac{v_3}{v_2} - 1\right) v_2^2 + \frac{P_3}{P_0} \tag{3}$$

$$\frac{h_2}{RT_0} + \frac{1}{RT_0}\,\frac{v_2^2}{2} = \frac{h_3}{RT_0} + \frac{1}{RT_0}\,\frac{v_2^2}{2}\left(\frac{v_3}{v_2}\right)^2 \tag{4}$$

where the subscript o refers to reference conditions, and 2 and 3 refer to upstream and downstream of the normal shock, respectively. A fourth equation, the caloric equation of state, is available only in graphical form as a Mollier diagram for equilibrium air (1, 2). A simple iteration procedure was used to obtain the solution to these equations for the IRBM and ICBM trajectories at their extremes. After determining the conditions behind the normal shock, stagnation conditions were determined from the Mollier chart by assuming an isentropic recompression of the gas. Figure 3 shows the complete gas "path" on the Mollier chart, starting at point 1, the heater stagnation condition, then by isentropic expansion to point 2, which represents the free stream values. The normal shock (nonequilibrium) transition is shown by the broken line from point 2 to point 3. The return to stagnation by isentropic recompression is shown by the short line between points 3 and 4. The conditions at point 4 are those found at the stagnation point on the test object. The difference in the values at the stagnation point and behind the shock is very small; in fact it is possible to assume that the two are equal with very little error. By using a magnetic driver instead of the simple expansion system, the point 1 may be bypassed, and the initial process may start at much lower pressure and temperature and still arrive at point 2 with the necessary specific total enthalpy.

CIRCUMVENTION OF STATIC PRESSURE: Having determined the postshock conditions, it is now possible to determine environments other than exact altitude simulation which might also produce them. In this way it may be possible to eliminate the requirement of very high arc chamber pressures and temperatures. Since the total enthalpy should remain constant, any improvement

will necessarily have to come from moving to the right along the isenthalp which represents the required total enthalpy, or toward lower temperature and pressure. Thus, a lower mach number is required to lessen the entropy increase across the normal shock.

A means of determining suitable combinations of preshock variables is suggested by the relative numerical values of the various terms in the normal shock conservation equations. At these high mach numbers the preshock total pressure is contributed almost entirely by the velocity term $\rho_2 v_2^2$, while the postshock total pressure consists primarily of static pressure. Thus, if ρ_2 and v_2 are varied so as to keep $\rho_2 v_2^2$ at the desired value, the postshock pressure will have, to first order, the proper value for the altitude and flight velocity in question.

Another property of the conservation equations in this application is that the postshock specific enthalpy is very nearly equal to stagnation enthalpy, that is, the $\frac{1}{2} v_3^2$ term is small. The stagnation enthalpy is to be kept at the value corresponding to the desired altitude and velocity, thus the postshock enthalpy has the correct value. Since the two thermodynamic variables h_3 and p_3 have approximately the correct values for simulating the desired flight velocity and altitude, it follows that the other thermodynamic variables also have approximately the correct values. For more accurate evaluation of these variables the iterative process mentioned previously is used, to take into account the small but not completely negligible values of the $\rho_3 v_3^2$, and $\frac{1}{2} v_3^2$ terms.

A relatively small fractional decrease in preshock gas velocity, v_2, results in a very substantial increase in h_2 for fixed stagnation enthalpy. This change more than offsets the required increase in ρ_2, with the result that the preshock entropy, S_2/R, may be increased by several units with only a moderate change in v_2 as illustrated in figures 4 and 5. In these figures gas properties in the test section and in the arc chamber are plotted as a function of the ratio of post shock velocity v'_3 to the velocity v_3 which applies to exact simulation. The decrease in the required arc chamber pressure is very pronounced for the first 10 percent or so change in gas velocity.

Such modifications of test parameters may, of course, require different design of some parts of the test system, such as vacuum chamber volume and nozzle exhaust area. Vacuum chamber requirements depend on mass flow rate, $\rho_3 v_3 A$. The variation of test parameters is to be such that ρ_3 remains the same as for exact simulation. The postshock velocity is determined by the conservation of mass

$$v_3 = \frac{\rho_2 v_2}{\rho_3} = \left[\frac{\rho_2 v_2^2}{\rho_3}\right] \times \frac{1}{v_2} \qquad (5$$

The variation of the test parameters ρ_2 and v_2 is to be such that the bracketed term remains constant. Thus v_2' varies inversely with v_2' where the primed quantities represent modified test parameters. For a fixed nozzle exhaust area, the required vacuum chamber volume is increased; however, power requirements for a given nozzle exhaust area are increased by the same factor. To meet the given limitations on the power input, the nozzle exhaust area must be reduced, with the result that the mass flow rate will be the same for operation with modified test parameters as for exact simulation. Since the normal shock recovery pressure is the same to first order, as for exact simulation, the vacuum chamber requirements are unchanged.

Modifications of test parameters as described above not only reduce arc chamber pressures to tolerable values, but also greatly ease the problem of erosion at the nozzle throat. A 40-percent reduction in preshock gas velocity reduces the nozzle exhaust to throat area ratio from a value of the order of 10^3 or more to a value of the order of 10.

Calculations were made to determine preshock and arc chamber conditions for the IRBM case, assuming that the postshock and stagnation point conditions are approximately equal and assuming a velocity of 3,340 m/sec. and an expansion to an appropriate temperature. Table II compares the values obtained in this manner with those of the exact simulation case. While it was not possible to match exactly the conditions after the shock, they do not differ by more than 8 percent, although the arc chamber temperature and pressure have been greatly decreased. Thus, it appears possible to get meaningful information from a process such as the one described. It should be noted, however, that this analysis is true only for the normal portion of the shock.

A preliminary study of the feasibility of using magnetic drivers indicates that, aside from the practical difficulties in achieving satisfactory operation, the driver sections may need to be quite long for some of the desired test section conditions. Even if a driver could be operated with no entropy increase, (no joule heating) the arc chamber pressure will be rather high for the lower altitudes considered. Thus, an arc chamber entropy of $S/R = 29$ (corresponding to 122,000-ft altitude and no joule heating in the magnetic driver) and temperature of 3,000°K (required to get minimum acceptable gas conductivity) corresponds to an arc chamber pressure of about 100 atmospheres.

POWER SUPPLY (8,000 kw): At this time one major component of the facility has been completed, the d.c. power supply, and another component, the vacuum system, has been designed and is under construction by a contractor. The power supply, rated nominally at 7,200-kw d.c. output, may be overloaded for the short

periods of operation which are anticipated, to give a full 8,000-kw output. The maximum short period output (1 minute or less) is limited by the local substation rather than the motor-generator sets. The long period (more than 2 hours) maximum power is limited by the continuous horsepower rating of the drive motors. This capacity is about 3,000 kw. The exciters are to be used in certain applications requiring fixed point operation or slowly varying operation. However, a complete program control system is being designed for rapid change control over the full range of available power.

VACUUM SYSTEM: The design of the vacuum system was determined by the pressure-mass flow characteristic expected for the facility. The quantity of gas to be pumped at a given pressure may be determined by using the mass flow rate ρvA. By using the altitude (pressure)-time relation for ICBM's and IRBM's, the quantity of gas to be pumped at a given time may be determined. From these relations the pumping requirements may be uniquely founds. The use of pumps alone appeared uneconomical (estimated cost $300,000 to $600,000). The use of a vacuum reservoir was studied and found to be satisfactory, especially since readily available vacuum pumps could be used to pump it down, so a 10,000-cubic foot vacuum tank at 0.5μ Hg of pressure was designed.

TEST CHAMBER: The test chamber design was based on consideration of the flow stream diameter, the anticipated model size, and the boundary layer thickness. The gas, heated in the arc chamber, will be driven by a pressure differential (table II) through the nozzle into the test chamber. The gas stream diameter at the nozzle exit will be approximately 0.3 meter. A model 0.1 meter in diameter can be tested in this stream. Field tests and theoretical calculations indicate that the gas stream behind the model is of interest as far as 30 to 60 body diameters downstream. This requirement sets the minimum length downstream from the model at six meters. The length of the upstream region, the section between the nozzle exit and the model, is not exactly prescribed, but this section must be long enough to prevent interaction between the model and the nozzle. Also, there must be room for a variety of measuring devices which will be used to determine properties in front of the model. The length of section between the nozzle and model was set at three meters to accommodate these factors. In the event larger models, say 0.2 meter diameter, are used, additional distance downstream will be needed since the length of this section is determined by model diameter. The overall length of the test section was set at 15 meters (three 5-meter sections). A large window is to be placed in each side of the test section at the model position. (The optical properties of Lucite have proven satisfactory even for

Schlieren systems, and 2-inch thick Lucite plates will be used in these windows).

The tank diameter, 2 meters, is determined from the expected boundary layer thickness. This thickness was calculated from the equation

$$\delta = 5.2L/\sqrt{R(L)} \qquad \text{(reference 5)}, \qquad (6$$

where δ is the boundary layer thickness, L the characteristic length, in this case the tank length, and R is the Reynolds number based on L and the aerodynamic properties at 91,000 meters altitude. This thickness plus other factors such as the stream diameter give a total diameter near two meters.

REFERENCES

1. Korobkin, I., and Hastings, S. M.: Mollier Charts for Air in Dissociated Equilibrium at Temperatures of 2,000°K to 15,000°K. NAVORD Report 4446, U. S. Naval Ordnance Laboratory, White Oak, Maryland, 1957.
2. Feldman, S.: Hypersonic Gas Dynamic Charts for Equilibrium Air, AVCO Research Report 40, AVCO Research Laboratory, Everett, Massachusetts, 1957.
3. Vaughan, W. W.: Numerical Tables of Atmosphere Parameters Based on the ARDC Model Atmosphere, 1957. Report DA-TN-197, Army Ballistic Missile Agency, Huntsville, Alabama, 1958.
4. Perry, R. W., and MacDermott, W. N.: Development of the Spark Heated, Hypervelocity, Blowdown Tunnel-Hotshot. AEDC-TR-58-6, ASTIA Doc. AD-157138, June 1958.
5. Enkenhus, K. R.: The Design, Instrumentation and Operation of the UTIA Low Density Wind Tunnel. UTIA Report 44, AD-148-061, June 1957.

TABLE I
COMPARISON OF TYPES OF FACILITIES FOR REENTRY SIMULATION
(Satisfactory operation is shown by xx)

Facility	Enthalpy	Velocity	Altitude	Test Time	Model
Shock tube	xx	xx	xx	No	No
Hyperthermal tunnel	xx	xx	xx	xx	xx
Wind tunnel	No	No	xx	xx	xx
Rocket exhaust	xx	No	No	xx	xx
Free flight ranges	xx	xx	xx	No	No
Hot shot	xx	xx	xx	No	xx

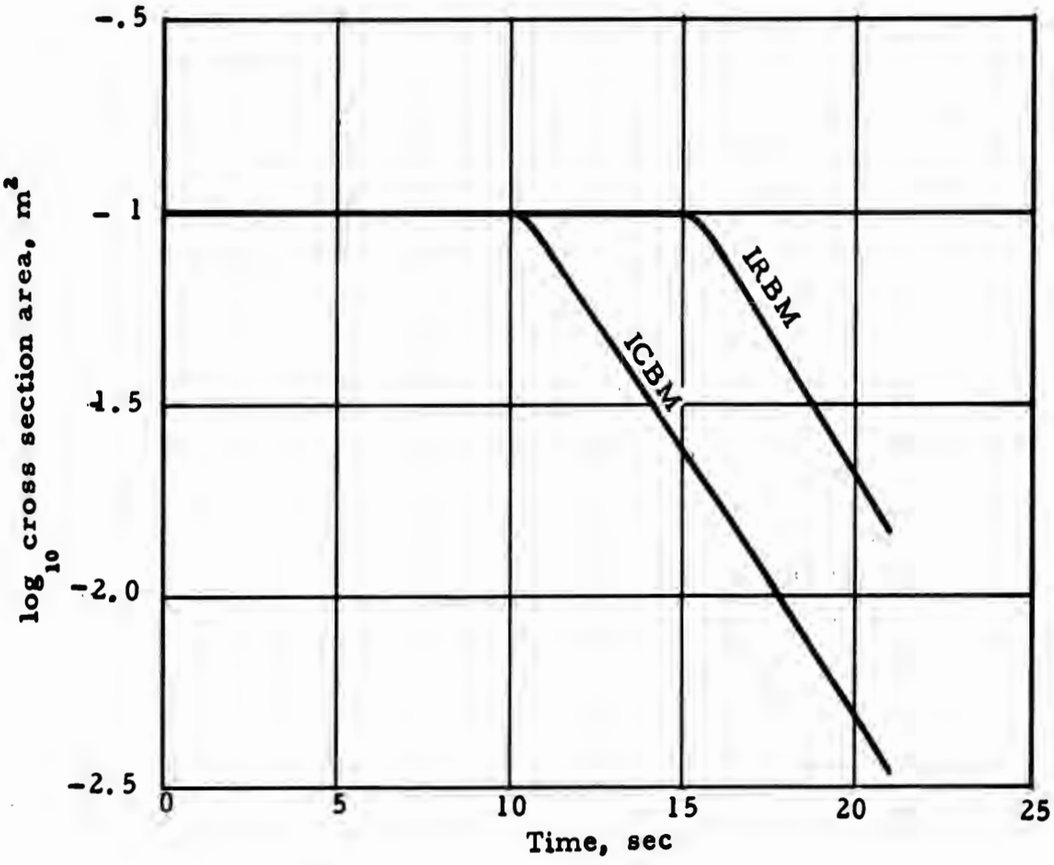


Figure 1 - Flow stream cross section area as a function of time. (ref. eq. 1a)

TABLE II

COMPARISON OF GAS PROPERTIES FOR THE REAL AND SIMULATED REENTRY ENVIRONMENTS FOR A HYPOTHETICAL IRBM

Real environment			Simulated environment		
Free stream conditions					
Gas property	Altitude = 37,500 m	Altitude = 91,100 m	Gas property	Altitude = 37,500 m	Altitude = 91,100 m
S_2/R	29.0	38.3	S_2'/R	37.3	45.8
h_2/RT_o	3.13	2.51	h_2'/RT_o	65.0	65.0
T_2 (°K)	254	197	T_2' (°K)	3,100	2,350
P_2/P_o	4.11×10^{-3}	1.85×10^{-6}	P_2'/P_o	.122	5.40×10^{-5}
ρ_2/ρ_o	4.67×10^{-3}	2.71×10^{-6}	ρ_2'/ρ_o	8.80×10^{-3}	5.20×10^{-6}
Postshock conditions					
S_3/R	39.6	48.8	S_3'/R	39.4	48.6
h_3/RT_o	137	136	h_3'/RT_o	135	134
T_3 (°K)	5,200	4,200	T_3' (°K)	5,100	4,100
P_3/P_o	1.14	6.76×10^{-4}	P_3'/P_o	1.16	6.70×10^{-4}
ρ_3/ρ_o	.0501	3.55×10^{-5}	ρ_3'/ρ_o	.0525	3.80×10^{-5}
v_3/v_2	.0933	.0763	v_3'/v_2'	.168	.144
Arc chamber conditions					
S_1/R	29.0	38.3	S_1'/R	37.3	45.8
h_1/RT_o	137	136	h_1'/RT_o	135	134
T_1 (°K)	7,000	5,300	T_1' (°K)	5,400	4,400
P_1/P_o	10^4	3.16	P_1'/P_o	6:67	6.31^{-3}
ρ_1/ρ_o	400	.132	ρ_1'/ρ_o	.282	3.16×10^{-4}

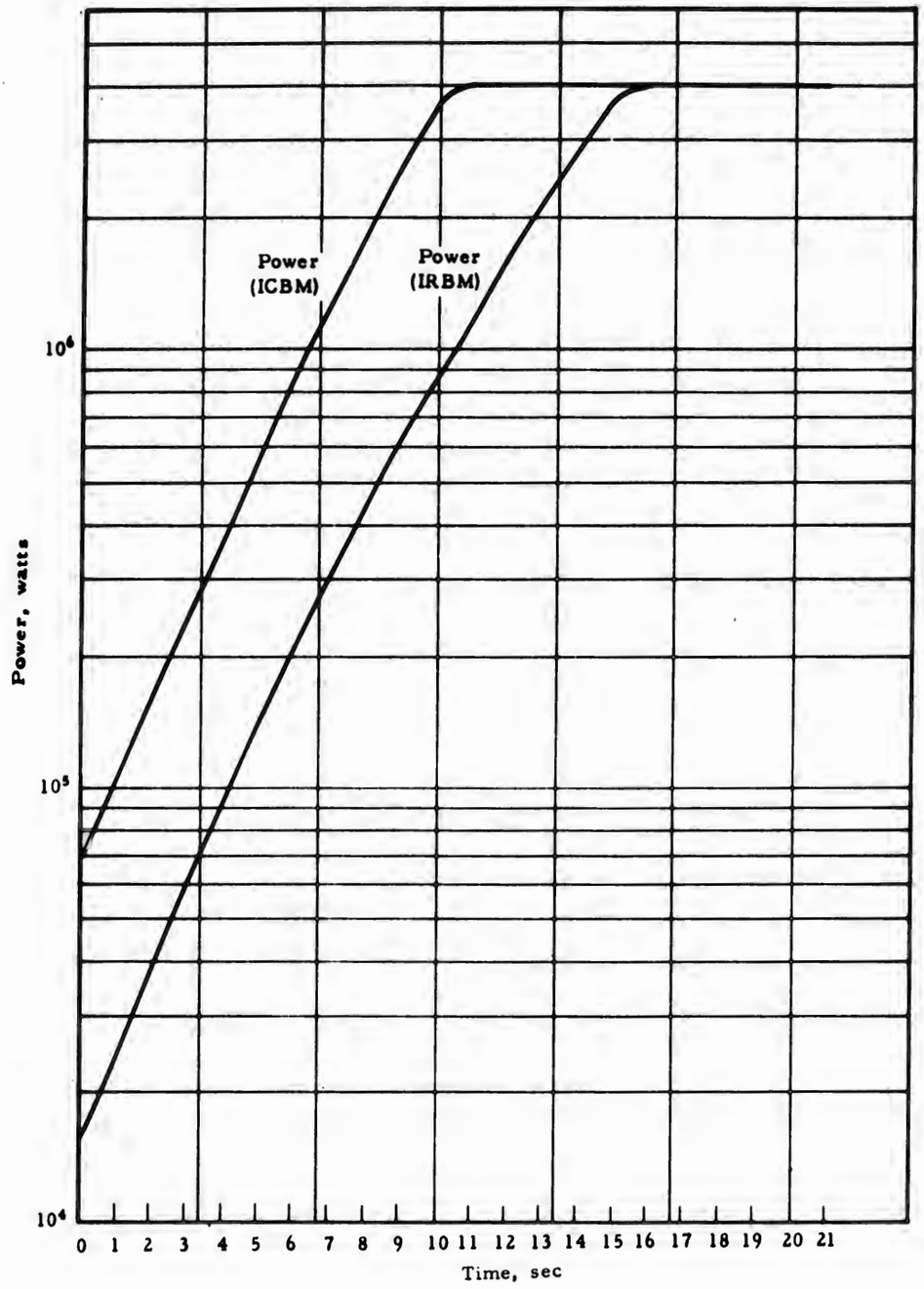


Figure 2 - Power versus time history for typical reentry trajectory (Limited by 4, 000 kw)

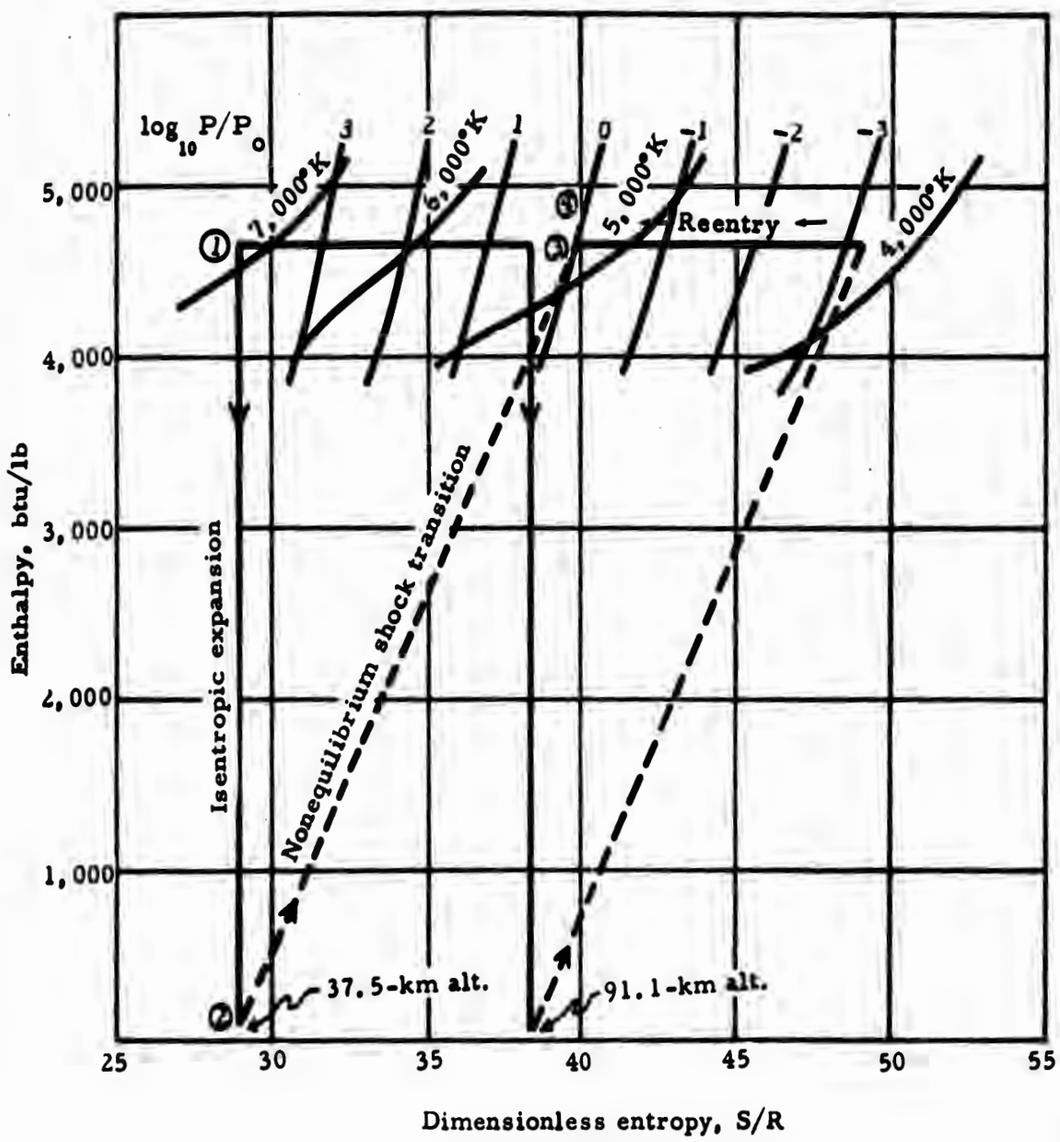


Figure 3. - Gas "path" on Mollier chart for IRBM reentry.

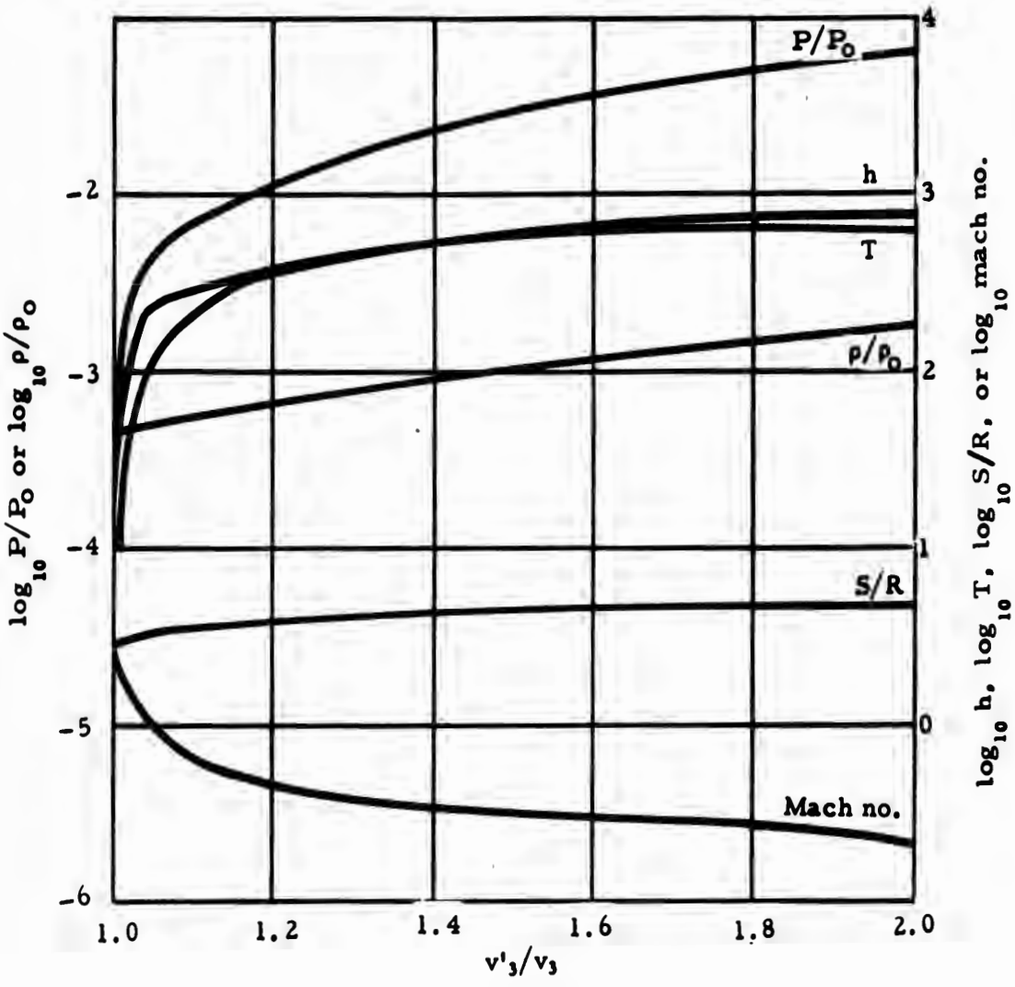


Figure 4. - Gas properties in test section as a function of v_3'/v_3 (ICBM, 37.5km)

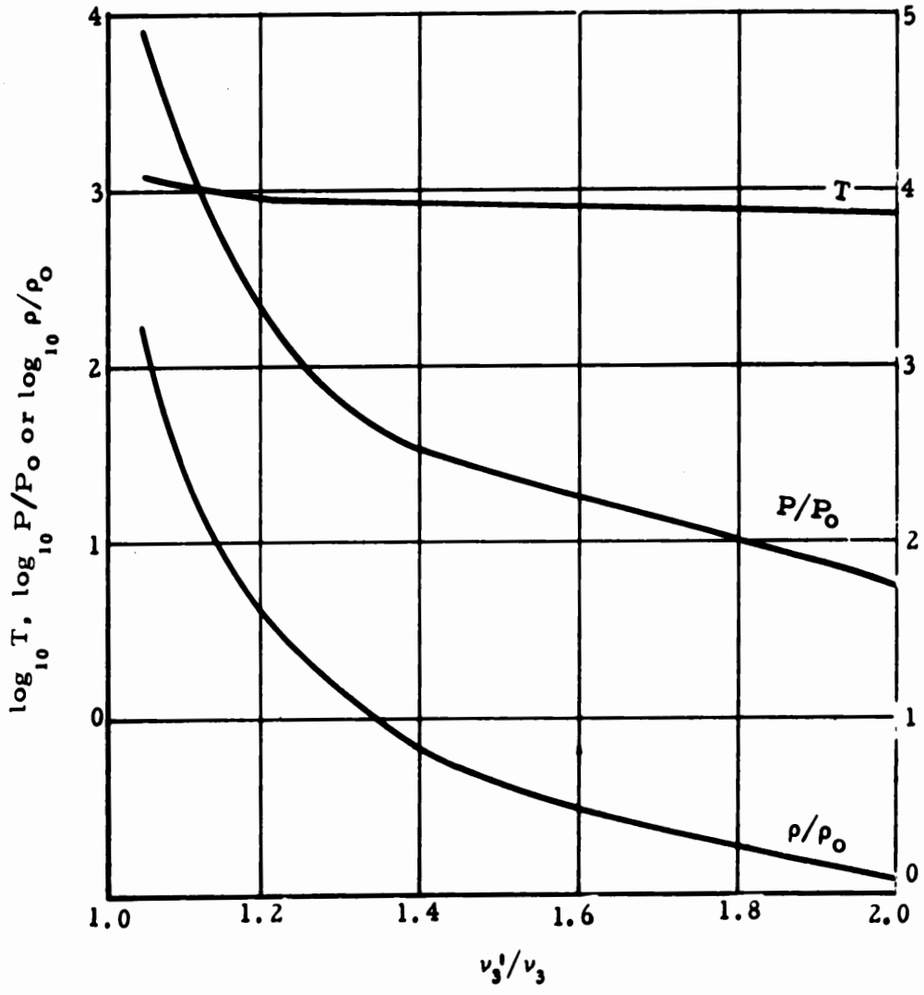


Figure 5. - Gas properties in the arc chamber as a function of ν_3'/ν_3 (ICBM, 37.5 km)

LIVE AND KILLED TULAREMIA VACCINES: EVALUATION IN ANIMALS AND MAN

H. T. EIGELSBACH, COL. W. D. TIGERTT, MC, S. SASLAW, AND F. R. McCRUMB, JR.
U. S. ARMY BIOLOGICAL LABORATORIES
FORT DETRICK, MARYLAND

Human tularemia is an acute incapacitating febrile illness resulting from infection with the bacterial species Pasteurella tularensis. In America the usual initial manifestation is an ulcer of the skin, which is followed by systemic disease lasting 2 to 4 weeks with disability for 8 to 12 weeks. Before the advent of antibiotics, if pneumonia were a part of this systemic disease, or if the infection had been acquired by inhalation, about half of such patients died (1).

Numerous infections that have occurred in laboratory workers as the result of inhalation of the organisms indicate a potential of this organism for use as a biological weapon. Various killed vaccines against tularemia have been used for years but the continued occurrence of infections in laboratory workers, despite repeated doses of vaccine, clearly indicated that protection, if any, was of limited degree.

Several years ago a program was designed to seek specific information on this problem through the use of volunteers. It was approved by the Department of the Army and the studies eventually led to the production and use of a live vaccine for the prevention of tularemia and to the investigation of various methods of administering it. This report considers these findings in man in sequential steps, together with the extensive corollary studies in animals.

CHARACTERISTICS OF THE VIRULENT STRAIN OF P. TULARENSIS

To examine the degree of protection afforded by the various vaccine regimens, animals and men were inoculated with a virulent strain of P. tularensis, designated SCHU S4. Each time it was employed the inoculum was checked to insure that maximum

virulence had been maintained, the requirement being that, under suitable conditions, 10 cells would produce death when administered to nonvaccinated monkeys, guinea pigs, and mice.

EVALUATION OF A KILLED P. TULARENSIS VACCINE

A killed vaccine, developed originally by Foshay (2), was used and subjects were examined to ascertain the degree of protection afforded thereby when the virulent strain was subsequently administered by inoculation directly into the skin (intracutaneous) or by inhalation.

Intracutaneous inoculation. Preparatory to the initial studies in volunteers, a series of investigations was performed using more than 200 monkeys (Macaca mulatta) to obtain data on the intracutaneous dose of P. tularensis required to infect, the protection afforded by repeated vaccination, and the effectiveness of streptomycin or chlortetracycline therapy. It was found that an intracutaneous dose of 10 organisms would infect all animals routinely and that the majority died within 30 days. Detailed clinical and laboratory study, including histological examination of the tissues of all animals that died or were sacrificed, revealed that the vaccinees were as readily infected with 10 organisms as were the controls. However, the vaccinees developed a modified disease with a slightly lower mortality. When 10^3 organisms were used for the intracutaneous challenge dose, there was no appreciable difference between vaccinees and controls. Streptomycin was highly effective and superior to chlortetracycline for treatment (Table 1). If treatment were begun approximately 28 hours after animals

TABLE 1. ANTIBIOTIC THERAPY OF THE MONKEY INFECTED INTRACUTANEOUSLY WITH PASTEURELLA TULARENSIS SCHU S4

ANTIBIOTIC	THERAPY SCHEDULE	MORTALITY WITHIN 30 DAYS, % *	ANIMALS HARBORING P. TULARENSIS OR SHOWING PATHOLOGY WHEN SACRIFICED, %	
			CULTURE ISOLATED	NECROTIC LESIONS
NONE		100		
STREPTOMYCIN	50 MG Q6H (IM)	0	0	0
STREPTOMYCIN	100 MG Q12H (IM)	0	0	0
CHLORTETRACYCLINE	125 MG Q6H (ORAL)	0	70	60

* 10 ANIMALS PER GROUP INFECTED WITH 10^3 ORGANISMS.

became febrile, there were no deaths. Data provided by these and other studies, and accumulated information from the naturally occurring disease and from laboratory infections of man, formed a broad base of preliminary experience for study of tularemia in volunteers.

In the initial experiment, volunteers were inoculated intracutaneously with approximately 10 virulent organisms. Local skin lesions developed in all of the 12 nonvaccinated men and in 17 of 19 previously vaccinated. Constitutional signs of infection were observed in 11 of the 12 controls and in 5 of 19 vaccinees. Only 2 of the vaccinees were as sick as the nonvaccinated controls. Under the conditions of the study, streptomycin therapy (2 g daily for 10 days) was instituted when signs of systemic infection became apparent; the clinical response to therapy was dramatic (3). It was concluded that vaccination resulted in increased ability to combat tularemia when the infection was initiated by a small intracutaneous challenge.

Inhalation. Studies with the monkey on tularemia resulting from inhalation of a homogeneous small-particle aerosol revealed that the minimal infecting dose for this animal was less than 10 organisms and that 25 organisms resulted in deaths in about half of the animals exposed. Streptomycin therapy was highly effective; no disease residua were observed upon thorough gross and microscopic examination of organs and tissues of sacrificed animals.

Subsequent to these animal experiments, vaccinated and nonvaccinated volunteers inhaled 15 to 48 organisms of virulent P. tularensis. The results are shown in Table 2. Clinical disease

TABLE 2. RESPIRATORY CHALLENGE OF VOLUNTEERS ADMINISTERED THE FOSHAY PHENOLIZED TULAREMIA VACCINE

TEST NO.	RESPIRATORY DOSE (NO. ORGANISMS)	VACCINATED (ILL/CHALLENGED)	NONVACCINATED (ILL/CHALLENGED)
1	15		2/2
2	17	1/2	2/2
3	22	1/4	1/2
4	27	3/4	2/2
5	48	3/4	1/2
TOTAL		8/14	6/8*

* BASED ON TESTS 2-5

occurred in vaccinated and nonvaccinated volunteers. The clinical syndrome and the incubation period were comparable in both groups. Early use of streptomycin effected prompt remission of disease symptomatology (4).

DEVELOPMENT AND EVALUATION OF A LIVE TULAREMIA VACCINE

Development of a live vaccine. The studies on the Foshay killed vaccine, and the knowledge that 98 persons at Fort Detrick had been hospitalized with tularemia even though they had received the Foshay vaccine (5,6), stimulated efforts to provide a more effective vaccine. Previous findings in our laboratories and other published reports (7-10) indicated that a live tularemia vaccine might be the solution. A USSR live tularemia vaccine culture was found to be comprised of a mixture of variants differing in morphology and in virulence for the white mouse. Since previous studies had indicated that P. tularensis organisms incapable of killing mice regularly were also incapable of inducing protection when inoculated into guinea pigs or monkeys, the next step was to select a stable variant and increase the virulence by passing it serially through white mice. From the blood of a moribund mouse a culture of suitable virulence was obtained and characterized. When handled under appropriate cultural conditions, it was shown to be stable (11). This derivative "live vaccine strain" (LVS) initially was administered to animals by intradermal inoculation in a fashion comparable to that generally used for small-pox vaccination. This produced a lesion at the site of administration but no evidence of systemic disease. The animals subsequently were found to have developed considerable protection. Inhalation of 75 cells of the virulent P. tularensis strain resulted in disease in all control and all vaccinated animals. However, most of the controls died, but all vaccinees survived. At a respiratory dose of 1×10^4 to 5×10^4 organisms (Table 3),

TABLE 3. SUMMARY DATA ON RESPIRATORY CHALLENGE OF MONKEYS ADMINISTERED LIVE TULAREMIA VACCINE

GROUP*	INFECTED, %	SURVIVAL, %
NONVACCINATED	100	0
VACCINATED	100	62

* 13 ANIMALS PER GROUP CHALLENGED WITH 1-5 × 10^4 ORGANISMS.

more than 50% of the vaccinated monkeys survived, but all nonvaccinated controls died (12). These animal data were sufficiently encouraging to warrant an extension of the study to man.

Vaccination of man. After a preliminary trial in 2 volunteers, the rehydrated lyophilized live vaccine was administered by multiple puncture to 22 young adult males without history of tularemic infection and who had not previously received a killed vaccine. A benign lesion was produced at the site of vaccination; only a pink scar remained 1 month after inoculation. The appearance of the lesion at 4 and 8 days after vaccination is shown in Figure 1. No significant systemic reactions occurred. There was slight enlargement of the regional lymph nodes and agglutinin response was good.

Another group of volunteers was vaccinated in the same manner and the vaccine again proved innocuous. Two to 4 months later these and nonvaccinated volunteers were exposed to respiratory challenge with virulent P. tularensis (4). Each man inhaled 12 to 47 organisms contained in aerosols comparable in physical characteristics to those described previously. Eight of 10 nonvaccinated controls developed overt disease; only 3 of 18 vaccinees showed any evidence of systemic infection. All were promptly treated with streptomycin (Table 4). One of the 3 vaccinees receiving antibiotic therapy was mildly ill and was treated only as a precautionary measure.

In a subsequent investigation, approximately 1 year after vaccination, volunteers inhaled 200 to 20,000 organisms (13). All nonvaccinated volunteers developed a severe illness and required antibiotic therapy. However, a majority of the vaccinees exposed to 200 or 2,000 organisms escaped clinical illness. Results are summarized in Figure 2. Patients with hands on head represent vaccinees with transient illness not requiring therapy. Seated patients are those with greatly modified disease not requiring therapy. The vaccinee depicted in bed had unmodified disease comparable to that in the nonvaccinated controls except that he responded to a single short course of chloramphenicol without subsequent relapse, a feature observed in all of the nonvaccinated participants in this study.

Few vaccinees have ever been subjected to tests this severe, and it is evident that the LVS is a particularly potent immunizing agent for man. Only a single inoculation is required, in contrast to the several-dose schedule used for the killed product.

Routine use of LVS in Army personnel. The foregoing findings indicated that live tularemia vaccine was innocuous for man and afforded substantial protection against respiratory challenge. In June 1959 use of the live vaccine was begun for Fort Detrick personnel who were considered subject to occupational risk of laboratory

infection. Approximately 1,800 persons at Fort Detrick and elsewhere have now been given this LVS product. No lost time has resulted from the vaccination procedure, and laboratory-acquired tularemia requiring hospitalization has disappeared, although prior to the introduction of live vaccine, 10 to 12 individuals per year at Fort Detrick alone became clinically ill with tularemia. Continued study of personnel administered live vaccine indicates that the agglutinin titer and the intensity of the skin test reaction decrease substantially between 1 and 2 years after vaccination (14). Should experience indicate that revaccination is desirable, preliminary investigation has revealed that revaccination is accompanied by even less reaction than that produced by primary vaccination.

Further studies are suggested by the observation that animals that are given LVS and then survive an exposure to virulent P. tularensis are significantly more resistant if again exposed to the virulent organism. Vaccination with LVS followed by revaccination with a less attenuated strain is now underway in man.

RESPIRATORY VACCINATION AGAINST TULAREMIA

After preliminary studies had indicated that the LVS could be employed by inhalation, the clinical and immunological responses of animals vaccinated by this route were compared with those observed when LVS was given by intradermal inoculation.

Monkeys vaccinated aerogenically were as readily infected by a respiratory challenge of 750 organisms as those vaccinated intracutaneously but exhibited a milder disease with lower mortality.

Vaccination of the guinea pig prolonged survival after challenge but eventually most of the animals died. Nevertheless, in this animal aerogenic vaccination was superior to subcutaneous vaccination when tested either by respiratory challenge or by subcutaneous inoculation of 10^3 organisms (Table 5).

Study of the host-parasite relationship in monkeys. To elucidate the role of the route of vaccination on the host-parasite relationship, 10^5 cells of live tularemia vaccine were administered intracutaneously or by inhalation to groups of 24 monkeys. Animals were killed at intervals of 1 hour to 90 days thereafter for the gross and microscopic observations and for determination of the number of bacteria present. For comparison, the same studies were made with monkeys challenged via the respiratory route with virulent strain SCHU S4.

Strain SCHU S4 produced a fulminant, acute, necrotizing reaction that was in striking contrast to the subacute, non-necroti-

zing, and limited inflammatory reaction evoked by the live vaccine. Growth of SCHU S4 in the lungs of monkeys exposed aerogenically was characterized by rapid and persistent multiplication; within 2 days the viable population approached 10^9 organisms. The major sites of tissue reaction to the virulent strain were the bronchioles and adjacent pulmonic parenchyma, the intrapulmonic lymphoid tissues, the tracheobronchial lymph nodes, and to a lesser extent the liver and spleen. Death occurred within 3 days after exposure before the more granulomatous lesions of classical tularemia could become manifest. The viable count of SCHU S4 in the lungs of each of the 4 animals examined ranged from 3×10^8 to 5×10^9 organisms.

After aerogenic administration of the live vaccine, gross changes in the monkeys were limited to the lungs and tracheobronchial lymph nodes. Minute areas of erythematous pulmonic parenchyma were seen on the 7th day but had disappeared by the 14th day. The tracheobronchial nodes became hyperplastic, reaching a maximal dimension of 1 cm by the 10th day. Petechiae were seen but no necrosis. Lymphadenitis was not evident on the 28th day. Fluorescent antibody studies showed that P. tularensis organisms had impinged along the walls of the respiratory bronchioles and had initiated small foci of subacute, non-necrotizing inflammation during the 1st week. These were absent by the 14th day and on the 20th day no residua were seen except for a moderate increase in the plasmacytic population of the pulmonic lymphoid tissue.

After intracutaneous inoculation with the live vaccine, a gradually enlarging area of erythema and induration developed at the site and small central ulcers appeared during the latter part of the 2nd week. These re-epithelized, and inflammation subsided during the 3rd and 4th weeks. The axillary and deep cervical lymph nodes were moderately enlarged by the end of the 2nd week but did not become necrotic. Histological changes were characteristic of a very mild subacute inflammatory reaction without granuloma formation. Healing with minimal scarring ensued. In the spleen the changes were very subtle, and no miliary granuloma were seen in any organs.

The rate of growth of LVS at the site of intracutaneous vaccination and of LVS or SCHU S4 in the lungs of monkeys exposed aerogenically is shown in Figure 3. Within 1 hour after intradermal vaccination with 10^5 LVS, the mean viable population recovered at the site of inoculation was 5×10^2 organisms. At this time LVS was recovered from the regional lymph nodes. Evidence of multiplication at the site of dermal vaccination was obtained within 6 to 12 hours. Maximal viable populations of approximately 10^6 LVS at the inoculation site or in the lungs occurred within 3 days after dermal or aerogenic vaccination. LVS was cultured from either the dermal site or lung of one or

both vaccinees sacrificed at each interval through the 14th day but not on the 28th or 90th day. Twenty-eight days after vaccination, LVS could be isolated only from tracheobronchial lymph nodes of the animals vaccinated aerogenically and from the tracheobronchial and axillary lymph nodes of animals vaccinated intracutaneously. LVS was not isolated from tissues cultured on the 90th day.

As another method of comparing the 2 routes of vaccination, serological studies were used. The agglutinin response of monkeys after dermal (closed circles) or aerogenic vaccination (open circles) is shown in Figure 4. Agglutinins were first observed in the sera of both groups of animals 7 days after vaccination. Peak mean agglutinin titers, 1:1,720 for animals vaccinated dermally and 1:3,520 for animals vaccinated aerogenically, were reached within 21 to 28 days and gradually declined thereafter. The mean agglutinin titer was consistently higher for animals vaccinated aerogenically.

The mild tissue changes without granuloma formation, rapid resolution of the inflammatory response, and complete clearance of LVS within 90 days after intracutaneous or aerogenic vaccination of the monkey with 10^5 viable organisms depict the benign course of LVS infection in the monkey. In relation to the study of LVS in man these findings are of particular importance because the monkey is less capable of dealing with tularemia than is man. Thus LVS could be expected to produce no particular serious effect in man.

Respiratory vaccination of man. Thirty volunteers were exposed to aerosolized live tularemia vaccine to determine the clinical reaction and serological response. Inhaled doses ranged from 1,500 to 30,000 organisms. There was no clinically discernible reaction. The agglutinin response was as good as that observed in individuals vaccinated dermally.

In a second study, 24 volunteers inhaled 200 to 1,200 organisms. Again no clinically discernible reaction was observed and the serological data indicated that man was unable to discriminate between doses in this range; an agglutinin response was induced in 25 to 50% of the volunteers irrespective of dose. However, in general, agglutinin titers were lower than in the previous study.

A potential for mass aerogenic vaccination of man with live vaccine is indicated by these studies as well as by the reports of Aleksandrov and his colleagues in the USSR (15,16). However, this possibility must be tempered by the fact that the studies reported here were done on carefully selected healthy young male adults. Additional studies are indicated. In the near future, respiratory challenge of man with virulent organisms will be used to compare the effectiveness of intradermal and aerogenic vaccination.

For troop training and field tests the aerosolized live tularemia vaccine has potential as a simulant for biological weapons. Troops can be exposed without significant risk. A symptomless infection would develop which could be proved serologically or by skin test.

CONCLUSIONS

The extensive animal studies conducted as the essential preliminary to any investigation involving volunteers show that, for the disease tularemia, a suitable laboratory model is provided by monkeys (Macaca mulatta), provided that it be appreciated that these animals are less capable of dealing with this disease than is man and that they are less able to respond to immunizing procedures.

Clinical disease regularly results if man or monkey is exposed to 10 or more organisms of Pasteurella tularensis (SCHU S4) by either the intradermal route or by inhalation; the monkey dies as a result; man, even without treatment, may survive.

Vaccination of man or monkey with multiple doses of a killed vaccine against this disease affords little protection when measured by the subsequent inhalation of virulent organisms.

Both man and monkey have an extremely bland reaction when inoculated with the live vaccine developed as described in this report.

The administration of this live vaccine by inoculation into the skin produces modest protection in monkeys against the subsequent inhalation of large numbers of virulent P. tularensis organisms, but affords a remarkable protection to men vaccinated in the same manner, as demonstrated by these studies. This is the first bacterial vaccine to be so validated.

It has been shown, further, that this live vaccine can be administered to man and to monkeys by inhalation and that this procedure is without untoward effects. Animals so vaccinated have considerable protection when subjected to the inhalation of virulent P. tularensis organisms.

Although it must be shown by actual test to be true, it can be predicted that men vaccinated by inhalation with the viable vaccine will be found to be solidly protected against tularemia.

ACKNOWLEDGEMENT

These investigations represent the combined and coordinated effort of a large group of senior advisory and participating personnel whose individual support has been and will be acknowledged in detail elsewhere.

LITERATURE CITED

1. Stuart, B.M. and Pullen, R.L., Am. J. Med. Sci., 1945, v210, 223.

2. Foshay, L., Hesselbrock, W.H., Wittenberg, H.J., and Rodenberg, A.H., Am. J. Pub. Health, 1942, v32, 1132.

3. Saslaw, S., Eigelsbach, H.T., Wilson, H.E., Prior, J.A., and Carhart, S., Arch. Int. Med., 1961, v107, 121.

4. Saslaw, S., Eigelsbach, H.T., Wilson, H.E., Prior, J.A., and Carhart, S., Ibid., 1961, v107, 134.

5. Van Metre, T.E., Jr. and Kadull, P.J., Ann. Int. Med., 1959, v50, 621.

6. Overholt, E.L. and Tigertt, W.D., Kadull, P.J. and Ward, M.K., Am. J. Med., 1961, v30, 785.

7. Eigelsbach, H.T., Braun, W., and Herring, R.D., J. Infect. Dis., 1952, v91, 86.

8. Olsuf'ev, N.G. (Monograph Editor), Moskva, Izdatel'stvo Akademii Med. Nauk SSR, 1953.

9. Sil'chenko, U.S., Zhur, Mikrobiol. Epidemiol. Immunobiol., 1955, No. 10, 83.

10. Eigelsbach, H.T., Downs, C.M., and Herring, R.D., Bact. Proc., 1958, v1958, 74.

11. Eigelsbach, H.T. and Downs, C.M., J. Immunol., 1961, v87, 415.

12. Eigelsbach, H.T., Tulis, J.J., Overholt, E.L., and Gochenour, W.S., Jr., Bact. Proc., 1959, v1959, 87.

13. McCrumb, F.R., Jr., 1961, Bact. Revs., 1961, v25, 262.

14. Kadull, P.J., Personal Communication, 1962.

15. Aleksandrov, N.I., Gefen, N.Ye., Garin, N.S., Gapochko, K.G., Daal-Berg, I.I., Sergeyev, V.M., Voyenno-Med. Zhur., 1959, No. 12, 34.

16. Aleksandrov, N.I., Gefen, N.Ye., Garin, N.S., Gapochko, K.G., Sergeyev, V.M., Smirnov, M.S., Tamarin, A.L., and Schliakhov, E.N., Ibid., 1959, No. 8, 27.

TABLE 4. RESPIRATORY CHALLENGE OF VOLUNTEERS ADMINISTERED LIVE TULAREMIA VACCINE

TEST NO.	RESPIRATORY DOSE (NO. ORGANISMS)	VACCINATED (ILL/CHALLENGED)	NONVACCINATED (ILL/CHALLENGED)
1	12	0/2	1/2
2	48	1/4	2/2
3	25	1/4	2/2
4	11	0/4	1/2
5	47	1/4	2/2
TOTAL		3/18	8/10

TABLE 5. EFFECT OF ROUTE OF VACCINATION ON RESISTANCE OF THE GUINEA PIG TO RESPIRATORY CHALLENGE

ROUTE OF VACCINATION	LIVE VACCINE DOSE	SURVIVAL ON INDICATED DAY AFTER RESPIRATORY CHALLENGE, %*	
		7	15
RESPIRATORY	10^5	100	60
SUBCUTANEOUS	10^5	30	5
RESPIRATORY	10^3	75	35
SUBCUTANEOUS	10^3	45	0
NONVACCINATED CONTROLS		0	0

* 20 ANIMALS PER GROUP CHALLENGED WITH 10^3 CELLS.

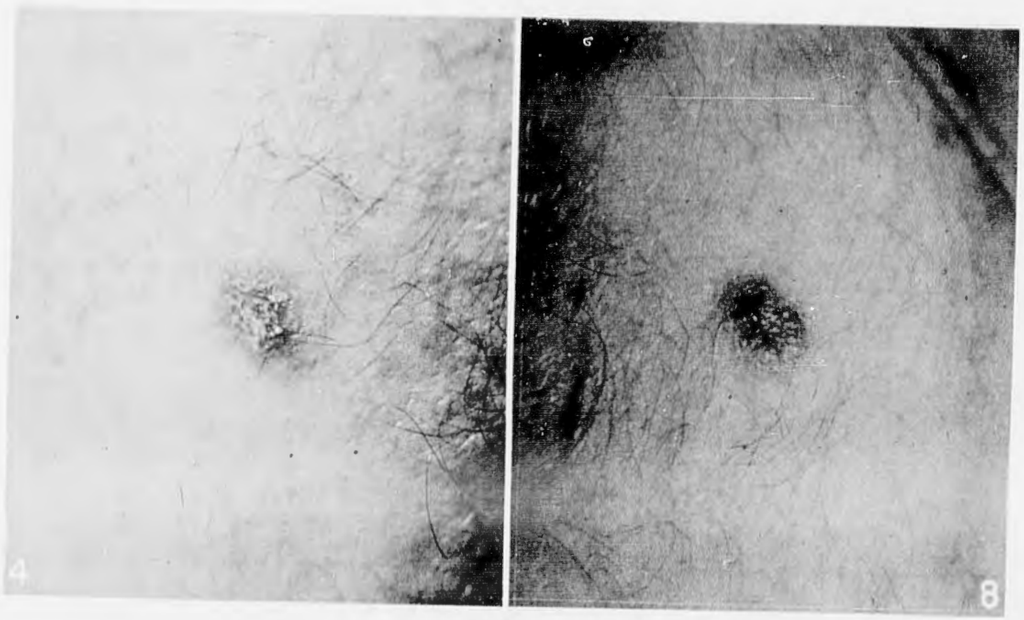


FIGURE 1. INOCULATION SITE 4 AND 8 DAYS AFTER ADMINISTRATION OF LIVE TULAREMIA VACCINE TO MAN BY ACUPUNCTURE.

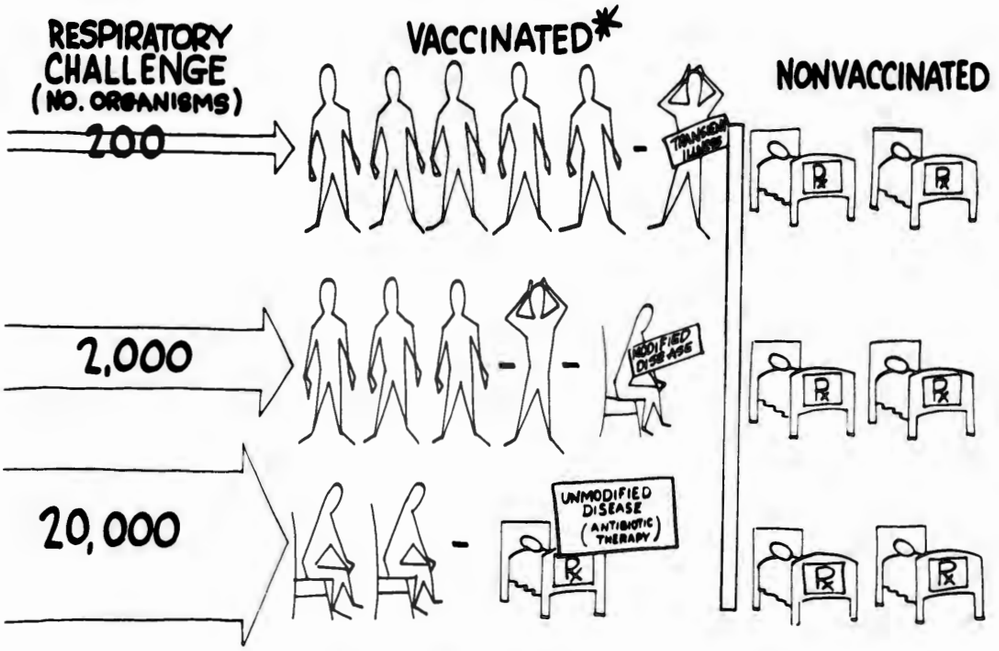


FIGURE 2. ILLNESS AFTER RESPIRATORY CHALLENGE WITH STRAIN SCHU S4.

*LIVE TULAREMIA VACCINE.

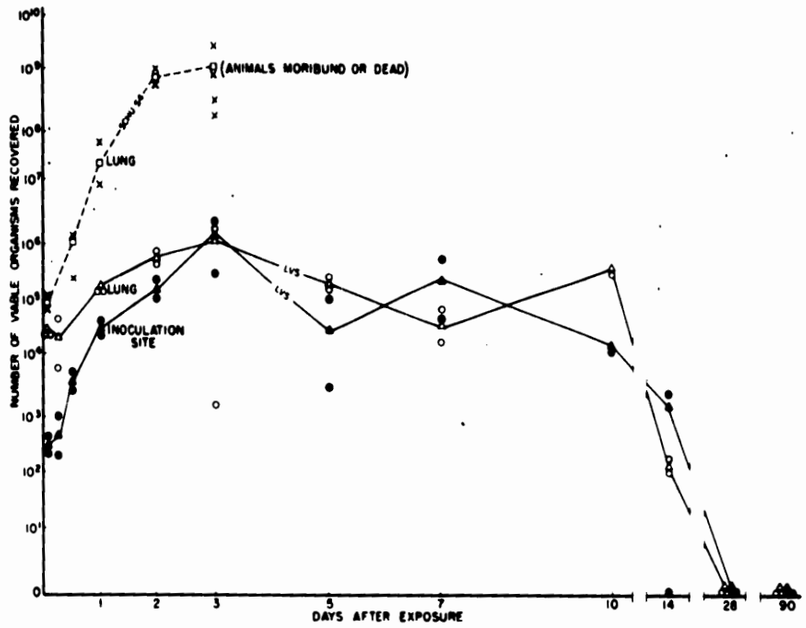


FIGURE 3. RATE OF GROWTH OF PASTEURELLA TULARENSIS LVS OR SCHU S4 IN THE MONKEY: RESPIRATORY EXPOSURE, O OR X; INTRACUTANEOUS INOCULATION, ●; MEAN VIABLE POPULATION, ▲—▲, ▲—▲, OR □----□.

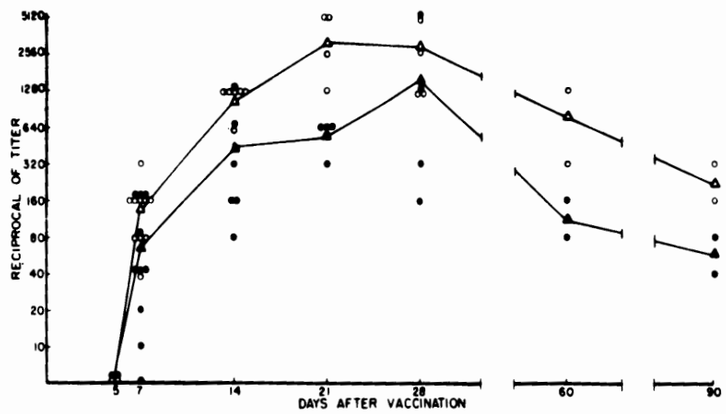


FIGURE 4. AGGLUTININ RESPONSE OF MONKEYS ADMINISTERED LIVE TULAREMIA VACCINE: INTRACUTANEOUS VACCINATION, ●; AEROGENIC VACCINATION, O; MEAN TITER, ▲—▲ OR ▲—▲.

DETERMINATION OF THE GEOMETRY OF HIDDEN DEFECTS BY ULTRASONIC PULSE ANALYSIS TESTING

O. R. GERICKE
WATERTOWN ARSENAL LABORATORIES
WATERTOWN, MASSACHUSETTS

INTRODUCTION

Defects concealed in metals or other solid materials can be detected by ultrasonic inspection. The test procedure predominantly employed for this purpose is generally referred to as the ultrasonic pulse echo method and is in principle quite similar to the Sonar echo sounding technique used in submarine warfare. However, the range of ultrasonic frequencies employed for material inspection is different and extends from about 0.2 to 100 megacycles.

In the ultrasonic pulse echo method, a single ultrasonic transducer acts as transmitter and as receiver of ultrasonic signals. The transducer, which is coupled to the surface of the test specimen, transmits a pulse of ultrasonic energy into the specimen and is then ready to receive any echoes reflected from internal defects or from the back surface of the specimen. The received echoes are amplified and displayed as vertical indications on a cathode-ray-tube. The X-deflection plates of the cathode-ray-tube are usually connected to a time base generator to provide a means for measuring the travel time of the initially transmitted ultrasonic pulse from the transducer to a specific defect. From the knowledge of the travel time the distance between the defect and the test surface can be deduced, if the velocity of ultrasonic waves in the tested material has previously been determined. The two remaining coordinates of the defect location can be derived by scanning the test specimen surface with the ultrasonic transducer and noting the position at which the defect echo amplitude is the largest. If the defect is much greater than the width of the ultrasonic beam used for scanning the test specimen, the procedure will even yield the geometric outline of the defect. In this case it is generally easy to assess the severity of a flaw on the basis of the ultrasonic test results. However, if the flaw is of the same order of magnitude or smaller than the beam width, the flaw geometry is not revealed by the pulse echo test and the interpretation of test results is, therefore, quite a problem.

In the past, numerous attempts have been made by researchers to determine the size of defects which are smaller than the ultrasonic beam width. (In the near radiation field, the beam width is equal to the ultrasonic transducer diameter). One such attempt is to compare the defect echo amplitude to indications derived from artificial flaws of known size. Although great efforts have been expended on this approach, in most cases the results are not satisfactory.

A different technique, which has received much less attention because of experimental difficulties, offers more promise as a solution to the problem. In this technique, pulse echo tests are carried out at various ultrasonic frequencies rather than at only one particular frequency. Thus, the specimen is subjected to ultrasound having different wavelengths. Since the reflection of ultrasonic energy from a defect depends on the ratio of the defect size to the ultrasonic wavelength, the defect echo amplitude is influenced by changes of the ultrasonic wavelength. A determination of the frequency dependence of the defect echo amplitude will, therefore, yield information on the defect geometry.

Unfortunately, the practical application of the aforementioned concept has certain difficulties. Although most commercial pulse echo test instruments are so designed that they can be adjusted for various ultrasonic frequencies, the successive application of different ultrasonic frequencies to a single test specimen would be very cumbersome. Retuning of the transmitter and/or amplifier would be required and, in addition, exchanging of the transducers would be necessary. The latter necessity is a most objectionable feature of the successive procedure because by changing transducers a variation of the transducer-to-specimen coupling conditions may be introduced and, thus, the reliability of the test may be impaired.

To overcome the difficulties encountered with procedures employing a successive application of different ultrasonic frequencies, a novel method was developed at Watertown Arsenal Laboratories which permits the simultaneous use of a wide range of ultrasonic frequencies in pulse echo testing.

THEORETICAL INVESTIGATION

The first step in the development of the novel multifrequency pulse echo test was to investigate how ultrasonic signals could be produced which would contain energy covering a wide range of frequencies. A theoretical study of this problem revealed that single polarity as well as carrier frequency pulses having a rectangular outline would be suitable as sources of multifrequency or, to borrow a term from optics, "white" energy. The spectral energy functions for two representative pulses of such type are given by the equations shown in Figure 1. The spectra derived from the formulae of Figure 1 for the two basic pulses and for a carrier frequency pulse that comprises three cycles are illustrated in Figure 2. Absolute spectral

amplitudes are plotted on a logarithmic scale versus frequency on a linear scale. The reason for selecting this type of plot was that curves thus obtained, to a large extent, resembled displays produced by the electronic spectrum analyzer available for the experimental work.

Comparison of the spectra depicted by Figure 2 indicates that the third pulse, consisting of three cycles of the carrier frequency, contains higher frequencies at about the same level as the two other pulses. Although the spectral amplitude in the vicinity of the carrier frequency is relatively larger, the 3-cycle pulse can still be regarded as a fairly "white" signal. This result is of practical importance because the generation of an ultrasonic signal consisting of three cycles proved to be feasible.

PRELIMINARY EXPERIMENTS

By applying rectangular voltage pulses of single polarity to barium titanate transducers it was possible to produce "white" ultrasonic pulses similar to the last pulse illustrated by Figure 2. Other piezoelectric materials which were also examined did not yield equally favorable results. The upper trace of Figure 3 shows an ultrasonic signal and its spectrum obtained with a transducer having less desirable characteristics. The pulse exhibits a build-up and decay of the oscillation slower than demonstrated by the lower trace of Figure 3 which was derived from the specially selected barium titanate transducer. Correspondingly, the spectra of the two pulses show significant differences. While the upper spectrum can be considered as almost monochromatic, the lower spectrum exhibits a relatively large content of higher frequencies. Although the energy distribution in the spectrum differs somewhat from the theoretical case illustrated by Figure 2, the signal obtained from the barium titanate transducer proved to be quite useful as a "white" pulse for the purpose of relative defect geometry analysis.

APPLICATION

"White" ultrasonic pulses of the type illustrated by the lower trace of Figure 3 were employed for experiments on specimens with fabricated flaws of various geometries. The experiments were carried out with compressional ultrasonic waves and relatively large specimens. The velocity of ultrasonic waves could, therefore, be regarded as a material constant. Hence, the ultrasonic wavelength was inversely proportional to the frequency. Since aluminum specimens were used, the wavelength of a 10 megacycle signal, for instance, was 0.024 inch.

The purpose of the first series of experiments was to demonstrate that "white" ultrasonic pulses were superior to monochromatic test signals with regard to their ability to determine

the configuration of hidden defects. The same pair of test specimens was, therefore, subjected to an inspection by the two different test systems schematically outlined by Figure 4. The system depicted by Figure 4A possesses a relatively narrow band width and, hence, uses essentially monochromatic ultrasonic energy. The pulse associated with the narrow band width system is relatively long and has a round contour. A test system which utilizes wide band width components is shown by Figure 4B together with the much shorter ultrasonic pulse it produces.

Figures 5 and 6 illustrate the results obtained from the inspection of two aluminum test specimens provided with cylindrical holes of 1/8 and 1/32 inch diameter respectively. (In this and in all the following experiments, the amplifier gain was so adjusted that the peak amplitudes of the defect echoes were always the same. Thereby, comparisons of pulse shapes and pulse spectra were facilitated.)

Figure 5 which shows defect echoes obtained with the narrow band width system indicates that no significant differences in pulse contour could be observed if the 1/8 inch and 1/32 inch holes reflected monochromatic pulses. In contrast to this result, the pulse shapes and spectra of "white" pulse reflections, which are illustrated by Figure 6, show significant differences. The distinguishing characteristic of the pulse shapes are the amplitudes of the second cycles indicated by markers on Figure 6. In addition, the pulse spectra exhibit pronounced differences, a fact which is not surprising because, according to the Fourier theory, the time function can be strictly correlated to the frequency function.

In evaluating the results illustrated by Figure 6, it is important to note that the two cylindrical flaws, which closely represent the practical case of stringer inclusions, had equal length and differed only in diameter. Hence, the fact that it is possible to distinguish between these flaws is quite significant.

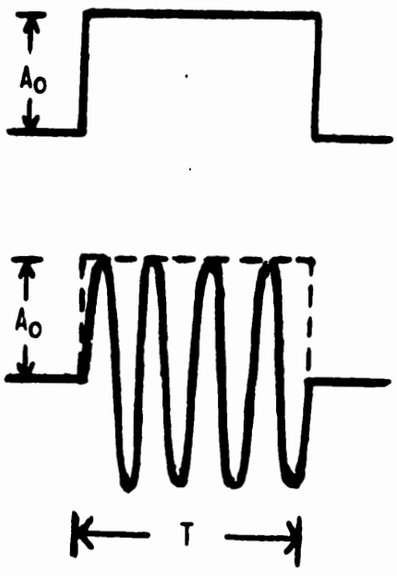
Further experiments were carried out by applying the "white" pulse test method to aluminum cylinders provided with axially drilled flat-bottom holes as indicated by Figure 7. These artificial defects were introduced into the specimens to simulate internal porosity. The pulse shapes and spectra derived from such defects are shown by Figure 7. On the basis of these test results, it is possible to distinguish between a porosity consisting of a large number of small holes and a porosity consisting of a small number of large holes, although the total reflecting cross-section in each case is about the same. Information concerning the size of the individual cavities which can thus be obtained is very important for practical quality inspection because small holes can sometimes be eliminated by a subsequent forging process while large holes would be a basis for rejection of the part under inspection.

As a final example of applications for the ultrasonic pulse analysis method the inspection for internal cracks is considered. The interpretation of ultrasonic test results obtained from specimens having internal cracks is extremely difficult if conventional monochromatic methods are employed. This difficulty exists because by such methods only the amplitude of the reflection is available as a criterion. The echo amplitude, however, is strongly influenced by the orientation of the crack with reference to the surface of the specimen to which the transducer is coupled. A relatively large crack for instance is liable to be misinterpreted as being an insignificant defect if it is so oriented that at the specific test frequency only a small amount of ultrasonic energy is reflected back to the transducer. In the pulse analysis technique, however, errors of this nature are less likely to occur because amplitude differences between individual frequency components of a "white" ultrasonic signal yield additional information.

To investigate the effectiveness of the pulse analysis method for the inspection of internal cracks, the aluminum specimens illustrated in Figure 8 were prepared. The specimens contained cuts at angles of 10, 20 and 30 degrees to the test surface to simulate internal cracks of various orientations. Figure 8 shows also the defect echo and echo spectrum obtained from each specimen and it is noted that pulse shapes and spectra exhibit very significant differences for the various defect orientations.

CONCLUSIONS

The results of the experiments conducted indicate that the effectiveness of ultrasonic pulse echo testing can be greatly enhanced by the introduction of multifrequency signals and defect echo analysis. The main benefit to be derived from this innovation is that differences in the configuration or orientation of concealed defects can be determined.



$$A(f) = \frac{A_o T}{\pi} \frac{\sin \pi T f}{\pi T f}$$

$$A'(f) = \frac{A_o T}{2\pi} \frac{\sin \pi T (f - f_o)}{\pi T f}$$

FIGURE 1 - Spectral functions A (f) and A' (f) of single polarity and carrier frequency pulses with rectangular envelopes. f_o = carrier frequency, f = frequency, T = pulse length, A_o = pulse amplitude.

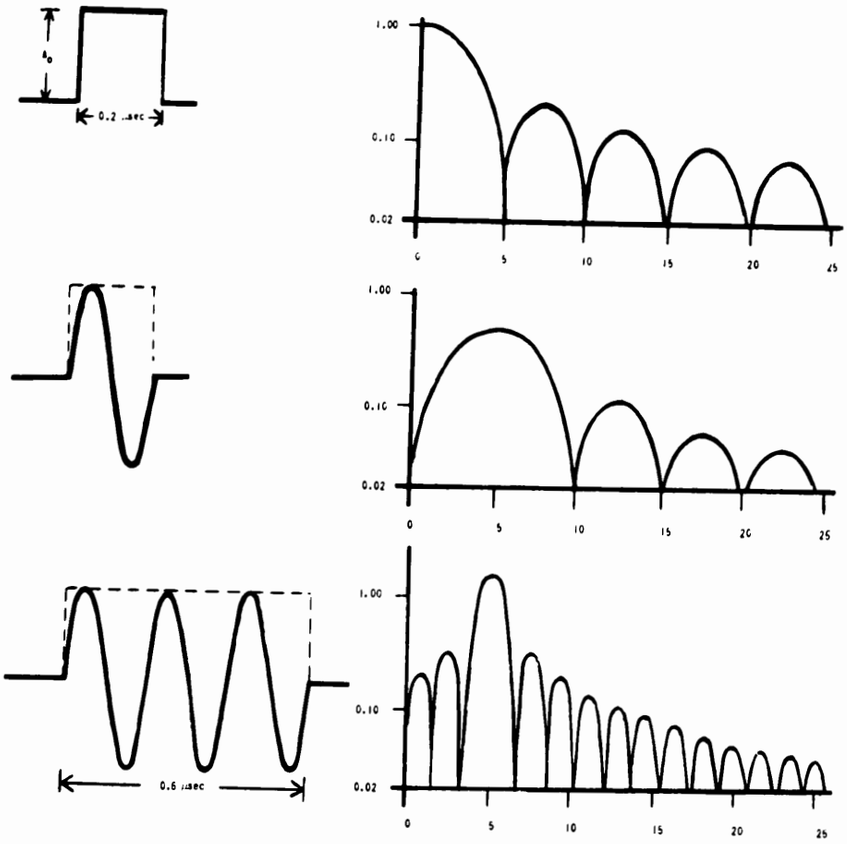


FIGURE 2 - Spectra of single polarity and carrier frequency pulses with rectangular envelope. Frequency range 0-25 megacycles.

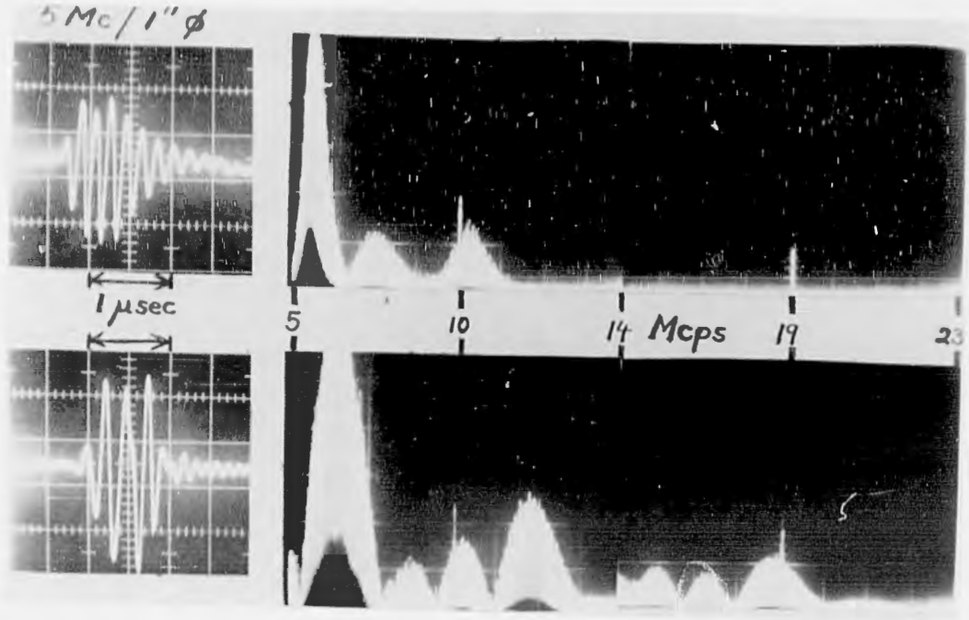


FIGURE 3 - Spectra of ultrasonic pulses with round and with rectangular envelope.

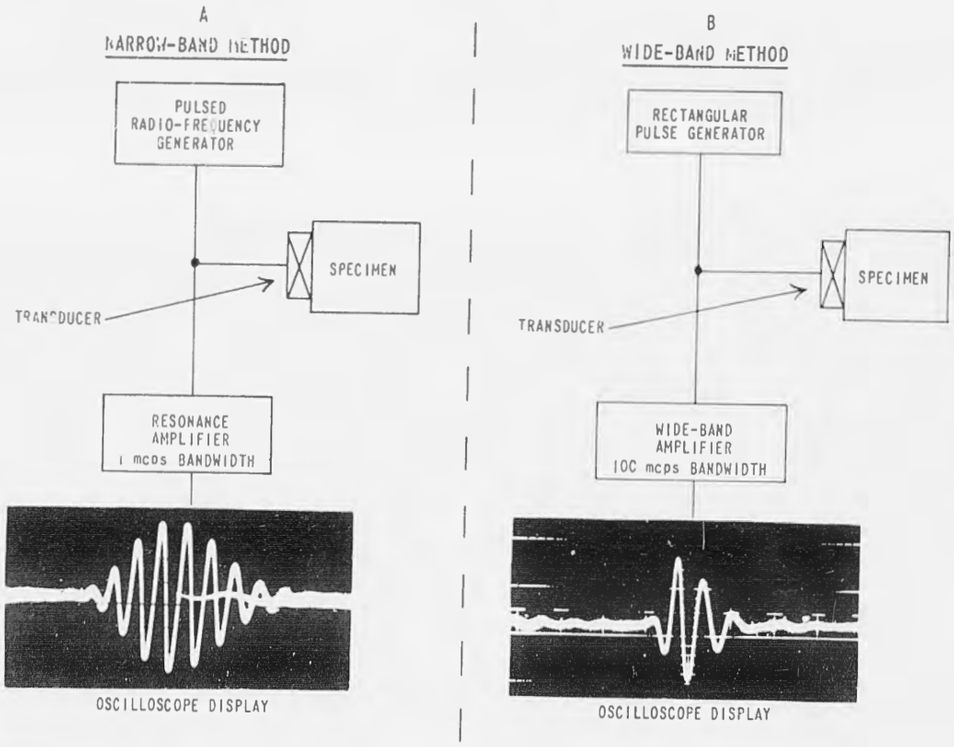


FIGURE 4 - Block diagrams of conventional narrow-band (A) and novel wide-band (B) ultrasonic test systems and resultant pulse shapes.

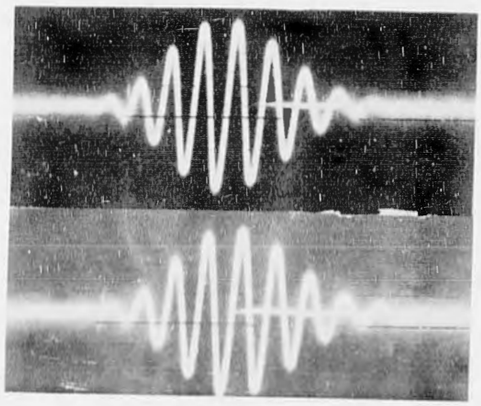


FIGURE 5 - Ultrasonic reflections obtained from 1/8 inch and 1/32 inch holes with a narrow-band test instrument.

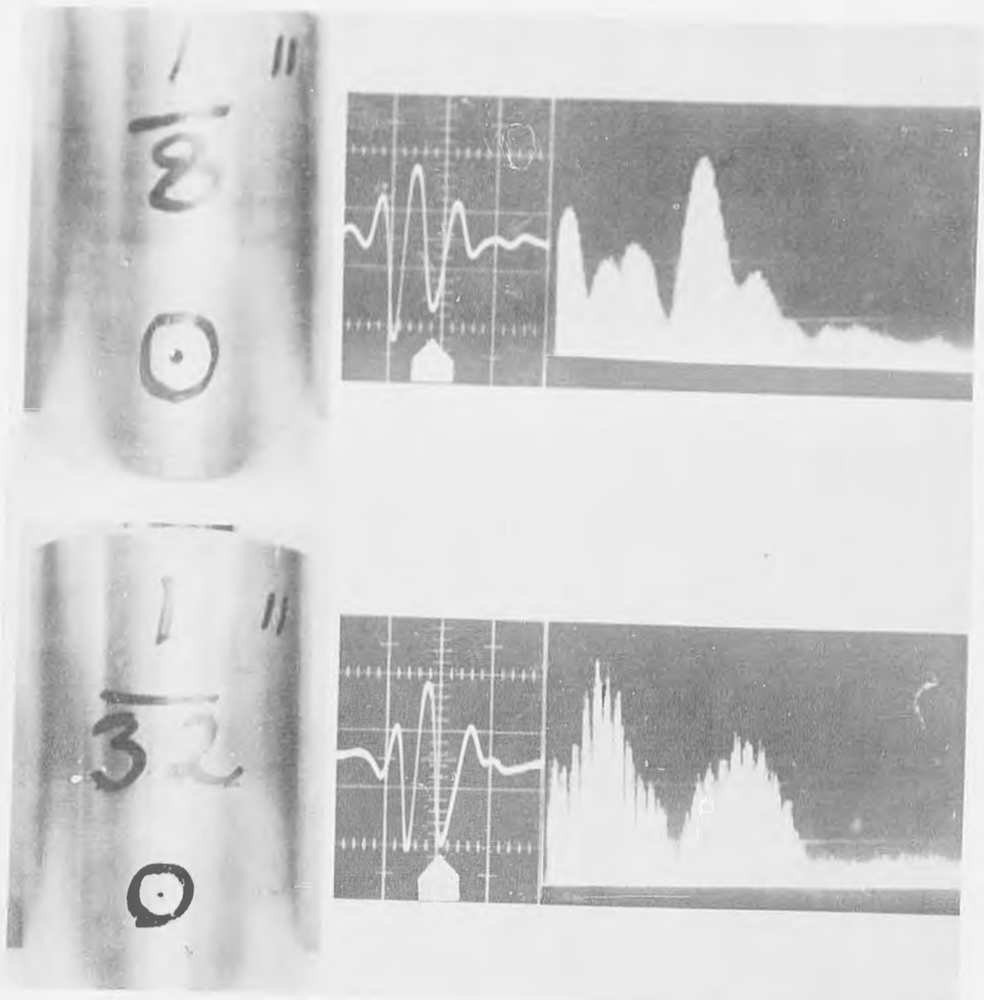


FIGURE 6 - Ultrasonic reflections obtained from 1/8 inch and 1/32 inch holes with a wide-band test instrument.

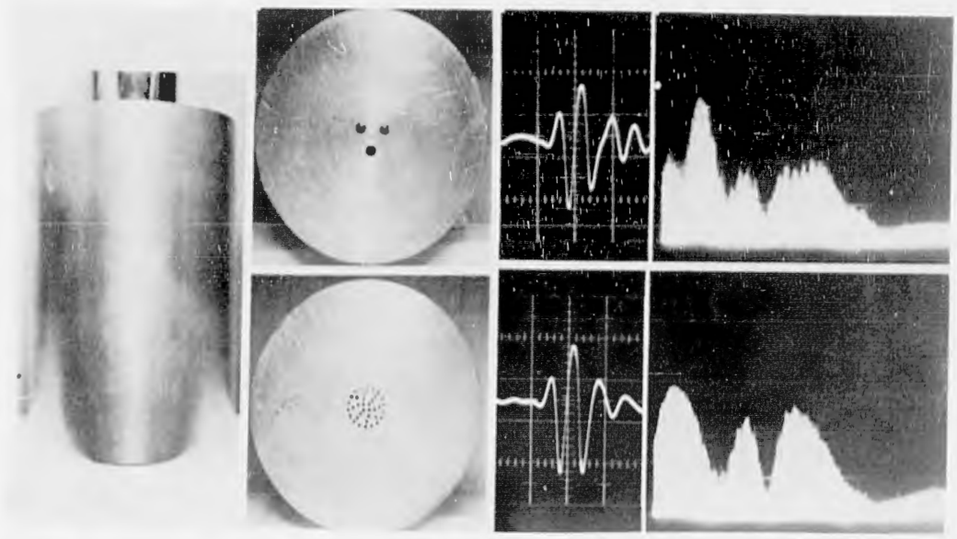


FIGURE 7 - Specimens with artificial porosities, defect echoes and echo spectra.

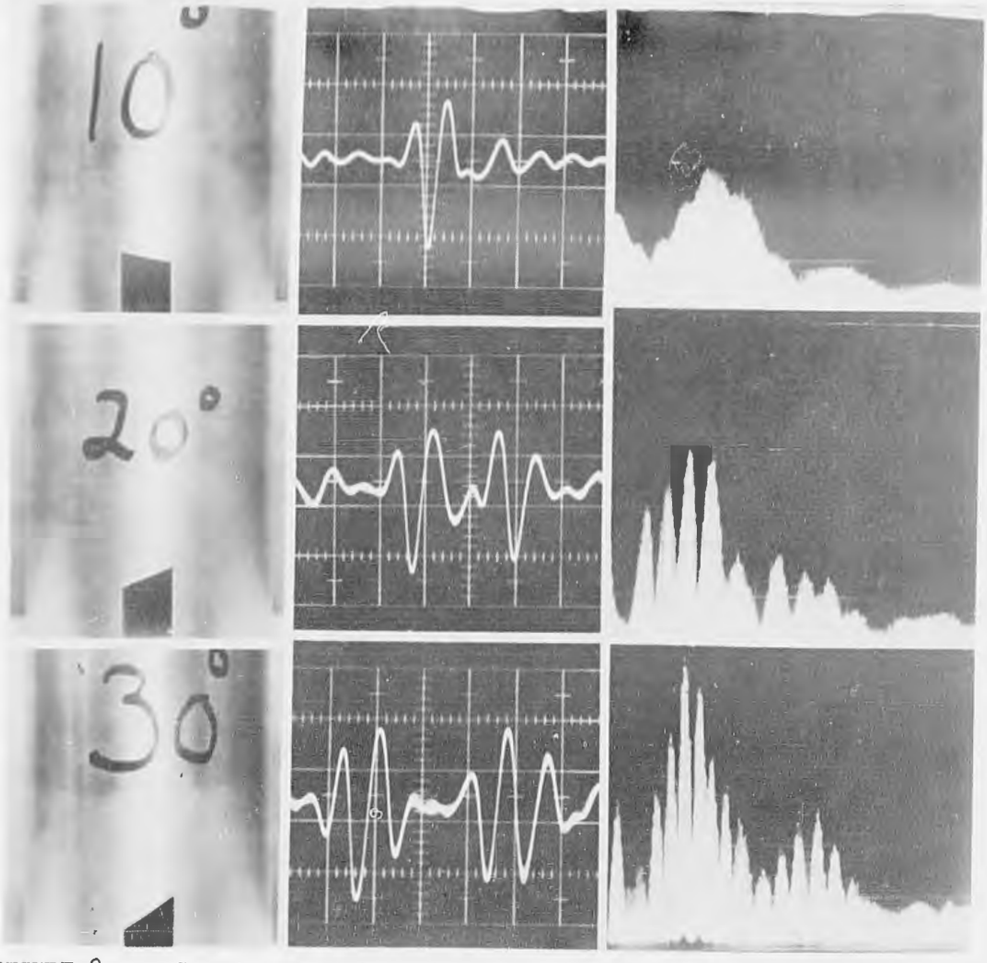


FIGURE 8 - Specimens with angle cuts, defect echoes, and echo spectra.

ULTRA-PRESSURE HIGH-TEMPERATURE RESEARCH: A FRONTIER IN THE QUEST FOR NEW ELECTRONIC MATERIALS

A. A. GIARDINI and J. A. KOHN
U. S. ARMY SIGNAL RESEARCH AND DEVELOPMENT LABORATORY
FORT MONMOUTH, NEW JERSEY

INTRODUCTION: Early in 1955, the synthesis of diamond was announced by The General Electric Co. (1). Thus the last of nature's minerals to defy laboratory preparation was finally added to man's repertoire. Scientists had been striving for a hundred years to accomplish this feat; the principal and most controversial endeavors were those of Moissan (2) and Hannay (3) around the turn of the century. Since diamond is a commercially strategic material for which this country had been totally dependent upon imports, the synthesis itself was a major accomplishment. The total significance of the feat, however, was considerably more far-reaching. It demonstrated that means could be devised for using high-strength materials beyond their ultimate yield strengths in the design of chambers which could sustain pressures in excess of 1,000,000 psi concurrent with temperatures of 2000°C and higher. The door was thus opened to a host of research investigations concerning the effect of such environmental extremes on known materials.

Of more glamorous appeal, and perhaps greater significance, is the possibility of synthesizing entirely new materials unknown in nature, with novel, practical properties. Since the dimensional yield of such substances is expected to remain small for some time, unusual electronic applications appear particularly inviting, especially in view of the present trend toward microminiaturization.

Although the potentialities heralded by diamond synthesis were recognized early, process and instrumentation details were not released until late 1959 (4), resulting in a loss of almost six years of valuable research time. The USASRDL superpressure program, initiated in 1958, achieved successful diamond synthesis in November 1959 using instrumentation independently designed and constructed (5).

Since then, exhaustive studies have been carried out on a variety of metal-carbon systems, culminating in a basic understanding of the mechanism by which diamond forms from such environments. The practical consequence of this fundamental knowledge is a demonstrated ability in the controlled synthesis of pure, electronic grade, diamond single crystals.

Extra dividends accruing from the fundamental diamond studies include the synthesis of several hitherto unknown metal carbides, a new cobalt-containing garnet as well as iron and manganese garnets, pure structurally-perfect graphite single crystals, and discovery of the transformation of pyrophyllite to the minerals coesite and kyanite.

This paper describes (a) the diamond studies and their by-products, (b) the synthesis of cubic boron nitride, a diamond analogue unknown in nature, together with a new structural form of tantalum nitride, (c) the pressure dependence of gallium arsenide resistivity and silver halide ionic conductivity. The latter has demonstrated the use of solid-state, ionic, electrochemical cells as practical pressure measuring devices.

INSTRUMENTATION: Pressure apparatus used in conducting the studies reported herein has been described in detail previously (5,6) and is reviewed here only briefly. As shown in Fig. 1, the chamber consists of two sets of symmetrically opposed, staged, cemented-carbide pistons (items A-C, Fig. 1), and a double-bore, cemented-carbide cylinder assembly (items 1 and 2) surrounded by a series of hardened-steel bands (items 3-5). Component 6 is a mild steel safety ring; finally, hardened-steel hold-down plates (item 7) act to counter lateral shear. The high-pressure, high-temperature reaction cavity consists of the volume enclosed by items 1 and C. An intact sample recovered from the cavity is shown later (Fig. 3) to illustrate a typical arrangement of reaction components.

Working force is applied to the piston assemblies by a 1000-ton hydraulic press. Electrical leads, attached to the external pistons, furnish low-voltage, high-current power for resistance heating of the specimen capsule. This apparatus is capable of sustained, near-hydrostatic pressures to at least 100,000 bars* (~1,500,000 psi) with temperatures exceeding 2500°C. Pressures within

*1 bar = 10^6 dynes/cm^2 ~1 atmosphere

the sample cavity can be determined by electrical resistance measurements on pure bismuth, thallium, barium, and manganin alloy. Temperatures are measured by use of thermocouples and pressure-corrected melting points of pure gold, nickel, and platinum.

DIAMOND SYNTHESIS AND REACTION MECHANISMS: The bulk of the physical-chemical investigations on diamond synthesis conducted at this Laboratory has been concerned with metal-carbon systems. Examples of the latter are reactions of carbon with the elements Pt, Ni, Co, Fe, Mn, Cr, and Ta (or alloys thereof). In addition, studies have been made on systems of carbon with several nonmetallic salts.

In the past, for want of a true understanding of the nature of synthetic diamond growth, "general purpose" terminology such as "catalysis" has been used to describe reactions found to be successful. The present studies have shown that the mechanism of diamond synthesis under conditions of high pressure and high temperature is simply a solution of carbon in any suitable solvent to a state of supersaturation. The thermodynamic instability then causes decomposition and subsequent crystallization of the precipitated free carbon. The latter crystallizes as graphite or diamond, depending upon the pressure-temperature environment. Theoretical data on diamond-graphite equilibria, derived by Berman and Simon (7), have been experimentally verified.

The behavior of the metal carbon solvents listed above falls into two general categories: Those prone to form stoichiometric carbon compounds (carbides), and those which are not. Examples of the former are Fe, Mn, Cr, and Ta; Pt, Ni and Co exemplify the non-carbide-formers. On the basis of diagnostic evidence from residual end products, the mechanism of diamond synthesis in the case of non-carbide-forming solvents is that of simple binary solutions carried to supersaturation. The mechanism for carbide-forming solvents is found to be the sequential formation of carbon compounds with increasing stoichiometric carbon content. The sequence proceeds, in the presence of excess carbon, to the formation of an "ultimate carbide" having the highest possible stoichiometry of carbon. The latter then decomposes to the next lower carbide, liberating free carbon atoms. Typical examples for each class of solvent are given in Fig. 2. The two flow diagrams schematically illustrate the carbide-forming system manganese-carbon, and the simple solution of carbon in nickel.

Fig. 3 is an axial cross-section of a reacted specimen capsule taken directly from the pressure apparatus. The original components consisted of a chemically-pure carbon disk, A, sandwiched between two pure iron disks, B, and a surrounding cylinder of pyrophyllite, C. The latter serves both as electrical and thermal insulation, and as a pressure-transmitting medium. The pressure-calibrating resistance circuit is visible at D. Zone E is an

alteration area in which pyrophyllite undergoes a transformation to the minerals coesite and kyanite. Diamonds, visible in the regions between A and B, are embedded in a mixture of "ultimate carbide" (FeC), Fe_3C, and unreacted carbon. These grade into unreacted iron through a zoned sequence of phases having consecutively lower carbon content. Freedom from distortion in the capsule (Fig. 3) attests to the near-hydrostatic conditions during reaction.

It has been found that crystals synthesized from the solvents Ni and Co invariably contain an internal contamination of the metal elements. The degree of contamination (and also crystal morphology) is dependent upon the temperature during crystallization. The impurities can readily be detected through visual discoloration in shades of yellow, green or brown; or more quantitatively by the degree of magnetism, analytical optics, or x-ray diffraction. Diamonds synthesized from reactions involving Fe and Mn, on the other hand, can be made colorless, chemically-pure, and well-crystallized.

The fundamental cause for this difference in purity is the degree of structural compatibility with diamond of phases present during crystallization. For example, Ni or Co, containing limited amounts of carbon, possess structures which fit precisely in symmetry and dimensions with that of diamond; consequently they are readily and systematically incorporated within diamond as dispersed, oriented inclusions. The carbide-formers, on the other hand, do not have this structural correspondence and are completely rejected from the diamond crystal, leaving the latter pure. Crystals of this type are shown in Fig. 4.

This information now provides the knowledge and basic technology for the synthesis of diamond having a chemical purity not found in nature nor otherwise attainable. Such a capability constitutes a necessary prerequisite for the controlled creation of both chemically-pure, insulating diamond and semiconducting diamond. The latter can be derived by the injection of small quantities of select, elemental impurities into the diamond structure. This has been demonstrated under conditions of high pressure and high temperature without damaging the diamond crystals.

BY-PRODUCTS OF DIAMOND STUDIES: As frequently observed during the pursuit of fundamental knowledge, extra dividends are realized. In this case, during the research on diamond synthesis, the following new compounds have been established: MnC, Mn_3C, FeC, and $Co_3Al_2(SiO_4)_3$ (a cobalt garnet). The carbides are of fundamental significance to ferrous metallurgy. Furthermore, FeC possesses both a markedly high mechanical strength and resistance to chemical attack, and may very well point the way to superior steels formed by high-pressure, high-temperature technology.

During experimentation designed to confirm the supersaturation theory of diamond growth, reactions were carried out to precipitate metal constituents of solutions and thus induce a controlled, premature state of carbon supersaturation. It was found that this could be readily accomplished by the formation of insoluble silicates, especially garnet compounds. The success of this work has economic significance in establishing that carbon supersaturation can be induced at pressures and temperatures significantly lower than previously found possible. Some of the garnet compounds encountered are $Mn_3Al_2(SiO_4)_3$ (spessartite), $Fe_3Al_2(SiO_4)_3$ (almandite), and the previously mentioned cobalt garnet. In the latter, the magnetic ion cobalt has been completely incorporated for the first time into the garnet structural family, a group of materials currently significant in microwave electronics.

In studies to establish additional solvents of a nonmetallic nature, it was found that ammonium salts, e.g. halides, sulphates, and phosphates, were indeed capable of dissolving carbon at elevated pressures and temperatures. Although these reactions have not yet yielded diamond, pure, well-formed, single crystals of graphite have been obtained. The latter are believed to be especially suited for precision micro-resistor elements.

Finally, it has been observed that the pyrophyllite used within the reaction cavity of the apparatus altered, under sufficiently high pressures and temperature, to the minerals coesite and kyanite (refer zone E, Fig. 3). This has potential geochemical and geophysical value in interpreting the intensity of past geological environments (8).

BORON NITRIDE STUDIES: The existence of a boron nitride compound having a structure and properties analogous to diamond was predicted (9,10) years before its synthesis by man (11). Thus far it is unknown in nature. Cubic (β) boron nitride was synthesized at this Laboratory in 1960 (12). Its properties have been found to equal those of diamond in all respects, proving even superior in thermal stability. A summation of properties is given in Table 1. Fig. 5 is a photograph of several cubic boron nitride crystals.

The crystals are most readily prepared by reacting lithium nitride and hexagonal (normal) boron nitride in a resistive heating element (usually tantalum) above approximately 45,000 bars (~660,000 psi) and 1400°C. For a given pressure, the largest crystals grow most readily at a temperature slightly exceeding the minimum required for successful synthesis. Unfortunately, such crystals are also the most impure and deeply colored. Crystallization carried out at higher temperatures yields crystals of smaller size but greatly improved clarity and purity. This association of temperature with color is believed to be related to an imbalance of stoichiometry. The most deeply colored crystals appear to contain an excess of boron, whereas colorless specimens approach ideal stoichiometry. The

apparent relationship between temperature and stoichiometry provides an easy method for control of electrical properties. Consequently, cubic boron nitride may prove to be a more manageable material than diamond for precision tailoring of electronic properties. This is currently under study.

Table 1 - Properties of Cubic Boron Nitride

Size	Up to 0.5 mm, thus far
Habit (shape)	Generally tetrahedral
Color	Clear to amber to red-brown
Crystal structure	Cubic, $\bar{F}43$ m
Cell dimension	3.615±0.001Å (11)
Twinning	(111), common
Index of refraction	>2.0
Density	3.45 g/cm^3 (11)
Hardness	~Equivalent to diamond
Thermal stability	~1800°C., in air
Chemical stability	Inert to solutions of common acids and bases.
Electrical behavior	Insulating when pure and stoichiometric; semiconducting otherwise

During the course of studies carried out on reaction products from the synthesis of cubic boron nitride, a heavy, acid-resistant, often gold-colored material was observed. A combination of microscopy, x-ray diffraction, emission spectroscopy, and x-ray fluorescence showed the unknown phase to be a new form of tantalum nitride, TaN. The latter is cubic (common TaN is hexagonal) and has a crystal structure identical to that of cubic boron nitride. The pressures and temperatures required for its formation are the same order of magnitude as those for boron nitride. Studies designed to investigate crystallization relationships showed that cubic TaN forms independent of BN. The new phase simply results from a direct reaction of Li_3N with the tantalum heating element. Single crystals are dendritic and thus far up to 0.1 mm in size. An observed range in the cubic unit cell dimension of TaN, from 4.315Å to 4.375Å, demonstrates a corresponding variation in stoichiometry. The properties of this material are currently under further study.

CONDUCTIVITY AS A FUNCTION OF PRESSURE: In addition to the physical-chemical studies discussed above, the resistivity of semiconducting gallium arsenide and the ionic conductivity of silver halides have been studied as a function of pressure.

Several oriented, single-crystal specimens of n-type GaAs were carefully prepared with ultrasonically soldered, ohmic contacts of indium. The crystals were placed in cylinders of silver chloride sheathed by normal boron nitride for maximum hydrostaticity of applied pressure.

The results are shown by the curve plotted in Fig. 6. The monotonic increase of resistance with pressure and the apparent leveling off between 50,000 and 60,000 bars is explained in general terms by a scheme of two conduction bands whose separations from the valence band are independent functions of pressure. The remarkable increase in resistivity, by a factor of approximately 600, immediately brings to mind application as extremely sensitive strain gages in the range of 25,000 to 50,000 bars.

It has been observed that the ionic conductivities of AgCl, AgBr, and AgI are highly responsive to pressure up to approximately 30,000 bars. In the latter range, both structural and related chemical defects are essentially eliminated, resulting in a depression of conductivity to a near-vanishing point. Similar to gallium arsenide, the potential value of this material for stress measurement is apparent.

To exploit such behavior, a miniature, solid, self-powered, ionic, electrochemical cell was designed and constructed. The power output of the cell as a function of pressure, as well as its chemical make-up, is given in Fig. 7. The maximum at approximately 6500 bars represents a pressure-induced structural phase transition in the AgI and serves as a convenient reference point. The other silver halides studied showed no maximum output but simply a steady decrease. Comparison of Figs. 6 and 7 shows a convenient overlap of the pressure-sensing characteristics of GaAs and the silver halide cells. Thus the two provide a mechanism for continuous pressure measurement from atmospheric to approximately 50,000 bars.

SUMMARY AND CONCLUSIONS: The results of the studies described herein are of both fundamental and practical significance. Foremost is an understanding of the growth mechanism in synthetic diamond, based upon the principle of supersaturation. The consequences of this knowledge are many-fold. For example, it furnishes the ability to create well-crystallized diamond in a chemically-pure, insulating form, or as semiconducting crystals. It also provides the means to induce supersaturation by use of secondary chemical reactions, thus permitting diamond growth at lower temperatures and pressures than otherwise possible. The economic value of this principle is obvious. In addition, the supersaturation theory contributes toward a clarification of several genetic problems related to the occurrence of both terrestrial and meteoric diamond.

The new carbides which have been established during this study on synthetic diamond not only provide additional fundamental knowledge to the metallurgical field, but also have practical significance. A case in point is FeC, whose observed superior mechanical and chemically-resistive properties imply an ability to create improved steels by high-pressure, high-temperature methods. The inventory of new materials for practical exploitation is further increased by the cobalt garnet discovered during this research.

The preparation of cubic boron nitride constitutes the first independent confirmation of the synthesis of this man-made diamond analogue. Studies indicate that the growth of cubic boron nitride is more susceptible to control than diamond in providing desired electronic properties. A new form of tantalum nitride has been discovered during the boron nitride studies. Although no practical significance can yet be associated with this finding, it is a contribution to man's fund of knowledge in crystal chemistry.

The conductivities of GaAs and silver halides have been studied and a marked variation observed as a function of pressure. This behavior provides a practical mechanism for continuous pressure measurement up to 50,000 bars.

The results derived from this modest beginning in high-pressure, high-temperature research demonstrate the enormous potential not only for creating new materials for man, but also in studying fundamental processes and properties under conditions of extreme environment.

ACKNOWLEDGMENTS: The authors are particularly grateful to J. E. Tydings and D. W. Eckart for their support in various phases of this research. We also appreciate the efforts of Frank Becker, G. E. Tomes, and F. M. Plock, whose contributions helped to make this study possible.

REFERENCES

(1) "Man Made Diamonds," General Electric Rsch. Info. Serv., March 1955.

(2) Moissan, H., Compt. Rend. 118, 320 (1894); 123, 206 (1896).

(3) Hannay, J. B., Proc. Roy. Soc. 30, 188, 450 (1880).

(4) Chem. Engr. News, November 23, 1959; Bovenkerk, H. P., Bundy, F. P., Hall, H. T., Strong, H. M., and Wentorf, R. H., Jr., Nature 184, 1094 (1959).

(5) Giardini, A. A., Tydings, J. E., and Levin, S. B., Am. Mineral. 45, 217 (1960).

(6) "High-pressure, High-temperature Device and Method," Giardini, A. A., and Tydings, J. E., U. S. Patent #2,995,776, 15 August 1961.

(7) Berman, R., and Simon, F., Z. Elektrochem. 59, 333 (1955).

(8) Giardini, A. A., Kohn, J. A., Eckart, D. W., and Tydings, J.E., Am. Mineral. 46, 976 (1961).

(9) "Diamond Substitutes," U. S. Bur. of Mines internal report, 27 April 1951.

(10) "Boron Nitride," U. S. Bur. Mines Info. Circ. 7664 (1953).

(11) Wentorf, R. H., Jr., J. Chem. Phys. 26, 956 (1957).

(12) Giardini, A. A., AAAS Research Conf., Meriden, N. H. (1960).

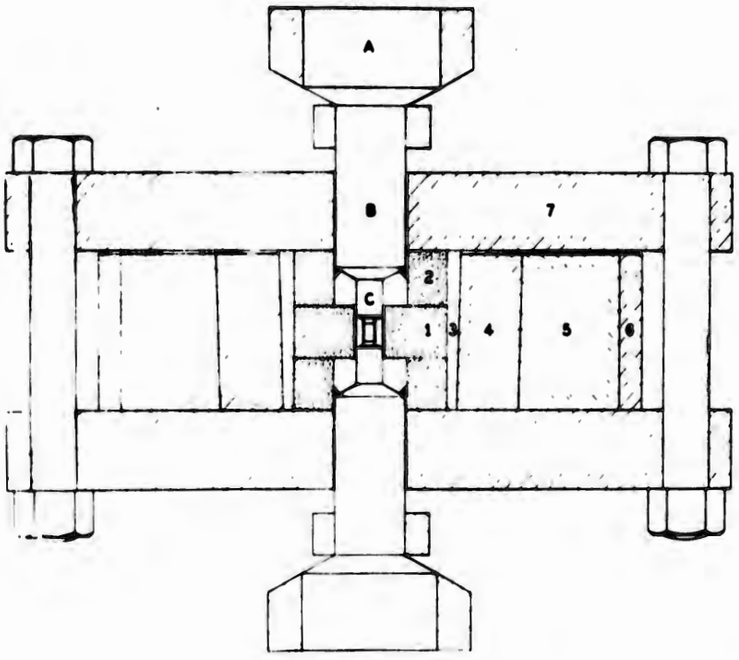
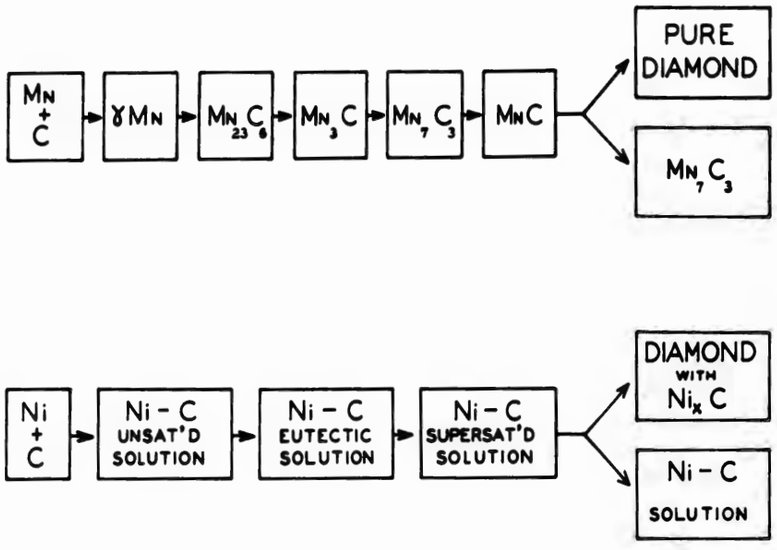


Fig. 1 - High-pressure, high-temperature apparatus ($\frac{1}{2}$ scale).



TYPICAL DIAMOND SYNTHESIS REACTIONS

Fig. 2 - Diamond formation flow diagrams.

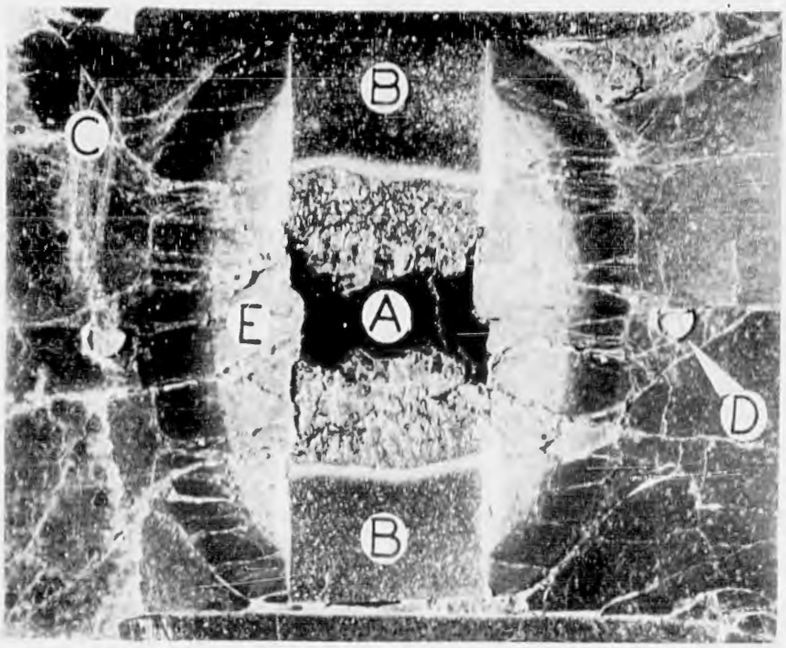


Fig. 3 - Cross-section of reacted capsule (details in text; 10X).

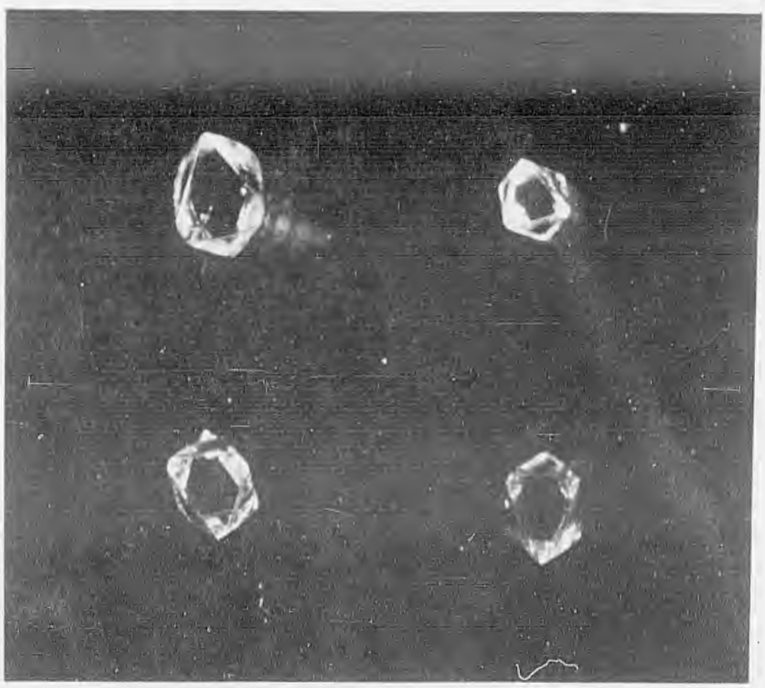


Fig. 4 - Pure, synthetic diamonds from Mn-C system (20X).

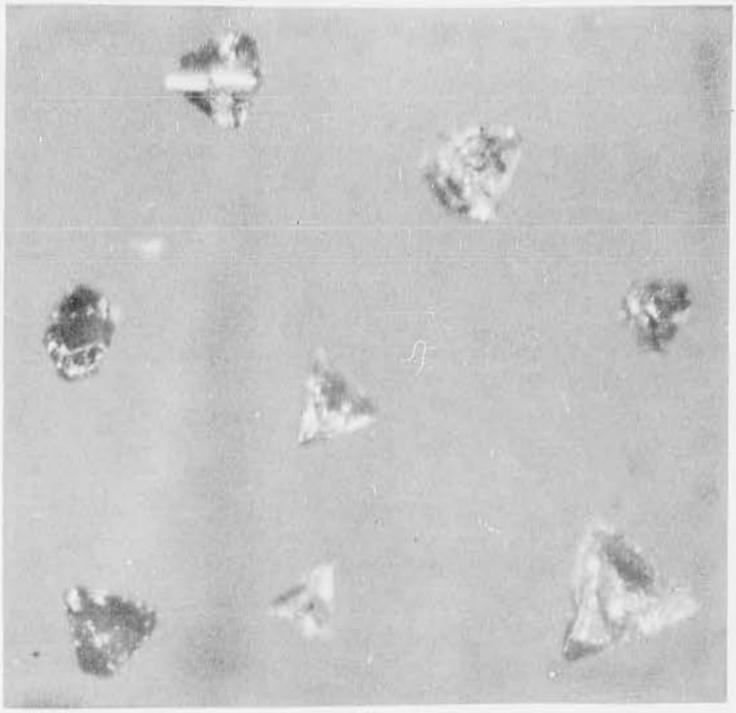


Fig. 5 - Crystals of cubic boron nitride (80X).

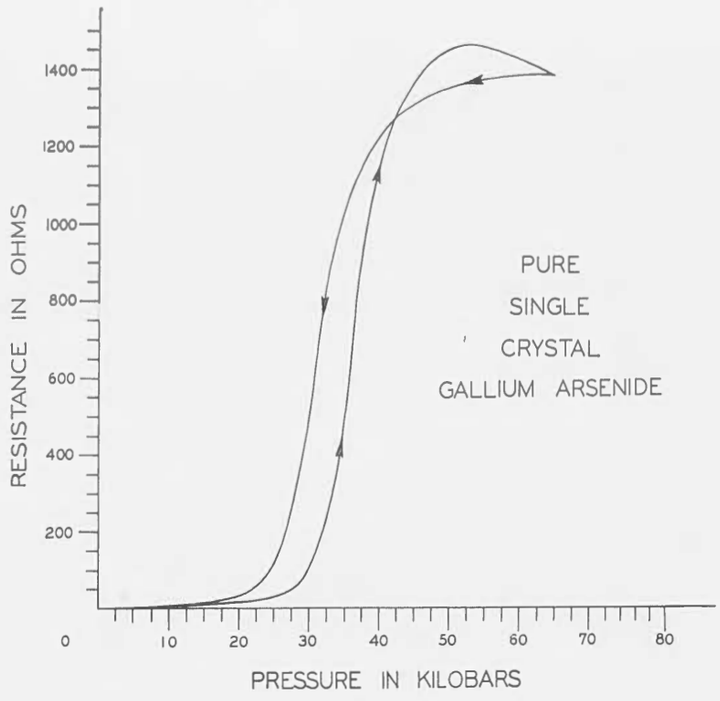


Fig. 6 - Resistance as a function of pressure in gallium arsenide.

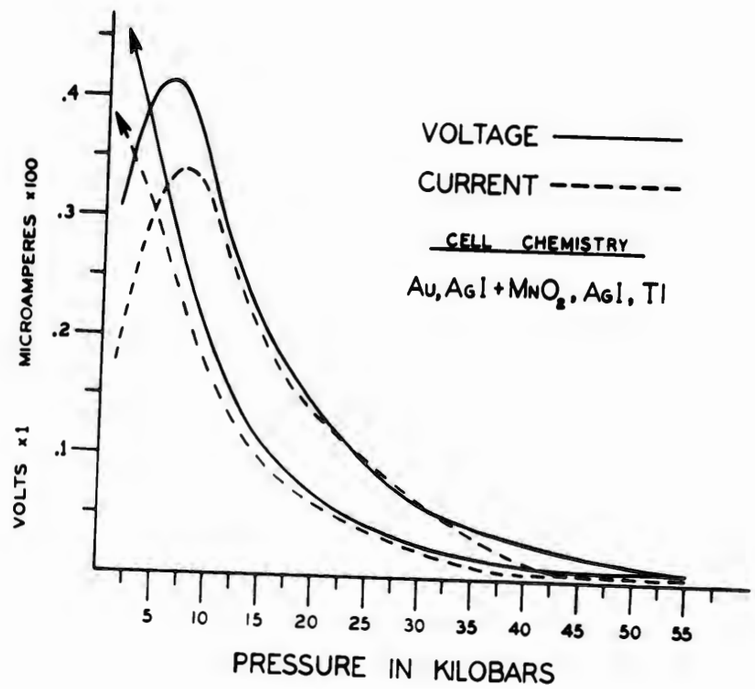


Fig. 7 - Voltage and current characteristics as a function of pressure in a solid, ionic, electrochemical cell.

INVESTIGATIONS ON A BEAM WAVEGUIDE FOR OPTICAL FREQUENCIES

G. GOUBAU & J.R. CHRISTIAN*
INSTITUTE FOR EXPLORATORY RESEARCH
U. S. ARMY SIGNAL RESEARCH & DEVELOPMENT LABORATORY
FT. MONMOUTH, N. J.

With the recent development of the optical MASER (LASER) an extremely large frequency range has been made available to communication. Unfortunately the utilization of this frequency range is seriously hampered by the vulnerability of light propagation through the atmosphere by fog, rain and snow. An obvious way to overcome this difficulty is to provide a protected light path. An efficient long-distance transmission system would, however, not result if the light beam were simply passed through a pipe. Although the beam emitted by a LASER has very small spread angle, it still is wide enough to hit the wall of the pipe within a relatively short distance even if the pipe were perfectly straight. This would cause not only substantial transmission loss but also severe delay distortions because of reflections on the guide wall.

During the past few years the authors have studied a novel kind of waveguide for millimeter and submillimeter waves, which appears to be applicable also to the transmission of coherent light. This waveguide differs from conventional waveguides in that the propagated field is not a wave mode in the usual sense, but a wave beam with special field configuration. This waveguide has been termed, "Beam Waveguide" in a paper on this subject presented at the 1959 Army Science Conference. (1)

There are two kinds of beam waveguides which are illustrated schematically in Fig. 1: the iris-type and the lens-type. The "iris-type" simply consists of uniformly spaced irises, while the "lens-type" comprises in addition lens-shaped dielectric phase correction plates. Both are based on the existence of so-called "reiterative wavebeams," this means, wave beams whose field distribution can be re-established at periodic intervals. In the iris-type guide the iteration is accomplished by diffraction. The beam when passing through an iris is somewhat restricted in diameter by the aperture of the iris. The associated field distortion compensates

for the field transformation that takes place during the journey of the beam from one iris to the next. In the lens-type guide the iteration is primarily produced by refraction. Each lens resets the cross-sectional phase distribution in the beam to the distribution which exists at the preceding lens. Since the lenses do not act as optical image forming devices the names "phase correction plates" or "phase transformers" are more appropriate from the functional point of view. Diffraction by the lens aperture contributes little to the iteration process though it affects iteration loss.

In both cases very low iteration loss can be achieved if the apertures are sufficiently large compared to the wave length. This is apparent from Fig. 2 which shows curves of the theoretical losses per iteration of the fundamental beam modes in the two types of beam guides as a function of the aperture area divided by the product of wavelength and iteration path length D. The curves have been verified experimentally at various frequencies, (2), (3), (4) although not within the optical frequency range. But there is no reason to assume that they are not applicable to this range. Of course, the iteration loss in the lens-type guide is increased by inherent losses of the lenses, i.e., reflection and absorption losses. With modern techniques these can be limited to less than 1/10 db per lens.

In the millimeter wave range the iris type guide is impractical because it requires an extremely large aperture-to-wavelength ratio in order to be efficient. Although the loss per iteration can be reduced by shortening the iteration pathlength D, the loss per unit length is little changed because of the increased number of iterations required per unit guide length. In the optical range an aperture to wavelength ratio of the order of 10^4 is readily obtainable, and therefore the iris-type beam waveguide also warrants consideration.

At first sight the iris-type guide appears to be simpler in construction than the lens-type guide. However, a more thorough study reveals that the opposite is true. Assume, for instance, an optical beam of 5.5×10^{-5} cm wavelength which is iterated by irises of R= 1 cm aperture radius at intervals, D of 2 meters. $A=R^2/2D$ then has a value of 90 which yields a loss per iteration of approximately 0.002 db estimated from curve I of Fig. 2. The transmission loss per kilometer is therefore 500 times as large or 1 db.

For a lens-type guide with the same aperture, but an iteration path length of D= 100 m, A is 1.8. The iteration loss due to diffraction is less than 0.001 db per iteration (see curve II on Fig. 2) and is therefore negligible. Assuming a reflection and absorption loss of 0.1 db/lens the total transmission loss per km guide length is also 1 db.

Experimental investigations of the iris-type guide (3) at millimeter waves indicated that slight curvatures of the light path as may be caused by sags of the guide between supports would greatly increase the transmission loss. Therefore the iris guide requires a very rigid supporting structure over its entire length. In contrast to this the lens-type guide which has a very much greater iteration pathlength requires rigid supports only for mounting the phase transformer units, this means at intervals of say 100 m. The pipe connecting these units may have bends and sags as long as its diameter is large enough to ensure unobstructed passage of the beam from one lens to the next. Of course, the lenses must be aligned. But since the effect of a misalignment is readily predictable, the alignment should not be too difficult. Small tilts of the lenses or displacements in the direction of the beam have no appreciable effect on the transmission loss. This fact has also been verified in previous experiments [2].

If a lens is displaced perpendicular to the beam axis by an amount ξ the beam is slightly deflected (see Fig. 3a) and proceeds on a zig-zag path as indicated in Fig. 3b, with a maximum deviation from the straight path by $\pm 2\xi$. The bending effect by lateral displacements of the lenses can even be used to correct for curves in the layout of the guide. Additional diffraction losses which occur when the beam does not pass through the center of the lenses are negligible if the lens diameters are sufficiently large.

Although there is no doubt from the theoretical point-of view that beam waveguides can be applied to the transmission of coherent light there are a number of practical questions which can only be answered by experimental studies. It was therefore decided to construct an experimental model long enough to be representative for long distance transmission. From the foregoing considerations it was concluded that the lens-type guide should be much less costly in construction and offer a better chance for success than the iris-type guide.

The questions which are of primary interest are the following:

1. Can the alignment of the phase transformers be practically realized and maintained, considering the fact that any support, no matter how rigidly it is constructed will be subjected to the movements of the ground?
2. If the guide is filled with air or any other gas, how serious is the effect of turbulance in the gas on the propagation, and what are the tolerable temperature and pressure gradients to keep the turbulance at an acceptable level?
3. What precision is required in the manufacturing of the phase transformer units, in particular, what tolerance is acceptable in the construction of the lenses?
4. What minimum transmission loss can be practically achieved?

A representative length for a lens-type beam waveguide should comprise at least 10 iterations, assuming that the beam is pre-shaped to match approximately the field distribution of the desired reiterative beam mode. For a lens radius of 1 cm and an iteration length of 100 m the total length should therefore be approximately 1 km.

It is quite obvious that the alignment of the phase transformer requires remotely controlled movements. After the guide is aligned, small displacements of the supports should not require corrections in the alignment, if the aperture radius is made larger than necessary from the point of diffraction loss only.

Turbulence could be best minimized by burying the guide several feet in the ground. However, this would have been quite expensive, since all the parts would have to be made waterproof to prevent flooding. It furthermore would have necessitated bends in the guide since the profile of the available real estate varied over 7 feet in height. Although bends can be easily accomplished by inserting glass prisms into the path of the beam, it is felt that this would be an undesirable complication for the very first experimental set-up.

The basic structual arrangement was made as simple as possible with the provision that any modification found necessary during the experiments could be incorporated with a minimum of effort. Fig. 4 shows the overall layout. As the inclosure for the light path a 6" aluminum irrigation pipe was used and supported by utility poles about 2' to 8' above ground, depending upon the profile of the terrain, and 40 feet apart. Additional X-shaped supports made of two by fours, were afterwards inserted half way between the poles because the pipes developed mechanical oscillation in heavy wind. The phase transformer units are mounted separately and connected to the pipe with flexible coupling hoses to avoid the transmission of mechanical forces from the pipes to their units. Their spacing is 97 m to accommodate ten sections within the total length of 970 m of available real estate. At both ends, approximately 48 meters from the first and last phase transforming units steel frames were fastened on the ground for mounting source and receiver equipment.

The lenses used for the phase transformation have a diameter of 2 cm and a focal length of one half the lens spacing. They were ground by personnel of the Frequency Control Division of this Laboratory. The quarter-wave coating used for reflection compensation was designed for wavelengths around 5500 Å. At these wavelenths a lens radius of 0.7 mm suffices to keep the diffraction loss below 0.01 db per iteration, which is still small compared to the inherent lens loss. A displacement of the light beam of up to $\pm$3mm from the center of the lens should, therefore, not cause any substan-

tial increase in diffraction loss.

A phase transforming unit is shown in Fig. 5. The lens is mounted in a frame which can be moved laterally in two perpendicular directions by stepping-solenoids with suitable gearing mechanism. Each step produces a movement of the lens by an amount of 1/20 of a millimeter.

The movements are remotely controlled with a plug-in control box which can be operated at each end of the guide and at the locations of any one of the transformer units. The control box allows the movement of any single lens or any combination of two lenses. The combined movement of lenses proved advantageous in the alignment procedure. If from three successive lenses the first and the last one are moved by the same amount ξ, the beam is displaced by 2ξ at the middle lens, but remains in position at all the other lenses not belonging to the group (see Fig. 6).

Although a beam waveguide can only guide coherent beams, an incoherent light source may be used for measurements. It is well known from elementary optical interference experiments that a small amount of coherent light can be derived from incoherent sources. It can be shown that the maximum amount of energy which is available for a single beam mode is given by $I=\sigma\lambda^2/2\pi$ where σ is the intensity per unit area of the source.

The source used is a high pressure mercury lamp in conjunction with a filter to pass only the two spectral lines 5461Å and 5770Å for which the quarter wave coat on the lenses is most effective. The light is modulated by a mechanical chopper consisting of a disc with holes rotated by a synchronous motor, yielding a modulation frequency of 990 c/s. Since the source is mounted 48 m before the first lens and the arc size of the mercury lamp is only 0.3mm x 0.3 mm, the beam entering the first lens is practically coherent. Of course, the field distribution deviates substantially from that of the reiterative beam. In other words, there are higher beam modes simultaneously developed at the first aperture, which also propagate along the guide.

There is an optimum aperture size for which the ratio between the fundamental mode and the higher modes is a maximum. The iteration loss of the higher modes is substantially greater than that of the fundamental mode. Therefore, measurements of the transmission loss of the guide will require the design of a "launching device" which develops a beam with a field distribution close to that of the desired beam mode. It shall be mentioned in this connection that the beam modes are orthogonal with respect to power (5). This means the total power of a beam is equal to the sum of the powers of the various modes.

The receiver consisted of RCA photomultiplier tube type 7265 in conjunction with a narrow band amplifier. The equipment was made portable so that it could be moved by a vehicle from phase transformer to phase transformer to aid in the alignment of the beam.

It is not surprising that the initial construction could not be operated during daytime, particularly in clear weather. The temperature gradient inside the pipe can cause deflections of the beam of up to 8 cm per 100 m., because a temperature gradient of only 0.1° C/cm causes bending of a beam of light of 5 cm. Even during the night turbulence was frequently extremely high because the air inside the pipe was coupled to the air outside. Observations could be made only during quiescent weather conditions. Other problems were condensation of water vapor on the lenses and fog developing inside the pipe.

In spite of these difficulties this construction was adequate to answer the first and most fundamental question concerning the alignment. It was found that in favorable weather conditions the alignment of the phase transformers did not present serious difficulty. It was also found that whenever the beam was once aligned the alignment was substantially maintained indefinitely. However, the intensity fluctuations of the beam caused by turbulence were too great to make any loss measurements. Still the result was encouraging enough to warrant improvements of the set-up, particularly to reduce turbulence and to provide a viewing device for observing and photographing the beam as it passes through the various phase transformers. The purpose of photographing the beam at each lens is to obtain calibrated field distribution curves. From these it should be possible to determine the amount of higher modes present as the beam progresses down the line. Information on the quality of each lens should also be obtainable by this method. The reduction in turbulence was achieved by inserting a 4" pipe coaxially into the 6" pipe supported by plastic foam spacers. Air exchange is prevented by air-tight connection between the pipe sections and the phase transforming units. The viewing device, designed to observe and photograph the beam, is shown in Fig. 7. It consists of an adjustable mirror which can be controlled from the outside so that the beam can be diverted through a viewing flange. However, it was found that the modification brought substantial improvement in the stability of the beam, though it still appears questionable whether operation during sunlight is possible or whether further shielding of the light channel is required. Unfortunately, the completion of these modifications was delayed so that no data have been available at the time of the writing of this report.

ACKNOWLEDGEMENT

The writers wish to thank Mr. W. Schneider for his design of the precise lens control units, and to Mr. Walter Merkl for the

grinding of the lenses. Acknowledgment is extended also to Mr. James Arnold for his efforts in the construction and operation of the experimental structure.

REFERENCES

(1) "A New Waveguide for Millimeter Waves," G. Goubau & J. R. Christian--presented at the 1959 Army Science Conference June 1959.

(2) "Experimental Studies on a Waveguide for Millimeter Waves," G. Goubau & J. R. Christian--I.R.E. TRANSACTIONS ON ANTENNAS & PROPAGATION: Vol AP-9 No 3, May 61 PP 256-263.

(3) "Some Measurements on an Iris Beam Waveguide," J. R. Christian & G. Goubau--Proceedings of IRE (correspondence) Vol 49, No 11, November 1961.

(4) "Measurements of Resonators formed from Circular plane and confocal paraboloidal mirrors," E. L. Scheibe--Proceedings of IRE (correspondence) Vol 49 Page 1079 June 61.

(5) "On the Guided Propagation of Electromagnetic Wave Beams," G. Goubau & F. Schwering--IRE TRANSACTIONS ON ANTENNAS AND PROPAGATION Vol. AP 9 No 3, May 1961.

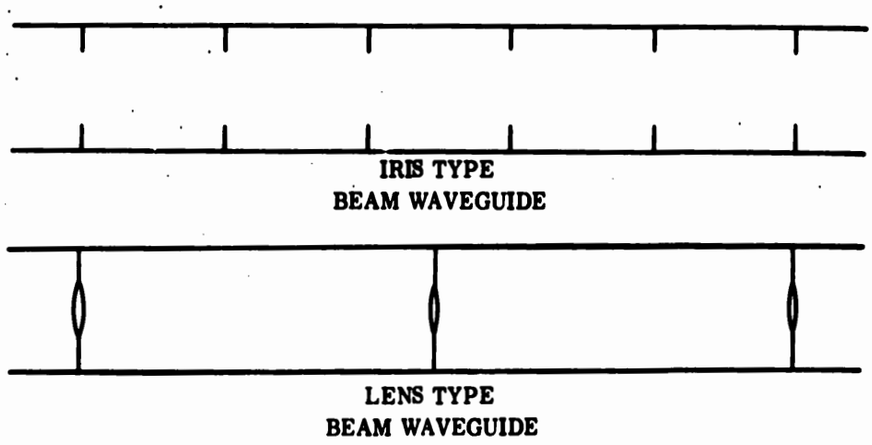


Fig. 1 - SCHEMATIC OF BEAM WAVEGUIDE TYPES

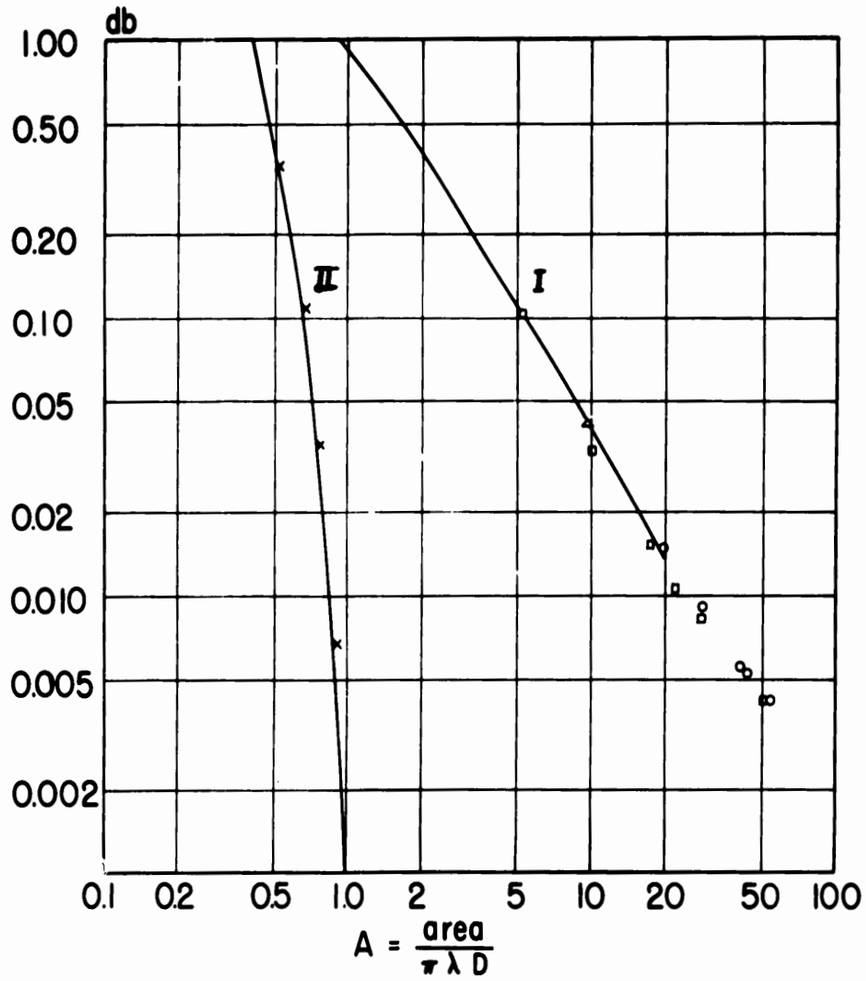


Fig. 2 - THEORETICAL DIFFRACTION LOSS VS. A=area/$\pi\lambda$D.

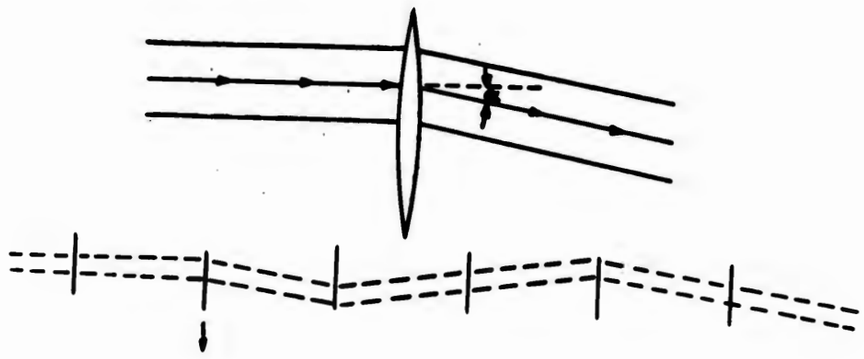


Fig. 3 -A, ILLUSTRATION OF BEAM DEFLECTION BY LENS WHICH IS OFF-AXIS
B, ILLUSTRATION OF BEAM DIRECTION IF ONE LENS IS DISPLACED OFF-AXIS BY A SMALL AMOUNT

Fig. 4 - PHOTOGRAPH OF EXPERIMENTAL OPTICAL BEAM WAVEGUIDE.

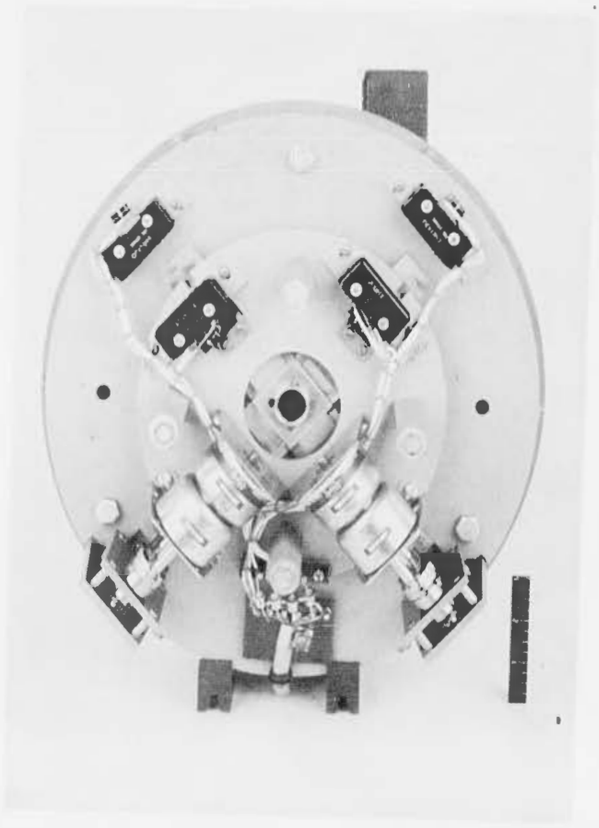


Fig. 5 - PHOTOGRAPH OF PHASE TRANSFORMER UNIT (WITH COVER OFF), SHOWING LENS AND STEPPING-SOLENOIDS.

Fig. 6 - ILLUSTRATION OF BEAM DEFLECTION IF TWO LENSES ARE DISPLACED AN EQUAL AMOUNT IN THE SAME DIRECTION.

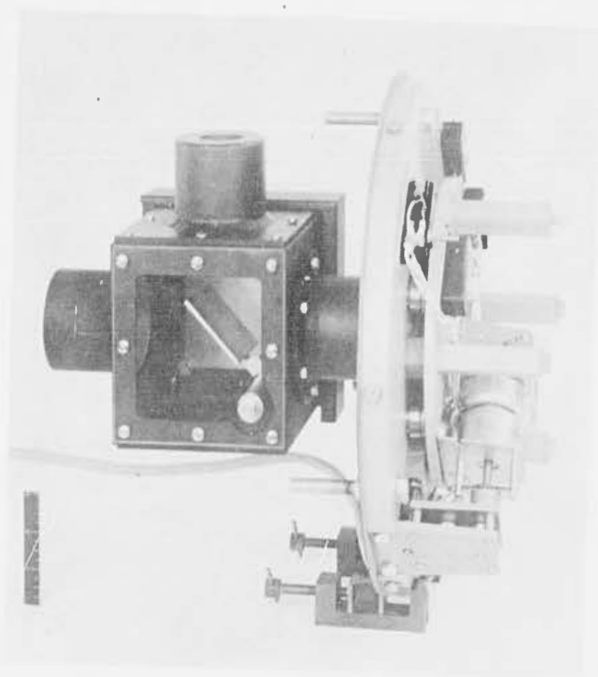


Fig. 7 - PHOTOGRAPH OF VIEWING DEVICE (TRANSPARENT WINDOW ADDED SO THAT INSIDE OF DEVICE CAN BE SEEN).

REGENERATION OF FOOD FLAVORS THROUGH ENZYMATIC ACTION

TORSTEN HASSELSTROM, S. DAVID BAILEY, and ELWYN T. REESE
QUARTERMASTER RESEARCH & ENGINEERING COMMAND
PIONEERING RESEARCH DIVISION
NATICK, MASSACHUSETTS

The problem of restoring to processed foods the aroma characteristics associated with the fresh product is one which is well recognized by anyone who has been subjected to Army rations. It is generally recognized that the major factor in the acceptance of food is based on the sensory perception of odor, taste and texture. Of these, we are most sensitive to odor.

The amount of compounds giving the characteristic odors associated with the fresh fruits and vegetables is extremely small, represented in many cases by parts per million. Thus the study of the original food flavors, and the problem of restoring flavor are of considerable difficulty. One of the factors that have made this particular type of work successful has been the recent application of new, extremely sensitive means of analytical detection of volatile compounds (1, 2). The small quantities of volatile substances, which we associate with aroma, are related to less volatile compounds which we call flavor precursors. These are more stable and less likely to be lost or destroyed during processing treatments. The bulk of the food gives texture or body and the fundamental food nutrients such as protein, carbohydrates, fats, etc.

When a food is processed we can distinguish between the original fresh flavor and an altered flavor, which may or may not be desirable. In many cases the processed flavor has now become accepted and is characteristic of the processed food without relating it to the fresh flavor. As an example, canned peas have a flavor which is entirely different from that of the fresh product, yet most people find the canned product completely acceptable. It is, in fact, a different product. An off-flavor is an aroma we do not associate with a certain food. Yet it may normally occur in other foods. For example, acetaldehyde is a normal constituent of most fruit aromas, but it is definitely an off-flavor in butter where it sometimes results from microbial action. Hydrogen sulfide and dimethyl sulfide are important

components of beer flavor where they are present in amounts below the threshold of olfactory sensitivity. In other foods they might be evidence of putrefaction.

The aroma components of foods are usually associated with the following classes of organic compounds: aliphatic acids, esters, aldehydes, ketones and alcohols. In addition, minute quantities of nitrogen- or sulfur-containing chemicals and terpenes may be present.

The problem with which we are concerned is how can we restore to a processed food all or part of the aroma associated with fresh food. In the first slide

SLIDE 1

is shown schematically the concept of flavor regeneration which we have developed. The first block shows the presence in the fresh product of a considerable amount of natural flavor along with latent flavor precursors and active enzymes. Through processing treatments, these volatile natural flavors are largely lost and in part are replaced by altered flavor as shown in the second block. The flavor precursors are more stable and survive the processing treatment. The activity of the enzymes is lost by the heat treatment involved. If we can demonstrate that certain enzymes will now react with the flavor precursors, it should be possible to restore some of the natural flavor to the processed food. This has been done and is shown in the third block.

One could very quickly ask why process foods if this is going to result in the loss of the original fresh flavor. We believe the answer is obvious. The processing of food is basically a stabilizing process. The stabilization is achieved by substantially destroying or inhibiting microorganisms and inactivating enzymes, either through freezing, heat-sterilization, the use of ionizing radiation, dehydration, or simply by drying, the latter being the method handed down to us from prehistoric times. These stabilizing processes prevent or retard chemical changes, and deterioration through bacterial action. Thus while we achieve the goal of stabilizing food, we have done so at the price of freshness.

In this study we have made the basic observation that potential flavor, in the form of flavor precursors, may be converted into fresh flavor through the action of the proper enzyme added at the time of reconstitution of the food, or in its preparation for consumption (3, 4). In order to obtain the maximum flavor regain, it is necessary to understand the nature of the enzyme precursor reaction. Where a dehydrated product is used, water is required, and the reaction must proceed at the proper pH and temperature. The next slide

SLIDE 2

demonstrates the reconstitution of flavor in dehydrated cabbage. The gas chromatograph of head space above freshly minced cabbage shows the presence of allyl isothiocyanate and butyl isothiocyanate. These are important aroma constituents of the fresh product. The identity of these compounds has been established by gas chromatography and mass spectrometry. In the case of the dehydrated cabbage, these compounds are absent. The sensitivity of the method is such that the absence of a peak is also evidence for absence of detectable odor. Now when an enzyme preparation is added to the dehydrated cabbage, the two isothiocyanates reappear (1). This provides analytical evidence for restoration of two important flavor compounds. The next slide

SLIDE 3

shows other isothiocyanates and sulfides which have been found in cabbage, (1, 5, 6) through analysis by gas chromatography, mass spectrometry and by conventional isolation and characterisation methods. It may be of interest to note that in securing the data shown in this slide approximately 100 grams of cabbage were required while by contrast our earlier studies required 100 lbs of cabbage in order to extract and separate out the flavor constituents.

The enzymes used for the reconstitution of flavor may be isolated from the fresh food, or from one of its waste products such as leaves, stems, etc. In the establishment of a commercially successful system for reconstituting flavor, an economical source of enzymes must be obtained. As part of this study, we have found that the fungus, Aspergillus sydowi, is capable of producing the required enzyme (a thioglucosidase), and that the yields of enzyme are much better than those obtained from higher plants. This fungal enzyme has been shown to be effective in the release of isothiocyanates from the thioglucoside flavor precursors (7).

While the basic principle is established that enzymes are involved in the release of certain flavor components, it is by no means certain that all flavor components arise in this fashion. As shown in this slide (slide 3) we have observed 20 volatile sulfur compounds in fresh cabbage (1). With the exception of the isothiocyanates, and possibly certain sulfides, we have yet to show the enzymatic origin of the other components. We are confident, however, that the isothiocyanates play the major role in the fresh aroma characteristic of cabbage and other members of that family because of their pungent odor. It is known that there are about 25 naturally occurring isothiocyanates (mustard oils). The precursors of these isothiocyanates is the family of compounds called thioglucosides. The in vitro enzymatic hydrolysis of these compounds has been studied in detail by Ettlinger et al. (8) who have made the interesting observation that Vitamin C acts as a co-enzyme. The reactions are shown in the next slide

SLIDE 4

The addition of the vitamin increases the release of the isothiocyanate. This slide shows the typical equations for thioglucosides undergoing enzymatic hydrolysis, through two intermediates, to form isothiocyanates and nitriles. The ascorbate-activated enzyme cleaves the β-glucoside (I) to form the unstable thichydroxamate ion (II) which is spectrophotometrically detectable and in turn undergoes Lossen rearrangement at neutral pH to form isothiocyanates and sulfate. At more acid pH, nitrile and sulfur appear as decomposition products of the thiohydroxamic acid (III).

The isothiocyanates vary in their aroma characteristics. Some are pungent as for example the allyl isothiocyanate, while others are mildly pleasant such as the methyl-thiomethylisothiocyanate (9). The subtle characterization of any food aroma is probably due not to any single chemical entity but rather to a mixture of volatile compounds.

As part of this work a general synthesis of thioglucosides (10) permitting the synthesis of new compounds has been achieved as shown in the next slide

SLIDE 5

In what might be regarded as the simplest possible scheme for synthesizing thioglucosides (I), an aldoxime can be chlorinated and without isolation the resulting hydroxamyl chloride sulfonated with sulfur trioxide in dimethylformamide to a hydroxamyl chloride-O-sulfonate, V, isolated as a potassium salt. Condensation of the chloride with sodium glucopyranosyl mercaptide in aqueous solution formed the thioglucoside (I) anion isolated by ion exchange conversion to a crystallizable salt.

Another flavor enzyme system in cabbage is that which converts the precursor, a sulfoxide, to the volatile flavor constituent disulfides. These sulfides along with the isothiocyanates represent the important characteristics of the aroma of cabbage. The enzymatic- and acid-hydrolyses of S-methyl-cysteine sulfoxides (2, 11, 12) are shown in the next slide

SLIDE 6

Starting with the sulfoxides intermediates are shown to occur prior to the formation of the dimethyldisulfides.

In the last slide

SLIDE 7

are shown other processed foods which have been similarly treated with enzymes, and which have been shown by sensory evaluation to have

enhanced flavor regeneration (4). From this table it can be seen that a large number of common foods may have flavor regeneration systems comparable to those in cabbage. It is now necessary to establish the nature of the enzymes in these foods in which the release of flavor has been shown organoleptically, and to determine the nature of the flavor precursors involved.

REFERENCES

1. Bailey, S. D., M. L. Bazinet, J. L. Driscoll, and A. I. McCarthy, ibid., 26, 163 (1961).

2. Dateo, G. P., R. C. Clapp, D. A. M. MacKay, E. J. Hewitt and T. Hasselstrom, Food Research, 22, 440 (1957).

3. Hewitt, E. J., D. A. M. MacKay, K. Konigsbacker and T. Hasselstrom, Food Technology X, 487 (1956).

4. Hewitt, E. J., T. Hasselstrom, D. A. M. MacKay, and K. S. Konigsbacker, U.S. Patent No. 2,924,521; French Patent No. 1,175, 251; German Patent No. 1,052,221, E 14, 196 IVa/53k.

5. Clapp, R. C., L. Long, Jr., G. P. Dateo, F. H. Bissett and T. Hasselstrom, J. Am. Chem. Soc., 81, 6278 (1959).

6. Altamura, M. R., L. Long, Jr., and T. Hasselstrom, Biol. Chem. J., 234, No. 7 (1939).

7. Reese, E. T., R. C. Clapp and M. Mandels, Arch. Biochem. Biophys., 75, 229 (1958).

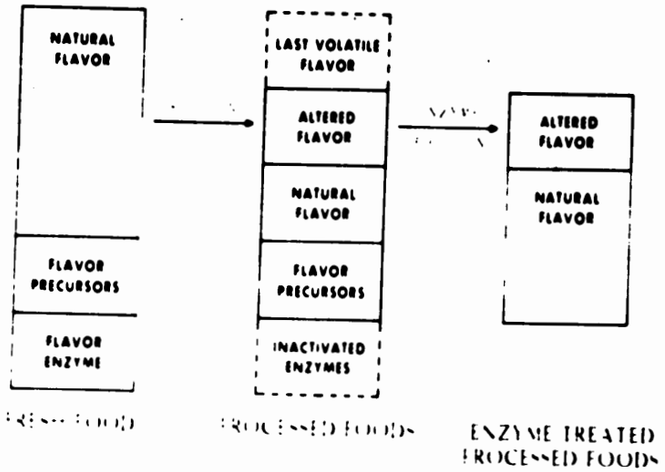
8. Ettlinger, M. G., G. P. Dateo, B. W. Harrison, T. J. Mabry and C. P. Thompson, Proc. Nat. Acad. Sci., 47, 1875 (1961).

9. Hasselstrom, T., R. C. Clapp and L. Long, Jr., J. Org. Chem., 26, 3026 (1961).

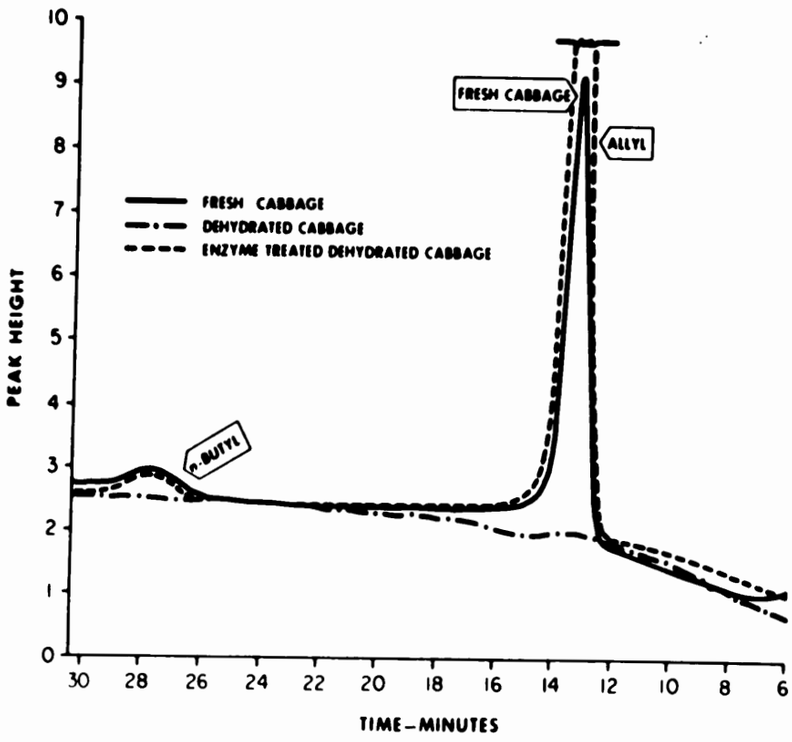
10. Ettlinger, M. G., and G. P. Dateo, Final Report, DA19-129-QM-1059. Project 7-84-06-032, Simplified Food Logistics, 1961.

11. Ostermayer, F., and D. S. Tarbell, J. Am. Chem. Soc., 78, 1605 (1960); Quartermaster Research Contract No. DA19-129-QM-1076.

12. Stoll, A., and E. Seebeck, Helv. Chim. Acta. 32, 866 (1949).



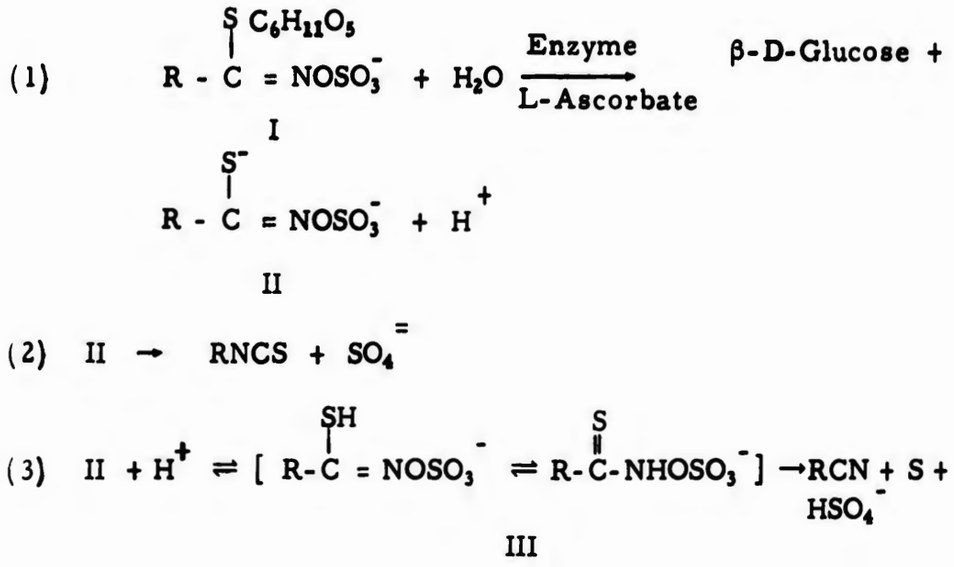
Slide 1. Regeneration of Natural Flavor



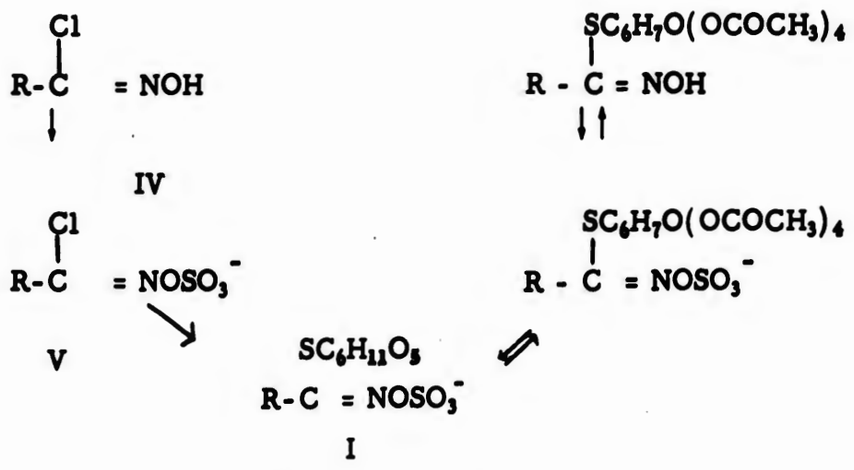
Slide 2. Flavor Release by Enzymes

ISOTHIOCYANATES	DISULFIDES
Methyl	Carbon
n-butyl	Dimethyl
Butenyl	Methyl ethyl
Allyl	Diethyl
Methylthiopropyl	Ethyl propyl
SULFIDES	Dipropyl
Hydrogen	Propyl butyl
Carbonyl	Propyl allyl
Dimethyl	Diallyl
Diethyl	**TRISULFIDES**
Dibutyl	Dimethyl

SLIDE 3. VOLATILE SULFUR COMPONENTS OF FRESH CABBAGE

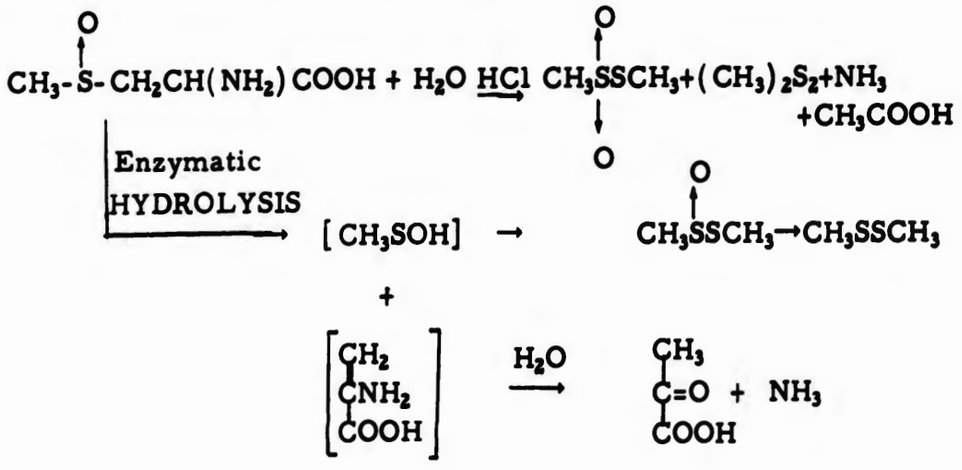


SLIDE 4. HYDROLYSIS OF THIOGLUCOSIDES BY ENZYMES



R = Alkyl or Aryl

SLIDE 5. SYNTHESIS OF MUSTARD OIL GLUCOSIDES



SLIDE 6. ENZYMATIC AND ACID HYDROLYSIS OF S-METHYL CYSTEINE SULFOXIDE

Celery	Carrots	Oranges
Cabbage	Onions	Strawberries
Mustard	Garlic	Pineapple
Parsley	Milk	Raspberries
Horseradish	Tomatoes	Radish
Spinach	Bananas	Watercress, etc.

Slide 7. Processed Foods Improved in Flavor by Addition of Enzymes

FAST WAVE RADIATORS

LEONARD HATKIN
U. S. ARMY SIGNAL RESEARCH AND DEVELOPMENT LABORATORY
FORT MONMOUTH, NEW JERSEY

Leaky wave antennas of various types have been described in the literature (1,2,3,4,5). These are special cases of what will be referred to hereafter as fast wave radiators. This paper will discuss fast wave radiation in general, and one particular type of leaky wave antenna which appears to have certain advantages.

Before proceeding with the treatment of fast wave radiation, let us examine some of the properties of open waveguides. An open waveguide is one which has some part of the transverse cross section exposed to free space, so that the fields which exist in the waveguide have a tangential component on a free space interface. Typical examples of open waveguides are conventional balanced, unshielded transmission lines and dielectric slab waveguides. Both these examples support only "bound" waves. An example of an open waveguide which supports leaky waves is a conventional rectangular waveguide, operating in the dominant mode, but with a long longitudinal slot in the narrow wall.

The characteristics of the two types of waveguides are quite different. In the case of those which support only the "bound" wave, the field components decay in a direction perpendicular to the axis of the waveguide. Hence, essentially all the energy carried by the wave is "bound" to the vicinity of the guiding structure. Also there either are no components of both the electric and magnetic fields at the free space interface which are mutually perpendicular to each other and to a vector perpendicular to the guide axis, or if these components do exist they are in phase quadrature. In any case, there is no component of the Poynting Vector which points perpendicularly outward from the guide axis, and power is transported only in the direction of the guide axis or possibly inward to account for guide losses. The phase velocity of the wave travelling down this type of waveguide is equal to or less than the velocity of light in free space. The "bound" wave is therefore sometimes referred to as a "slow" wave.

For the guides which support the leaky waves, the fields in

free space at the interface actually increase exponentially as we move away from the axis of the guide. This situation, of course, does not go on indefinitely but exists only in the immediate vicinity of the interface. At the interface itself there is some component of the Poynting Vector which is perpendicular to the guide axis and in an outward direction. In other words, if the wave travelling down the guide is pictured as consisting of two criss-crossing waves, the transverse components of these waves see a resistive component of impedance at the guide-air interface and some power is transferred through the interface. Thus a portion of the power being transported by the wave is continuously "leaked off." This may be compared with the situation for the bound wave, where the transverse impedance at the interface is zero, infinite or reactive so that the transverse wave is completely reflected at the interface and no real power is transferred across it. The phase velocity of the leaky wave is greater than that of light in vacuo, hence, the leaky wave is often referred to as a "fast" wave.

The presence of a Poynting Vector component perpendicular as well as parallel to the axis of the waveguide results in radiation at an angle to the guide axis. This angle, which for a uniform medium is the same as that made by the criss-crossing wave, is given by

$$\phi = \cos^{-1} \frac{c}{v_p} \qquad 1.$$

where c is the velocity of light in vacuo and v_p is the phase velocity along the guide. Since the power lost through radiation, per unit length, is proportional to the relative strength of the outward perpendicular component of the Poynting Vector it can be seen that as ϕ increases from zero, which is the endfire direction, to 90° (broadside), the attenuation constant of the waveguide due to radiation loss also increases. The maximum attenuation occurs when the path of the criss-crossing waves is entirely transverse to the axis of the guide. Now, none of the power is carried down the axis of the guide, but is all radiated in a broadside direction. The existence of only transverse waves is the condition for cut-off in conventional waveguides. Referring to Equation 1, it is seen that broadside radiation occurs with infinite phase velocity. This is also consistent with the cut-off condition for waveguides.

Aside from interest in the fundamental phenomena of fast wave radiation and the place of leaky modes in the overall modal analysis of waveguides, activity in the field of fast wave radiators stems from the number of practical advantages that this type of antenna appears to offer. In general, the attenuation constant and phase velocity of fast wave antennas are individually subject to precise control. This means that the radiation patterns of these antennas may be very accurately predicted, or conversely a specified pattern may be closely synthesized. Since the direction of radiation is a function of the phase velocity and this in turn is dependent on frequency, the radiated beam may be made to scan as a function of

frequency. At a given frequency, the beam may be made to scan by electrically or mechanically changing the phase velocity along the antenna. Also, in many practical embodiments the fast wave radiator can be made lighter, more compact and more efficient than conventional antennas of comparable electrical performance.

At this point it may be well to consider how a TEM mode on a conventional transmission line can be converted into a fast or leaky mode. The phase velocity of a wave on a transmission line is determined by the reciprocal of the square root of the product of the series inductance per unit length of line and the shunt capacitance per unit length, or

$$v_p = 1/\sqrt{LC} \qquad 2.$$

with v_p being the phase velocity, L, the series inductance per unit length, and C, the shunt capacitance per unit length. For a TEM wave the phase velocity is always equal to the velocity of light in the medium between the conductors of the line. In particular, if the medium is free space, the phase velocity along the line is the velocity of light in free space. If now, the series inductance per unit length, and/or the shunt capacitance per unit length is reduced from its TEM value, the phase velocity along the line will be greater than that of light in free space. Since the wavelength is equal to the phase velocity divided by the frequency, the wavelength along the line is increased above the free space or TEM value in proportion to the increase in phase velocity. Physically, this increased wavelength along the line may be explained by assuming the wave propagating down the line to be made up of criss-crossing plane waves with the criss-crossing waves travelling with the velocity of light. This picture is often used to explain the behavior of conventional waveguides which can be regarded as TEM lines loaded with the guide walls so as to produce a high phase velocity. If in addition the transmission line or waveguide is open, or partially open as defined previously, each time the criss-crossing wave strikes the open boundary, a portion of the energy in the wave is leaked off into free space. An attenuation constant along the guide therefore exists due to the constant proportional loss of energy due to radiation. This situation is shown in Figure 1.

Since the fast wave is no longer a TEM wave it must have a low frequency cut-off. If we assume that the L or C of the TEM wave as given by Equation 2 is reduced by adding series capacitance or shunt inductance respectively, cut-off occurs when the added reactance resonates with the original reactance. At this frequency the phase velocity and guide wavelength become infinite. Below this frequency, the line or guide acts as an attenuating filter, since the series and shunt reactances in the transmission line equivalent circuit are both of the same type, either inductive or capacitive. This results in an imaginary characteristic impedance and real propagation constant, which corresponds to exponential attenuation. As frequency is increased above cut-off the effect of the loading reactance is

decreased, so that the phase velocity approaches the velocity of light. The attenuation also decreases with increasing frequency, theoretically approaching zero. The behavior of the ideal fast wave line is shown in Figure 2.

The particular embodiment chosen for intensive study consisted of a parallel plate balanced transmission line with inductive shunt loading. The loading was accomplished by means of conducting pins connecting the parallel plates and arranged along the mid-plane of the parallel plates. Different spacings between the pins were tried to determine the effect of varying the loading inductance per unit length. A coaxial balance to unbalance transformer was fitted to the line at either end so that coaxial equipment could be used at the input and output of the line. The fast wave line is shown in Figure 3 and schematically in Figure 4. Table I shows the pertinent parameters of the line.

TABLE I

Plate Width,	2a	0.50 in.,		
Plate Separation,	b	0.05 in.,		
Pin Separation,	d	0.39 in.,	0.79 in.,	1.57 in.,
Pin Diameter,	2r	0.013 in.,		
Antenna Length,	L	15.62 in.,		

The particular dimensions chosen were determined by a number of compromises which permitted relative ease of fabrication, single mode operation of the antenna in the frequency ranges of interest, and the use of previously derived expressions to facilitate theoretical analysis.

The theoretical determination of propagation constants was made by the transverse resonance method (3,4). This method is based on the fact that in any guiding structure

$$k_g^2 = k^2 - k_t^2 \qquad 3.$$

where the g subscript refers to the direction along the guide axis, the t subscript to the direction transverse to the guide axis and the absence of subscript refers to TEM propagation.
In general,

$$k_g = -j\gamma_g = \beta_g - j\alpha_g \qquad 4a.$$

$$k = -j\gamma = \beta - j\alpha \qquad 4b.$$

$$k_t = -j\gamma_t = \beta_t + j\alpha_t \qquad 4c.$$

where γ is the propagation constant; β the phase shift constant or $2\pi/\lambda$; and α is the attenuation constant. Thus the real part of k_g gives us the phase constant along the guide axis, from which the guide phase velocity can be found, while the imaginary part is the attenuation constant. Note that for physical and mathematical reasons, the imaginary part of k_t is positive. In order for a fast wave to exist, the real part of k_t must have a non-zero value, which means that a finite wavelength exists in the transverse plane. Since there is at least partial reflection in the transverse plane, a standing

wave exists and the transverse plane is resonant. This fact allows us to use the transverse resonance equation which may be stated in one form as

$$\overleftarrow{Y} + \overrightarrow{Y} = 0. \qquad 5.$$

$\overrightarrow{Y}$ refers to the admittance looking to the right at a given reference plane, while $\overleftarrow{Y}$ refers to the admittance looking to the left at the same plane. The solution of Equation 5 in general will yield k_t which when substituted into Equation 3 will give us k_g. Cut-off will occur approximately when the real part of k is equal to the real part of k_t. If k_t is a constant, as is essentially the case in many structures of interest including the line under consideration, then it can be seen that the cut-off wave length λ_c equals λ_t, and the cut-off wave length can be directly determined by the solution of Equation 5. This transverse wavelength λ_t, is the wavelength of the standing wave which is permitted to exist in the transverse plane in terms of geometry of the line or guide.

In order to construct a valid equivalent circuit of the transverse plane certain assumptions must be made. First, we assume that the array of discrete loading pins forms a continuous inductive wall down the center of the guide. This assumption holds as long as the pin spacing remains reasonably small in terms of wavelength. It breaks down as the spacing approaches a half-wavelength. Next, we assume that no standing waves exist in the direction perpendicular to the parallel plates. The small spacing used insured that this is true in the frequency range of interest. Third, we assume that the apertures on either side of the parallel plate line are independent, i.e., the fringing fields do not extend from one aperture to the other. Also we hope that the reactance of the pin grid is not affected by the presence of the apertures. The plate width and construction were chosen so as to make these assumptions reasonably true.

Figure 5 shows the equivalent circuit of the transverse section. This consists of a section of line of length 2a, terminated in an admittance $G_a + jB_a$ on either side. This admittance is that due to an aperture between two parallel plates as seen by a wave propagating between the plates. An inductive susceptance, $-jY_s$ located in the center plane of the equivalent transmission line represents the reactive effect of the grid of pins on the transverse wave. The transmission line has a characteristic admittance of Y_o which in this case can be considered to be the same as for free space, and a complex phase shift constant k_t. It should be noted that although all the conducting surfaces are considered lossless, k_t must be complex in order to account for radiation loss, schematically represented as the partially resistive terminations on the transverse line. The transverse wave being investigated is of the simplest type. The electric field is perpendicular to the parallel plates with no variation in this direction, and the standing wave pattern consists only of a single loop.

By symmetry, the equivalent circuit may be bisected with no loss in generality. At the central plane $\overrightarrow{Y} = \overleftarrow{Y}$. When combined with

Equation 5, this results in $\vec{Y} = \overleftarrow{Y} = 0$. Thus, the entire left hand half of the equivalent circuit may be replaced by a magnetic wall, or open circuit. This action results in the circuit of Figure 6 with $Y_s' = Y_s/2$.

Fortunately, the equivalent admittances have been previously worked out. The expression for the aperture admittance is in Reference 6 and that for the array of pins in References 7 and 8. These are, subject to approximations permitted by the geometry,

$$Y_s'/Y_o \approx \frac{\pi}{k_t d \ \ln(d/2\pi r)} \qquad 6.$$

$$G_a/Y_o = k_t b/4 \qquad 7a.$$

and

$$B_a/Y_o = \frac{k_t b}{2\pi} \ln \frac{4\pi e}{1.781\ k_t b} \qquad 7b.$$

with e being the Napierian base, 2.718.

If we now chose as the reference plane the aperture itself, substitution in Equation 5 yields,

$$j \frac{Y_s'/Y_o - \tan k_t a}{1 + (Y_s'/Y_o)\tan k_t a} = \frac{G_a}{Y_o} + j \frac{B_a}{Y_o} . \qquad 8.$$

The left hand side represents the negative of the input admittance to the line at the aperture, while the right hand side is the aperture admittance itself. It can be readily seen that substitution of the values for the admittances results in a transcendental equation of rather complicated form. However, it does turn out that under the assumptions made, k_t is a constant.

The details of the calculation will not be presented here since they are rather tedious and lengthy. However, the methods used will be briefly indicated. In the first approach, Equation 8 with the proper substitutions was programmed for the Burroughs 220 Computer and solved for the real and imaginary parts of k_t. The second method was based on a simplified physical analysis in which the inductance is replaced by a section of shorted line. The third method used was an iterative, linear extrapolation, in which a test value of k_t is tried in Equation 5 and a correction term obtained from the first term in a Taylor expansion of the Y's as a function of k_t. This procedure may be repeated as often as necessary to obtain the required accuracy.

The results obtained by the approximate methods, particularly the last, agreed very closely with those from the complete machine solution, thus showing that desk calculators could be used for design problems.

From the machine computed data and Equation 3, k_g was calculated. The values of β_g and α_g thus obtained have been converted into relative phase velocity and total attenuation of the antenna in db. These values are plotted in Figure 7 and 8. It should be noted

that the curves for each of the spacings is similar except displaced in frequency. The low frequency cut-off is apparent, with a lesser amount of loading resulting in a lower cut-off frequency.

To determine the validity of the theoretical analysis, both relative phase velocity and total attenuation were experimentally measured. The relative phase velocity measurement was done in two ways. In the first method, the direction of maximum radiation from the antenna was measured. From Equation 1 we see that the radiation should emerge at an angle from the axis whose cosine is $1/\delta$ where δ is the relative phase velocity. Although there are factors present which modify this direction slightly in the actual case, the relationship is essentially correct except very close to end-fire. The second method consisted of short circuiting the load end of the line so as to set up a standing wave along the antenna. The near field along the antenna was then probed to determine the wavelength of the standing wave. This wavelength divided by the wavelength of free space is the relative phase velocity. The comparison of calculated and measured values of relative phase velocity for the 2 cm. pin spacing is shown in Figure 9. It should be noted that the divergence of the calculated and measured velocity is greatest as cut-off is approached. It is precisely in this region that the measured values are least accurate. This is so because in the radiation pattern method the beam becomes quite broad and it is difficult to accurately determine the direction of the maximum. In the standing wave method, the high attenuation near cut-off masks the standing wave pattern and it is hard to determine the wave length. However, the agreement away from cut-off appears quite good.

The total attenuation was experimentally determined by measuring the insertion loss due to the antenna in a matched coaxial system. The comparison between the measured and calculated attenuation for the 2 cm. pin spacing is shown in Figure 10. The slight difference in these curves can be accounted for by the loss in the coaxial to fast wave line transitions at either end of the antenna, and the small but finite ohmic loss present in the antenna.

Measured radiation patterns at three typical frequencies, for the 2 cm. pin spacing are shown in Figure 11. It can be seen that as frequency increases from cut-off, the beam tilts away from broadside toward the endfire direction. Also as frequency increases, the beamwidth narrows since the attenuation is decreased and more of the antenna aperture is illuminated. The lower attenuation is also manifested by more side lobes and deeper nulls. The theoretical antenna pattern was calculated by assuming the parallel line to consist of two anti-phase magnetic current line sources separated by the width of the parallel plates. The values of relative phase velocity and attenuation used were those obtained from the machine solution. A typical theoretical pattern is the dashed curve in Figure 11 which may be compared with the pattern measured at the same frequency.

Since agreement between measured and calculated values of relative phase velocity, attenuation constant and radiation patterns appeared to be very good for the three degrees of loading

investigated, it was felt that an understanding of the principles of operation of the antenna rested on a sound theoretical basis. It appeared reasonable therefore, to investigate the means by which the advantages of this type of antenna could be incorporated into practical embodiments. Usually it is desirable that an antenna produce a single unidirectional beam rather than the "rabbit ears" radiated by the parallel plate line antenna. By locating a flat strip reflector parallel to and adjacent to one aperture of the line, the beam impinging on the reflector could be made to radiate in the same direction as and reinforce the directly radiated beam. Proper spacing of the reflector resulted in a clean, unidirectional beam. It was found that the presence of the reflector had only a very minor effect on the relative phase velocity, so that the beam was radiated at the angle indicated by theory. With the 4 cm. pin spacing, a well formed beam could be supported below 2500 Mc/s. A conventional slot antenna would require a backing waveguide at least three inches wide to operate at the same frequency. When the bulky accessory waveguide to coaxial adapters are considered, the savings in size and weight offered by the fast wave parallel line can well be imagined.

A high gain aperture antenna may be formed if two of these unidirectional antennas are arranged in the form of a "V" and fed in parallel from the vertex as shown in Figure 12. The condition for a uniform phase front across the aperture of the "V" is that the half-angle at the vertex be equal to the arc-cosine of the reciprocal of the relative phase velocity. It can thus be seen that an antenna with a relatively large vertex angle can be made to have a uniform phase distribution across the aperture. A sectoral horn of comparable aperture would have to be long and with a small vertex angle in order to approximate a uniform phase distribution. Figure 13 shows the pattern obtained with an aperture of approximately eleven wavelengths and a vertex angle of 120°. A sectoral horn of the same aperture would have a vertex angle of about 20° in order to approximate the same pattern data. In addition to being larger, the sectoral horn would be completely enclosed by metal thus adding significantly to its weight.

It should be noted that the correct vertex angle for the fast wave "V" antenna depends only on the relative phase velocity and not on the length of the arms. This is in contrast to the case for the sectoral horn and the low frequency long wire "V" antenna where the optimum vertex angle is a complicated function of frequency and arm length. Even when this optimum angle is obtained the aperture phase distribution has appreciable curvature with resultant loss in gain.

The antennas investigated thus far have all had uniform loading with the result that the field intensity along the individual fast wave antenna drops off exponentially as distance from the feed point increases. Very often it is desirable for the illumination to have a symmetrical taper about the center of the antenna. Means for calculating the proper attenuation constant in terms of the required aperture distribution are given in Reference 3. Physically, the proper attenuation constant may be obtained by varying the spacing between the plates of the parallel plate line, since this variation has a

major effect on the attenuation constant and only a minor effect on the relative phase velocity. Any effect on the relative phase velocity may be compensated for by adjusting the pin spacing locally so as to maintain the phase velocity uniform along the length of the antenna.

In addition to the embodiments mentioned, many practical applications of the fast wave radiator may be imagined. Among these are methods of scanning the radiated beam. The use of frequency variation to control the beam pointing, or frequency diversity to receive from a multiplicity of directions is obvious from the analysis, since beam pointing is a function of relative phase velocity, and this in turn depends on frequency. At a given frequency, the beam may be steered by mechanical means either by changing the separation between the pins and/or changing the pin diameter. Electrically, the beam steering may be accomplished by replacing the pins with hollow dielectric cylinders filled with ionized gas. The change in loading is obtained by varying the degree of ionization by means of a d.c. potential between the plates. Clearly, considerable research would have to be done before such a scheme could be brought to a practical state. In any case, this type of antenna would be limited to low power levels.

Leaky wave antennas of the general kind discussed also have application where it is desired to synthesize beam patterns of a required complex shape. The precise and independent control of both attenuation and relative phase velocity are exactly the requirements necessary for accurate beam synthesis.

The place of leaky wave antennas in the field of standing wave antennas still remains to be explored. Often it is desired to have a linear antenna of the broadcast or F.M. type with the gain and directivity of an aperture several half-wavelengths long, but without the phase reversals that accompany each half wavelength loop of current, since the reversals reduce the gain and distort the pattern. By using a low attenuation, high phase velocity antenna which is physically several half-wavelengths of the operating frequency long, an antenna with only one-half wavelength loop of current may be obtained.

In conclusion, a fast wave antenna has been described which allows accurate analysis of attenuation constant and relative phase velocity. The advantages offered by this type of antenna in the way of light weight, compactness, efficiency and flexibility indicate many possible applications in the field of military and commercial antennas.

REFERENCES

1. V. H. Rumsey, "Travelling Wave Slot Antennas," Journal of Applied Physics, Vol. 24, pp. 1358-1365, November, 1953.

2. W. Rotman and S. Nauman, "The Design of Trough Waveguide Antenna Arrays," Air Force Cambridge Research Center Report, TR-58-154, ASTIA Doc. No. 152427, June, 1958.

3. R. C. Honey, "A Flush Mounted Leaky Wave Antenna with Predictable Patterns," IRE Transactions on Antennas and Propagation,

Vol. AP-7, pp. 320-328, October, 1959.

4. L. O. Goldstone and A. A. Oliner, "Leaky Wave Antennas I: Rectangular Waveguides," IRE Transactions on Antennas and Propagation, Vol. AP-7, pp. 307-319, October, 1959.

5. L. O. Goldstone and A. A. Oliner, "Leaky Wave Antennas II: Circular Waveguides," IRE Transactions on Antennas and Propagation, Vol. AP-9, pp. 280-290, May, 1961.

6. N. Marcuvitz, "Waveguide Handbook," McGraw-Hill Book Co. Inc., New York, Toronto and London, 1951, p. 179.

7. N. Marcuvitz, op. cit, p. 286.

8. G. G. MacFarlane, "Surface Impedance of an Infinite Parallel-Wire Grid at Oblique Angles of Incidence," Journal of the Institute of Electrical Engineers, Vol. 93, Part IIIa, pp. 1523-1527, 1946.

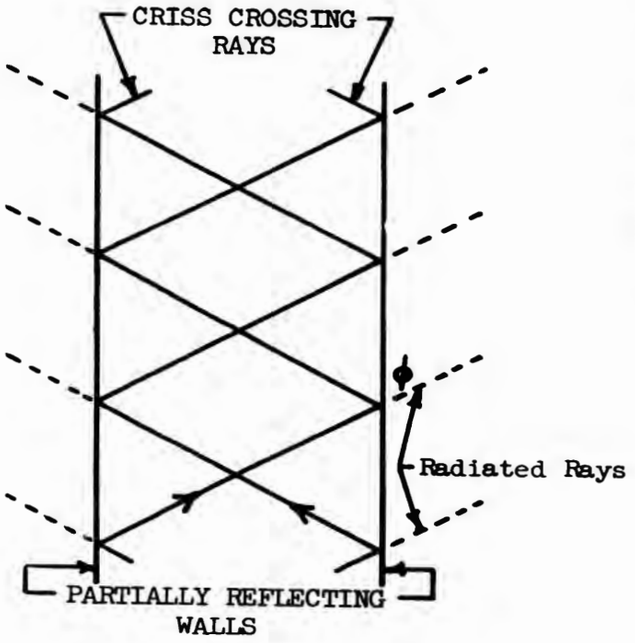


FIG. 1 - LEAKY WAVE MECHANISM

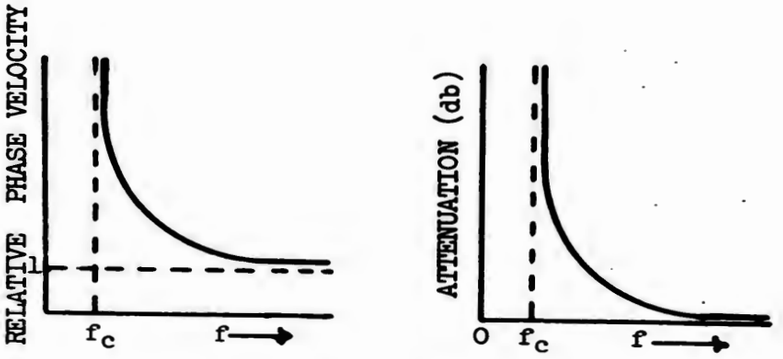


FIG. 2 - IDEAL FAST WAVE LINE BEHAVIOR

FIG. 3 - FAST WAVE LINE

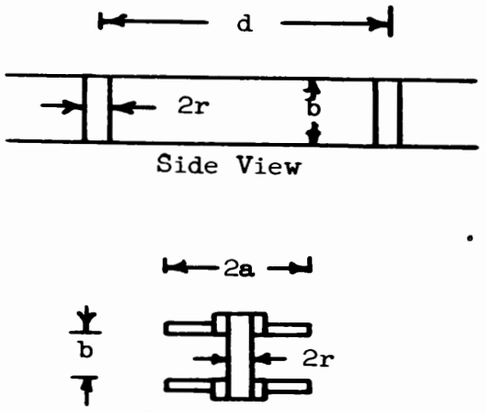


FIG. 4 - SCHEMATIC OF FAST WAVE LINE

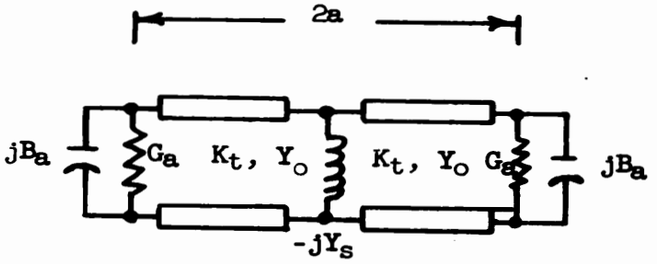


FIG. 5 - EQUIVALENT CIRCUIT OF TRANSVERSE SECTION

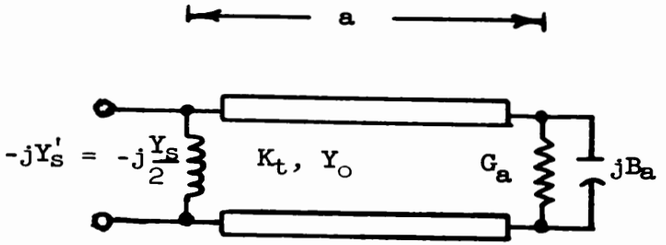


FIG. 6 - BISECTED EQUIVALENT CIRCUIT

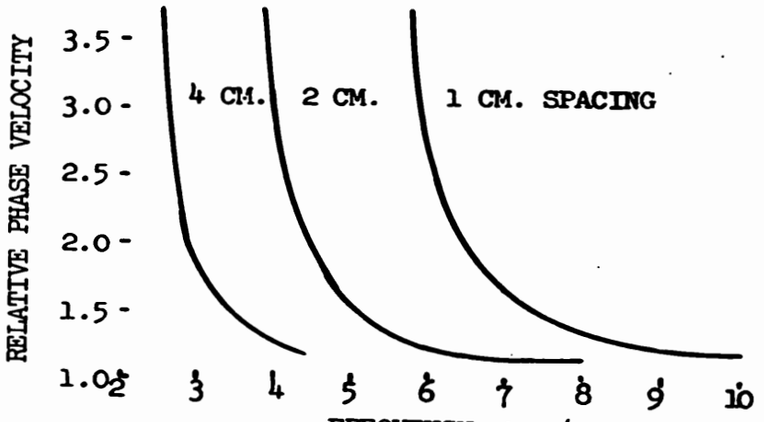


FIG. 7 - RELATIVE PHASE VELOCITY vs FREQUENCY

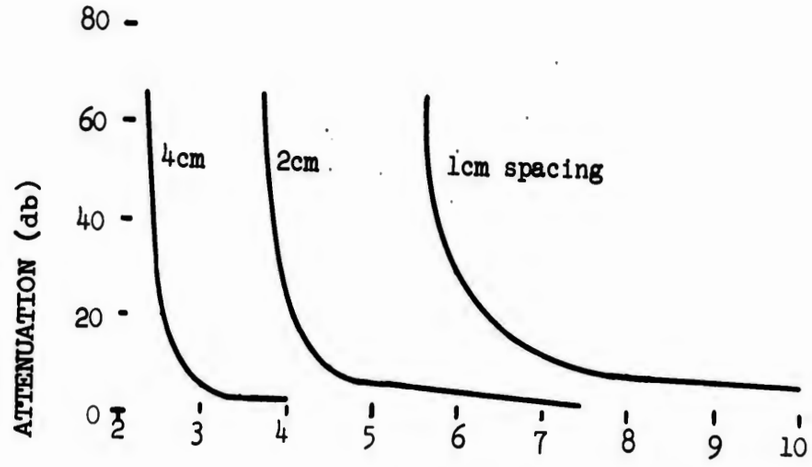


FIG. 8 - TOTAL ATTENUATION vs FREQUENCY

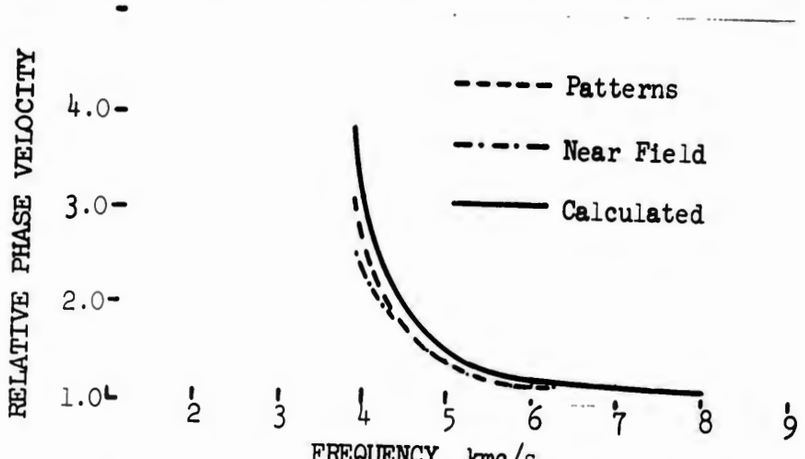


FIG.9-COMPARISON OF MEASURED AND CALCULATED RELATIVE PHASE VELOCITY

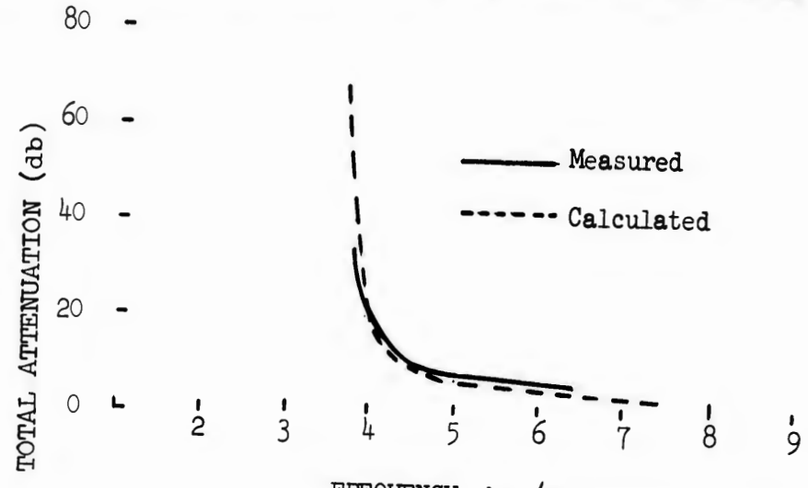


FIG.10 - COMPARISON OF MEASURED AND CALCULATED ATTENUATION

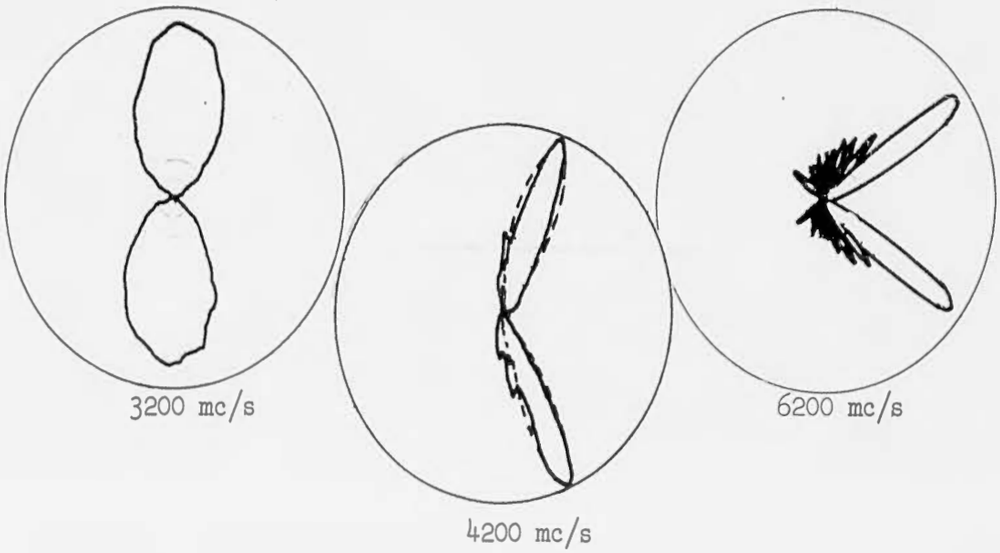


FIG. 11 - TYPICAL RADIATION PATTERNS

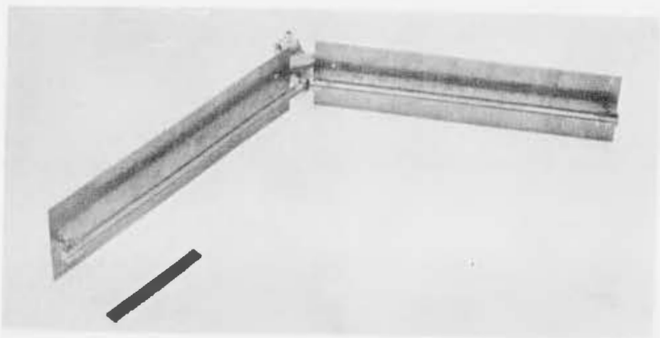


FIG. 12 - FAST WAVE "V" ANTENNA

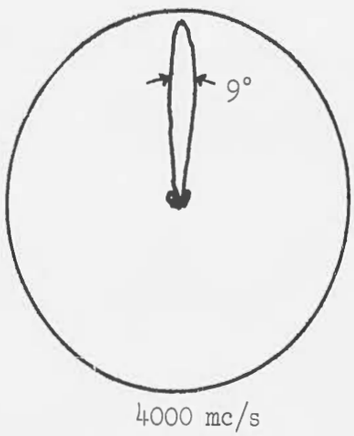


FIG. 13 - "V" ANTENNA RADIATION PATTERN

ORGANOMETALLIC RESEARCH USING GROUP IV ELEMENTS

MALCOLM C. HENRY, JAN G. NOLTES, WENZEL DAVIDSON, and ADOLF KREBS [a]
QUARTERMASTER RESEARCH & ENGINEERING COMMAND
NATICK, MASSACHUSETTS

The business of metal organic chemicals has recently been described as a billion dollar a year operation (1). Research in this field, broadly termed organometallics, is diversified and complex. This is a result of the variety of scientific disciplines being applied to a field which may involve any one of the elements in the periodic table. Generally speaking, it is more or less common for organically oriented scientists to concentrate either on the group IVB or VB elements which form stable covalent bonds with carbon or to be interested in the alkali or alkaline earth elements by means of which organic intermediates may be formed. Inorganic chemists, on the other hand, tend to concentrate on other groups of elements within the periodic table. Transition metal chemistry or the study of pi-complexes, for example, are more attractive to scientists oriented in classical inorganic techniques. In a similar manner physical and analytical chemists join the organometallic "elite" and their contribution to the field centers around areas in which their specific talents may be most profitably extended.

The program at the Quartermaster Laboratories has been organically oriented, due first of all to our basic interests in organic chemistry, and, secondly, because of the obvious potential application of this area of chemistry to Quartermaster end items. In cooperation with the Germanium Research Committee, the Lead Industries Association, and others, we have been for several years concerned with a basic program dealing with certain aspects of the organic chemistry of germanium, tin and lead.

Although it is true that organic silicon chemistry has attracted the major share of research activity for a number of years, recent interest in germanium, tin and lead chemistry has shown signs of intensifying. This may partially be attributed to the large

(a) J. G. N., WD, & AK, Research fellows, sponsored by Am. Zinc Inst., Germanium Res. Committee, & Lead Ind. Assoc., respectively.

differences in biological activity when comparing similar compounds of silicon, germanium, tin and lead. Silicon compounds in general are non-active biologically whereas all lead compounds are potentially active. In between these two extremes it is expected to find an entire spectrum of activities. Several organotin compounds, for example, have already been developed as commercial items in the fungicide field.

One of the most interesting recent developments in the chemistry of these IVth group elements has been the reaction between the alkyl and aryl metal hydrides and olefinic and acetylenic linkages. This reaction has made it possible to take the organic chemistry of these elements out of the isolated realm of solely metal-hydrocarbon chemistry. This work, pioneered by G. J. M. van der Kerk and J. G. Noltes of the Organic Chemistry Institute, TNO, Utrecht, Holland, was originally concerned with organotin compounds. Under the auspices of a Secretary of the Army Research Fellowship, we (MCH) were able to participate in some of the later stages of this development as a part of Prof. van der Kerk's research team. The reaction and the versatility of this reaction have been amply demonstrated (2), [Slide # 1]. Extensions of this work have resulted in successfully adding mono- and dihydrides to olefinic and acetylenic linkages yielding many new types of monomers (3), polymers(4), ring compounds (4, 5) as well as a series of compounds containing different group IV elements (4, 6) [Slide # 2].

Considerable work has now been done with silicon hydride additions (7) using catalysts, with germanium hydrides both with (8) and without catalysts(9) as well as with organolead hydrides(10). The later case, however, is limited in scope due to the instability of the organolead hydrides.

Extensions of our initial work in organogermanium chemistry led us to attempt the synthesis of some new organogermanium-sulphur compounds. The extension of a known organosilicon reaction, namely, that between triphenylsilyl lithium and elemental sulphur, was found to be complex and unsatisfactory. Since an alternate route was indicated, we felt it might be possible to synthesize some organo-sulphur compounds if sodium triphenyl-germanethiol could be initially prepared from the reaction of triphenylbromogermane and sodium sulphide. Subsequent reaction with RX type compounds would yield the desired compounds, namely, of the type $\equiv GeSR$ [Slide 3]. The successful preparation and isolation of sodium triphenylgermanethiol as a pure compound has made it possible to synthesize a variety of new types of organogermanium-sulphur compounds (11). The reaction with alkyl halides and aryl halides with sodium triphenylgermanethiol is a clean, general reaction that yields thioalkyl- and thioaryl-substituted triphenylgermanes. Reaction of the sodium salt with organic sulphonyl halides results in bis-triphenylgermanium di-

sulphide and the reaction of two equivalents of the sodium salt with sulphur dichloride yields the yellow crystalline compound, bis-triphenylgermanium trisulphide. Reaction of two moles of the sodium salt with diphenyldibromogermane yields the octaphenyl compound containing three germanium atoms and two atoms of sulphur, bis-triphenylgermanethio diphenylgermane. The synthetic possibilities of this reaction are currently being investigated and extended.

The analagous reaction could not be satisfactorily extended to the case of lead due to the instability of the initial sodium salt of triphenylleadthiol. [Slide #4] This salt even in solution decomposes to bis-triphenyllead sulphide. Other possibilities exist for synthesizing organolead-sulphur compounds as for example the reaction of triphenyllead hydroxide or tetraphenyllead with mercaptans. These two possibilities are not as good experimentally as the lead mercaptide method. Via this method, namely, the reaction between triphenyllead chloride and suitable lead (II) mercaptides, a series of organolead mercaptides has been prepared of the type Ar_3PbSR where R is methyl-, ethyl-, propyl-, butyl-, phenyl-, benzyl-, acetyl- and benzoyl- (12). With the exception of the butyl- compound which is a liquid, these compounds are all reasonably stable, colorless compounds having a slight odor reminiscent of the initial mercaptan. In all cases the lead sulphur bond is readily cleaved by mineral acids to form the corresponding triphenyllead halide and the free mercaptan.

A third area of interest has been in synthesizing organolead arsonates containing basically the sequence $\equiv PbOAs(O)-$ similar to inorganic lead arsonate but containing organic groupings attached to both the lead atom and the arsenic atom. [Slide #5] Organometallic compounds containing M-O-M units are well known for molecules where M atoms are identical and with M-O-M' units where M and M' are dissimilar but of the same group or subgroup in the periodic table. Less well known are compounds containing a group IV element connected via oxygen to a group V element, nitrogen, phosphorous and carbon excluded.

The synthesis of M-O-As type compounds has been described for the cases where M= silicon, germanium and tin but was unknown for the case where M = lead. This latter type of compound may be synthesized by reaction of triphenyllead chloride with organic substituted arsonic or arsinic acids or their salts. The products are substituted organolead arsonates and arsinates respectively. [Slide #6] Disubstituted lead diarsonates may be synthesized by reaction of tetraphenyllead and free organic substituted arsonic acids in methanol or between triphenyllead chloride and substituted arsonic acids in pyridine. Reaction of diphenyllead dichloride with disodium methyl arsonate yields an organolead arsonate of uncertain structure containing the repeating

unit $\left[(C_6H_5)_2\ PbOAs(O)(CH_3)\right]_x$. These compounds are all non-melting solids (up to 250°) and are insoluble in common organic solvents (13).

The results of our basic research program are currently being brought to the attention of testing laboratories for possible application. In this regard the potential value of these compounds with respect to biological activity, as lubricants, adhesives, impregnating agents, accelerators and herbicides is to be evaluated. New ideas for possible applications are being continually sought so that these new substances do not remain unknown laboratory curiosities.

REFERENCES

(1) A. R. Anderson, Chemical Week, Sept. 17, 1960 - "Abundance of Metal Organics".

(2) G. J. M. van der Kerk, J. G. A. Luijten and J. G. Noltes. Chem. & Ind., 352 (1956); id, J. Appl. Chem., 7, 356 (1957); id, Angew. Chem., 70, 298 (1958).

(3) G. J. M. van der Kerk and J. G. Noltes, J. Appl. Chem., 9, 106 (1959).

(4) J. G. Noltes and G. J. M. van der Kerk, Rec. trav. chim., 80, 623 (1961); id, ibid. 80, 1292 (1961); id, Chimia, 16, in press (1962).

(5) M. C. Henry and J. G. Noltes, J. Am. Chem. Soc., 82, 561 (1960).

(6) M. C. Henry and J. G. Noltes, J. Am. Chem. Soc., 82, 558(1960).

(7) C. Eaborn, Organosilicon Compounds, Academic Press (1959).

(8) J. Satgé, Annales de Chemie, 6, 519 (1961).

(9) M. C. Henry and M. F. Downey, J. Org. Chem., 26, 2299 (1961).

(10) W. E. Becker and S. E. Cook, J. Am. Chem. Soc., 82, 6264 (1960).

(11) M. C. Henry and W. E. Davidson, J. Org. Chem., 27, in press (1962).

(12) M. C. Henry and A. Krebs - unpublished results.

(13) M. C. Henry, J. Org. Chem. - in press (1962).

ORGANOTIN HYDRIDE REACTIONS WITH OLEFINS

$R_3SnH + CH_2 = CHR^1 \rightarrow R_3SnCH_2CH_2R^1$

$R_2SnH_2 + 2CH_2 = CHR^1 \rightarrow R^1CH_2CH_2Sn(R)_2CH_2CH_2R^1$

R = Alkyl, Aryl

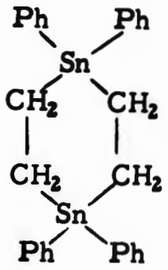
R^1= Alkyl, Aryl, -CN, -COOH, -COOR, $-CONH_2$, -OH, -OCOR, $-CH=CH_2$

SLIDE 1

ORGANOTIN HYDRIDE ADDITIONS

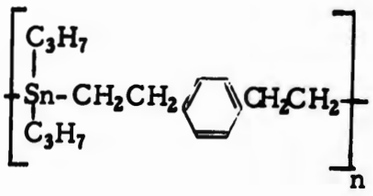
Ring Compound Formation

$(Ph)_2SnH_2 + (Ph)_2Sn(CH=CH_2)_2 \rightarrow$



Polymer Formation

$(C_3H_7)_2SnH_2 + CH_2 = CH$ $CH=CH_2 \rightarrow$



Compounds with 2-IVA Elements

$Ph_2SnH_2 + 2Ph_3MCH=CH_2 \rightarrow Ph_3MCH_2CH_2SnCH_2CH_2MPh_3$ (Ph)

SLIDE 2

SODIUM TRIPHENYLGERMANE: Some Reactions

$Ph_3GeBr + Na_2S \rightarrow Ph_3GeSNa + NaBr$

$Ph_3GeSNa + RX \rightarrow Ph_3GeSR$

R = $-CH_3$, $-C_6H_5$, $-C_2H_5$, $-CH_2C_6H_5$, $-C(O)$ NO_2

$-Ge(C_6H_5)_3$, $-CH_2SCH_3$

$$2\ Ph_3GeSNa + Ph_2GeBr_2 \rightarrow Ph_3GeS\underset{\displaystyle Ph}{\overset{\displaystyle Ph}{Ge}}SGePh_3$$

$2\ Ph_3GeSNa + SCl_2 \rightarrow Ph_3GeSSSGePh_3$

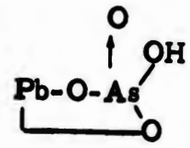
SLIDE 3

POSSIBLE REACTIONS LEADING TO $(C_6H_5)_3Pb\ SR$-COMPOUNDS

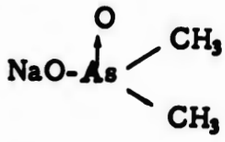
1. $(C_6H_5)_3\ PbSNa + RX \rightarrow (C_6H_5)_3\ Pb\ SR + NaX$
2. $(C_6H_5)_3Pb\ OH + RSH \rightarrow (C_6H_5)_3\ Pb\ SR + H_2O$
3. $(C_6H_5)_4\ Pb + RSH \rightarrow (C_6H_5)_3\ Pb\ SR + C_6H_6$
4. $2(C_6H_5)_3\ PbX + Pb(SR)_2 \rightarrow (C_6H_5)_3\ Pb\ SR + PbX_2$

SLIDE 4

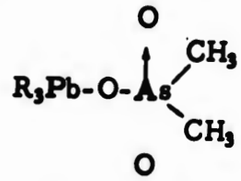
Lead Arsonate



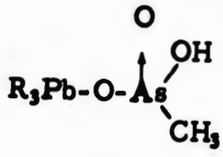
Selective Herbicides (e.g. Dimet)



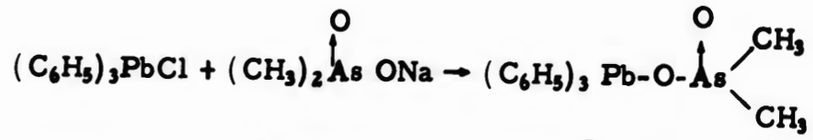
Organolead Arsinate

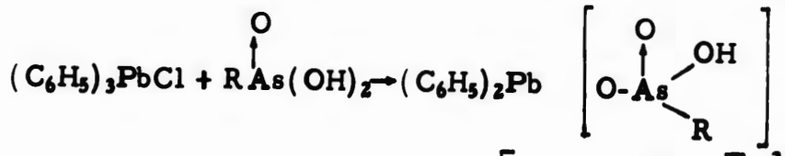


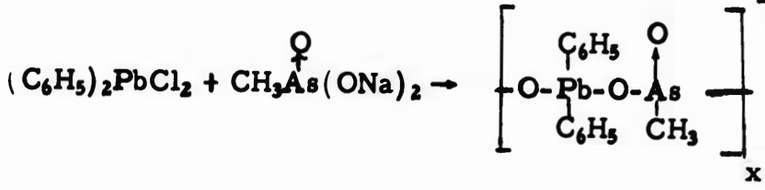
Organolead Arsonate



SLIDE 5







SLIDE 6

BEHAVIOR OF NON-NEWTONIAN FLUIDS UNDER CONDITIONS OF HIGH PRESSURE RAPID ACCELERATION AND HIGH VELOCITY

B. A. HOWARD, JR.
HEADQUARTERS
ORDNANCE WEAPONS COMMAND
ROCK ISLAND, ILLINOIS

Pseudo Plastic and Dilatant Fluids

The simple shear flow of pseudo-plastic and dilatant fluids is expressed with considerable exactness by Waele-Ostwald's formula as follows:

$$\frac{du}{dy} = \frac{\tau^n}{\mu_{psu}} \tag{1}$$

where μ_{psu} is the quantity similar to the viscosity of Newtonian fluid, τ is the shearing stress, du/dy is the velocity gradient, n is the rheological constant, The flow of pseudo-plastic fluids is expressed by $n > 1$, and that of dilatant fluids is given by $n < 1$ in the above equation, respectively.

The relation between the stress components and rate-of-strain components for non-Newtonian fluids can be expressed as:

$$p_{ik} = \eta e_{ik} - p\delta_{ik} \tag{2}$$

Where p_{ik} is a component of the stress tensor (p_{ik}), as indicated in the following array:

$$(p_{ik}) = \begin{bmatrix} p_{xx} & p_{xy} & p_{xz} \\ p_{yx} & p_{yy} & p_{yz} \\ p_{zx} & p_{zy} & p_{zz} \end{bmatrix} \tag{3}$$

and e_{ik} is a component of the rate-of-strain tensor (e_{ik}) which is expressed as:

$$(e_{ik}) = \begin{bmatrix} e_{xx} & \frac{1}{2}e_{xy} & \frac{1}{2}e_{xz} \\ \frac{1}{2}e_{yx} & e_{yy} & \frac{1}{2}e_{yz} \\ \frac{1}{2}e_{zx} & \frac{1}{2}e_{zy} & e_{zz} \end{bmatrix} \tag{4}$$

These tensors are symmetric tensors, that is:

$$p_{ik} = p_{ki}\ ,\quad e_{ik} = e_{ki} \qquad (i \neq k)$$

p_{ik} is a normal stress if $i = k$, a tangential or a shear stress if $i \neq k$, p is the pressure, which is the mean of the normal stresses over three planes mutually at right angles. δ_{ik} is the Kronecker delta ($\delta_{ik} = 1$ if $i=k$, $\delta_{ik} = 0$ if $i=k$; $i,k = x,y,z$). u, v, w are parallel velocity components to the coordinate axes of rectangular Cartesian coordinates x,y, z, respectively. e_{xx}, e_{yy}, etc. are

$$e_{xx} = \frac{\Delta u}{\Delta x}\ ,\quad e_{xy} = \frac{\Delta v}{\Delta x} + \frac{\Delta u}{\Delta y} \quad \text{etc.} \quad (\Delta \text{ denotes partial derivative})$$

η is an arbitrary scalar function of the three invariants of the rate-of-strain, and may be expressed as

$$\Theta = e_{xx} + e_{yy} + e_{zz} \qquad (5)$$

$$e = e_{yy}e_{zz} + e_{zz}e_{xx} + e_{xx}e_{yy} - \tfrac{1}{4}(e_{yz}^2 + e_{zy}^2 + e_{xy}^2) \qquad (6)$$

$$|e_{ik}| = e_{xx}e_{yy}e_{zz} + \tfrac{1}{4}\, e_{yz}e_{zx}e_{xy} - \tfrac{1}{4}(\, e_{xx}e_{yz}^2 + e_{yy}e_{zx}^2 + e_{zz}e_{xy}^2\,) \qquad (7)$$

If we treat the fluid as being incompressible, the first invariant Θ vanishes, or

$$e_{xx} + e_{yy} + e_{zz} = 0 \qquad (8)$$

Consequently, η can be expressed as an arbitrary scalar function of the two invariants e, $|e_{ik}|$. In rectilinear flow and in two-dimensional flow, the third invariant $|e_{ik}|$ is identically zero. Hence η can be expressed as a scalar function of the second invariant e only in three-dimensional flow. Taking account of eq. (8), the second invariant e becomes

$$e = -\tfrac{1}{4}[\, 2(e_{xx}^2 + e_{yy}^2 + e_{zz}^2) + e_{yz}^2 + e_{zx}^2 + e_{xy}^2\,] \qquad (9)$$

If we generalize Waele-Ostwald's formula by assuming that the stress components are, in general, expressible by the 1/n th power functions of the rate-of-strain components for non-Newtonian fluids, the relations between the stress components and the rate-of-strain components may be expressed as

$$\begin{aligned} p_{ii} &= -p + 2\mu_{psu}^{1/n}\, e_{ii}\, \Theta, \quad (i = x, y, z) \\ p_{ik} &= \mu_{psu}^{1/n}\, e_{ik}\, \Theta, \quad (i,k, = x, y, z;\ i \neq k) \end{aligned} \qquad (10)$$

$$\Theta = [2(e_{xx}^2 + e_{yy}^2 + e_{zz}^2) + e_{yz}^2 + e_{zx}^2 + e_{xy}^2\,]^{(1-n)/2n} \qquad (11)$$

If we include the kinetic reactions and the external forces, and consider the equilibrium of an infinitesimal rectangular parallelepiped with its centroid at a point P(x, y, z,) and its edges parallel to the coordinate axes, then by resolving the resultant forces along

the axes, we obtain

$$\rho \frac{Du}{Dt} = \rho X + \frac{\Delta p_{xx}}{\Delta x} + \frac{\Delta p_{xy}}{\Delta y} + \frac{\Delta p_{xz}}{\Delta z}$$

$$\rho \frac{Dv}{Dt} = \rho Y + \frac{\Delta p_{yx}}{\Delta x} + \frac{\Delta p_{yy}}{\Delta y} + \frac{\Delta p_{yz}}{\Delta z} \qquad (12)$$

$$\rho \frac{Dw}{Dt} = \rho Z + \frac{\Delta p_{zx}}{\Delta x} + \frac{\Delta p_{zy}}{\Delta y} + \frac{\Delta p_{zz}}{\Delta z}$$

where X, Y, Z are the components of the external force acting on the fluid per unit mass, ρ is the density, and Du/Dt, Dv/Dt, and Dw/Dt are the components of the acceleration of a particle of the fluid. Substituting eq. (10) in eq. (12), then making use of eq (8) we have the fundamental equations of motion for non-Newtonian fluid in the form

$$\rho \frac{Du}{Dt} = \rho X - \frac{\Delta p}{\Delta x} + \mu_{psu}^{1/n} \left[\Theta \nabla^2 u + 2 e_{xx} \frac{\Delta \Theta}{\Delta x} + e_{xy} \frac{\Delta \Theta}{\Delta y} + e_{xz} \frac{\Delta \Theta}{\Delta z} \right]$$

$$\rho \frac{Dv}{Dt} = \rho Y - \frac{\Delta p}{\Delta y} + \mu_{psu}^{1/n} \left[\Theta \nabla^2 v + e_{yx} \frac{\Delta \Theta}{\Delta x} + 2 e_{yy} \frac{\Delta \Theta}{\Delta y} + e_{yz} \frac{\Delta \Theta}{\Delta z} \right] \qquad (13)$$

$$\rho \frac{Dw}{Dt} = \rho Z - \frac{\Delta p}{\Delta z} + \mu_{psu}^{1/n} \left[\Theta \nabla^2 w + e_{zx} \frac{\Delta \Theta}{\Delta x} + e_{zy} \frac{\Delta \Theta}{\Delta y} + 2 e_{zz} \frac{\Delta \Theta}{\Delta z} \right] [$$

where $\nabla^2 \equiv \frac{\Delta^2}{\Delta x^2} + \frac{\Delta^2}{\Delta y^2} + \frac{\Delta^2}{\Delta z^2}$

The equation of continuity is $e_{xx} + e_{yy} + e_{zz} = 0$ (8')

This is treating the fluid as being incompressible. Parameters should be theorized for variation in density for fluid under pressure and eqs. (12) and (13) should be solved with the appropriate terms inserted. This will be investigated in another paper.

If n = 1, eq. (13) becomes the Navier-Stokes equation for Newtonian fluid. The rheological constant n has the different values for varied fluids and n, in general, seldom has an integral value. The approximate solution by the energy method may be applied for non-Newtonian flow. The dissipation of energy in non-Newtonian flow may be calculated by

$$I = \int\int\int [e_{xx} p_{xx} + e_{yy} p_{yy} + e_{zz} p_{zz} + e_{yz} p_{yz} + e_{zx} p_{zx} + e_{xy} p_{xy}] [dxdydz]$$

Substituting eq. (10) into this equation and then using eq (8') we get

$$I = \mu_{psu}^{1/n} \int\int\int [2(e_{xx}^2 + e_{yy}^2 + e_{zz}^2) + e_{yz}^2 + e_{zx}^2 + e_{xy}^2]^{(n+1)/2n} dxdydz \qquad (14)$$

so that the dissipation function ϕ, which is the rate of dissipation of energy per unit time per unit volume is obtained as

$$\phi = \mu_{psu}^{1/n} [2(e_{xx}^2 + e_{yy}^2 + e_{zz}^2) + e_{yz}^2 + e_{zx}^2 + e_{xy}^2]^{(n+1)/2n} \qquad (15)$$

Now if the velocity distribution of flow gets the minimum dissipation

of energy, that is $\delta I = \delta \iiint \Phi \, dxdydz = 0$ (16)

It may be proved that the velocity distribution also satisfies the equation of motion in which inertia terms are neglected and external forces derived from a potential or zero, that is

$$\begin{aligned}
&\Theta \nabla^2 u + 2e_{xx}\frac{\Delta\Theta}{\Delta x} + e_{xy}\frac{\Delta\Theta}{\Delta y} + e_{xz}\frac{\Delta\Theta}{\Delta z} = \frac{\Delta H}{\Delta x} \\
&\Theta \nabla^2 v + e_{yx}\frac{\Delta\Theta}{\Delta x} + 2e_{yy}\frac{\Delta\Theta}{\Delta y} + e_{yz}\frac{\Delta\Theta}{\Delta z} = \frac{\Delta H}{\Delta y} \qquad (17) \\
&\Theta \nabla^2 w + e_{zx}\frac{\Delta\Theta}{\Delta x} + e_{zy}\frac{\Delta\Theta}{\Delta y} + 2e_{zz}\frac{\Delta\Theta}{\Delta z} = \frac{\Delta H}{\Delta z}
\end{aligned}$$

Where H is a function of x, y, z. From eq. (15) we have

$$\begin{aligned}
\delta\Phi = \frac{n+1}{n}\mu_{psu}^{1/n}\Theta[&2\frac{\Delta u}{\Delta x}\frac{\Delta\delta u}{\Delta x} + 2\frac{\Delta v}{\Delta y}\frac{\Delta\delta v}{\Delta y} + 2\frac{\Delta w}{\Delta z}\frac{\Delta\delta w}{\Delta z} + (\frac{\Delta w}{\Delta y} + \frac{\Delta v}{\Delta z})(\frac{\Delta\delta w}{\Delta y} + \frac{\Delta\delta v}{\Delta z}) \\
&+ (\frac{\Delta u}{\Delta z} + \frac{\Delta w}{\Delta x})(\frac{\Delta\delta u}{\Delta z} + \frac{\Delta\delta w}{\Delta x}) + (\frac{\Delta v}{\Delta x} + \frac{\Delta u}{\Delta y})(\frac{\Delta\delta v}{\Delta x} + \frac{\Delta\delta u}{\Delta y})\,]
\end{aligned}$$

Therefore the variation δI of I becomes

$$\begin{aligned}
\delta I = \frac{n+1}{n}\mu_{psu}^{1/n}\iiint\Theta[&2\frac{\Delta u}{\Delta x}\frac{\Delta\delta u}{\Delta x} + 2\frac{\Delta v}{\Delta y}\frac{\Delta\delta v}{\Delta y} + 2\frac{\Delta w}{\Delta z}\frac{\Delta\delta w}{\Delta z} + (\frac{\Delta w}{\Delta y} + \frac{\Delta v}{\Delta z}) \\
&(\frac{\Delta\delta w}{\Delta y} + \frac{\Delta\delta v}{\Delta z}) + (\frac{\Delta u}{\Delta z} + \frac{\Delta w}{\Delta x})(\frac{\Delta\delta u}{\Delta z} + \frac{\Delta\delta w}{\Delta x}) + (\frac{\Delta v}{\Delta x} + \frac{\Delta u}{\Delta y})(\frac{\Delta\delta v}{\Delta x} + \frac{\Delta\delta u}{\Delta y})\,]\,dxdydz \qquad (18)
\end{aligned}$$

Considering that δu, δv, δw are zero at the body surface add infinity and then using the integration by parts, the differential terms of δu, δv, and δw with respect to x become

$$\begin{aligned}
&\iiint\Theta[2\frac{\Delta u}{\Delta x}\frac{\Delta\delta u}{\Delta x} + (\frac{\Delta v}{\Delta x} + \frac{\Delta u}{\Delta y})\frac{\Delta\delta v}{\Delta x} + (\frac{\Delta u}{\Delta z} + \frac{\Delta w}{\Delta x})\frac{\Delta\delta w}{\Delta x}\,]dxdydz \\
&= \iint\Big\{\Big|\Theta[2\frac{\Delta u}{\Delta x}\delta u + (\frac{\Delta v}{\Delta x} + \frac{\Delta u}{\Delta y})\delta v + (\frac{\Delta u}{\Delta z} + \frac{\Delta w}{\Delta x})\delta w]\Big| - \int\frac{\Delta\Theta}{\Delta x}[2\frac{\Delta u}{\Delta x}\delta u + (\frac{\Delta v}{\Delta x} + \frac{\Delta u}{\Delta y})\delta v \\
&+ (\frac{\Delta u}{\Delta z} + \frac{\Delta w}{\Delta x})\delta w]dx - \int\Theta[\frac{\Delta^2 u}{\Delta x^2}\delta u + \frac{\Delta}{\Delta x}(\frac{\Delta v}{\Delta x} + \frac{\Delta u}{\Delta y})\delta v + (\frac{\Delta}{\Delta x})(\frac{\Delta u}{\Delta z} + \frac{\Delta w}{\Delta x})\delta w]dx\Big\}\,dydz \\
&= -\iiint\frac{\Delta\Theta}{\Delta x}[2\delta u\frac{\Delta u}{\Delta x} + \delta v(\frac{\Delta v}{\Delta x} + \frac{\Delta u}{\Delta y}) + \delta w(\frac{\Delta u}{\Delta z} + \frac{\Delta w}{\Delta x})]\,dxdydz \\
&- \iiint\Theta\,[2\delta u\frac{\Delta^2 u}{\Delta x^2} + \delta v\frac{\Delta}{\Delta x}(\frac{\Delta v}{\Delta x} + \frac{\Delta u}{\Delta y}) + \delta w\frac{\Delta}{\Delta x}(\frac{\Delta u}{\Delta z} + \frac{\Delta w}{\Delta x})]dxdydz
\end{aligned}$$

Since the differential terms of δu, δv, and δw wtih respect to y and z are transformed in the same way, eq. 18 may be written as

$$\begin{aligned}
\delta I = -\frac{(n+1)}{n}\mu_{psu}^{1/n}\iiint[&\delta u(\Theta\nabla^2 u + 2e_{xx}\frac{\Delta\Theta}{\Delta x} + e_{xy}\frac{\Delta\Theta}{\Delta y} + e_{xz}\frac{\Delta\Theta}{\Delta z}) + \delta v(\Theta\nabla^2 v \\
&+ e_{xy}\frac{\Delta\Theta}{\Delta x} + 2e_{yy}\frac{\Delta\Theta}{\Delta y} + e_{yz}\frac{\Delta\Theta}{\Delta z}) + \delta w(\Theta\nabla^2 w + e_{zx}\frac{\Delta\Theta}{\Delta x} + e_{zy}\frac{\Delta\Theta}{\Delta y} + 2e_{zz}\frac{\Delta\Theta}{\Delta z})
\end{aligned}$$

$$+ \Theta\left(\delta u\frac{\Delta}{\Delta x} + \delta w\frac{\Delta}{\Delta z} + \delta v\frac{\Delta}{\Delta y}\right)\left(\frac{\Delta u}{\Delta x} + \frac{\Delta v}{\Delta y} + \frac{\Delta w}{\Delta z}\right)\Big]\, dx dy dz$$

Utilizing eq. (8') and eq. (16) the above equation becomes

$$\iiint [\delta u(\Theta \nabla^2 u + 2e_{xx}\frac{\Delta\Theta}{\Delta x} + e_{xy}\frac{\Delta\Theta}{\Delta y} + e_{xz}\frac{\Delta\Theta}{\Delta z}) + \delta v(\Theta \nabla^2 v + e_{yx}\frac{\Delta\Theta}{\Delta x} + 2e_{yy}$$

$$\frac{\Delta\Theta}{\Delta y} + e_{yz}\frac{\Delta\Theta}{\Delta z}) + \delta w(\Theta \nabla^2 w + e_{zx}\frac{\Delta\Theta}{\Delta x} + e_{zy}\frac{\Delta\Theta}{\Delta y} + 2e_{zz}\frac{\Delta\Theta}{\Delta z})]\, dxdydz = 0 \quad (19)$$

Since the variations δu, δv, and δw must be satisfied by eq. (8) so that

$$\frac{\Delta\delta u}{\Delta x} + \frac{\Delta\delta v}{\Delta y} + \frac{\Delta\delta w}{\Delta z} = 0$$

Multiplying this equation by $\alpha(x, y, z)$ and then adding its results to eq. (19) produces the following equation: (20)

$$\iiint [\delta u(\Theta \nabla^2 u + 2e_{xx}\frac{\Delta\Theta}{\Delta x} + e_{xy}\frac{\Delta\Theta}{\Delta y} + e_{xz}\frac{\Delta\Theta}{\Delta z}) + \delta v(\Theta \nabla^2 v + e_{yx}\frac{\Delta\Theta}{\Delta x} +$$

$$2e_{yy}\frac{\Delta\Theta}{\Delta y} + e_{yz}\frac{\Delta\Theta}{\Delta z}) + \delta w(\Theta \nabla^2 w + e_{zx}\frac{\Delta\Theta}{\Delta x} + e_{zy}\frac{\Delta\Theta}{\Delta y} + 2e_{zz}\frac{\Delta\Theta}{\Delta z}) + \alpha(\frac{\Delta\delta u}{\Delta x} + \frac{\Delta\delta v}{\Delta y} + \frac{\Delta\delta w}{\Delta z})]\, dxdydz$$

equals zero. Since δu, δv, and δw are zero at the body surface and infinity, using integration by parts gives

$$\iiint \alpha\frac{\Delta}{\Delta x}\delta u\, dxdydz = \int\int [|\alpha\delta u| - \int\frac{\Delta\alpha}{\Delta x}\delta u\, dx]\, dydz = -\iiint \frac{\Delta\alpha}{\Delta x}\delta u\, dxdydz$$

Therefore eq. (20) becomes

$$\iiint [(\Theta \nabla^2 u + 2e_{xx}\frac{\Delta\Theta}{\Delta x} + e_{xy}\frac{\Delta\Theta}{\Delta y} + e_{xz}\frac{\Delta\Theta}{\Delta z} - \frac{\Delta\alpha}{\Delta x})\delta u + (\Theta \nabla^2 v + e_{yx}\frac{\Delta\Theta}{\Delta x} + 2e_{yy}\frac{\Delta\Theta}{\Delta y}$$

$$+ e_{yz}\frac{\Delta\Theta}{\Delta z} - \frac{\Delta\alpha}{\Delta y})\delta v + (\Theta \nabla^2 w + e_{zx}\frac{\Delta\Theta}{\Delta x} + e_{zy}\frac{\Delta\Theta}{\Delta y} + 2e_{zz}\frac{\Delta\Theta}{\Delta z} - \frac{\Delta\alpha}{\Delta z})\delta w]\, dxdydz = 0$$

As δu, δv, and δw may take arbitrary values in the domain of flow, the coefficients of δu, δv, and δw must be zero in order to always satisfy the above equation. Then, the same equation as eq. 17 is obtained as follows:

$$\Theta \nabla^2 u + 2e_{xx}\frac{\Delta\Theta}{\Delta x} + e_{xy}\frac{\Delta\Theta}{\Delta y} + e_{xz}\frac{\Delta\Theta}{\Delta z} = \frac{\Delta\alpha}{\Delta x}$$

$$\Theta \nabla^2 v + e_{yx}\frac{\Delta\Theta}{\Delta x} + 2e_{yy}\frac{\Delta\Theta}{\Delta y} + e_{yz}\frac{\Delta\Theta}{\Delta z} = \frac{\Delta\alpha}{\Delta y}$$

$$\Theta \nabla^2 w + e_{zx}\frac{\Delta\Theta}{\Delta x} + e_{zy}\frac{\Delta\Theta}{\Delta y} + 2e_{zz}\frac{\Delta\Theta}{\Delta z} = \frac{\Delta\alpha}{\Delta z}$$

Thus, the velocity distribution that satisfies eq. (16) is found to be the solution of eq. (17), if the inertia terms are neglected and external forces derived from a potential or zero. Therefore, if the flow satisfies the conditions mentioned above, the approximate solution of flow may be obtained by using this minimum dissipation of energy method: the velocity distribution containing some undetermined coefficients and satisfying the boundary conditions is appropriately assumed, and then the undetermined coefficients are determined so that the dissipation energy becomes minimum.

VISCO-ELASTICITY

Some of the characteristics of visco-elastic substances are desirable in the operation of hydraulic systems of aircraft, weapon systems and transportation equipment. In effect, there is a need to absorb energy without a shock result. In visco-elastic substances, the ratio of lateral to longitudinal strain in uni-axial stressing is not a constant but a variable parameter depending on time and the conditions of the test. A theoretical analysis of this parameter under various assumptions concerning the visco-elastic response of the material and the testing conditions as developed by Freudenthal and others is given. The Maxwell body more closely approximates the desired reaction. Promise of further approximation is shown in the characteristics of some long chain polymeric fluids which show a high bulk modulus under high pressures. The final solution to the problem will be accomplished by development of the desirable characteristics from a molecular structure approach.

The stress and strain components σ_{ij} and ε_{ij} may be written as S_{ij}, e_{ij} and σ_{kk}, ε_{kk}, respectively, so that

$$\sigma_{ij} = S_{ij} + 1/3\ \delta_{ij}\sigma_{kk} \quad \text{and} \quad \varepsilon_{ij} = e_{ij} + 1/3\ \delta_{ij}\varepsilon_{kk} \tag{1.1}$$

where δ_{ij} is the same Kronecker delta. The general linear visco-elastic medium is defined by the linear operator equations

$$P\, S_{ij} = 2GQ\, e_{ij} \quad \text{and} \quad P'\sigma_{kk} = 3KQ'(\varepsilon_{kk} - 3\alpha T) \tag{1.2}$$

where T denotes the temperature difference with respect to a constant reference temperature and α the coefficient of thermal expansion. The constants of the operators

$$\begin{aligned} P &= a_o + a_1 \frac{\Delta}{\Delta t} + \ldots\ldots\ldots\ldots a_m \frac{\Delta^m}{\Delta t^m} \\ Q &= b_o + b_1 \frac{\Delta}{\Delta t} + \ldots\ldots\ldots\ldots b_n \frac{\Delta^n}{\Delta t^n} \\ P' &= a'_o + a'_1 \frac{\Delta}{\Delta t} \ldots\ldots\ldots\ldots\ldots + a'_r \frac{\Delta^r}{\Delta t^r} \\ Q' &= b'_o + b'_1 \frac{\Delta}{\Delta t} + \ldots\ldots\ldots\ldots b'_s \frac{\Delta^s}{\Delta t^s} \end{aligned} \tag{1.3}$$

are combinations of relaxation times, retardation times and shear moduli.

Combining eqs. (1.2) under the assumption of isothermal conditions, the visco-elastic stress-strain relations may be obtained in the alternativeforms:

$$\sigma_{ij} = \left(3K \frac{Q'}{P'} - 2G \frac{Q}{P}\right)(1/3)\varepsilon_{kk}\,\delta_{ij} + 2G \frac{Q}{P}\,\varepsilon_{ij}$$

and

$$\varepsilon_{ij} = \left(\frac{1}{9K} \frac{P'}{Q'} - \frac{1}{6G} \frac{P}{Q}\right) \sigma_{kk}\,\delta_{ij} + \frac{1}{2G} \frac{P}{Q}\,\sigma_{ij} \tag{1.4}$$

Poisson's ratio ν for the elastic medium related to K and G by the relation

$$\nu = (3K - 2G)/\,2(3K + G) \tag{1.5}$$

can be transformed and becomes, in linear visco-elastic theory, the

OPERATOR

$$\nu = \frac{3KP\,Q' - 2GQ\,P'}{2(3KP\,Q' + GQ\,P')} \tag{1.6}$$

Dividing the numerator and denominator by 18 KG, eq. 1.6 takes the form

$$\nu = \frac{\frac{1}{9K}\frac{P}{Q'} - \frac{1}{6G}\frac{P}{Q}}{\frac{1}{9K}\frac{P'}{Q'} + \frac{1}{3G}\frac{P}{Q}} \tag{1.6a}$$

which can be obtained by evaluating the second eq. (1.4) for the condition. Thus

$$\varepsilon_{11} = \left(\frac{1}{9K}\frac{P'}{Q'} + \frac{1}{3G}\frac{P}{Q}\right)\sigma_{11}$$

and

$$\varepsilon_{22} = \left(\frac{1}{9K}\frac{P'}{Q'} - \frac{1}{6G}\frac{P}{Q}\right)\sigma_{11} = \varepsilon_{33} \tag{1.7}$$

the time dependent ratio $\nu(t)$ is obtained as the ratio $\nu(t) = -\varepsilon_{22}/\varepsilon_{11} = \nu(\sigma_{11})$

Laplace transforms are used to evaluate this ratio by use of the stress-strain relaions (1.2) and (1.4) and the operator equations (1.3)

$$\overline{P}(p) = \sum_{k=1}^{k=m} a_k p^k \qquad \overline{Q}(p) = \sum_{k=1}^{k=n} b_k p^k \tag{1.8}$$

and

$$\overline{P}'(p) = \sum_{k=1}^{k=r} a'_k p^k \qquad \overline{Q}'(p) = \sum_{k=1}^{k=s} b'_k p^k$$

The ratios

$$\frac{1}{\overline{M}}(p) = \frac{\overline{P}(p)}{\overline{Q}(p)} \qquad \text{and} \qquad \frac{1}{\overline{M}'}(p) = \frac{\overline{P}'(p)}{\overline{Q}'(p)} \tag{1.9}$$

are quotients of two polynomials, the inverse transforms of which can usually be expressed in the form of series of negative exponentials.

The transformed eqs. (1.7) can be directly written in the form

$$\overline{\varepsilon}_{11} = \left(\frac{1}{9K\,\overline{M}'(p)} + \frac{1}{3G\,\overline{M}(p)}\right)\overline{\sigma}_{11}$$

and

$$\overline{\varepsilon}_{22} = \left(\frac{1}{9K\,\overline{M}'(p)} - \frac{1}{6G\,\overline{M}(p)}\right)\overline{\sigma}_{11} \tag{1.10}$$

Poisson's ratio $\nu(t)$ is obtained as the inverse transform

$$\nu(t) = -\frac{L^{-1}\left(\left[\frac{1}{9K\,\bar{M}'(p)} - \frac{1}{6G\,\bar{M}(p)}\right]\bar{\sigma}_{11}\right)}{L^{-1}\left(\left[\frac{1}{9K\,\bar{M}'(p)} + \frac{1}{3G\,\bar{M}(p)}\right]\bar{\sigma}_{11}\right)} \qquad (1.11)$$

When $\varepsilon_{11} = \varepsilon_{11}(t)$ and $\bar{\varepsilon}_{11}(p)$ is given, with $\sigma_{22} = \sigma_{33} = 0$, $\bar{\sigma}_{11}$ is evaluated with the aid of the first eq. (1.10) and introduced into the second. Hence the ratio

$$\nu(t) = L^{-1}\left\{\left[\frac{\frac{1}{9K\,\bar{M}'(p)} - \frac{1}{6G\,\bar{M}(p)}}{\frac{1}{9K\,\bar{M}'(p)} + \frac{1}{3G\,\bar{M}(p)}}\right]\bar{\varepsilon}_{11}\right\} \Big/ \varepsilon_{11}(t) \qquad (1.12)$$

Eq. (1.11) and (1.12) will now be evaluated for the Maxwell type of idealized linear visco-elastic behavior in volume-constant distortion. If we assume elastic volumetric deformation M' reduces to unity. The testing conditions investigated are : constant stress σ_{11} = const., constant stress-rate σ_{11} = ct where c is an arbitrary constant, constant strain ε_{11} = const. and constant strain-rate $\dot{\varepsilon}_{11}$ = ct. For the actual evaluation the arbitrary relation $3K = 8G$ has been introduced, for which the frequently used elastic value $\nu = 1/3$ is obtained.

Evaluation of "Poisson's Ratio for Simple Elastically Compressible Visco-Elastic Media--Maxwell Body

The mechanical behavior of a Maxwell body is determined by

$$\bar{P} = \frac{1}{\tau} + p \qquad \bar{Q} = p$$

$$\bar{P}' = 1 \qquad \bar{Q}' = 1$$

where $\tau = \eta / G$

Constant stress σ_{11} is specifically

$$\sigma_{11} = c\,H(t), \qquad \bar{\sigma}_{11} = c\,\frac{1}{p}$$

From eq. (1.11)

$$\nu(t) = \frac{L^{-1}\left[\left(\frac{1+\tau p}{6G\,\tau\,p} - \frac{1}{9K}\right)\frac{c}{p}\right]}{L^{-1}\left[\left(\frac{1+\tau p}{3G\,\tau\,p} + \frac{1}{9K}\right)\frac{c}{p}\right]}$$

$$= \frac{1}{2} \frac{\frac{t}{\tau} + 1 - \frac{2G}{3K}}{\frac{t}{\tau} + 1 + \frac{G}{3K}}$$

The diagram of $\nu(t)$ for $G/K = 3/8$ is shown in figure 1, curve $\sigma_{11} = c$. At $t = 0$, the value of $\nu = 1/3$ and as $t \to \infty$, $\nu = 1/2$. These limits may be explained by the fact that the instantaneous response of a Maxwell body is purely elastic and consequently ν is controlled by the ratio G/K only. Under long time loading a Maxwell body behaves as a viscous fluid which results in $\nu \to 1/2$, with the elastic compressibility becoming unnoticeable under the uniaxial testing conditions.

The constant stress rate is specified by

$$\sigma_{11} = ct\ H(t) \qquad \text{and} \qquad \bar{\sigma}_{11} = c\frac{1}{p^2}$$

From eq. (1.11)

$$\nu(t) = \frac{L^{-1}\left[\left(\frac{1+\tau p}{6G\tau p} - \frac{1}{9K}\right)\frac{c}{p^2}\right]}{L^{-1}\left[\left(\frac{1+\tau p}{3G\tau p} + \frac{1}{9K}\right)\frac{c}{p^2}\right]} = \frac{\frac{t}{2\tau} + 1 - \frac{2G}{3K}}{2\left(\frac{t}{2\tau} + 1 + \frac{G}{3K}\right)}$$

The corresponding diagram is shown in figure 1, curve $\sigma_{11} = ct$

Constant strain is specified by

$$\varepsilon_{11} = c\,H(t) \qquad \text{and} \qquad \bar{\varepsilon}_{11} = c\frac{1}{p}$$

From eq. (1.12)

$$\nu(t) = L^{-1}\left[\frac{3K\left(\frac{1}{\tau} + p\right) - 2Gp}{2\left[3K\left(\frac{1}{\tau} + p\right) + Gp\right]p}\right]$$

$$= \frac{1}{2} + \left(\frac{3K - 2G}{3K + G} - 1\right)\exp\left(-\frac{3K}{3K + G}\frac{t}{\tau}\right) \qquad (1.13)$$

This is shown by curve $\varepsilon_{11} = c$ in figure 1

Constant strain rate is specified by

$$\dot{\varepsilon}_{11} = c\,t\,H(t) \qquad \text{and} \qquad \bar{\dot{\varepsilon}}_{11} = c\frac{1}{p^2}$$

From eq. (1.12)

$$\nu(t) = L^{-1}\left\{\frac{3K\left(\frac{1}{\tau} + p\right) - 2Gp}{2\left[3K\left(\frac{1}{\tau} + p\right) + Gp\right]p^2}\right\} \Big/ t \qquad (1.13)'$$

$$= \frac{1}{2} - \frac{G\tau}{2K}\left[1 - \exp\left(-\frac{3K}{3K+G}\,\frac{t}{\tau}\right)\right] \qquad (1.14)$$

This is shown by curve $\epsilon_{11} = c\,t$ in figure 1

But $\tau^n = \mu\frac{du}{dy}$

Then eq. (1.13) becomes

$$\nu(t) = \frac{1}{2} + \left(\frac{3K - 2G}{3K + G} - 1\right)\exp\left(-\frac{3K}{3K+G}\,\frac{t}{\left(\mu\frac{du}{dy}\right)^{1/n}}\right)$$

and eq. (1.14) becomes

$$\nu(t) = \frac{1}{2} - \frac{G\left(\mu\frac{du}{dy}\right)^{1/n}}{2Kt}\left[1 - \exp\left(-\frac{3K}{3K+G}\,\frac{t}{\left(\mu\frac{du}{dy}\right)^{1/n}}\right)\right]$$

which gives the ratio of lateral to longitudinal strain in uni-axial stressing in terms of viscosity and fluid velocity.

The paper has proposed a generalization of the Navier-Stokes equation for non-Newtonian fluids in which the stress is proportional to a power n of the rate of strain. The ratio of the transverse to the longitudinal strain is not a constant but a function of time or of strain and depends strongly upon the visco-elastic response of the medium as well as on the loading conditions.

REFERENCES:

1. Non-Newtonian Fluids, W. L. Wilkinson
2. Advanced Mathematics for Engineers, Reddick and Miller
3. Fundamental Formula of Non-Newtonian Flow, Tomita
4. Advanced Calculus for Engineers, Hildebrand
5. Advanced Calculus, Kaplan
6. Poisson's Ratio in Visco-Elastic Media, Freudenthal-Henry
7. Recent Results in the Continuum Theory of Visco-Elastic Fluids, Coleman and Noll
8. Steady Flow of Slightly Visco-Elastic Fluids, Langlois and Rivlin

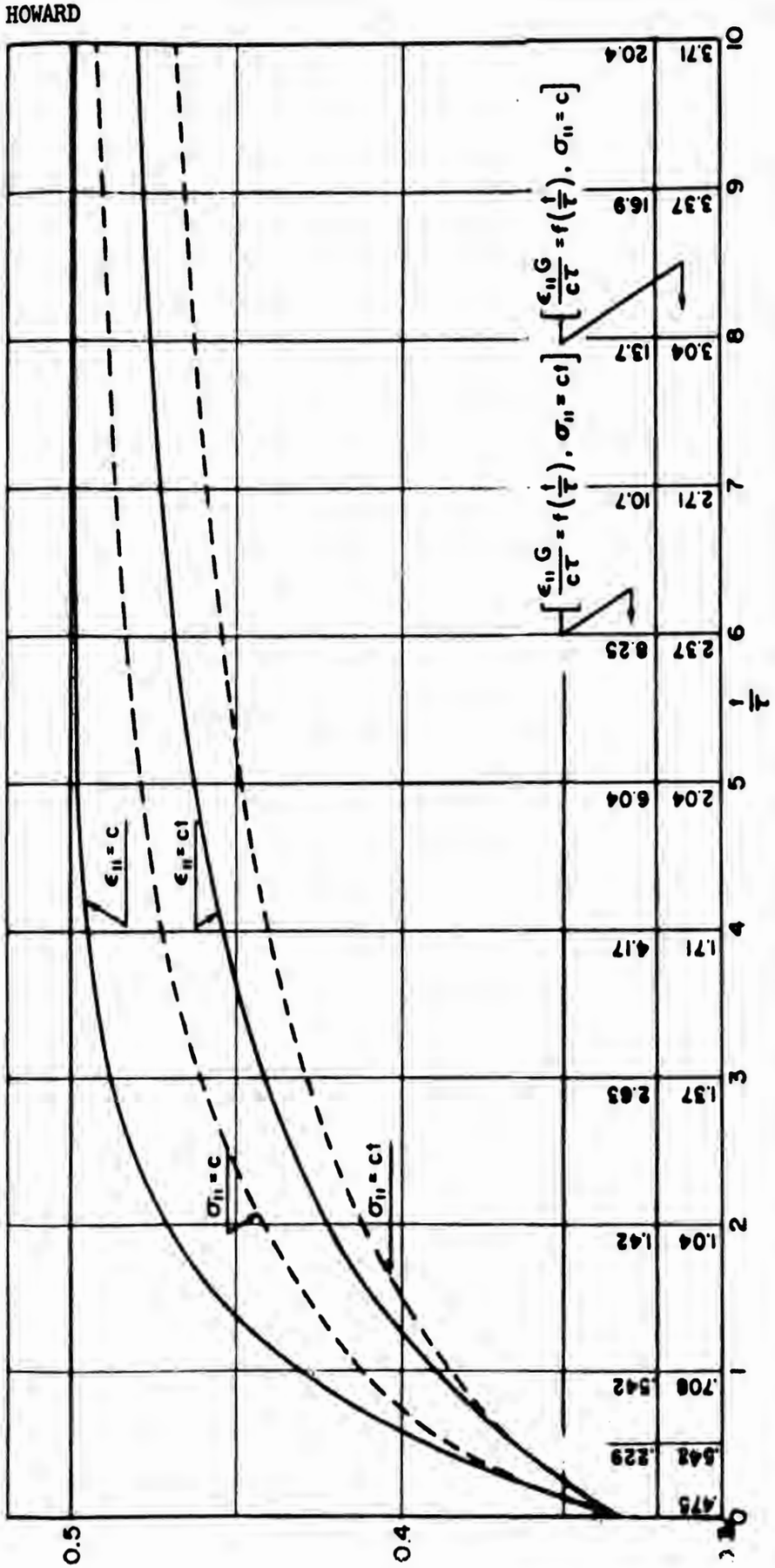


Fig. 1. Poisson's Ratio ν as a Function of Time $\nu(t/C)$ or of strain for Elastically Compressible Maxwell Body.

Notations

p_i Incident pressure, psi

p_o Ambient pressure, psi

p_r Reflected pressure, psi

t Time, msec

W Charge weight (TNT), lb

λ_c Reduced height of burst, $ft/lb^{1/3}$

λ_t Reduced Mach stem (height), $ft/lb^{1/3}$

λ_x Reduced horizontal distance, $ft/lb^{1/3}$

AIR BLAST PHENOMENA IN AN ARCTIC CLIMATE

L. F. INGRAM
U. S. ARMY ENGINEER WATERWAYS EXPERIMENT STATION
VICKSBURG, MISSISSIPPI

Background

In recent years the Corps of Engineers has been increasingly interested in arctic construction, particularly on the Greenland Ice Cap. Because of the protection afforded against the elements, it is expedient to construct ice cap installations under the snow surface. The "cut-and-cover" method developed by the U. S. Army Snow Ice and Permafrost Research Establishment (SIPRE) has become standard practice in construction of undersnow installations. This method results in a tunnel roof of processed snow capable of self-support.

With the development of this unique building method, the need arose for a knowledge of the vulnerability of undersnow structures to damage from explosively-generated loads. That is, an understanding of the response of structures to blast loads in this type of environment was needed to supplement protective construction design methods.

Measurements made by the Waterways Experiment Station (WES) in 1958 revealed that explosively-generated stress waves are attenuated much more rapidly with distance traveled in natural snow than in air. This suggests that the most effective damage can be caused by direct air-blast loading of the target rather than by indirect ground- or snow-shock induced loading. A survey of the literature on blast loading of surfaces revealed considerable high explosive and some nuclear test data for detonations over concrete, soil, and water, but none for blasts over snow. Since the physical properties of snow are unlike the more rigid surfaces previously investigated, the necessary data could only be obtained through a study of shock wave reflection phenomena over snow. This paper describes the procedures for and results of tests made on the Greenland Ice Cap during the summer of 1960 for this purpose.

Objectives

The overall objective of the WES arctic research program is to study the effects of nuclear weapons on terrain and engineering structures in arctic regions. The objectives of these tests were to study the phenomena of blast wave reflection over snow surfaces, and to determine the surface loading for the most significant above-surface charge positions. The determination of the Mach stem formation and propagation should contribute to the basic understanding of shock wave phenomena.

Reflection Phenomena

Inspection of fig. 1 will aid in understanding the reflection phenomena. Spherical shock waves emanate from the explosive source and progress to the interface. For a rigid boundary the wave is reflected as shown. Because of preheating of the air in the region of the boundary at the intersection of the incident and reflected shocks, the reflected shock overtakes the incident shock front and as the wave progresses the shocks combine to form the Mach stem. The intersection of the Mach stem and the incident and reflected shocks is called the triple point.

Scope

Forty 4- and 32-lb TNT spheres were fired at scaled burst heights from zero (surface) to 12 $W^{1/3}$. Gage arrays were used to measure pressures at the surface of the snow and at various distances above the surface at reduced horizontal distances ranging from zero (directly under the charge) to 20 $W^{1/3}$. The reduced vertical distances of the gage height ranged from zero (surface of the snow) to 5.7 $W^{1/3}$.

Coordination and Support

Planning and execution of the experimental work were coordinated with SIPRE. Logistical support including transportation, food, clothing, shelter, and special arctic equipment was provided by the U. S. Army Polar Research and Development Center, whose Engineer Research and Development Detachment provided direct field support by assignment of scientific and engineering enlisted personnel to assist civilian project personnel.

Tests

Test site

The tests were conducted on the ice cap near Camp Fistclench

which is located 250 miles east of Thule, Greenland. The test site is at elevation 6800 ft above mean sea level and is located about one mile north of Camp Fistclench. A 16- by 24-ft Jamesway building provided instrument shelter and office space for the group. A 15-kw, diesel generator supplied electrical power for the instruments, lights, and power tools. The generator was placed in a tarpaulin-covered trench alongside the Jamesway building to shelter it from the blowing snow. Weasels pulling 1-ton sleds were used for local transportation.

Environmental conditions

The Camp Fistclench area is covered to a depth of 100 ft or so with stratified layers of snow and thin ice lenses; below approximately 100 ft solid ice is found. Density of the snow varies with depth in the top 100-ft layer. Measurements in the test area showed an average density of 0.28 gm/cm^3 in the top 1-ft layer and 0.35 gm/cm^3 in the top 6-ft layer.

During the test series the air temperature ranged from 7 to 31 F at shot times. With the exception of two shots fired on calm days, the wind speed varied from 6 to 15 knots at shot times. Wind velocities and temperatures were recorded for each shot.

Test equipment

Piezoelectric blast gages using tourmaline, barium titanate, and lead zirconate were used as pressure transducers. Gages were placed either at the surface to obtain surface pressures (for the height-of-burst curves) or at several locations above the surface to locate the Mach triple point at several horizontal ranges. Coaxial cables connected the gages to associated electronic equipment in the shelter.

Pressure-time traces were recorded on a cathode-ray recorder, a compact unit containing eight dual-beam cathode-ray tubes whose beams are focused on a continuous photosensitive paper strip. The unit contains all required amplifiers, power supplies, optical system, and recording-paper transport mechanism. A recording-paper speed of 400 in./sec was used for all shots. Auxiliary equipment included cathode-follower preamplifiers, a calibration-step generator, a detonator firing supply, and a sequence timer for controlling events. The recording paper was developed and stabilized with a commercial record processor.

Photographic equipment

A 16mm camera, operated at 4000 frames/sec, was used to photograph the shock waves in an attempt to follow the formation and progression of the Mach stem. Plywood boards painted with alternate black and white diagonal strips were set up as a backdrop against which the shock waves were photographed. Refraction of light from

the shock front caused an apparent fringe-shift pattern in the backdrop. Documentary, 16mm color motion pictures were taken as well as 35mm black-and-white and color photographs.

Test layout

Gages were mounted on wooden 2- by 4-in. stakes driven in the snow. All gage coordinates were referenced to ground zero. However, the surface gages were, in reality, mounted approximately 1 in. above the surface to avoid partial burial of the sensing elements. The axes of the pencil-type gages were oriented horizontally and pointed toward ground zero. TNT spheres (32 lb) were fired from surface elevation to 18 ft above the surface, and 4-lb TNT spheres were detonated at higher reduced charge-height positions. The 32-lb spheres were mounted atop 4- by 4-in. pedestals for all above-surface positions, and the 4-lb spheres were suspended from a wire drawn taut between two poles. Fig. 2 is a postshot view of a 16-gage array (8 tourmaline and 8 pencil gages) for locating the Mach triple point.

Firing sequence and procedure

After the gages had been mounted and the snow smoothed around the gage area, a final "squeeze test" of each gage was made to ensure that the proper cable connections had been made and every gage was working properly. Just before the charge was fired, final checks of the amplifier gains were made by manual application of a calibration voltage to all recording channels. After the charge was armed and a warning siren had sounded, the sequence timer was started which performed automatically the following functions:

a. Started the recording paper of the cathode-ray recorder at a speed of 25 in./sec.

b. Applied the calibration voltage to all channels.

c. Changed the recording paper speed to 400 in./sec.

d. Started the 16mm, high-speed camera.

e. Fired the charge.

f. Stopped the 16mm, high-speed camera and the cathode-ray recorder.

g. Reset for the next cycle.

The pressure-time records were processed immediately after each shot, and preliminary reduction of the records was made before setting up the next shot. This procedure made possible the adjustment of amplifier gains to obtain optimum records and also provided

information with which to determine desired gage-charge geometry. Thus, complete data was ensured.

Test Results

Instrument performance

Overall instrument performance was exceptionally good. One complete record was lost (on the first shot) because of malfunctioning of the paper-transport mechanism. However, after this was corrected very little difficulty was encountered with this equipment. Relatively few gages sustained serious damage or failure. It is noteworthy that no serious difficulty was experienced with the diesel generator.

Photography

Results of the high-speed shock wave photography with the 16mm camera were inconsistent. Part of the records were excellent whereas others were poor. Facilities for developing the film were not available in the field; thus it was impossible to evaluate the setup procedures from shot to shot. Also, the lighting conditions varied considerably because of weather and photographic geometry (angle of the sun, etc.).

The definition of the better film records was sufficient for studying the expansion of the primary shock wave, but because of poor definition of the reflected shock, the triple point could not be located. It might have been possible, by use of proper lighting and backdrop orientation, to have photographed the reflected shock. It appears that best results are obtained with the backdrop lines inclined toward ground zero.

Record interpretation

Peak pressure, positive duration, impulse, and Mach-triple-point loci were the most significant parameters deduced from the records. Fig. 3 shows typical pressure-time records and the parameters derived therefrom. Peak pressure values shown are the maximum values read from the records. Impulse was derived by integration of the pressure-time curves.

Pressure-time records from gages near the surface of the snow exhibited a rounding of the incident peak which was more pronounced at farthest distances from ground zero. The maximum pressure usually occurred within 1 msec of the incident shock arrival for gages at horizontal ranges of 60 ft for a 32-lb charge.

Paths of the triple point were determined from the pressure-time records. At each horizontal gage station the vertical placement

of the gages provided for four or more measurements in the region of regular reflection (above the triple point). Mach stem height was determined by extrapolation to zero time interval the plot of time interval between incident and reflected shocks versus gage height.

Blast Parameters

Although considerable data were obtained in these tests, which have been tabulated and plotted in various forms, figs. 4, 5, and 6 illustrate the pertinent findings. Fig. 4 shows the familiar "height-of-burst" curves for pressure over natural snow. This figure presents complete surface pressure information for all burst heights investigated and can be used to determine optimum burst height, reflection coefficients (for various angles of incidence), etc.

Fig. 5 compares the reflection (or absorption) characteristics of snow with a rigid boundary. From this figure it can be seen that the surface pressure on snow is approximately 35 percent less than on a rigid surface for an equivalent test geometry.

Fig. 6 shows the paths of the Mach triple point for the burst heights investigated.

Conclusions

Height-of-burst data for high-explosive charges detonated over natural Greenland snow have been established for reduced burst heights from zero (surface) to 12 $W^{1/3}$ and reduced horizontal ranges from zero (ground zero) to 20 $W^{1/3}$. The peak pressure range for these geometries is from 2 to 100 psig.

Surface pressure measurements show that considerable energy is absorbed by the snow and that snow cannot be regarded as an ideal reflecting medium. Although the Mach stem formation over snow is not essentially different geometrywise from that over a rigid boundary, the magnitude of the reflected pressure waves is appreciably less. Therefore, it is concluded that snow affords considerable protection to buried installations from air-blast loading by attenuation of the shock strength.

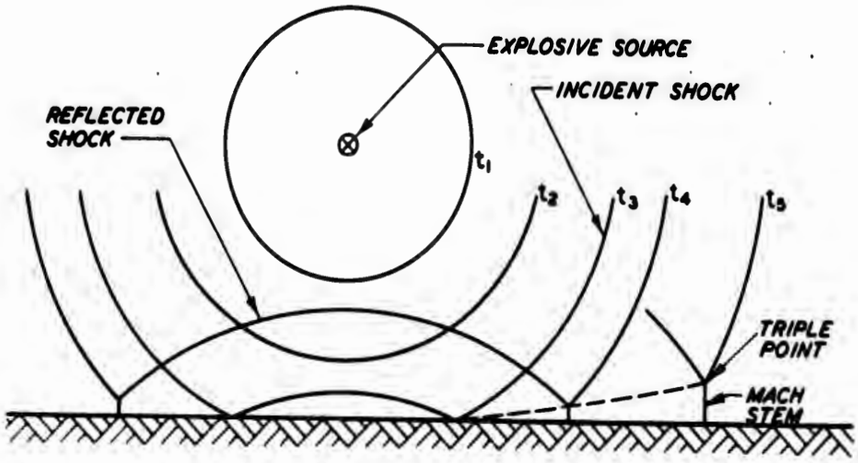


Fig. 1. Schematic of a shock wave reflecting from a plane interface

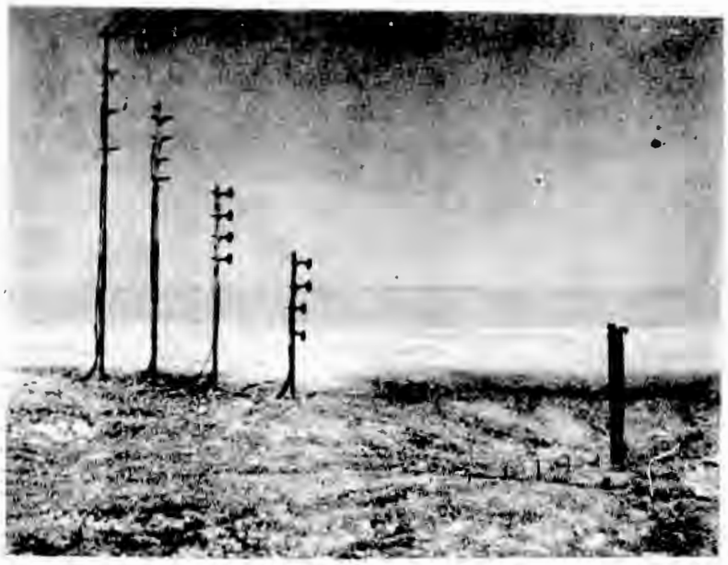


Fig. 2. Gage array for Mach stem measurement, postshot

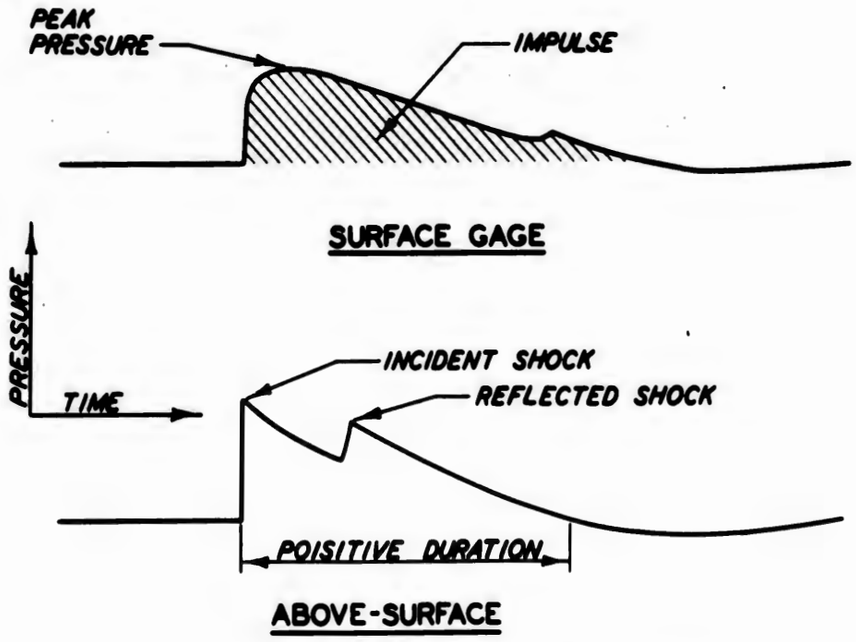


Fig. 3. Typical pressure-time records

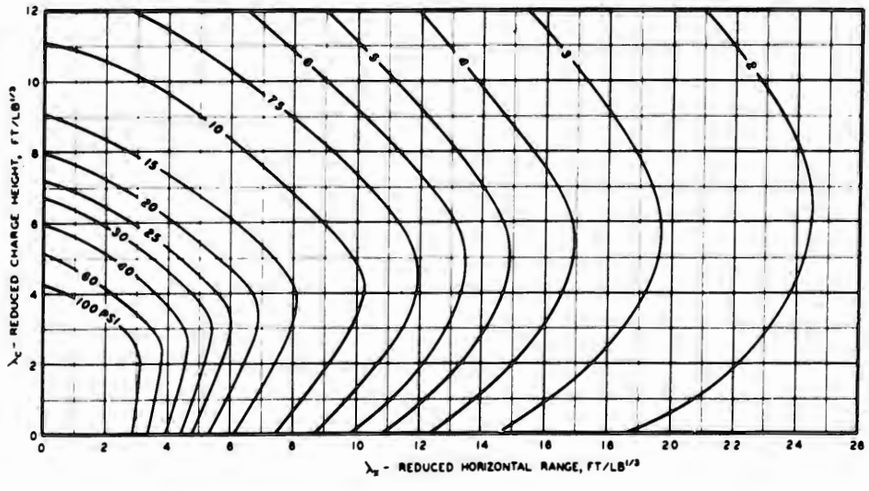


Fig. 4. Peak pressure on the natural snow surface

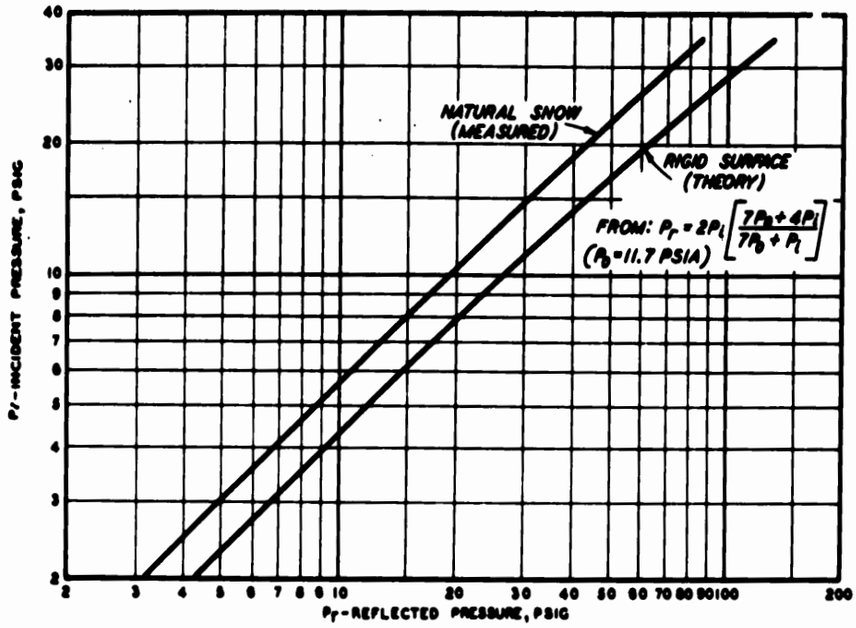


Fig. 5. Reflected versus incident pressure for normal incidence

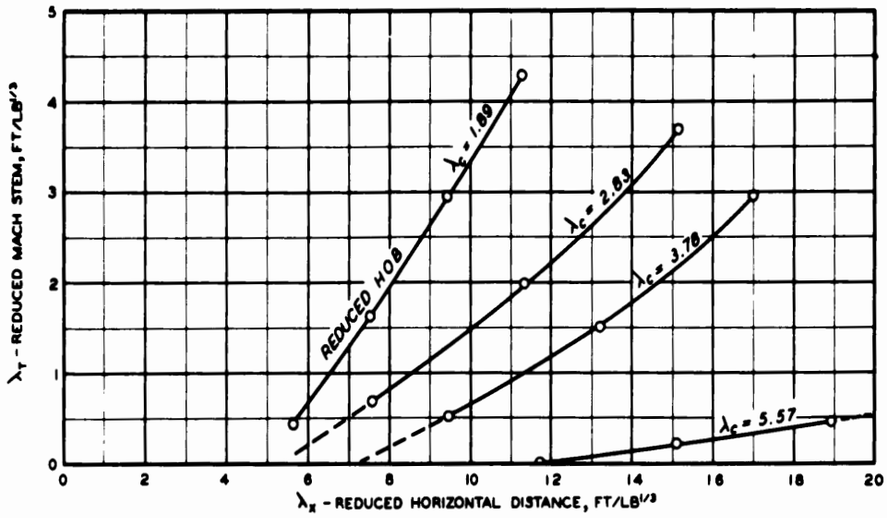


Fig. 6. Mach-triple-point loci for detonations over natural snow

THE INTERACTION OF ELECTROMAGNETIC RADIATION AND SEMICONDUCTORS

H. JACOBS, F. A. BRAND, J. D. MEINDL, S. WEITZ

U. S. ARMY SIGNAL RESEARCH AND DEVELOPMENT LABORATORY
FORT MONMOUTH, NEW JERSEY

INTRODUCTION

In the study of the interaction of microwaves with bulk semiconductors, one can find a means of determining many electrical properties of materials. Several recent reports have indicated that as electromagnetic energy is transmitted through a semiconductor medium, the presence of excess free carriers can be detected by measurements of transmission, reflection or absorption of the electromagnetic power. Gibson,[1] Benedict[2] and Shockley, Harrick[3] and Ramsa[4] have suggested techniques of microwave and infrared transmission to measure such properties as dielectric constant, mobility, conductivity, and lifetime. In addition, others[5] have suggested possible device applications inherent in the interaction of microwaves and bulk semiconductors. In what follows, a description will first be given of more recent experiments in the study of physical properties of semiconductors and second, it will be shown that the same techniques can be applied in the development of new devices for microwave and infrared applications.

Consider reflections in air due to the presence of a semiconductor medium infinite in the x, z plane and in the -y direction. The reflections here are analogous to the reflections from a waveguide filled with a semiconductor in the -y region and air in the +y region as shown in Fig. 1, or with a transmission line which has a propagation constant, Γ_{01} and impedance Z_{01} in the +y region and Γ_{02}, and Z_{02} in the -y region. Now if an incident wave is partially reflected, the magnitude and phase of the reflected wave will depend upon the conductivity of the material amongst other factors. We note that if the conductivity is changed by the creation of excess minority carriers due to light or direct current injection, the amount of reflection will change. Next consider the three medium problem (Fig. 2) where a wave approaches from the right through air and is

normally incident upon an infinite sheet of semiconductor of thickness l_2. Multiple reflections will occur in medium 2, some power will be reflected back to the right, some will be transmitted to medium 3 which consists of air and some will be absorbed. The analogue for a waveguide with a germanium plug inserted in a section can be developed as a transmission line. If high resistivity semiconductor for medium 2 are used, changes in conductivity will cause large changes in both amplitude and phase of the transmitted wave.

RESISTIVITY AND DIELECTRIC CONSTANT BY REFLECTION

To measure resistivity and/or dielectric constant, one of the simplest approaches is the reflection system in Fig. 1. We shall offer some quantitative estimates for a typical case at 10,000 Mc/second. Let us assume an infinite rod of semiconductor material surrounded by the waveguide structure with its own waveguide impedance and propagation constant, Z_{02} and Γ_{02} respectively. The basic equation for the reflection coefficient in a waveguide is given by,

$$r_R = \frac{Z_{02} - Z_{01}}{Z_{02} + Z_{01}} . \tag{1}$$

To obtain Z_{01} we use the following relations,

$$\Gamma_{01} = j \frac{2\pi}{\lambda_1} \sqrt{\left[1 - \left(\frac{\lambda_1}{\lambda_0}\right)^2\right]} , \tag{2}$$

$$Z_{01} = \frac{j\omega\mu}{\Gamma_{01}} , \tag{3}$$

where $\omega = 2\pi \times 10^{10}$ radions/second, $\lambda_0 = 4.56 \times 10^{-2}$ meters, the cut-off wavelength at X-band, $\lambda_1 = \frac{3.0 \times 10^8}{10^{10}} = 3.0 \times 10^{-2}$ meters, the wavelength in free space, $\mu = 4\pi \times 10^{-7}$ henry per meter. To obtain Z_{02} we use,

$$\Gamma_{02} = j \frac{2\pi}{\lambda_2} \sqrt{\left[1 - \left(\frac{\lambda_2}{\lambda_0}\right)^2\right]} - j \frac{\sigma}{\omega\epsilon} , \tag{4}$$

and

$$Z_{02} = \frac{j\omega\mu}{\Gamma_{02}} , \tag{5}$$

where $\omega = 2\pi \times 10^{10}$ radions/second, $\lambda_0 = 4.56 \times 10^{-2}$ meters, $\lambda_2 = 3.0 \times \frac{10^{-2}}{\sqrt{\epsilon_r}}$ meters, the wavelength in lossless germanium, $\mu = 4\pi \times 10^{-7}$ henry per meter, $\epsilon = \epsilon_0 \epsilon_R$ farad/meter. In addition, the conductivity, σ, will be varied throughout the useful range of semiconductors.

In Fig. 3 we see the computed effects of varying the dielectric constant $\epsilon_R = \frac{\epsilon}{\epsilon_0}$, and at the same time changing the conductivity of the semiconductor. In the region of higher conductivity, that is above 10 (ohm meter)$^{-1}$, the reflection coefficient magnitude is independent of the dielectric constant and hence does not depend upon the material. In the region below approximately 10 (ohm meter)$^{-1}$, the reflection coefficient magnitude is more dependent upon the dielectric constant. This indicates that if one would like to measure the dielectric constant for a relatively pure material, the determination of the reflection coefficient would provide a rather accurate means. In Fig. 4 the magnitude of the reflection coefficient is given as a function of various dielectric constants with values ranging from 1 to 20. What is proposed then is as follows: the conductivity and/or dielectric constant can be determined at a frequency of 10^{10} cycles per second by inserting the semiconductor rod in the waveguide and measuring the reflection coefficient. A further approximation can be made in placing a large block, or ingot of semiconductor, over the open end of the air filled section of waveguide. Since the dielectric constant of germanium and silicon is so high, the calculations using a material filled waveguide gives very nearly the same results as the case of reflections from the large crystal not enclosed by waveguide walls.

Experiments have been conducted to test the validity of the calculations. Here, using the technique of placing a large ingot adjacent to the opening of the air filled waveguide, reflection measurements were made to determine the resistivity of crystals previously measured with four point probe techniques. Furthermore, in the resistivity measurements by microwave reflection, the values are not critically dependent upon surface preparation. In these cases, ϵ_R, was taken at 16 for germanium and 12 for silicon.

It was found that these techniques could be applied to lifetime measurements. In this case the surfaces were polished, etched, washed and dried. A pulsed light irradiated the crystal, either at the point of reflection or nearby. By measuring the decay in reflection changes, a measure of lifetime could be obtained.

This technique can be applied to frequencies or temperatures where inertial effects come into play. The equations which give the effective conductivity, σ', and permittivity, ϵ', are given

in Eqs. 6 and 7. Here, using the Drude-Zener theory of complex conductivity,

$$\sigma' = \sigma_o \frac{1}{1 + \omega^2\tau^2}, \tag{6}$$

$$\epsilon' = \epsilon \left(1 - \frac{1}{\omega\epsilon} \sigma_o \frac{\omega\gamma}{1 + \omega^2\tau^2}\right). \tag{7}$$

It turns out that one can substitute ϵ' and σ' directly into the nomograph shown in Figs. 3 and 4, or stated conversely, if one measures the reflection coefficient and σ' in the region of inertial effects one can determine ϵ'.

RESISTIVITY AND TRANSMISSION

Resistivity can also be measured by transmission. Consider a transmission system consisting of a waveguide, in which electromagnetic energy is being propagated in the TE_{01} mode. Assume that a slab of semiconductor is inserted in the waveguide as shown in Fig. 2, where the large surface of the semiconductor slab is oriented perpendicularly to the direction of propagation. The following theory will apply to the measurement of E_o/E_{in}, or the ratio of transmitted electric field to the incident electric field. This ratio is given by,

$$\frac{E_o}{E_{in}} = r_t \left(\cosh \Gamma_{02} l_2 - \frac{Z_{02}}{Z_{ab}} \sinh \Gamma_{02} l_2\right), \tag{8}$$

and

$$r_t = \frac{2Z_{ab}}{Z_{ab} + Z_{01}}, \tag{9}$$

$$Z_{ab} = Z_{02} \left(\frac{Z_{03} + Z_{02} \tanh \Gamma_2 l_2}{Z_{02} + Z_{03} \tanh \Gamma_2 l_2}\right), \tag{10}$$

where Z_{01} is the impedance of the line (waveguide) in air and is equal to Z_{03},
Z_{02} is the impedance of the waveguide filled with germanium and is a function of conductivity and wavelength,
Z_{ab} is the impedance of the germanium at the front surface,
Γ_{01} is the propagation constant in air in the waveguide,
Γ_{02} is the propagation constant in germanium in the waveguide,
l_2 is the thickness of the slab of semiconductor.

Typical computations of $\left|\frac{E_o}{E_{in}}\right|$ at 10,000 megacycles are shown in Fig. 5. The use of this type of curve together with measurements of thickness and frequency allow the determination of conductivity. By varying the temperature of the sample in the waveguide by means of an oven heating a section of the guide, conductivity has been measured as a function of temperature. Results are shown in Table I indicating the applicability of the new techniques.

At very reduced temperatures σ' and ϵ' in (6) and (7) can be substituted directly, (8), (9) and (10) to take into account inertial effects of free carriers.

LIFETIME AND SURFACE RECOMBINATION VELOCITY

The principle of the measurement can be explained with the aid of Fig. 6. Here a germanium or silicon sample is cut so that it is rectangular in shape and can be fitted into an X-band waveguide. It is then placed in the thermal control unit and the microwave power is transmitted through the semiconductor block at normal incidence to the largest face. As a light pulse is flashed on the side edge of the semiconductor, excess carriers are created resulting in an absorption of microwave power. At the cessation of light, the carriers recombine and the power transmitted is increased. The time constant of the increase in power as viewed on an oscilloscope is assumed to be the lifetime of excess minority carriers. Correlation was established in the microwave techniques and more conventional methods (see Table II). The thermal control unit as shown in Fig. 6 was made up of the following elements. A and D represent a large brass block, C is a sliding holder for the semiconductor sample B. E represents a quartz rod through which the light pulse can pass and irradiate the end of the semiconductor block. It can be noted that for a large block the light does not strike the large planar surfaces and hence after the carriers diffuse into the sample, the microwave energy is absorbed largely by carriers in the bulk material. The temperature was controlled by a container filled with fluid (not shown) soldered to the brass block A. The microwave system was closed by teflon windows inserted in the waveguide and flushing gases or a vacuum could be utilized. In general, thin samples were prepared for surface recombination studies and from the same crystal, thicker samples were prepared in order to determine the bulk lifetime. In following this procedure, two objectives were to be accomplished. First, a check could be made of the work of previous research in the field and second, new physical effects could be studied.

A wide variety of samples was examined and in general, all results were surprisingly consistent with the Shockley-Read theory of bulk lifetime and the Stevenson and Keyes modification for surface recombination. A typical run for the variation of the product of

surface recombination velocity and temperature as a function of the reciprocal of temperature is shown in Fig. 7. Here the ambient is oxygen and the microwave technique is used. Now the Shockley-Read theory predicts that the lifetime for bulk semiconductors is given by,

$$\tau = \frac{\tau_{po}(n_o + n_1) + \tau_{no}(p_o + p_1)}{n_o + n_p}, \qquad (11)$$

where τ is the measured lifetime, τ_{po} is the lifetime of holes injected into highly n-type material, τ_{no} is the lifetime of electrons injected into highly p-type material, and n_o, p_o, n_1 and p_1 are the equilibrium density of free electrons, free holes, the density of free electrons if the Fermi level were at the same energy level as the trapping centers and the density of holes if the Fermi level were at the same level as the trapping centers respectively. The application of this theory to experimental results give rise to the data in Table III where energy levels are determined as a result of temperature tests. The energy levels ϵ_t are interpreted as follows. For p-type silicon, $\epsilon_t = E_t - E_v = .065$ ev is the difference in the level of the recombination centers and the valence band. For thin p-type silicon slices $\epsilon_t \cong .15$ ev $= E_t - E_v$ at the surface of the particular silicon sample in oxygen. For thick n-type germanium, $\epsilon_t = E_c - E_t \cong .20$ ev, the difference in energy level of the conduction band and the trapping centers and for thin n-type germanium in oxygen and vacuum, $\epsilon_t = E_c - E_t \cong .25$ ev and .12 ev represents the differences in energy levels at the surfaces respectively. We conclude here that the microwave techniques represent a simple, electrodeless technique for measuring recombination processes.

DEVICES

In arrangements similar to that shown in Fig. 8, modulators of microwave carrier frequencies have been developed by biasing the semiconductor so that as a diode injects excess minority carriers, the transmission or reflection can be varied to give either amplitude or phase modulation or both, depending on the geometry of the semiconductor block. The limit of frequency is determined by the magnitude of $\omega\tau$. Preliminary calculations indicate that these devices should operate up to frequencies of 10^{13} cycles per second.

CONCLUSION

A general theory of microwave transmission and reflection in semiconductors has been formulated. New techniques for measurement of physical properties of materials has been described i.e., pertaining to resistivity, dielectric constant, lifetime, and surface recombination velocity. In addition, device possibilities from the microwave region to the far infrared regions of the frequency spectrum are suggested. A new family of devices consisting of amplitude

modulators, phase modulators or combinations have been developed by controlling the geometry of the semiconductor block and injection of excess minority carriers into the medium to vary its conductivity.

ACKNOWLEDGMENT

Acknowledgment should be made to Dr. W. Caldwell, Bendix Semiconductor Division, Holmdel, New Jersey and to Mr. R. Benjamin, Monmouth College, West Long Branch, New Jersey for their assistance and suggestions during the course of this work.

REFERENCES

(1) A. F. Gibson, "Progress in Semiconductors," John Wiley & Sons, Inc., New York, N. Y., Vol. 2, pp. 229-245; 1957.

(2) T. S. Benedict and W. Shockley, "Microwave Observation of Collision Frequency of Holes in Germanium," Phys. Rev., Vol. 91, pp. 1565-1566; 15 September 1953.

(3) N. J. Harrick, "Metal to Semiconductor Contacts: Injection or Extraction for Either Direction of Current Flow," Phys. Rev., Vol. 115, pp. 876-882; 15 August 1959.

(4) A. P. Ramsa, H. Jacobs and F. A. Brand, "Microwave Techniques in Measurement of Lifetime in Germanium," J. Appl. Phys., Vol. 30, pp. 1054-1060; July 1959.

H. Jacobs, F. A. Brand, J. D. Meindl, M. Benanti and R. Benjamin, "Electrodeless Measurement of Semiconductor Resistivity at Microwave Frequencies," Proc. IRE, Vol. 49, No. 5, pp. 928-932; May 1961.

(5) H. Jacobs, F. A. Brand, J. D. Meindl and R. Benjamin, "Multiple Reflections of Microwaves Propagating Through a Semiconductor Medium," Proc. IRE, Vol. 49, No. 11, pp. 1683-1684; November 1961.

A. F. Gibson, "Infrared and Microwave Modulation Using Free Carriers in Semiconductors," J. Sci. Instr., Vol. 35, pp. 273-278; August 1958.

TABLE I

ENERGY BAND GAP AND RESISTIVITY

Trial Run	Experimental Band Gap Near 300°K in ev	Experimental Resistivity in ohm cm at 300°K	Band Gap Near 300°K as Shown in Literature	Resistivity in ohm cm at 300°K as Shown in Literature
1	.70	46.1	.72	47.0
2	.75	45.3	.72	47.0
3	.75	52.5	.72	47.0
4	.75	46.4	.72	47.0
5	.74	48.4	.72	47.0

TABLE II

COMPARISON OF LIFETIME DATA BY METHODS OF DC PHOTOCONDUCTIVITY, MICROWAVE TRANSMISSION VARIATIONS AND MICROWAVE REFLECTIONS

Sample No. and Type	Lifetime by Microwave Transmission in Microseconds	Lifetime by dc Photo-conductivity in Same Holder in Microseconds. Bias Current at 5 ma	Lifetime by Reflection of Microwaves in Micro-seconds	Lifetime by dc Photoconductivity in Same Holder in Microseconds. Bias Current at 2.5 ma
No. 3 - Ge .257x1.035x2.17 cm 47 ohm cm, n-type	680	800	700	700
No. 4 - Ge .362x1.035x2.15 cm 47 ohm cm, n-type	1200	1300	1200	1250
No. 5 - Ge .550x1.210x2.15 cm 47 ohm cm, n-type	1250	1250	750 to 900	1225
No. 6 - Ge .550x.963x2.08 cm 47 ohm cm, n-type	450	450	475	450 to 475
No. 10 - Ge .241x.750x2.10 cm 8 ohm cm, p-type	90	90	100	110

TABLE III

BULK LIFETIME AND SURFACE RECOMBINATION

Description	Ambient	ϵ_t in ev	Magnitude	Test Run Number
Silicon (thick) .447 cm	Oxygen	.065	τ_o = 620 μsec at 300°K	S-23
Silicon (thin) .077 cm	Oxygen	.142	ST = 3.6×10^5 at $\frac{1}{T} = 3 \times 10^3$	S-9
Silicon (thin) .059 cm	Oxygen	.153	ST = 4.1×10^4 at $\frac{1}{T} = 3 \times 10^3$	O-14
n-type, 20 ohm cm Germanium (thick) .435 cm	Oxygen	.203	τ_o = 1350 μsec at 300°K	O-21
n-type, 20 ohm cm Germanium (thick) .455 cm	Vacuum	.192	τ_o = 1280 μsec at 300°K	A-5
n-type, 20 ohm cm Germanium (thin) .0283 cm	Oxygen	.265	ST = 6.5×10^4 at T = 300°K	A-26
n-type, 20 ohm cm Germanium (thin) .0354 cm	Oxygen	.244	ST = 4.95×10^4 at T = 300°K	A-23
n-type, 20 ohm cm Germanium (thin) .064 cm	Vacuum	.121	ST = 3.3×10^4 at T = 300°K	T-12

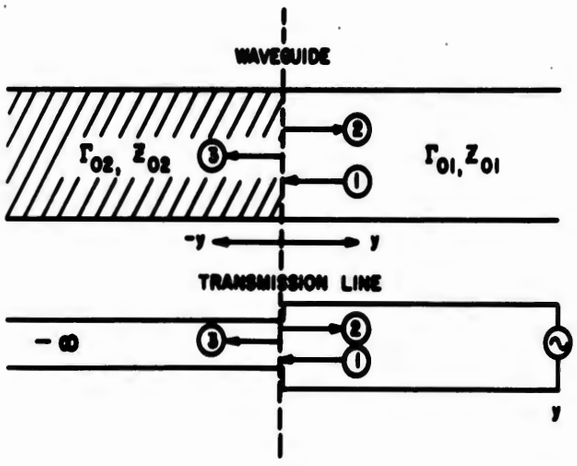


Fig. 1 Reflection of electromagnetic radiation in a waveguide and transmission line analogue.

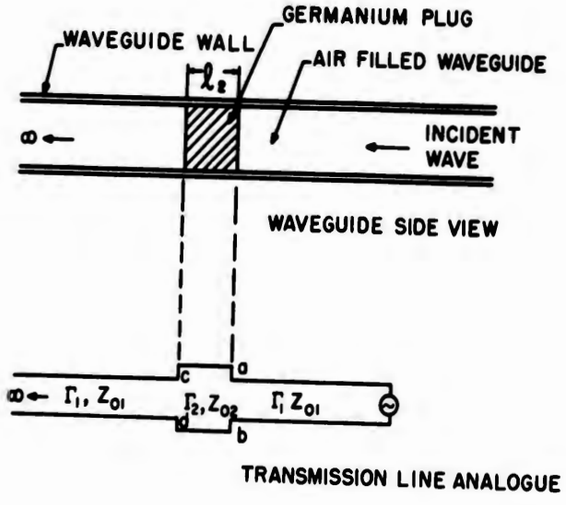
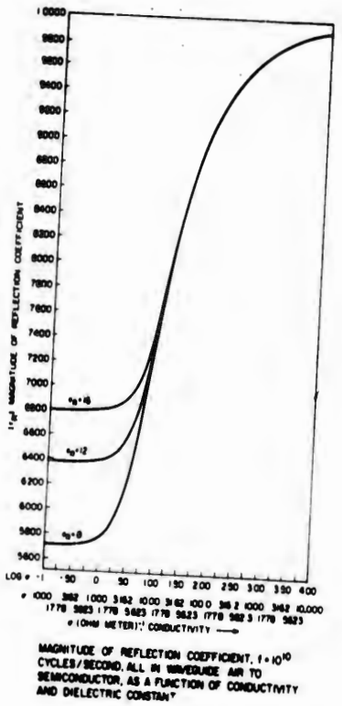
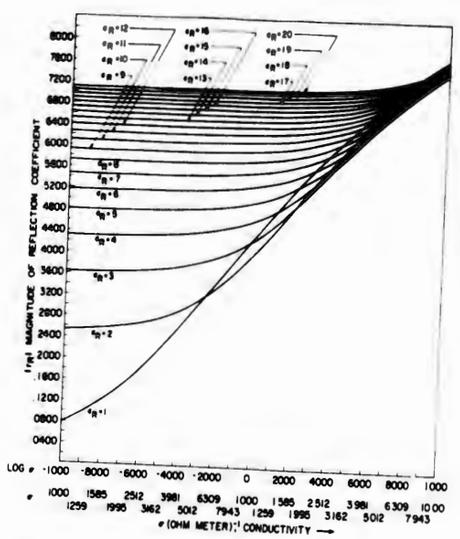


Fig. 2 Transmission of microwaves through a semiconductor plate.



MAGNITUDE OF REFLECTION COEFFICIENT, $f = 10^{10}$ CYCLES/SECOND, ALL IN WAVEGUIDE AIR TO SEMICONDUCTOR, AS A FUNCTION OF CONDUCTIVITY AND DIELECTRIC CONSTANT

Fig. 3 Magnitude of reflection coefficient as a function of conductivity at $f = 10^{10}$ cycles/second.



MAGNITUDE OF REFLECTION COEFFICIENT FROM AIR TO DIELECTRIC IN WAVEGUIDE AT $f = 10^{10}$ CYCLES/SECOND AS A FUNCTION OF DIELECTRIC CONSTANT AND CONDUCTIVITY

Fig. 4 Variations in reflection coefficient as a function of dielectric constant at $f = 10^{10}$ cycles/second.

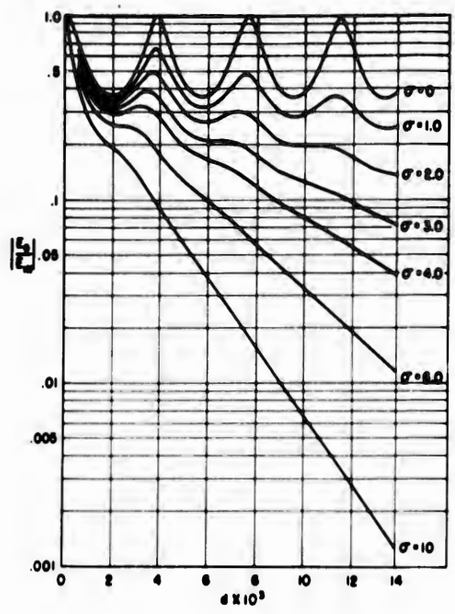
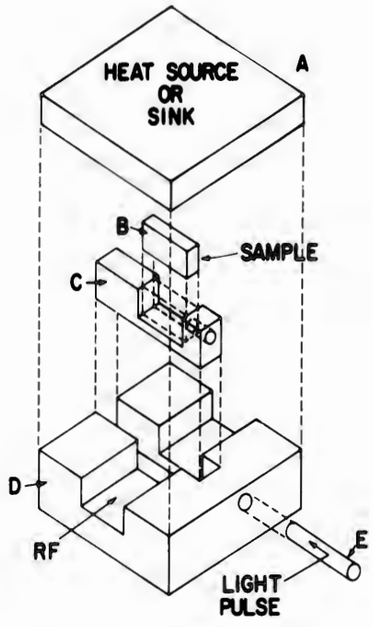


Fig. 5 Magnitude of transmission coefficient as a function of conductivity for germanium in X-band waveguide at f = 10^{10} cycles/second.



THERMAL CONTROL UNIT

Fig. 6 Thermal unit for control of temperature of semiconductor sample in waveguide.

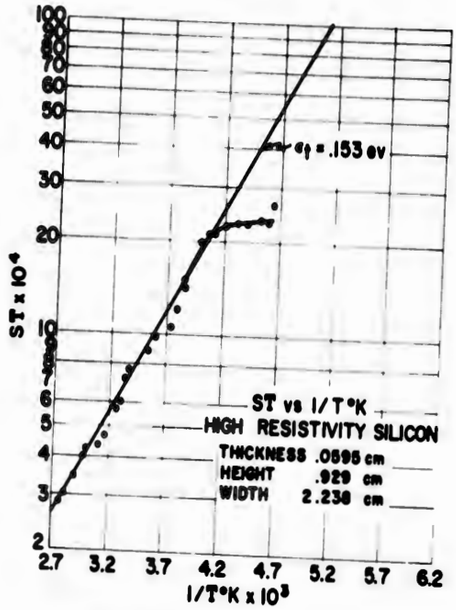


Fig. 7a Product of surface recombination velocity and temperature as a function of the reciprocal of temperature for high resistivity silicon.

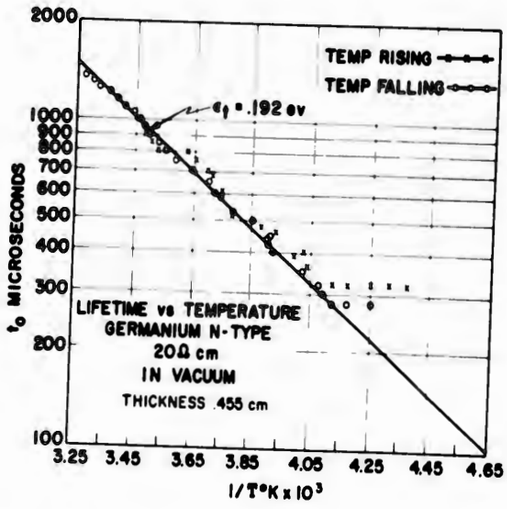


Fig. 7b Lifetime versus reciprocal of temperature for 20 ohm cm n-type germanium.

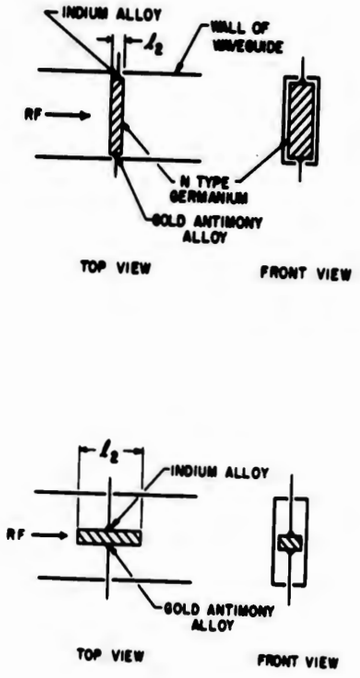


Fig. 8 Arrangement for microwave modulator for control of amplitude and/or phase.

LABORATORY STUDIES OF UPPER ATMOSPHERE CHEMICAL REACTIONS

FREDERICK KAUFMAN and JOHN R. KELSO
BALLISTIC RESEARCH LABORATORIES
ABERDEEN PROVING GROUND, MARYLAND

Spectacular progress has recently been made in man's conquest of the regions of the upper atmosphere and of interplanetary space. Immensely complicated and costly programs are under way to take human observers into satellite orbit, to the moon, and perhaps to other planets. Concurrently, a large and varied program of probing and measuring of the earth's gaseous envelope has gathered important data on the density, temperature, and composition of the upper atmosphere; on the magnetic fields surrounding our planet; on the distribution in frequency and in intensity of the impinging radiation from the sun; and on many other physical and chemical observables. The purpose of this scientific effort is to achieve a fuller understanding of the physics and chemistry of the upper atmosphere and not merely to produce convenient catalogues of its properties.

It is through such a fundamental understanding that reliable predictions can be made for problems of military importance, as in the re-entry of missiles or in disturbed conditions due to a nuclear explosion at high altitude. Moreover, any advance towards a fuller understanding represents scientific progress and is often accompanied by unexpected dividends in related fields.

Many of the important processes of the upper atmosphere are best studied in the laboratory where the gas temperature, composition, and pressure can be closely controlled and where energetic species such as free atoms, electronically excited species, or unstable molecules can be produced by light absorption, by thermal decomposition, or by electrical gas discharges. The many advantages of laboratory research in this field are not properly appreciated. Its relatively low cost, ease of experimental changes, and great specificity in the study of particular processes warrant much greater support than it has received up to now.

During the past five years, such a laboratory research program has been carried out in the Interior Ballistics Laboratory.

In this paper, the experimental methods of that work will be described, some of its earlier results will be summarized, and new data will be presented for two important reactions of oxygen atoms. The work deals mainly with the reactions of oxygen atoms or nitrogen atoms with other constituents of the upper atmosphere such as O_2, O_3, NO, and NO_2; with the emission of light from some of these reactions; and with the measurement of rate constants for such processes. The principal experimental method consists of producing the active species (O- or N-atoms) in microwave gas discharges under steady-state flow conditions, pumping the low-pressure gas mixture through a long flow tube with or without additions of other reactants, and measuring the spatial decay of a visible afterglow whose intensity is associated with the atomic species. These measurements are complemented by "atom-cleanup" or "titration" experiments in which these afterglows are sharply brought to extinction by the addition of known flows of certain molecules which react simply and rapidly with the free atoms.

EXPERIMENTAL

a. The Discharge-Flow System

A simple diagram of the present apparatus is shown in Fig. 1. The flow tube, F, is a precision-bore pyrex tube of 2.540 $\pm$ 0.003 cm i.d (Fischer and Porter Company), 90 cm in length. It is connected through large traps to an oil diffusion pump and a large mechanical pump (International, Model 116, 450 lit/min) with a ballast volume of 3 liters between pump and tube to reduce pressure fluctuations. At its upstream end it is sealed to a quartz discharge tube, D, (0.9 cm o.d.) in which the microwave discharge is excited. The flow tube is fitted with two inlets, A and B, for metered addition of gases whose flows are controlled by stainless steel needle valves and measured by the gas handling system, N, indicated above the tube in Fig. 1. This system consists of various flasks of known volume from which gases are admitted to F. The resultant pressure decrease in the closed volume is measured by the transducer, T, (Consolidated Controls, Model DCS4-5D, range 250 mm Hg differential, output 0 to 2.5V, D.C.) whose output is recorded by a pen recorder. Care is taken that the pressure decrease is sufficiently small that it does not affect the rate of gas addition during the measurement.

A second, independent gas handling system, displayed below F in Fig. 1, supplies the main flow of gases to the discharge and to by-pass C. Flows are controlled by stainless steel needle valves, and measured on ball-type flowmeters or by the pressure drop in a closed volume as read on a dial pressure gauge (Wallace and Tierman, Model 145). The pressure at either end of the flow tube or ahead of the discharge tube is measured by a calibrated McLeod gauge (Fig. 1, MC).

The microwave discharges are excited by a magnetron oscillator (Raytheon, QK-390) at 2450 mc/sec operating from a moderately filtered anode supply (6000 V $\pm$ 10%, 0 to 0.3 amp) and delivering 100 to 700 W of microwave power. Between the magnetron and the tapered, slotted discharge cavity, E, are placed a variable water attenuator

and a standing wave ratio meter. Two movable photomultiplier tubes, P_1 and P_2 (type 1P28) serve to measure the light intensity of afterglows in the flow tube at any point. They are fitted with collimating slits so as to receive light from well defined planes perpendicular to F and with glass filters to isolate certain wavelength regions. Their output is amplified (Leeds and Northrup, Model 9836-A) and recorded by a pen recorder (Leeds and Northrup, Speedomax Type G). This system accurately records light intensities over a relative range of 100,000, i.e. photocurrents of about 10^{-10} to 10^{-5} amp.

b. Materials

The purity of the gases entering the discharge is of great importance. Cylinders of oxygen are specially selected for their low nitrogen content as indicated by a very weak afterglow in the absence of added NO. The greenish-yellow air afterglow is due to electronically excited NO_2 made from O-atoms and NO. Since the discharge converts a fraction of the nitrogen impurity to NO, the brightness of the afterglow simply indicates the nitrogen content of the oxygen. Cylinders with less than 0.01% N_2 are chosen. Condensible impurities such as water vapor are removed by a trap cooled to -195° C immediately upstream of the discharge. When necessary, very pure oxygen is prepared by the thermal decomposition of potassium permanganate at 200 to 250° C. This oxygen is condensed at -195° C and a middle fraction of the liquid used. Best commercial cylinder grades of nitrogen (Matheson, prepurified), helium, argon, nitrous oxide and carbon dioxide are used. Nitric oxide is purified by passing cylinder NO through large tubes filled with Ascarite and condensing the gas to a pale blue liquid. NO_2 is prepared from purified NO with excess O_2 and condensed to a pure white solid at -77° C.

c. Measurement of O-atoms Kinetics

At the heart of the experimental work on O-atom reactions lie the two complementary methods of measuring O-atom concentrations which were developed in this laboratory (1, 2). The intensity, I, of the air afterglow has been found to be proportional to the product of the concentrations of O-atoms and of NO, i.e. $I \propto (O)(NO)$. Moreover, any NO_2 formed by the light-producing reaction $O + NO \rightarrow NO_2 + h\nu$ or by the faster, termolecular step $O + NO + M \rightarrow NO_2 + M$, where M is any third body, is quickly destroyed by the very fast reaction $O + NO_2 \rightarrow O_2 + NO$. Thus, within one experiment under steady-state conditions of flow, (NO) remains constant and I is proportional to (O) where the brackets denote concentrations. Any decrease of I along the flow tube, dI/dx, directly measures $d(O)/dx$. In steady-state flow, $\bar{v} = dx/dt$, where $\bar{v}$ is the average flow velocity, and therefore $\bar{v}\, dI/dx = d(O)/dt$, the rate of decay of O-atoms. Refinements have also been made (3, 4) which take into account axial and radial gas diffusion and the viscous pressure drop along the flow tube by simple mathematical analysis.

The light intensity method only measures the relative rate of decay of (O). The absolute magnitude of (O) is conveniently measured by a NO_2 "titration". The rapid reaction between O and NO_2 assures the presence of the air afterglow as long as O is in excess over NO_2, because the remaining, unreacted O finds itself in the presence of much NO which is produced in the reaction. When the flow of added NO_2 is increased to a point where it is equal to the flow of O-atoms, all the O-atoms react within a small mixing zone and the afterglow is sharply extinguished. The smallest flow-rate of NO_2 which will barely produce this extinction is therefore equal to the flow-rate of O-atoms. Their concentration, (O), is then easily calculated from the known total flow-rates and the measured pressure. In most of the experiments, pressures range from 0.5 to 5 mm Hg, total flow-rates of gases from 20 to 200 micromoles per second, and O-atom flow-rates from 0.3 to 10% of total flow.

RESULTS AND DISCUSSION

a. Nature and Intensity of Afterglows

In discharged oxygen, nitrogen, or mixtures of both, three principal afterglows can be distinguished: 1) The greenish-yellow air afterglow emitted in the combined presence of O and NO; 2) The orange-yellow Lewis-Rayleigh nitrogen afterglow emitted in the presence of N-atoms; and 3) The blue nitric oxide afterglow emitted in the combined presence of O- and N-atoms. These will now be briefly described, particularly where work in this laboratory has contributed to their understanding.

The air afterglow has been the object of many studies since 1910 (5). Its frequency distribution ranges from 3900 Å well into the near infrared region with maximum intensity emitted around 5500 to 6000 Å. The spectrum consists of some banded structure, corresponding to known features of the absorption or emission of NO_2, superimposed on a broad continuum whose intensity represents a large fraction of the total emitted power.

Some controversy still exists as to the relative importance of the bands and of the continuum (6) and further spectroscopic experiments are planned. The authors' early rough estimate of the rate constant of the light-producing reactions, $O+NO \rightarrow NO_2 + h\nu$, $k^* = 1.5 \times 10^7$ cm^3 $mole^{-1}$ sec^{-1} (1) was recently found to be somewhat too large (7). A more accurate value of 8.5×10^{-6} was reported. The important question which still needs to be resolved is this: Does most of the emitted light originate from a bound and stabilized electronically-excited state of NO_2 or does it arise as a true recombination continuum from an unstabilized state during the mutual approach of O and NO? (8) To clarify this point, we have measured the intensity of the emitted light in the presence of five added gases, M = He, Ar, O_2, CO_2, or N_2O. If the rate constant, k^*, depends strongly on the added gas, the former explanation is likely to be correct; if not, the latter. The

results presented in Table 1 were obtained by adding about ten-fold excess of M to discharged oxygen, adding a small, known flow of NO, and measuring (O) and the emitted intensity. Since the intensity is here expressed in arbitrary units, the results are reported as relative rate constants compared to k^* for $M = O_2$.

Table 1

Relative Efficiency of Added Gases, M, in the Light-Producing Reaction

	M = He	Ar	O_2	N_2O	CO_2
$k_M^*/k_{O_2}^*$	1.1	1.2	1	0.9	0.7

Though k^* does appear to be a function of M, its dependence on M is weak. Normally, one would expect much greater differences in the rates of collisional quenching of electronically-excited states between, say, Ar and N_2O than a factor of 4/3. It would appear that both processes, the direct recombination (continuum) and the stabilized NO_2^* emission (banded structure), are concurrently taking place and that the M-dependence of the total emission reflects the much stronger M-dependence of the latter process. Further experimental work is under way.

The orange-yellow, Lewis-Rayleigh nitrogen afterglow has been the object of a still larger number of investigations. Its spectrum consists of certain bands of the First Positive System of nitrogen, a known transition between two electronically-excited states of N_2, ($B\ ^3\Pi - A\ ^3\Sigma$). The unusual intensity distribution of the various bands is now thought to arise by the recombination of N-atoms into an intermediate molecular state (probably $^5\Sigma$) of low stability from which the excited triplet states are populated by collisions with other molecules. A contribution was made in this laboratory to the general understanding of the energetic species in discharged nitrogen by the discovery and measurement of vibrationally-excited ground-state N_2 ($^1\Sigma$) in large concentrations (9).

The blue NO afterglow is observed when a little NO is added to discharged nitrogen. With increasing NO addition the afterglow changes from yellow to mauve, purple, and blue; then, for a certain critical NO flow-rate, the afterglow is extinguished, but with further addition, the air afterglow is emitted. In the mauve, purple, and blue stages, both the nitrogen afterglow and an emission from electronically-excited NO molecules (the β, γ, and δ systems) are observed simultaneously. This sequence of afterglows and the mechanism of excitation of NO were explained by isotopic substitution experiments in this laboratory (10, 11). By addition of ^{15}NO to

normal, i.e., ^{14}N, active nitrogen it was shown that the emitted NO-radiation came from ^{14}NO. This means that the added ^{15}NO is rapidly destroyed by reaction with N-atoms, $NO + N \rightarrow N_2 + O$, and that the emitting NO is formed by the recombination $N + O + M \rightarrow NO + M$. The sequence of afterglow colors is explained as follows: With small NO additions, the N-atom concentration, (N), is lowered and some O-atoms are formed. The color changes from orange-yellow to mauve as the sum of yellow and blue. Increasing the NO flow further decreases (N), increases (O) and shifts the color towards blue. When the added NO equals the N-atoms, another "titration" point is reached, all N-atoms are removed, and the afterglow is extinguished. Further addition produces the greenish-yellow air afterglow because O and NO are both present in the gas stream. The above description, titration reaction and explanation of afterglows have been generally accepted and the techniques are now used by many investigators (12, 13, 14).

b. The Recombination of O and NO

We turn next to the determination of reaction rate constants by laboratory experiments. It must be understood that when the mechanism and rate constant of a certain reaction have once been established, its rate can be calculated for any value of pressure, composition, and temperature. The reaction $O + NO + M \rightarrow NO_2 + M$ has been studied in this laboratory and in others, and good agreement has been reached in three studies for $M = O_2$ as the third body. It was important, however, to investigate the efficiency of different third bodies, and this new work is reported here. The justification for such a study if twofold: 1) The reaction occurs in upper atmosphere seeding experiments where large quantities of NO are released from rockets or balloons; 2) there is an urgent need for detailed information on the rate of such simple, elementary reactions as a guide to better understanding of chemical kinetics and to development of improved theories.

The experimental study is simplified by two pieces of good luck: 1) As stated above, the reaction product, NO_2 rapidly reacts with O to regenerate NO, i.e., the concentration of NO remains constant within one experiment and the observed rate equals twice the rate of $O + NO + M$; 2) The reaction is much faster than the recombination of O with another O or with O_2, so that small additions of NO produce a large increase in the rate. This greatly facilitates the calculations of the rate constant, k_1, for $O + NO + M$.

An example of such a calculation is given in Fig. 2 and Table 2. In Fig. 2 are shown six decays along the flow tube of the emitted light intensity, I, due to the air afterglow. The decay is logarithmic in I, i.e., obeys first-order kinetics. The slope of each line, $d(\log I)/dx$, is multiplied by $\bar{v}$ to give $d(\log I)/dt$ which equals $k^1/2.303$. k^1 is the first-order rate constant defined by $-d(O)/dt = k^1(O)$. When some NO is present, $k^1 = k^o + 2k_1 (NO)(M)$, where k^o measures the slow, residual decay in the absence of NO due to

recombination at the wall or by reaction with O and O_2. From a knowledge of (NO) and (M), k_1 can be calculated as is shown in Table 2 for M = N_2O.

Table 2

Determination of k_1 for M = N_2O

Pressure = 3.71 mm Hg; $\bar{v}$ = 191 cm/sec; T = 25° C

$p_{NO} \times 10^3$	k^1	$\Delta k^1/p_{NO}$	$\Delta k^1/(p_{NO}p_M) = 2k_1$ corrected
mm Hg	sec^{-1}	$(mm\ Hg)^{-1}\ sec^{-1}$	$(mm\ Hg)^{-2}\ sec^{-1}$
0	1.69	----	----
3.44	5.79	1190	338
5.97	8.83	1195	340
11.0	15.0	1209	344
12.3	16.0	1160	329
17.3	22.7	1211	345

$$k_1 = 5.9 \times 10^{16}\ cm^6\ mole^{-2}\ sec^{-1}$$

The Δk^1 in the third column of Table 2 is the k^1 for the particular partial pressure of NO less the k^1 in the absence of NO. The correction in column 4 takes into account the small amount of O_2 which is always present. Table 2 is the result of six kinetic experiments at one pressure for M = N_2O. Such experiments were repeated at several pressures for M = He, Ar, O_2, N_2O, and CO_2, and yielded rate constants whose average values are shown in Table 3.

Table 3

Rate Constant for O + NO + M

	M = He	Ar	O_2	N_2O	CO_2
$k_1 \times 10^{-16}$ $cm^6\ mole^{-2}\ sec^{-1}$	2.7	3.1	3.2	5.6	4.1

There is a surprisingly weak dependence of k_1 on the nature of M. As the efficiency of M increases only by a factor of two between helium and nitrous oxide, it seems clear that the internal degrees of freedom of the third body, M, play only a small role in the stabilization of the collision complex of O and NO. This is very different from the published results of iodine atom recombination where greater complexity of M is accompanied by a large increase of the rate constant. In that case it is thought that I-atom recombination proceeds through the intermediate formation of loose collision complexes

between I and M which subsequently react with another I-atom. No such reaction path should therefore be postulated for the recombination of O and NO. The small range of k_1 may also be taken as evidence against the importance of excited NO_2 as an intermediate in the recombination, because the quenching of such a molecule should be greatly speeded by more complex third bodies.

c. The Recombination of O with O_2

The three-body reaction, $O + O_2 + M \rightarrow O_3 + M$, is one of the steps which control the steady-state concentrations of O and O_3 in the upper atmosphere. Its reverse is the initial reaction in the thermal decomposition of ozone, a much investigated reaction (15). The forward step has been studied in this laboratory and in others, but poor agreement still prevails. The reaction is not easily studied in discharge-flow systems, because it is very slow and therefore requires an accurate knowledge of all other processes in which O-atoms are recombined or destroyed: Wall recombination, reaction with NO, direct recombination (O + O + M), and, particularly, catalytic processes of hydrogen-containing discharge products from wet or impure oxygen.

Experiments were therefore performed in this laboratory with very pure oxygen made by the thermal decomposition of potassium permanganate. Yet, in very pure oxygen, discharges are strangely ineffective in producing dissociation as has been shown by us (16). This further complicates the experiments since both I and (O) are very small. At a given total pressure and flow of O_2, (O) is determined by NO_2-titration and the O-atom decay is measured by very small, downstream addition of NO so as not to accelerate the decay. The experiments are repeated, first at constant flow-rate of O_2 and increasing pressures, and then at higher or lower flow-rates. If (O) is sufficiently small that the contribution of $O + O + M \rightarrow O_2 + M$ may be neglected, the decay is again logarithmic, i.e., first-order in (O), and the measured rate constant $k^1 = k_W + 2k_2(O_2)^2$, where k_W is the first-order rate constant for surface recombination and k_2 is the rate constant for $O + O_2 + M \rightarrow O_3 + M$ with $M = O_2$. The factor of two arises from the assumption of a steady-state in which the ozone concentration, (O_3), remains very much lower than (O) because of the rapid reaction $O + O_3 \rightarrow 2\ O_2$. The plot of k^1 vs. $(O_2)^2$ has a slope of $2k_2$ and an intercept of k_W. As a result of many experiments a value of $4 \pm 1 \times 10^{13}$ cm^6 mole^{-2} sec^{-1} for k_2 is recommended. This value is considerably lower than other published estimates and it raises the interesting and yet unanswered question why the association reaction of O with O_2 should be almost 1000 times slower than that with NO, though both reactions are exothermic, apparently do not have an

activation energy, and should be allowed on grounds of quantum mechanical selection rules.

Very recent work by Schiff (17) and by us has pointed up further difficulties with this reaction scheme. Approximate measurements of the steady-state concentration of ozone are in poor agreement with the above simple scheme and show the wrong dependence on the O_2 concentration. The interesting explanation has been advanced (17) that vibrationally-excited O_2 molecules from the concurrent, direct recombination O + O + M are continuously destroying O_3 molecules and are keeping the O_3 concentration well below its normal steady-state value. Experiments are being set up to test this hypothesis and rapid progress can be expected in this important research area.

Space limitations do not permit the proper description of related experimental studies in our laboratories. These include the study of fast radical reactions such as H + O_3 and H + NO_2 by absorption spectroscopy (18, 19). This work will further clarify the mechanism of the night airglow in which vibrationally-excited OH radicals emit radiation. Lastly, an investigation of high-frequency discharges is under way in which the concentration of free electrons is measured by microwave diagnostic methods. By simultaneous measurements of atom and radical concentrations, and later of ionic species it is hoped that information will be obtained on important ion-electron and ion-molecule reactions.

REFERENCES

(1) F. Kaufman, Proc. Roy. Soc. A247, 123 (1958).

(2) F. Kaufman and J. R. Kelso, 7th Symposium on Combustion, Butterworths, 1959, p. 53.

(3) F. Kaufman and J. R. Kelso, Symposium on Chemical Reactions in the Lower and Upper Atmosphere, Interscience, 1961.

(4) F. Kaufman, Progress in Reaction Kinetics, Vol. I, Chapter 1, Pergamon Press, 1961.

(5) R. J. Strutt (Lord Rayleigh) Proc. Phys. Soc. 23, 66, 147 (1910).

(6) H. P. Broida, H. I. Schiff, and T. M. Sugden, Trans. Farad. Soc. 57, 259 (1961).

(7) A. Fontijn and H. I. Schiff, Symposium on Chemical Reactions in the Lower and Upper Atmosphere, Interscience, 1961.

(8) G. Herzberg, Spectra of Diatomic Molecules, Van Nostrand, 1950, p. 400.

(9) F. Kaufman and J. R. Kelso, J. Chem. Phys. 28, 510, (1958).

(10) F. Kaufman and J. R. Kelso, J. Chem. Phys. 27, 1209, (1957).

(11) F. Kaufman, J. Chem. Phys. 28, 992, (1958).

(12) L. Elias and H. I. Schiff, Can. J. Chem. 38, 1657 (1960).

(13) R. R. Reeves, Jr., G. Mannella, and P. Harteck, J. Chem. Phys. 32, 632, (1960).

(14) M. A. A. Clyne and B. A. Thrush, Trans. Farad, Soc. 57, 69, (1961).

(15) S. W. Benson and A. E. Axworthy, Jr., J. Chem. Phys., 26, 1718, (1957).

(16) F. Kaufman and J. R. Kelso, J. Chem. Phys. 32, 301, (1960).

(17) H. I. Schiff, private communication.

(18) F. Kaufman and F. P. Del Greco, J. Chem. Phys. 35, 1895, (1961).

(19) F. P. Del Greco and F. Kaufman, Farad. Soc. Discussion, 1962, in press.

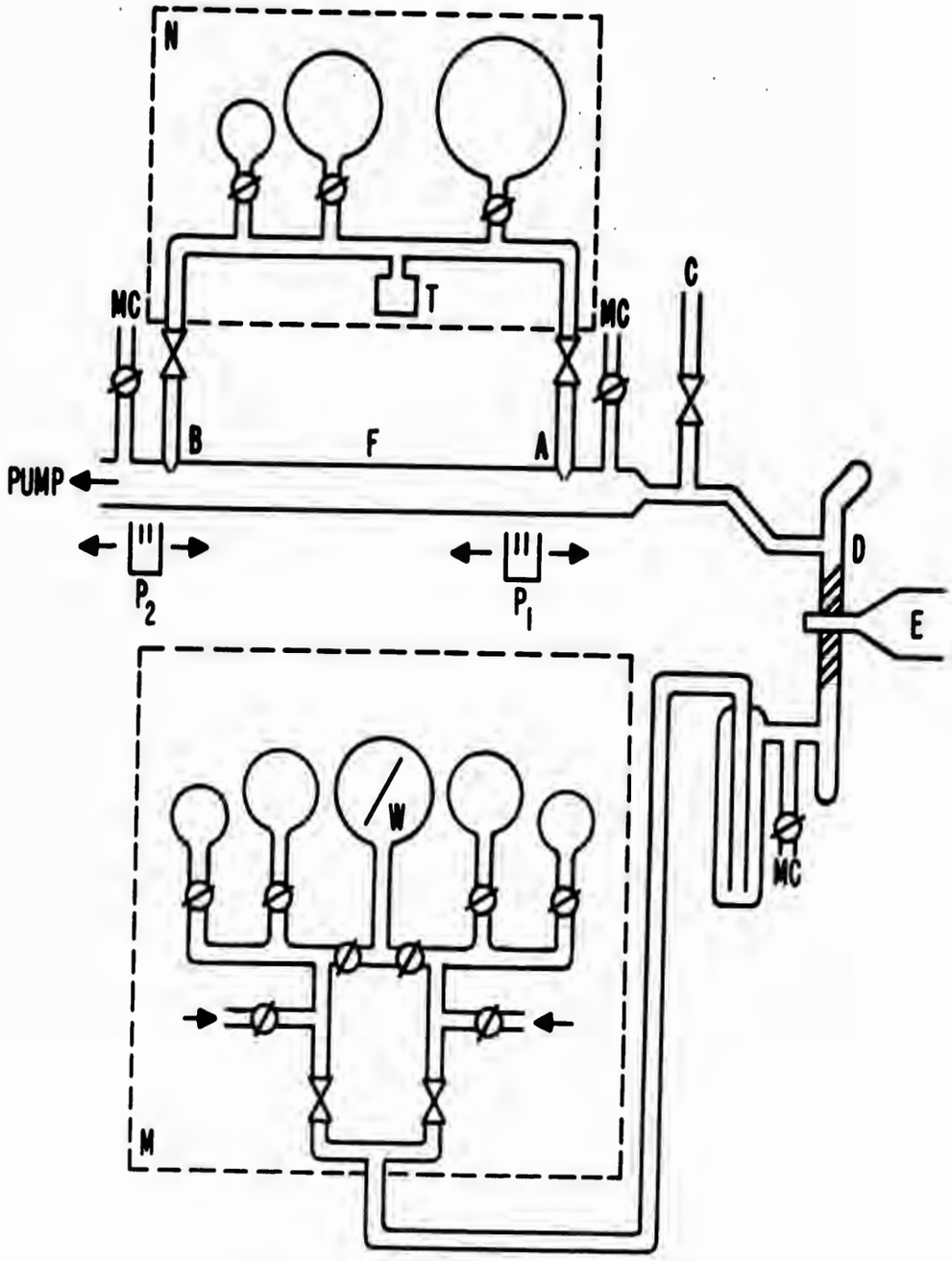


FIG. 1 DIAGRAM OF OXYGEN ATOM REACTION APPARATUS.

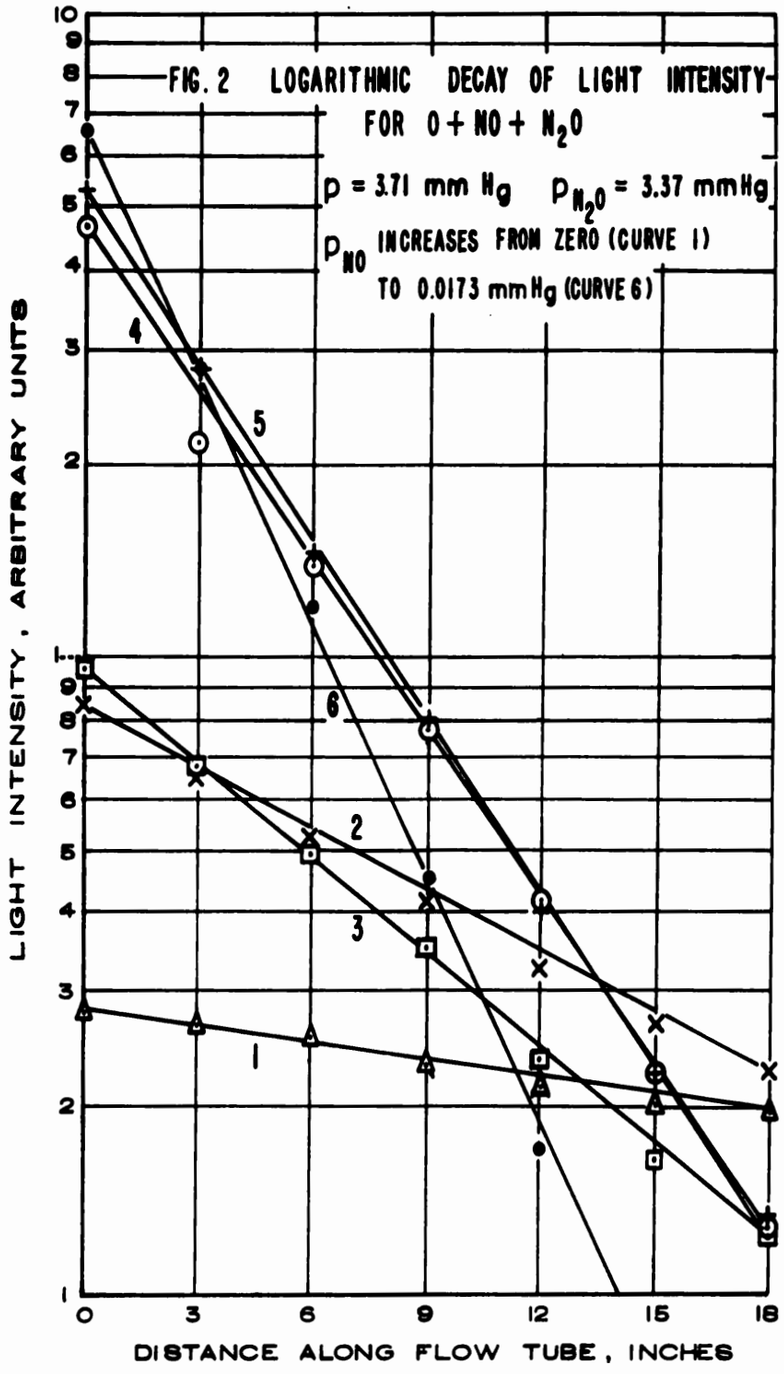


FIG. 2 LOGARITHMIC DECAY OF LIGHT INTENSITY FOR O + NO + N2O
p = 3.71 mm Hg pN2O = 3.37 mmHg
pNO INCREASES FROM ZERO (CURVE 1) TO 0.0173 mmHg (CURVE 6)
LIGHT INTENSITY, ARBITRARY UNITS
DISTANCE ALONG FLOW TUBE, INCHES
0
3
6
9
12
15
18
1
2
3
4
5
6

EVIDENCE THAT RHESUS MONKEYS DIE FROM ANTHRAX TOXIN

F. KLEIN, LT. D. R. HODGES (VC), B. G. MAHLANDT
W. I. JONES, B. W. HAINES and
R. E. LINCOLN

U. S. ARMY CHEMICAL CORPS
FORT DETRICK, FREDERICK, MARYLAND

Some of the earliest evidence of toxin production during the growth of Bacillus anthracis, although apparently unrecognized as such, were presented by Sobernheim (8) who described anthrax as producing edema and hemorrhage and interferring with blood clotting and carbohydrate metabolism. Later, Howie and Cruickshank (3) showed that they could increase the lethal effect of anthrax by injecting shock-producing substances into animals previously challenged with B. anthracis. Smith et al. (7) observed that the terminal phase of anthrax is accompanied by secondary shock which they reported to be the major cause of death. Their work, partly confirmed by Ross (6), lead to the demonstration of a specific toxin isolated from the plasma of guinea pigs which had died of anthrax. The plasma, sterilized by filtration, was shown to produce an edema when injected into the skin and death when injected intravenously into other guinea pigs. Recently, Stanley and Smith (9) and Beall and Taylor (1) independently discovered that "toxin" is composed of at least 3 components, identified by the latter authors as protective antigen, guinea pig edema factor and rat lethal factor. They also showed that rat lethal factor is toxic only in association with protective antigen.

It is our purpose to report on work demonstrating: (1) the presence of toxin in blood and lymph of the Rhesus monkey (2)quantitative increases of toxin in blood and lymph prior to death and (3) other evidence of toxemic death.

In order to obtain body fluids for these studies workers in our laboratories (Hodges and Rhian, 2) developed a survival surgical procedure for cannulation of the thoracic and right lymph ducts and the jugular vein of the Rhesus monkey. The procedure involves opening the chest to enter the superior mediastinum by bisecting the sternum thus exposing for cannulation both right and thoracic lymph

ducts. These ducts are pictured diagrammatically in (Fig. 1). Simultaneously a venous cannula is placed in the heart by jugular vein cannulation. Following this procedure the monkeys recover rapidly as shown in this picture of a monkey 24 hours post surgery. (Fig. 2) Lymph is collected continuously and blood samples can be drawn as desired. For this study, changes in concentration of microorganisms and toxin in the lymph and blood were measured following challenge with B. anthracis.

Quantitative assay for toxin was based on two procedures. The first, that of Beall and Taylor (1) based on the susceptibility of Fischer rats to anthrax toxin, and the second, based on the agar diffusion technique described by Thorne and Belton (10) to measure the precipitation of antibody and antigen.

Our studies are based on the following observations: (1) toxin is present in the terminal blood of monkeys dying of anthrax, (2) toxin increases in concentration in both the lymph and blood reaching a maximum at death, (3) sterile toxin alone causes death of the monkey and (4) rapid death occurs in monkeys following an injection of 10^{11} B. anthracis spores.

During the course of work reported elsewhere, in which the course of anthrax and the development of septicemia in the Rhesus monkey was studied, we tested the terminal blood of 15 Rhesus monkeys for toxin. In 12 of these terminal blood samples toxin was demonstrated by death of the rat.

Increasing concentrations of toxin, to a maximum at death, were obtained from two monkeys studied to date. The thoracic lymph duct and the jugular vein of the first monkey were cannulated. Twenty four hours post-surgery the monkey was inoculated intradermally with 10^5 virulent anthrax spores. The site of inoculation drained to the popliteal lymph node. The animal died 41.5 hours after challenge. Lymph was collected regularly during that time and the samples were kept separated so that changes in the lymph could be correlated with time. Changes in organisms and toxin levels are shown in Table 1. The results of three assays of each sample are shown. For instance, the lymph collected from this monkey at 17 hours prior to death was found to contain $< 0.2 \times 10^6$ orgs/ml while lymph collected 3 hours prior to death contained 88×10^6 orgs/ml. During this same period the toxin titer increased from 0 to a 1/5 dilution. When 1 ml of the 3-hour lymph sample was injected intravenously into each of two rats they survived for an average time of 144 minutes. From this table it can be seen that the lymph collected near the death of the monkey was both more antigenic and more lethal to rats than lymph collected several hours prior to death. This observed increase in toxin is statistically significant ($P < .001$).

Sterile lymph collected before challenge was not toxic to rats. It was also proven that a combination of two parts of the toxic lymph and one part of anthrax antiserum* was non-toxic. The above tests demonstrate that death was caused by toxin produced by B. anthracis infection.

The second experiment was similar to the first except the right lymph duct, which drains a major portion of the lungs, was used instead of the thoracic duct. Also, the monkey was challenged by the aerosol route with 3×10^5 anthrax spores. The data from this monkey, which died in 57 hours, are shown in Table 2. Since the lymph flow was very slow only a small amount could be collected over any time period. Thus the toxin test used only one rat per point and challenge was with only 0.5 ml of lymph. Therefore, the times to death on this table do not compare directly to those on the preceding table. The regression of time prior to death of the monkey on time to death of the rat has a slight significant slope ($P < .10$). Although the level of significance is low, probably due to the small number of rats, there is an increase in toxin concentration during the period of observation. These observations parallel those made in the previous experiment. The agar precipitation test was not performed on any of these samples because of insufficient material.

The terminal blood of each of these two monkeys was assayed for toxin. That of the first monkey was found to be negative whereas the terminal blood of the second monkey was toxic for rats. From Fig. 3 it is seen by the dotten line that the concentration of organisms in the lymph of the first monkey built up rapidly to approximately 10^8 organisms per ml. There was no such buildup however in the blood (solid line) of this animal. It is shown in Fig. 4 that the concentration of organisms in both the blood and lymph of the second monkey increased to relatively high levels. As with the first monkey, however, the buildup of organisms in the lymph was greater than in the blood. It was also found that the terminal blood of the second monkey was toxic for rats which died four hours after being injected with 0.5 ml of the blood.

The fact that blood of the first monkey did not develop toxin to a demonstrable level is not unexpected. We have shown that the principal extravascular route of anthrax organisms and toxin into the blood is through the thoracic and right lymph ducts. Malek et al. (4) demonstrated in sheep that at the peak of inflammation of the lymph nodes a pathological secondary lympho-venous communication is established which allows direct passage of organisms into the blood.

* Anthrax spore antiserum from hyperimmunized horse (DH-1-4A)

Since we collected all thoracic and right lymph before it entered the blood stream, we may have removed most of the organisms and toxin not actually diffused by this secondary pathological communication. Anatomically there are fewer lymph nodes in the popliteal and inguinal region than in the region draining the lungs. Hence, it is a question of the number of pathological secondary lympho-venous communications between the two lymph systems that may account for the absence of toxin and low concentration of organisms in the intradermally challenged monkey.

The data show that toxicity is proportional to the concentration of organisms in the body fluids. Specifically the lymph of both monkeys showed parallel increases of both organisms and toxicity. The blood of the first showed a negligible increase of organisms and no toxicity, whereas the blood of the second showed a substantial increase in organisms accompanied by a demonstrable toxicity. This is the first demonstration for such a build-up of toxin in body fluids of monkeys and complements the British observation on toxin in terminal blood of guinea pigs.

The next step was to show that anthrax toxin alone is sufficient to cause the death of these animals. Sterile toxin was produced *in vitro* by the method of Thorne *et al.* (11) of a potency such that 1 ml doses killed rats in an average of 80 minutes. A dose of 250 ml of this toxin was administered intracardially to a Rhesus monkey weighing 16 pounds. The monkey received the toxin with no apparent ill effects. It died 21 hours post-challenge with terminal symptoms typically associated with toxic death in rats. (i.e., partial paralysis of upper extremities and extreme respiratory distress, accompanied by massive pulmonary edema).

In another trial three 8-pound Rhesus monkeys were challenged intravenously with 200 ml, 250 ml and 400 ml respectively, of sterile *in vitro* produced anthrax toxin. A fourth monkey received 250 ml of toxin followed by three 50 ml injections of antiserum given 0, 1/2 and 21 hours post-challenge. The animal receiving 200 ml survived; the animals challenged with 250 ml and 400 ml died in 60 and 33 hours respectively. The animal receiving anthrax antiserum survived. These data indicate a dose-response relationship between anthrax toxin and death of the Rhesus monkey and that toxin was neutralized by antiserum.

The final evidence of toxemic death of Rhesus monkeys was revealed when a dose response curve for challenge doses of 10^5 through 10^{11} spores was run. The expected dose response was observed for the doses 10^5 through 10^{10} spores with death occurring between 20 and 50 hours. Three monkeys given 10^{11} spores, however, died 2 hours after challenge with symptoms of toxemic death. We interpret these observations as showing that 10^{11} spores, on germination and outgrowth,

produce enough toxin to kill a monkey. The entirely different course of the disease as evidenced by short time to death and only symptoms of toxemia indicate that these three monkeys died of anthrax toxin.

In summary it has been shown that: (1) Anthrax toxin can be demonstrated in body fluids of the Rhesus monkey at death. (2) Anthrax toxin is present in increasing amounts in body fluids of the Rhesus monkey dying of anthrax. (3) In vitro produced, sterile anthrax toxin is capable of causing the death of Rhesus monkeys. (4) In the one instance in which it was tried, anthrax antiserum protected the Rhesus monkey against in vitro toxin produced by B. anthracis.

The importance of toxin in affecting the pathogenic course and therefore the treatment of the disease in man remains to be determined. It would seem probable that the effect of toxin in man is like that in other animals. This additional knowledge of toxin has a significant implication on the treatment of the disease. Since the introduction of antibiotics, all treatment is directed against the organism. It seems clear now that in order to be fully effective in treating systemic anthrax it will be necessary also to direct treatment against the toxin. Thus antibiotics alone do not constitute adequate treatment, except, of course, when begun in a very early stage of the disease. It now remains to develop an effective antitoxin since the laboratory and clinical methods available to us at this time are inadequate for early detection of the disease. Like conclusions were drawn by Plotkin et al. (5) following their observations on the New Hampshire epidemic of inhalation anthrax who state "In view of the importance of toxemia to the outcome of the experimental disease, the use of antitoxin seems logically indicated in human inhalation anthrax".

REFERENCES

1. Beall, F.A. and M.J. Taylor. 1962 Accepted by J. Bact. for publication.

2. Hodges, D.R. and M.A.Rhian. 1962 J. Expr. Med and Surgery, 20th Vol.

3. Howie, J.W. and J. Cruickshank. 1947 J. Path. and Bacteriol. 59: 127.

4. Malek, P., J. Kole and F. Zak. 1959 J. of Bact., Parasit. and Hyg.; Abt. I Orig. 174: 94.

5. Plotkin, S.H., P. S. Brachman and M. Utell. 1960 The Amer. Jour. of Med. Vol. XXIX: 992.

6. Ross, J.M. 1955 Brit. J. Expr. Path. 36: 336.

7. Smith, H., J. Keppie, J.T. Stanley and P.W. Harris-Smith. 1955 J. Gen. Microbiol. 17: 731.

8. Sobernheim, G. 1931 Handbuck de Pathogenen Mikro-organismen 3: 1041 Gustov Fischer, Jena.

9. Stanley, J.L. and H. Smith. 1961 J. Gen. Microbiol. 26: 49.

10. Thorne, C.B. and F.C. Belton. 1957 J. Gen. Microbiol. 17: 505.

11. Thorne, C.B., D.M. Molnar and R.E. Strange. 1960 J. Bact. 79: 450.

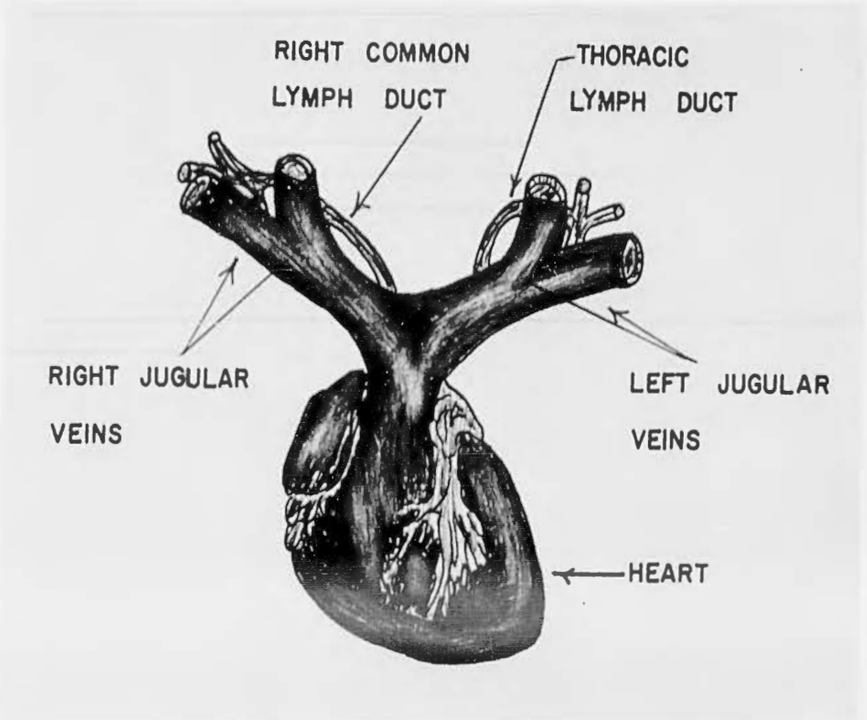


FIGURE 1. LOCATION OF THORACIC AND COMMON LYMPH DUCTS ANTERIOR VIEW.

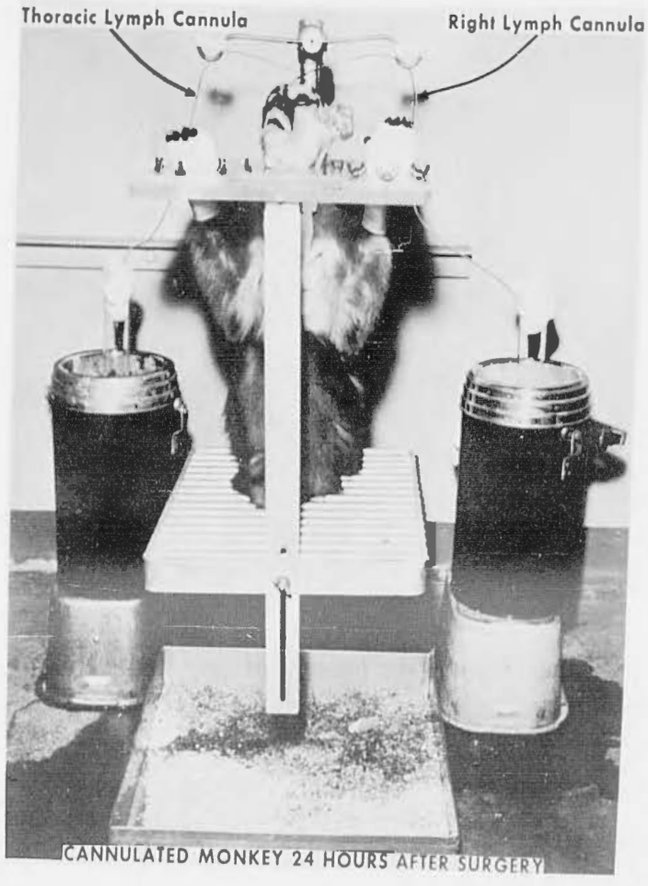

Thoracic Lymph Cannula
Right Lymph Cannula
CANNULATED MONKEY 24 HOURS AFTER SURGERY

FIGURE 2.

TIME PRIOR TO DEATH OF MONKEY (HRS.)	ORGS. IN LYMPH (10^6/ML)	TESTS FOR TOXIN	
		RAT (MIN. TO DEATH)	PRECIPITIN (TITER)
17	<.2	S	0
13	52	S	1/3
11	71	360	1/4
9	69	233	1/4+
7	168	143	1/5
5	85	126	*
3	88	144	1/5
DEATH	69	76	*

* NOT ENOUGH LYMPH COLLECTED TO RUN TITER.

S SURVIVED.

TABLE I. ANTHRAX ORGANISMS AND TOXIN IN LYMPH FROM MONKEY CHALLENGED BY INTRADERMAL ROUTE.

TIME PRIOR TO DEATH OF MONKEY (HRS.)	ORGS. IN LYMPH	TEST FOR TOXIN
		RAT
	10^6/ML	(MIN. TO DEATH)
8	15	S
7	18	240
5	19	237
3	15	180
DEATH	26	177

S SURVIVED

TABLE 2. ANTHRAX ORGANISMS AND TOXIN IN LYMPH FROM MONKEY CHALLENGED BY AEROSOL ROUTE.

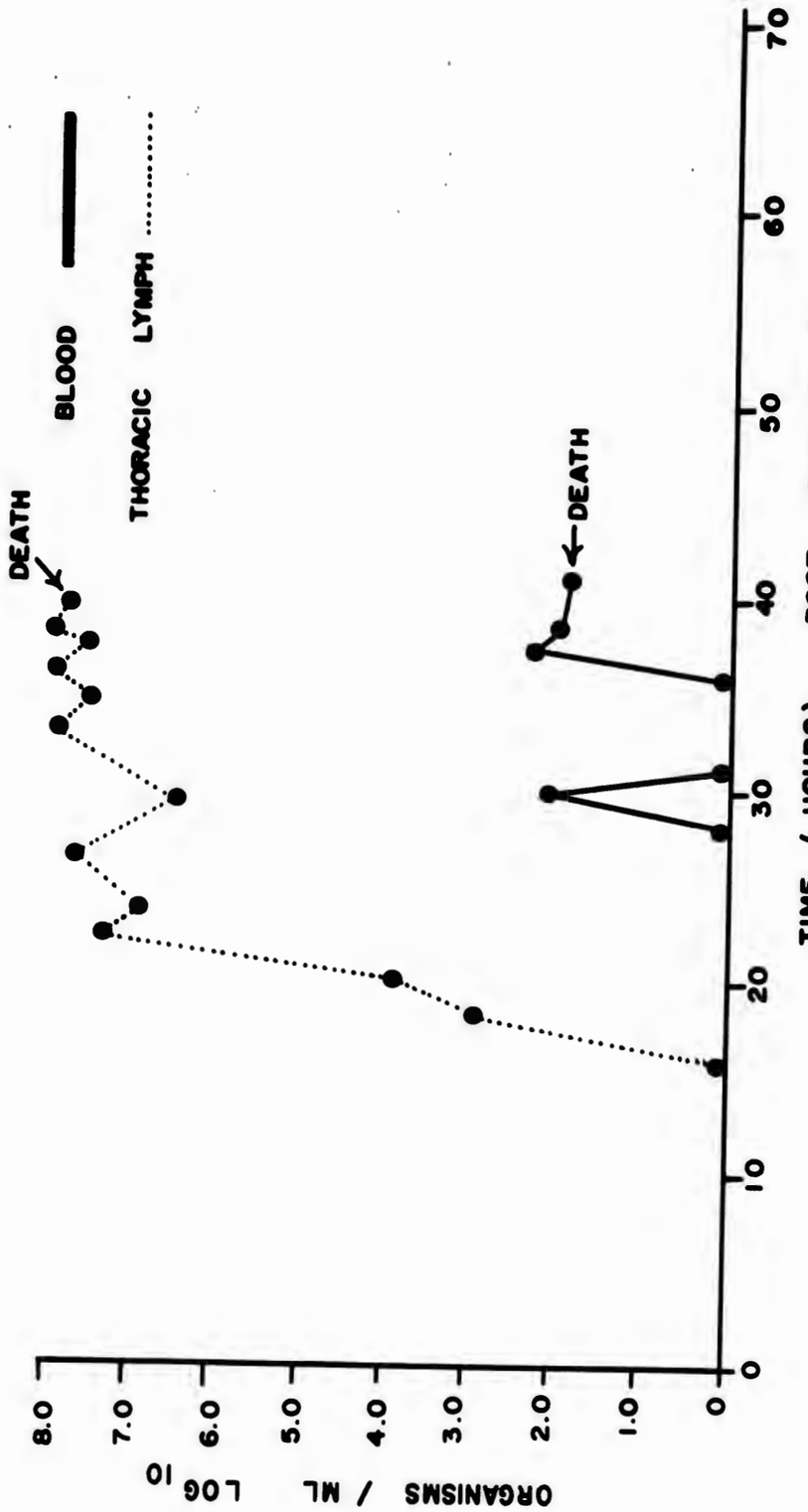


FIGURE 3. BACILLI IN THE BLOOD AND THORACIC LYMPH OF A MONKEY INTRADERMALLY CHALLENGED WITH SPORES OF B. ANTHRACIS.

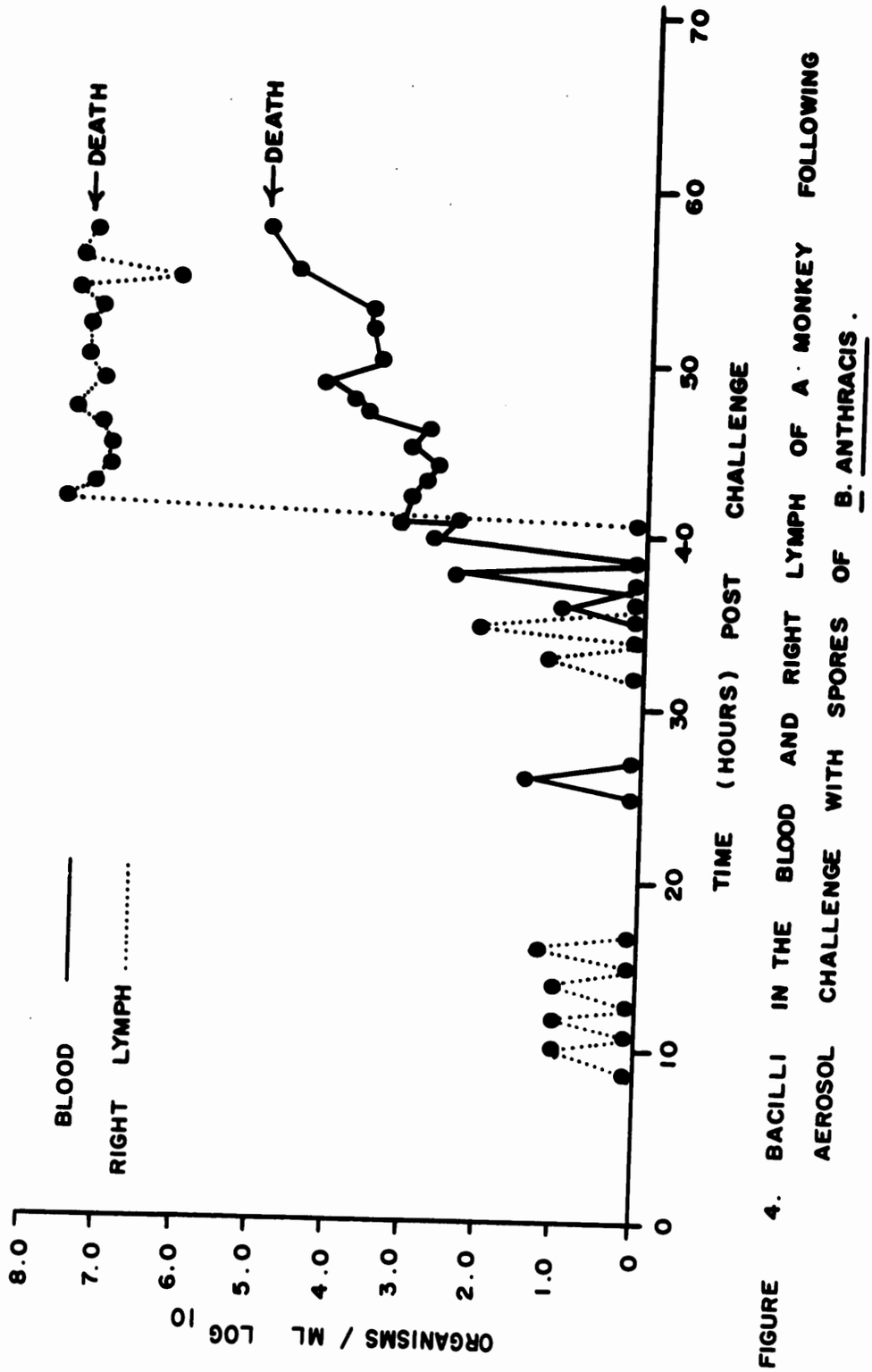


FIGURE 4. BACILLI IN THE BLOOD AND RIGHT LYMPH OF A MONKEY FOLLOWING AEROSOL CHALLENGE WITH SPORES OF B. ANTHRACIS.

CLASSIFICATION OF LANDSCAPE GEOMETRY FOR MILITARY PURPOSES

CHARLES R. KOLB
U. S. ARMY ENGINEER WATERWAYS EXPERIMENT STATION
VICKSBURG, MISSISSIPPI

Introduction

This paper outlines the method of classifying landscapes or terrain geometry developed at the U. S. Army Engineer Waterways Experiment Station, Corps of Engineers (WES). The classification was developed as part of a more comprehensive system of terrain description used to compare the terrain at the Research and Development Test Station at Yuma, Arizona, with the terrain of various desert areas throughout the world. Limited numbers of map-folio reports have been prepared which utilize this classification in mapping the desert regions of Northwest Africa, Northeast Africa, the Middle East, South Central Asia, Southwest United States, and Mexico. Details of the mapping techniques, the complete terrain classification, and the methods used to prepare terrain analogs are contained in WES Technical Report No. 3-506, A Technique for Preparing Desert Terrain Analogs, June 1959.

In this paper terrain is considered to be the sum of the various physical attributes of the land that describe an area. A terrain factor is an attribute of terrain such as slope, relief, etc. The effect of such a factor or of a combination of factors on an individual military activity, such as cross-country movement, air-drop potential, etc., is a terrain effect. A factor family is a convenient grouping of terrain factors. In the WES studies five factor families are considered: landscape geometry, microrelief, surface composition and consistency, hydrography, and vegetation. A mapping unit is a specific class of values or a definable subdivision of a terrain factor. For example, relief is a terrain factor and 0-10 ft of relief is a mapping unit.

Basis for Classification

General

Landscape geometry is the configuration or geometry of the landscape without regard to the vegetation that grows on it or the soil or rock type of which it is composed. It is an attribute of terrain that has long been qualitatively and imprecisely categorized in terms such as hilly, mountainous, rolling, etc. But what may be hilly to the mountaineer may be mountainous to the flatlander. Also, the same landscape may be described by one observer using one set of terms, and by another observer--or even the same observer at a later date--using an entirely different set of terms. Because there is no uniform system of description, the observer is forced to use qualitative terminology that varies to suit the occasion. Few deny the need for a precise, quantitative description of terrain. Such a description is essential if widely separated areas are to be compared, and if the effect of the configuration of the landscape on various military activities is to be critically assessed. This would in turn improve the application of operational and testing experience to other areas of the world, and make possible better estimations of operational capabilities and requirements.

Selection of terrain factors

The obvious advantages of a quantitative approach are its objectivity and the fact that mapping units can be rigorously defined. A more subtle, but equally important advantage is that quantitative factors offer the possibility of being manipulated mathematically in such a way that the effects of individual terrain factors acting in concert may be determined. In most instances such quantitative effects have evolved from studies aimed at determining terrain effects in special fields such as hydrology and agriculture. As a result, quantitatively expressed factors useful in presenting an aggregate or entire picture of terrain have not been explored to any great extent. Usually the process of quantitative analysis results in a multiplicity of quantitative measurements to express a simple descriptive qualitative term. Thus, the quantitative system gains in precision, but loses in simplicity. Quantitative terrain description has its limitation and should not be employed unless its contribution outweighs the complexities it introduces.

A major problem in developing the WES system of landscape classification was to choose simple descriptive terrain factors which could be measured and which describe a given landscape as completely as possible. There are many measurable properties relating to surface geometry that can be utilized in a quantitative system of terrain description. In fact, 37 such properties are employed by one analyst in describing fluvially eroded areas alone. Thus the problem is one of limiting the number of factors while still providing a reasonably complete picture of the terrain. Considerable effort was

spent in limiting the number of terrain factors while making sure that factors truly important in landscape description were not disregarded, and also in selecting and grouping terrain factors that, when considered in concert, are readily visualized and result in a minimum of cartographic complexity.

Of considerable importance in the selection of terrain factors is their "mappability" in poorly mapped or unmapped areas of the world. In other words, the factors chosen for mapping must be so chosen and so stratified that valid assumptions can be made concerning how they are to be mapped in areas where data are scarce. Thus, a knowledge of the basic rock type of an area often permits fairly accurate estimates of the depth of dissection, the density of deep drainageways, and whether the topographic highs are flat-topped or serrate. The general geologic structure often enables an evaluation of parallelism of the topographic forms and of relief and slope aspects. A knowledge of the geomorphic history and landform development of an area may permit the analyst to evaluate such terrain factors as areal extent of the topographic highs which rise above a plain or the slopes which are characteristic of these highs.

Selection of mapping units

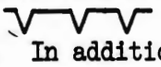
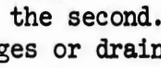
Similarly, naturally occurring groupings or ranges of terrain factors played an important role in establishing the various mapping units. For example, detailed reports dealing with alluvial fans conclude that the slope of a fan is always less than 17.5%, and in the vast majority of cases, less than 10%. Studies of barchan dunes indicate that windward slopes range from 9 to 25%. These, and additional data indicated that valid boundaries between slope-mapping units could be set at approximately 10 and 25%. Use of such naturalistic breaks permitted maximum use of landform-terrain factor ties or associations in mapping areas where little quantitative or detailed information was available.

Availability of data was another determinant in establishing mapping units. The same units of breakdown had to be utilized in mapping all the world desert areas. Unless this system was adhered to in the entire mapping program, the mapping of one area would have no relation to that of another, and no comparisons could be made. It should be emphasized, however, that availability of data was not permitted to unduly restrict the number of units with which each terrain factor was mapped. Individual terrain factors were subdivided into units sufficiently detailed and restrictive to reasonably define a given landscape from both a military and geomorphic standpoint. Where these ranges could not be mapped from available data, inference or judgment based on soils, landforms, or other associations was relied on.

To summarize, landscape geometry factors were chosen for mapping on the basis of (a) their importance, individually, as basic

elements of a given landscape, (b) their capacity, when viewed in concert, to provide a reasonably complete picture of a given terrain, and (c) their "mappability" in poorly mapped areas of the world. The selection of mapping units for each of these landscape factors was based on such considerations as (a) naturalistic breaks, and (b) adaptability of the unit to precise and, wherever possible, to quantitative definition.

Method of Classification

A landscape or segment of landscape can be identified in terms of slope, relief, and dissection or slope spacing. Thus, describing an area as having gently sloping, moderately spaced hills rising to considerable heights above the surrounding plain gives a reasonable picture of the landscape. When a region is described as having hills with slopes ranging between 17.5 and 35%, spaced about 1000 ft apart, and rising to heights of 50 ft, a more quantitative and clearer picture of the landscape is provided. Therefore, assigning a reasonable numerical range to these three factors and cataloging various combinations of ranges of these three factors are logical steps in landscape definition. A less tangible but equally important property necessary to complete this definition is the spatial distribution of these three factors, called, in this system of classification, the plan-profile. To illustrate, consider a gently sloping plain dissected by a number of deep, narrow drainageways. This would be mapped as an area of a given range of slopes, relief, and slope spacing. Consider another gently sloping plain with a series of narrow dikes or ridges crossing it. This would be mapped with exactly the same ranges of slope, relief, and slope spacing, yet the disposition of features composing the landscape in each instance would be different. A profile of the two landscapes would appear as  in the first instance, and as  in the second. In addition, it is desirable to know whether the ridges or drainageways are parallel or intersecting, continuous or discontinuous; in other words, a plan view of the area is needed. The plan-profile thus becomes a necessary part of landscape definition.

Characteristic slope

A slope may be defined as a surface designated in terms of its deviation from the horizontal. The amount of deviation is commonly expressed as a rate of vertical rise per horizontal interval, as a percentage, or in degrees. Slopes in nature are characterized by an infinite number of such deviations from the horizontal, and to map these changes in slope some standard of generalization must be adopted. Characteristic slope, therefore, is defined in terms of a set standard of generalization. Primarily, this standard is the contour interval found on available map coverage, the contour interval that determines the resulting topographic picture or "envelope."

For example, a map with a contour interval of 200 ft gives a much more generalized topographic picture of a region than a map of the same area utilizing a 25-ft contour interval. The map with the smaller contour interval will indicate many more slopes than its 200-ft counterpart, and therefore, the characteristic or most frequently encountered slope would differ. It follows that, up to a point, the smaller the contour interval the truer the resulting picture of the surface. It is impractical, however, to consider slope breaks apparent only on maps with 1-, 2-, or 3-ft contour intervals. Besides the almost insurmountable problems inherent in mapping surfaces in such minute detail, this detail often obscures slopes of major importance generated by larger contour intervals. For the system described in this paper, it was found convenient to arbitrarily set the limiting scale of generalization as that defining landscapes and associated slopes generated by a 10-ft contour interval. Irregularities apparent only on maps with contour intervals less than 10 ft are considered as microrelief and, as such, mapped separately as a distinctive factor family.

Characteristic slope is defined as the narrow range of slopes that predominates or is most common within a region based on an examination of the surface generated by a 10-ft contour interval. The utilized stratification of characteristic slope is shown in fig. 1. The breaks or boundaries are primarily "naturalistic" in that they help to define certain commonly recurring desert surfaces. Unit 1A (0 to 1%) is typically associated with very flat surfaces such as playas; unit 1B (1 to 3-1/2%) often identifies the very gently tilted valleyward peripheries of alluvial fans or alluvial flats common in desert regions; unit 2 (3-1/2 to 10%) brackets the slopes typically formed by alluvial fans; unit 3 (10 to 25%) represents the normal range of slopes forming the windward sides of sand dunes and consequently forms the characteristic slope of most dune fields; unit 4 (25 to 50%) characterizes many hill regions and the upper break approximates the limit of many subdued "angle of repose" deposits such as talus. The 100% break between units 5 and 6 is quite arbitrary; however, the characteristic slope of mountainous regions rarely exceeds this value.

Characteristic relief

Relief may be defined as the maximum difference in elevation per unit area. Although this definition is adequate in areas of fairly high characteristic slope, it is inadequate in regions exhibiting low characteristic slope or poorly developed drainage. If, for example, we define relief as the maximum difference in elevation per square mile, the surface of an alluvial fan with a slope of approximately 10% could exhibit "relief" of over 500 ft. In such cases the depth of incision of the drainageways is the important relief consideration rather than difference in elevation per unit area. Similarly, in sand dune areas the important relief consideration, the height of the dune above the interdune troughs, is usually missed when relief is expressed as elevation differential per small unit area. Consequently,

if certain important considerations are to be properly emphasized, a flexible definition of relief is needed.

A dual definition or classification of relief was, therefore, utilized in the WES system. Characteristic relief is considered to be (a) the modal vertical distance from interfluve crest to the immediately adjacent flow line in low-slope (less than 10%) areas, and (b) the maximum difference in elevation per square mile in high-slope (more than 10%) areas. Fig. 1 shows the ranges of relief values used in the mapping system.

Occurrence of slopes greater than 50%

Slope occurrence provides a measure of dissection or what might be termed "landscape compartmentation." For example, consider a well scored surface of an alluvial fan. The characteristic slope of the fan approximates 7% and the modal depth of the washes is 15 ft. As is usually the case, the wash-bank slopes are all steeper than 50%. In this case the landscape compartmentation consists of a series of flat-topped interfluves separated by pairs of steep slopes descending into narrow, shallow washes. Another type of compartmentation is exhibited in a field of complexly overlapping barchan dunes. The characteristic slope of the field falls between 10 and 25% (the windward slope of the dunes) and the modal relief is less than 50 ft. Here the landscape is "compartmented" by the steep (greater than 50%) slip-face slopes of the dunes. In contrast to the wash-scored alluvial fan surface, the steep slopes of the dune field do not occur in pairs. Consequently, in some cases it is desirable to describe compartmentation in terms of the frequency of occurrence of "compartmenting" slopes per unit of length, and in other cases in terms of average or modal distance between such slopes.

In the WES system, a compartmenting slope is considered to be any slope steeper than 50%. Here again the limit or boundary was established on the basis of naturalistic and military considerations. In desert regions most streams or wash banks and the surfaces of "angle of repose" deposits exceed 50%. This slope thus divides the landscape into compartments along such naturalistic breaks as stream banks, terrace escarpments, mountain slopes, dune slip faces, etc. Furthermore, except when wet, slopes less than 50% are of minor consequence from the standpoint of many military activities. Fig. 1 shows the ranges of occurrence values used in the mapping system.

Characteristic plan-profile

The three terrain factors discussed earlier are adequate measures of specific properties, but even taken in sum they fail to convey an impression of the geometry of the surface as a whole. It is possible, using any given combination of mapping units of these factors, to visualize a large number of quite different geometric configurations. Because overall configuration is vital to many military

activities, it is necessary to provide a descriptive fabric wherein the magnitude of slope, relief, and slope occurrence can be visualized.

Upon examination, it was determined that a minimum of four terrain factors was required to provide an adequate descriptive fabric. These four factors concern the spatial distribution and configuration in both plan and profile of the topographic highs and lows. The four factors selected are measures of: (a) the "peakedness" of the topographic highs; (b) the areal occupance of the highs; (c) the degree of elongation or planar shape of the highs; and (d) the orientation or degree of alignment exhibited by the highs.

It follows that for complete quantification of surface geometry the descriptive system should include four additional units, each indicating the distribution of one of the terrain factors mentioned above. However, the mental integration of these four factors into a visualization of the landscape is so difficult that they have been combined for the purposes of this study into a single composite terrain factor called the plan-profile. This has been done by classifying each of the four composite factors in such a way that, combined in the manner shown in fig. 2, they identify 24 basic variations of landscape configuration. Each of these 24 types, plus one other anomalous type, has been assigned a symbol and these types have been mapped as units.

Thus, the plan-profile is a composite terrain factor, and each component factor making up the composite is susceptible of quantitative treatment. It should be noted, however, that the final selection of the components of the plan-profile--even though each can be measured quantitatively--is a broad one; viz., the highs are either parallel or they are not, they are peaked or they are not, etc. Because of this broad stratification or classification, the choice of the plan-profile type can often be made by simple examination of an aerial photograph or a topographic map. It is only in those instances where one of the component factors, such as linearity, is doubtful that actual measurements must be made. The point is, that the decision as to whether the highs are linear or not can be made, wherever necessary, by quantitatively and objectively measuring them on topographic maps.

The dimensions of the landscape typified by the plan-profile are dependent upon the magnitude of relief and slope occurrence. For example, alluvial fans scored by steep-sided, shallow washes are mapped with the same plan-profile as extensive, high-standing, dissected plateaus. This is considered not only permissible but desirable, because with unrestricted dimensions, the plan-profile offers a convenient framework within which a knowledge of the values of slope, relief, and slope occurrence permits a readily assimilated mental image of the landscape.

Landscapes

A landscape map is prepared by simply combining plan-profile, slope occurrence, slope, and relief terrain factors. Through superposition of the various maps, combinations of these factors evolve, each representing a distinctive landscape. As the number of possible combinations of the four factors is rather appalling, it was considered impractical, during early stages of the WES studies, to try to develop a scheme for delineating individual landscapes. Multiplying the number of units of breakdown within each factor ($25 \times 6 \times 7 \times 7$) results in over 7000 mathematically possible combinations. Fortunately, however, nature has been far more selective, and superposition of the maps results in only 18 landscape types occurring at Yuma and only 78 in the entire North African Desert.

Thus the combining of the four basic geometry or form factors has resulted in a convenient method for cataloging landscapes. The method is believed to be as simple and straightforward as possible. Any landscape can be designated by a combination of four number of number-letter symbols, each representing a particular range of values of plan-profile, slope occurrence, slope, and relief. The combination shown in the upper diagram in fig. 3 (1L//-4-1b-2), for example, defines a plain with a slope of 1 to 3-1/2%, dissected by roughly parallel washes from 10 to 50 ft deep spaced from 1000 to 5000 ft apart. The lower diagram also defines a landscape type--in this instance type 5L//-2-4-6: a parallel ridge area with ridges from 2 to 10 miles apart, their heights ranging between 1000 and 400 ft, and their characteristic slopes between 25 and 50%.

The figure also illustrates conditions under which a dual or complex classification is mapped--the gross-component complex--which a moment's reflection will reveal as being necessary. Each landscape is composed of smaller landscapes and is, in turn, part of a larger or next order landscape. The lower diagram is referred to as a gross landscape. It is determined utilizing a 35-mile-diameter sampling circle, and is the landscape generated by a 100-ft contour interval. The upper diagram illustrates the restrictive landscape, a landscape determined by utilizing a 1-mile-diameter sampling circle with a 10-ft contour interval. Relief less than 10 ft falls into the category of microrelief or surface roughness. This is a convenient, if arbitrary, subdivision of terrain envelopes which has turned out to be of great value in the problem of landscape mapping. Wherever a gross-component relation exists, such as that shown in fig. 3 where the upper landscape is a smaller, restrictive, or component part of a larger or gross landscape, each is mapped with an appropriate symbol and identified.

Preliminary studies indicate that valid and worthwhile inferences can be made of the landscape geometry of a particular region from maps employing a 500-ft contour interval. These studies indicate that reasonably valid relations can be established among surfaces

defined by such maps, surfaces defined by 100-ft contour interval maps, and those defined by 10-ft contour interval maps. Detailed studies are needed to graph and compare geometry factor values of selected regions determined or mapped at different contour intervals and scales. Relations among the hierarchy of surface envelopes could then be computed for all areas mapped, and hypotheses derived and tested to explain significant variations in these relations--very possibly dependent upon lithology and climate.

Applications

Although the system of landscape classification described in this paper was developed for desert terrain, it has been used with reasonably good results in temperate and tropical areas. Some changes in the mapping units used for each of the factors may be necessary to reflect more closely naturalistic breaks common to humid areas. To date, the system has been used principally in developing desert analogs of the terrain at the Research and Development Test Facility at Yuma. The purpose of this analog development is to assess the suitability of Yuma as a desert testing site for men and materiel. After the method has been applied to the world deserts under consideration, it may be found that the Yuma Test Station lacks certain ranges or combinations of terrain-factor ranges found in other desert areas. Application of the technique to the desert of western United States should permit an evaluation of areas within the United States that may be more analogous to aggregate world desert conditions than the Yuma area, or which, when considered with Yuma, will cover a much more representative range of desert terrain. Such areas could well be utilized as supplementary testing sites. (It should be remembered that analog development and the study of the effect of terrain on military activities in general must be based on all the terrain factor families, not solely on the landscape geometry factors discussed in this paper.)

The determination of terrain analogs is only a preliminary step toward the important goal of predicting quantitative impacts of terrain on individual military activities. These activities are each concerned with different terrain factors and with varying degrees of refinement of the mapping units within each factor. Thus, after the various terrain factors have been modified to meet the needs of a particular problem, tests can be conducted in areas characterized by distinctive combinations of the devised mapping units. The important point is that the same system of characterizing the testing environment be used in all areas so that tests in one area can be related specifically to tests in another.

As an example, a comprehensive test program concerned with developing methods of predicting soil moisture for trafficability purposes is presently being undertaken by the Army Mobility Research Center of the WES. Soil-moisture-prediction sites are scattered throughout the United States and other parts of the world. One of the parameters affecting soil-moisture variation which is being considered is

the landscape type within which each site falls. Similarly, the landscape classification is being considered for selection of parachute-drop test sites and for estimating the impact of varying terrain types on airfield construction effort.

An interesting application presently (fiscal years 62-63) under way is the selection of courses at the Yuma Test Facility for testing the mobility of the Overland Train, a logistical cargo carrier built for the Transportation Corps and designed for cross-country movement. It has 10-ft-diameter tires, a length of 567 ft, and consists of 37-ft-long individual carriers capable of hauling a total of 136 tons. A number of test courses, approximately 200 ft wide and totaling more than 50 miles in length, were selected for the test runs. These courses cross terrain that is areally significant in deserts of North Africa, the Middle East, and South Central Asia. The test program is designed to study the effects of various types of terrain on the cross-country performance of the Overland Train. Performance data to be collected along each terrain type will include braking, steering, and driving energy; effects of vertical and horizontal acceleration on a cargo car and a supporting axle; fuel consumption; and dust densities developed. It is anticipated that grouping certain of the terrain mapping units within selected terrain factors will be required to account for the performance data. Conversely, it may be found that performance variations will occur in response to conditions within a single mapping unit. For example, it may be found that a certain slope value which lies within one of the established slope mapping units is critical in the performance of the Overland Train. In this event, it may be necessary to subdivide the existing mapping units in order to adequately describe the effects of terrain on performance. Ultimately, these data will be employed to construct a terrain effects map, in which ranges of effects on the Overland Train will be used as mapping units.

Summary and Conclusions

Landscape description to date has been almost entirely subjective, or where it is quantitative, often designed with a specific and fairly limited purpose in mind. The system outlined in this paper approaches an objective framework on which to hang terrain description wherever and for whatever purpose it is needed: determining the suitability of desert testing sites, assessing airfield construction difficulties, soil-moisture site classification, Overland Train route selection, to mention important military activities to which the method is presently being applied. Further, this framework is in the form of specific and definable physical properties or dimensions. No claims are made that this landscape classification is the ultimate in either sophistication or utility. Studies now under way, for example, promise considerable improvement in quantifying the classification. But the system as outlined is a possible approach and, it is hoped, will help to bring some order into the chaos of subjectivity with which landscape and terrain description, in general, must contend.

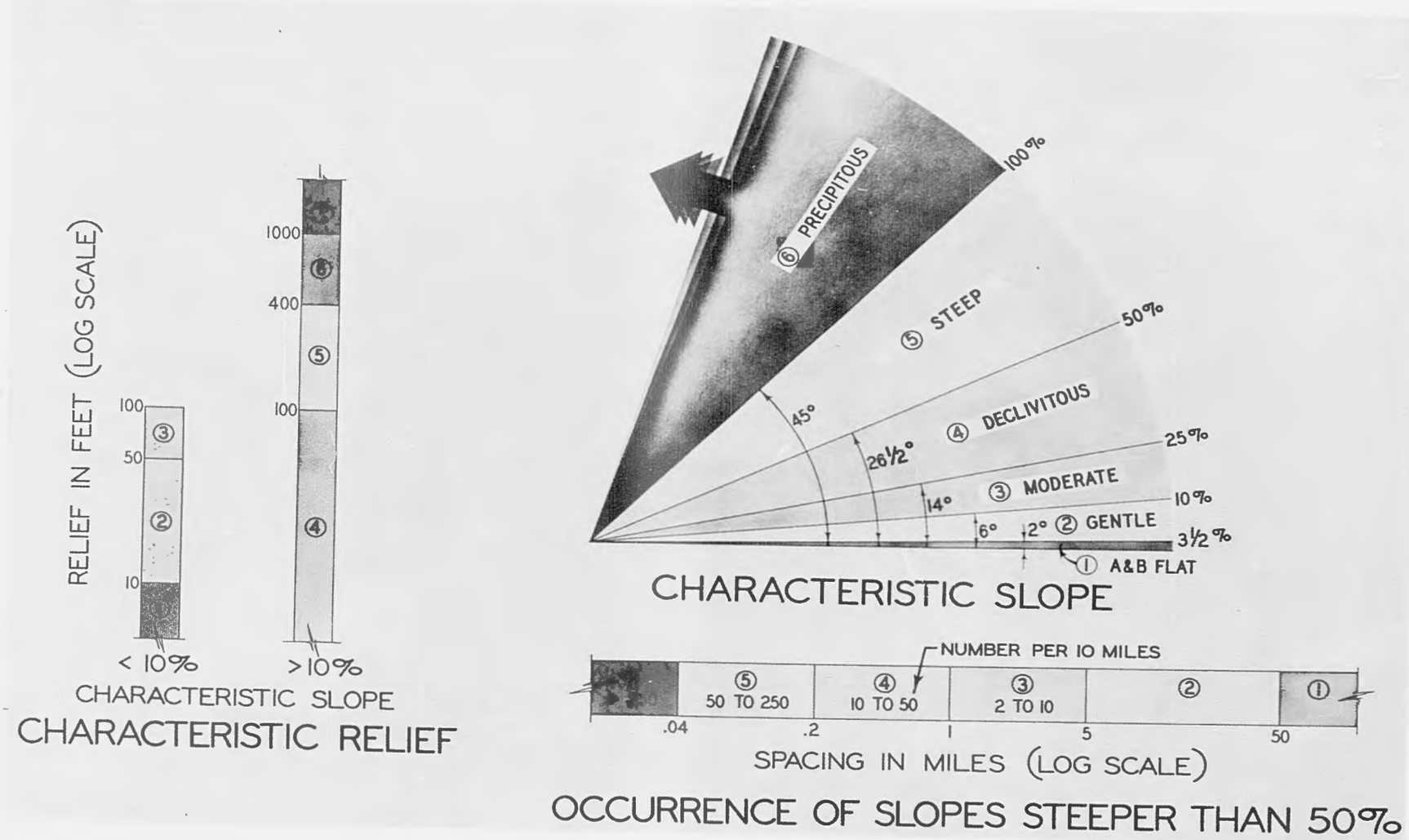
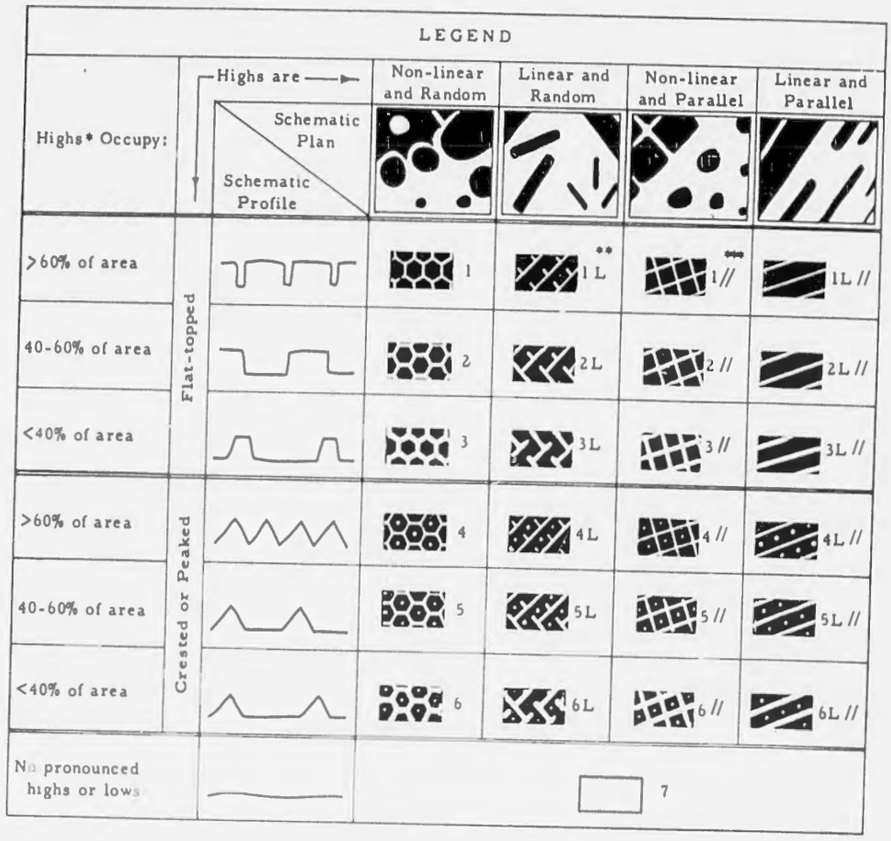

RELIEF IN FEET (LOG SCALE)
1000
400
100
100
50
10
⑥
⑤
④
③
②
< 10%
> 10%
CHARACTERISTIC SLOPE
CHARACTERISTIC RELIEF
⑥ PRECIPITOUS
100%
⑤ STEEP
50%
45°
26½°
④ DECLIVITOUS
25%
14°
③ MODERATE
10%
6°
2°
② GENTLE
3½%
① A&B FLAT
CHARACTERISTIC SLOPE
NUMBER PER 10 MILES
⑤ 50 TO 250
④ 10 TO 50
③ 2 TO 10
②
①
.04
.2
1
5
50
SPACING IN MILES (LOG SCALE)
OCCURRENCE OF SLOPES STEEPER THAN 50%

FIGURE 1

KOLB

CHARACTERISTIC PLAN-PROFILE

The characteristic plan-profile is the most commonly found plan-profile within a region. It may be either gross or restrictive. A gross plan-profile is one that can be subdivided into two restrictive component plan-profiles each exhibiting relief of a lower order than the gross plan-profile. Random sampling with circles 35 miles in diameter is used in determining the gross profile. Random sampling with circles 1 mile in diameter is used to determine the restrictive plan-profile. Local relief of less than 10 feet is not considered.



REPRESENTATIVE PLAN-PROFILES

Each of the following block diagrams illustrates a landscape representative of a specific plan-profile type. It should be emphasized that, within the defined limits of each type, a wide variety of landscape configurations are possible.

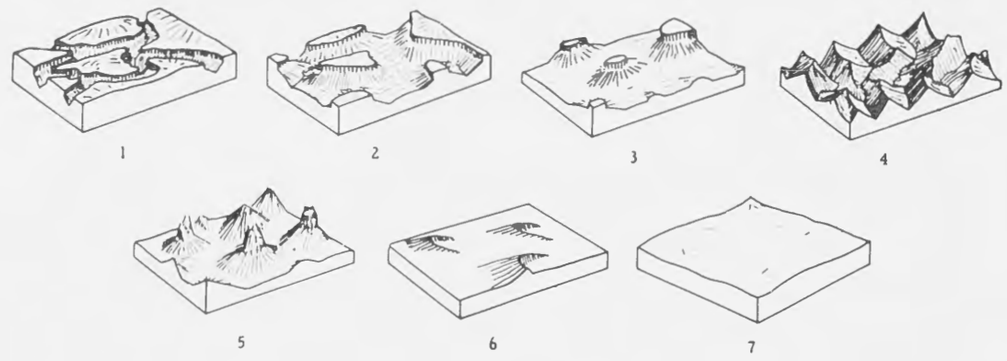
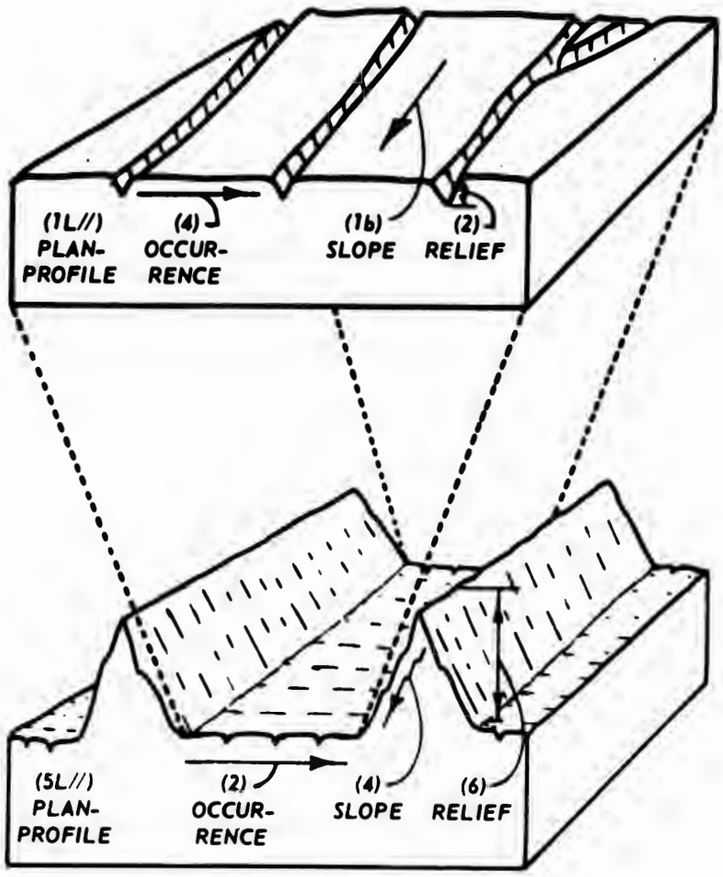


FIGURE 2

COMPONENT LANDSCAPE

A PLAIN WITH A 1 TO 3.5% SLOPE DISSECTED BY ROUGHLY PARALLEL WASHES FROM 10 TO 50 FT DEEP, SPACED FROM 1000 TO 5000 FT APART



GROSS LANDSCAPE

A PARALLEL RIDGE AREA WITH THE RIDGES FROM 2 TO 10 MILES APART, THEIR HEIGHT RANGING BETWEEN 400 AND 1000 FT, AND THEIR CHARACTERISTIC SLOPE BETWEEN 25 AND 50%

FIGURE 3

STUDIES OF NONDESTRUCTIVE TESTING TECHNIQUES TO CONTROL PROPERTIES OF CASE HARDENED STEEL COMPONENTS

R. D. KORYTOSKI, H. P. HATCH, and E. H. ABBE
SPRINGFIELD ARMORY
SPRINGFIELD, MASSACHUSETTS

In modern practice, the process of carburizing is applied to ferrous material with the object of increasing carbon content of the surface and in some cases the nitrogen content as well, so that when subjected to suitable heat treatment the surface portion thus carburized will be substantially harder and more wear resistant than the underlying metal. This combination of hard surface and softer interior is of inestimable value in modern engineering practice. By the proper selection and treatment of a carburizing alloy steel, a wide band of strength and toughness in the core can be combined with extreme surface hardness, resulting in a composite structure capable of withstanding certain kinds of stress to a high degree. Two such applications are the use of AISI 8620 resulphurized steel in the manufacture of M14 Rifle receivers and bolts (Figure 1).

STATEMENT OF PROBLEM

AISI 8620 resulphurized steel was selected in these applications as a compromise between strength and toughness as obtained with more highly alloyed steels and facility in machining and other production considerations as with plain carbon steels. However, a problem exists in achieving a homogeneous core structure to give the proper combination of strength and toughness. Axiomatically, construction alloy steels exhibit best strength and toughness properties when a fully quenched and tempered structure is developed. Reference to the isothermal transformation diagram for 8620 steel (Figure 2) indicates that this material is not capable of being fully quenched to eliminate undesirable transformation products, except in extremely small sections. The curve displays that an extremely rapid quench is required to obtain only 50% transformation of austenite to martensite.

The need for controls on M14 receivers and bolts was evident from results of evaluations made on the metallurgical characteristics of components involved in catastrophic field failures. Preliminary metallurgical examinations revealed that two major factors contributed to failures: the use of improper receiver material and inadequate heat treatment of bolts. Subsequent component investigations and observations of process methods employed by contractors indicated a most complex problem. Both gas and liquid carburizing methods were used in manufacture. Many process variables were present because of differences in equipment and heat treat practices at different plants. Process variables included differences in carburizing time and temperature, type and speed of quenching oil, degree of oil agitation, and choice of tempering temperature to meet hardness requirements. These conditions necessitated the development of nondestructive tests to segregate improper steel and to evaluate material core properties.

Improper steel, that of AISI 1330 composition instead of specified 8620, was segregated by well known and often used magnetic comparison permeability tests. This paper treats only of studies made in evaluation of material core properties of each component.

APPROACH TO PROBLEM

To clarify acceptance criteria, component specifications were reviewed and core hardness and structure requirements imposed in addition to the case requirements specified. Requirements which formed the basis for nondestructive tests are shown in Table I.

Table I. M14 Rifle Bolt & Receiver Specification Requirements.

	Bolt	Receiver
Case Hardness	Rc 54-59	Rd 61-69
Core Hardness	Rc 35-42	Rc 31-42
Case Depth	.015" - .020"	.012" - .018"
Temper Temperature	425°F MAX	450°F MAX
Core Structure	10% FERRITE	10% FERRITE

Many varied studies were undertaken in attempt to evaluate core characteristics. Measurements were made of basic magnetic properties of retentivity, flux density, and coercive force. Permeability tests were conducted comparing amplitude of the 60 cps. fundamental wave, amplitude of the third harmonic wave, and wave form distortion. In addition, the applicability of using mechanical hardness tests was studied. In conjunction with nondestructive tests, a large number of components were destructively examined to obtain core structure, core hardness, and case depth data. Each investigation required additional destructive examinations because once sectioned, the components with known properties could not be used.

MAGNETIC NONDESTRUCTIVE TEST STUDIES

Retentivity Measurements.

Retentivity, the flux density remaining after a magnetizing current sufficient to cause saturation has been applied and reduced to zero, was measured with laboratory equipment comprising a d-c power supply, a magnetizing coil and a gaussmeter. The arrangement used is shown in Figure 3. In additon, the opposing current required to bring the retentivity to zero was recorded. Measurements were extremely sensitive to probe position.

A graph of results (Figure 4) on specimens heat treated as a group indicated good correlation of opposing current with Rockwell C core hardness. Additional data gathered on material from various heat treat processes did not plot linearly. Numerous results did not fall on the curve. Components of Rockwell C44 - 46 hardness which had been retempered gave opposing current values of 350 - 370 milliamperes. Bolts which had not been retempered and with hardness Rockwell C37 - 40 measured 400 - 410 milliamperes. A group of components retempered at 530°F with core hardness Rockwell C41 - 43 had opposing current data in the vicinity of 265 milliamperes. It was thus concluded that the measurement was more affected by tempering temperature than core hardness.

A. C. Hysteresis Loop Measurements.

Alternating current hysteresis loop measurements were next made in an effort to correlate core hardness. A 60 cps. generator was used to induce an alternating current field in a "U" shaped, high permeability laminated core. A secondary winding encircled the component which was placed against the pole faces of the laminated core, thus completing the magnetic circuit. By means of a phase shifting network, a hysteresis loop was displayed on an oscilloscope. The experimental arrangement is shown in Figure 5.

Coercive force (H_c) and residual flux density (B_r) measurements were recorded, first at saturation flux density, then at a constant flux density below saturation. Largest changes were observed in measurements of coercive force. The greatest sensitivity was obtained using constant flux density below saturation.

Within individual lots of components, coercive force tended to increase with increased core hardness. However, as with retentivity measurements, it was shown that coercive force was more affected by tempering temperature than core hardness.

Magnetic Permeability Measurements.

Magnetic comparator type equipment (Figure 6) was used in magnetic permeability measurements. Basically, the equipment

contained a 60 cps generator, a pair of similar coils, an amplifier, and filter and detector circuits for indicating the resultant coil output voltage. Each of the similar coil units contained a primary and secondary winding. The primary winding of each coil applied an alternating current magnetizing field to the sample placed in the coil. The secondary windings were connected in series opposition so that only the difference voltage between the two secondaries was measured. In the actual test, a reference component was placed in one coil where it remained throughout the test. Components being compared were inserted in the remaining coil. The equipment was calibrated in such a way that increased permeability was shown by meter indications moving in the negative direction.

Studies indicated that many variables within components affected permeability. The fundamental 60 cycle wave was shown to be influenced by surface hardness, core hardness, core structure, and degree of tempering and retempering.

Surface hardness was considered a contributing variable because flux density is at a maximum on the surface of the component and decreases in an exponential manner into the core. The effect of surface hardness was studied through comparison of readings on components of essentially the same core hardness and structure but different in surface hardness (Table 2).

Table 2. Effect of Surface Hardness on Magnetic Permeability.

Magnetic Permeability Readings	Surface Hardness Rc	Core Hardness Rc
+25	62	40
-31	58	40
-62	55.5	40

Datum shows that a 6.5 point variance in surface hardness produced an 87 point change in reading. It also indicates that permeability increased with decreasing surface hardness.

Results on a similar study (Table 3) wherein surface hardness was maintained constant but core hardness was varied, showed that core hardness also influenced readings. Permeability increased with decreasing core hardness.

Table 3. Effect of Core Hardness on Magnetic Permeability.

Magnetic Permeability Readings	Surface Hardness Rc	Core Hardness Rc
-3	56	43.5
-20	56	38.5
-62	56	36.5

Core hardness measurements and metallurgical examinations of core structures indicated that decreasing hardness was accompanied by increasing percentages of ferrite and upper bainite transformation products, with the latter producing the greatest affect on hardness. As such, permeability was also affected by core structure.

Tempering or retempering components at relatively high temperatures produced the greatest changes in magnetic permeability. Table 4 shows the increase in permeability due to retempering. On initial retemper at 400°F, test readings were changed 8 to 15 points in the negative direction; a second retemper at 400°F altered readings 3 to 5 points further negative. A retemper at slightly higher temperature (425°F) moved readings 8 to 12 points more in the negative direction. Readings changed radically to minus off scale when retempered at 500°F. Two additional components retempered directly at 500°F gave minus off scale readings.

Table 4. Effect of Retempering.

Component Condition	Tempering	Magnetic Permeability Readings
Unparkerized		+30
	Retemper 1 hr @ 400°F	+15
	Retemper 1 hr @ 400°F	+10
	Retemper 1 hr @ 425°F	- 2
	Retemper 1 hr @ 500°F	-105
Parkerized		- 9
	Retemper 1 hr @ 400°F	-16
	Retemper 1 hr @ 400°F	-19
	Retemper 1 hr @ 425°F	-28
	Retemper 1 hr @ 500°F	- Off Scale
Parkerized		- 7
	Retemper 1 hr @ 500°F	- Off Scale
Parkerized		+ 2
	Retemper 1 hr @ 500°F	- Off Scale

Third Harmonic Amplitude Measurements.

Through use of a band pass filter and circuit, third harmonic amplitude measurements were made simultaneously with measurements of the fundamental wave. Calculations were made relating the amplitude of the 60 cps. wave to that of the 180 cps. wave.

Results of studies indicated a tendency for third harmonic amplitude to increase with a decrease in core hardness within individual lots or groups of components from each contractor. On combining data from all contractors, no direct correlation was shown to exist. Variables which affected fundamental wave measurements, likewise influenced third harmonic wave measurements, with the chief deterrent to correlation with core hardness that of degree of tempering and retempering.

Within individual heats, the ratio of amplitudes offered somewhat better correlation with core hardness than measurement with either wave amplitude alone. Highest ratios were obtained on components with lowest core hardness. Whereas 60 cps. and 180 cps. measurements were noticeably affected by retempering at a higher temperature, the ratio was not appreciably affected. As such, within individual processes, use of the ratio might possibly be used to evaluate core hardness regardless of tempering treatment.

Wave Form Pattern Studies.

During tests with the fundamental wave, it was observed that similar characteristic wave form patterns recurred. Three basic patterns were apparent. A pattern designated #1 displayed an undistorted sine wave; pattern #2 had third harmonic distortion; and pattern #3 indicated a marked distortion of the fundamental wave on the negative portion. Core hardness measurements and microstructure examinations of the core structure revealed that basic patterns offered some correlation with core hardness. Components with patterns #1 and #2 had core hardness within specification; components which displayed pattern #3 consistently had core hardness above specification requirements.

Approximately 150 oscillograms were taken of different degrees of distortion. Corresponding components were sectioned and metallurgically examined. Six obvious patterns (Figure 7) were correlated with material hardness and structure.

TEST METHOD EMPLOYED IN EVALUATION OF PROPERTIES

Based on results of studies, a combination test method employing Rockwell C surface hardness measurements, magnetic permeability readings and oscilloscope wave form patterns was developed. A Rockwell C hardness measurement was first taken.

The component was next subjected to magnetic test wherein amplitude was noted. Reference was then made to a chart (Figure 8) relating surface hardness and magnetic reading to determine expected core hardness without regard to wave form pattern. Finally, wave form pattern was observed and the combination evaluated.

By application of these procedures approximately 35,000 components were evaluated as to specification requirements. Components were segregated according to core hardness, core structure and degree of tempering. It was possible to determine the approximate percentage of free ferrite within the structure.

HARDNESS NONDESTRUCTIVE TEST STUDIES

In addition to magnetic nondestructive test studies, hardness investigations were made in attempt to evaluate material core hardness. Initial plots of superficial surface hardness measurements versus core hardness indicated no significant correlation. Direct Rockwell C measurements, influenced in some degree by the core, showed only slight correlation with extremely wide scatter. Plots of surface Rockwell C & D hardness versus magnetic permeability data, however, appeared to vary in some relation to core hardness. Decreasing hardness appeared to show somewhat a softer core condition. As a result, many different combinations of data were analyzed to determine whether a hardness method could be developed for prediction of core hardness for surface measurements.

The first combination which showed promise was a three-way plot of surface Rockwell C hardness in a component thin section, surface Rockwell C hardness on a heavy section and core Rockwell C hardness in the critical area. Data accumulated on 161 components resulted in a breakdown as shown in Table 5.

Table 5. Core Hardness Prediction Results Based on Rockwell C Hardness Measurements

Core Hardness Prediction	Prediction Accuracy
Within ± 2 Rc	83.5%
Within ± 3 Rc	93.5%
Within ± 4 Rc	97.5%
Within + 5 Rc	100%

Emphasis was next shifted to a method in which slope of hardness within components was studied. Method was well based in that hardness in any particular section decreases from the surface into the core and the slope of this hardness versus depth varies with hardenability and quenching rate. Measurements were made of surface hardness in one selected area. Two different loads, Rockwell D (100 kg) and Rockwell C (150 kg) and therefore, two penetration depths were employed. Results (Table 6) were most encouraging.

Table 6. Core Hardness Prediction Results Based on Rockwell C and D Hardness Measurements

Core Hardness Prediction	Prediction Accuracy
Within + 2 Rc	89%
Within + 3 Rc	97%
Within + 4 Rc	100%

A mathematical determination of this preliminary data took the following form:

Core Hardness = 2 (Rc - Rd + k) where factor k equalled 32.5. This factor apparently depends on section size and may vary with extreme changes in thermal process. Additional information is required to confirm these conjectures.

With data accumulated, a chart (Figure 9) was developed to predict core hardness of production components sectioned daily and metallurgically examined. At present, 170 components fabricated from six different heat lots have been evaluated with a more promising correlation (Table 7) indicated.

Table 7. Core Hardness Prediction Results on Production Components

Core Hardness Prediction	Prediction Accuracy
Within ± 2 Rc	93.5%
Within ± 3 Rc	99%
Within ± 4 Rc	100%

Approximately 60% of the components within the ±2 Rc group, were exactly predicted. Because predictions made related to a slightly different area than that sectioned and examined, it is felt that a more accurate determination would be obtained by taking measurements within the area sectioned, if possible.

SUMMARY

In summation, it is felt that adequate methods have been developed to evaluate core hardness and structure properties of M14 receiver and bolt components employing magnetic and hardness non-destructive tests. The techniques employed have application to the inspection of other case hardened components.

Material core evaluations required a combination of tests. One related a Rockwell C surface hardness measurement, magnetic permeability and wave form patterns. The other was concerned with the slope of the hardness gradient within selected sections and

required Rockwell C and D surface hardness measurements in the same section. Magnetic nondestructive methods did not correlate directly with component material properties on various heat lots from different contractors, although correlation of retentivity, coercive force and magnetic permeability with core properties was indicated within individual heats. Tempering time and temperature affected magnetic results to greatest extent and represents the chief variable opposing correlation of various magnetic tests with core hardness and structure.

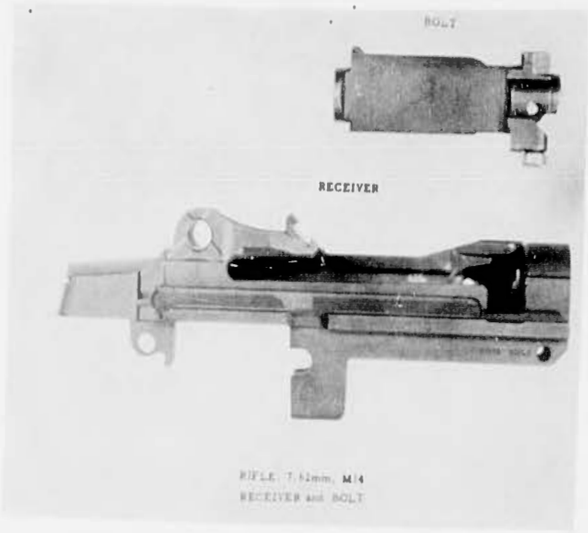


Figure 1

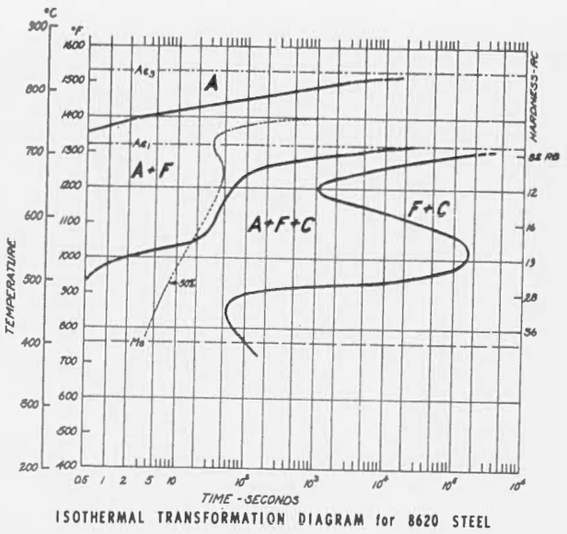
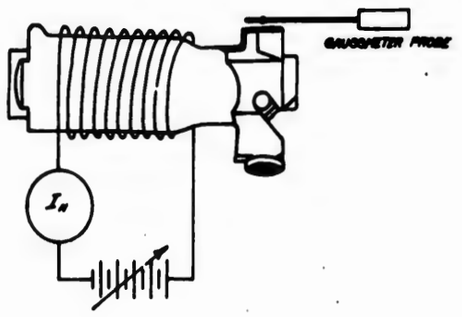
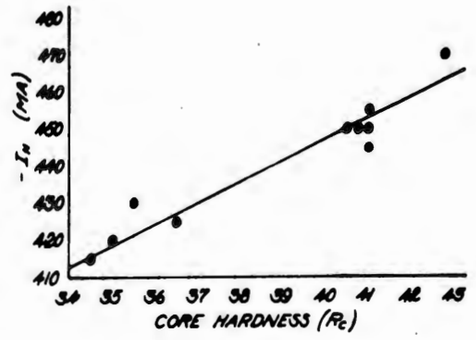


Figure 2



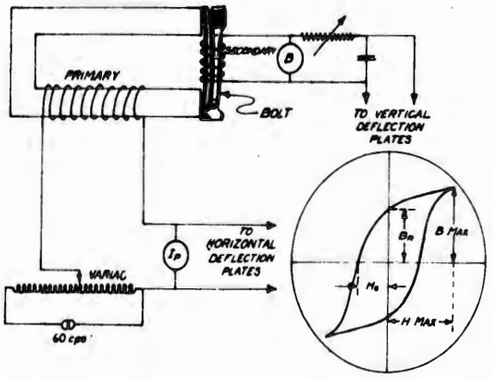
ARRANGEMENT for RETENTIVITY MEASUREMENTS

Figure 3



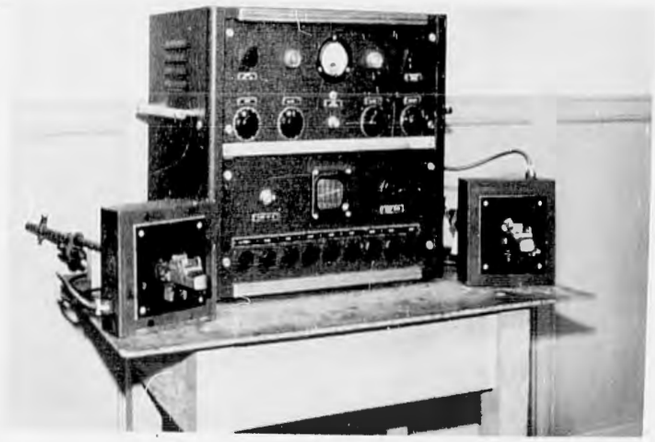
RETENTIVITY vs CORE HARDNESS MEASUREMENTS

Figure 4



EXPERIMENTAL ARRANGEMENT for A.C. HYSTERESIS LOOP MEASUREMENTS

Figure 5



MAGNETIC COMPARATOR-TYPE EQUIPMENT
Figure 6

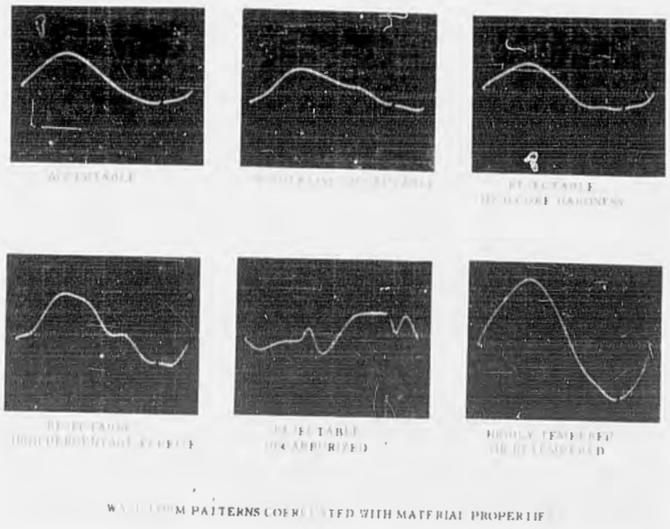
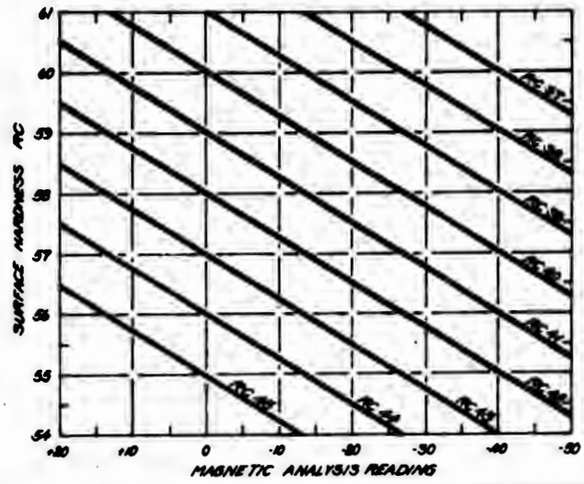
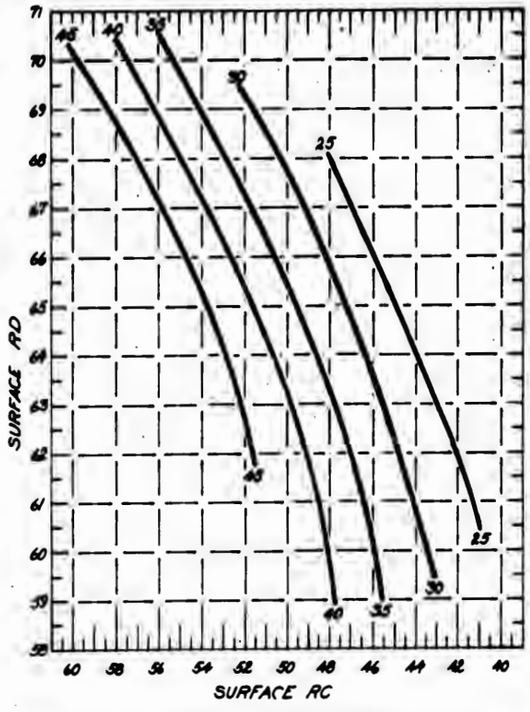


Figure 7



CORE HARDNESS PREDICTION CHART BASED on ROCKWELL C HARDNESS MEASUREMENTS

Figure 8



CORE HARDNESS PREDICTION CHART
BASED ON ROCKWELL C & D HARDNESS MEASUREMENTS

Figure 9

DISTRIBUTION LIST

DEPARTMENT OF DEFENSE

Director of Defense, Research & Engineering Washington 25, D.C.	1
Armed Services Technical Information Agency Arlington Hall Station Arlington 12, Virginia	10
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: Research Support Division	172

DEPARTMENT OF THE ARMY (Continued)

Chief US Army R&D Liaison Group (9983) (Europe) APO 757 New York, New York	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 R&D Group (Far East) (9984) APO 343 San Francisco, California	1
Commanding Officer US Army Personnel Research Office Washington 25, D.C.	1
Commanding Officer Army Research Office (Durham) Box CM, Duke Station Durham, North Carolina	1
Commanding General US Army Chemical Corps Materiel Command Army Chemical Center, Maryland	1
Commanding Officer US Army Biological Laboratories Fort Detrick, Maryland	1

Chemical Corps (Continued)

Commanding Officer — 1
US Army Chemical Corps Engineering Command
Army Chemical Center, Maryland

Commanding Officer — 1
US Army Chemical Warfare Laboratories
Army Chemical Center, Maryland

Commanding Officer — 1
US Army Chemical Nuclear Defense Laboratory
Army Chemical Center, Maryland

Commanding Officer — 1
US Army Chemical Corps Training Command
Fort McClellan, Alabama

Commanding Officer — 1
US Army Chemical Corps Proving Ground
Dugway Proving Ground
Dugway, Utah

Commanding Officer — 1
Rocky Mountain Arsenal
Denver 2, Colorado

Commanding Officer — 1
Pine Bluff Arsenal
Pine Bluff, Arkansas

Commanding Officer — 1
US Army Chemical Corps Intelligence Agency
Arlington Hall Station
Arlington 12, Virginia

Corps of Engineers

Chief of Engineers — 1
United States Army
Washington 25, D.C.

Commanding General — 1
US Army Engineer Division
Ohio River Division Laboratories
Cincinnati 27, Ohio

Director — 1
Cold Regions R&E Laboratory
Corps of Engineers
Hanover, N.H.

Director — 1
US Army Waterways Experiment Station
P.O. Box 631
Vicksburg, Mississippi

Director — 1
Engineer R&D Laboratories
Fort Belvoir, Virginia
Attn: Technical Documents Center

Corps of Engineers (Continued)

Director
US Army Engineer Geodesy Intelligence
and Mapping R&D Agency
Fort Belvoir, Virginia — 1

Commanding Officer
US Army Engineer (USAE) Experiment Station
Jackson Installation
P.O. Drawer 2131
Jackson, Mississippi — 1

Commanding Officer
US Army CRREL Keweenaw Field Station
Corps of Engineers
Houghton, Michigan — 1

Army Medical Service

The Surgeon General
United States Army
Washington 25, D.C. — 1

Commanding General
US Army Medical R&D Command
Washington 25, D.C. — 1

Commanding Officer
US Army Research Institute of Environmental Medicine
Natick, Massachusetts — 1

Commanding Officer
US Army Medical Research Laboratory
Fort Knox, Kentucky — 1

Commanding Officer
US Army Medical Research & Nutrition Laboratory
Fitzsimons Army Hospital
Denver 30, Colorado — 1

Commanding Officer
US Army Surgical Research Unit
Brooke Army Medical Center
Fort Sam Houston, Texas — 1

Commanding Officer
Medical Equipment Development Laboratory
Fort Totten, New York — 1

Director
Walter Reed Army Institute of Research
Walter Reed Army Medical Center
Washington 12, D.C. — 1

Commanding General
US Army Ballistic Missile Agency
Redstone Arsenal, Alabama 1

Commanding General
Aberdeen Proving Ground
Aberdeen, Maryland 1

Commanding General
Aberdeen Proving Ground
Aberdeen, Maryland
Attn: Human Engineering Laboratory 1

Commanding General
Army Ordnance Missile Command
Redstone Arsenal, Alabama 1

Commanding General
US Army Rocket & Guided Missile Agency
Redstone Arsenal, Alabama 1

Commanding General
Army Ordnance Missile Command
Redstone Arsenal, Alabama
Attn: ORDXM-XGR 1

Commanding Officer
Diamond Ordnance Fuze Laboratory
Connecticut Ave., & Van Ness Sts., N.W.
Washington 25, D.C. 1

Commanding General
Frankfort Arsenal
Bridge & Tacony Streets
Philadelphia 37, Penna. 1

Commanding General
Ordnance Materials Research Office
Watertown Arsenal
Watertown 72, Mass. 1

Commanding General
Ordnance Tank-Automotive Command
Detroit Arsenal, Center Line, Michigan 1

Commanding General
White Sands Proving Ground
New Mexico
Attn: Technical Library 1

Commanding Officer Picatinny Arsenal Dover, New Jersey	1
Commanding Officer Rock Island Arsenal Rock Island, Illinois	1
Commanding Officer Springfield Armory Springfield, Mass.	1
Commanding Officer Watertown Arsenal Watertown 72, Mass.	1
Commanding Officer Watervliet Arsenal Watervliet, New York	1
Commanding Officer Headquarters, Ordnance Weapons Command Rock Island, Illinois	1
Commanding Officer US Army Ordnance Special Weapons Ammunition Command Dover, New Jersey	1
Institute for Research, 1010 Frankfort Arsenal Philadelphia 37, Penna.	1
Commanding General Quartermaster R&E Command Natick, Mass.	1
Commanding Officer Quartermaster Food & Container Institute for the Armed Forces Chicago, Illinois	1
Commanding Officer Quartermaster Field Evaluation Agency Fort Lee, Va.	1

Signal Corps

Chief Signal Officer United States Army Washington 25, D.C.	1
Commanding General US Army Combat Surveillance Agency 1124 N. Highland Street Arlington, Virginia	1
Commanding General US Army Electronic Proving Ground Fort Huachuca, Arizona Attn: Library	1
Commanding General US Army Signal R&D Laboratories Fort Monmouth, New Jersey	1
Commanding Officer US Army White Sands Signal Agency White Sands Proving Ground New Mexico	1
XS, USASRDL Evans Area Belmar, New Jersey	1

Transportation Corps

The Chief of Transportation United States Army Washington 25, D.C.	1
Commanding Officer US Army Transportation Research Command Fort Eustis, Virginia	1
President US Army Transportation Board Fort Eustis, Virginia	1
Commanding General US Army Transportation Materiel Command St. Louis 66, Missouri	1
Commanding Officer US Army Transportation Combat Development Agency Fort Eustis, Virginia	1

Other Army Agencies

Army War College, Library-B-205 Carlisle Barracks Carlisle, Pennsylvania	1
Command and General Staff College Fort Leavenworth, Kansas Attn: Library	1
Commanding General US Continental Army Command Fort Monroe, Virginia	1
US Continental Army Command US Combat Development Experimentation Center Fort Ord, California	1
Office of Special Weapons Development Fort Bliss, Texas	1
Combat Operations Research Group Headquarters, Continental Army Command Fort Monroe, Virginia	1
President US Army Air Defense Board Fort Bliss, Texas	1
Combat Developments Office US Army Infantry School Fort Benning, Georgia	1
Superintendent US Military Academy West Point, New York	1

DEPARTMENT OF THE NAVY

Chief of Naval Operations Washington 25, D.C.	1
Chief of Naval Research Office of Naval Research (Code 407) Washington 25, D.C.	1
Director US Naval Research Laboratories Washington 25, D.C.	1
Director Weapons Systems Analysis Division Bureau of Naval Weapons Washington 25, D.C.	1

DEPARTMENT OF THE AIR FORCE

Headquarters, US Air Force (AFCOA) Washington 25, D.C.	1
Commander Air Force Systems Command Andrews Air Force Base Washington 25, D.C.	1

DEPARTMENT OF THE AIR FORCE (Continued)

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

National Institutes of Health 1
Bethesda, Maryland
Attn: Director

National Science Foundation 1
Washington 25, D.C.

The Library of Congress 1
Washington 25, D.C.

OTHER AGENCIES

British Joint Services Mission 1
Thru: Foreign Liaison Office
Assistant Chief of Staff, Intelligence
Department of the Army
Washington 25, D.C.

Canadian Joint Staff 1
Thru: Foreign Liaison Office
Assistant Chief of Staff, Intelligence
Department of the Army
Washington 25, D.C.

OTHER AGENCIES (Continued)

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, US Army University of Wisconsin Madison, Wisconsin	1

UNCLASSIFIED

UNCLASSIFIED